

Original Paper

Predicting the Risk of Total Hip Replacement by Using A Deep Learning Algorithm on Plain Pelvic Radiographs: Diagnostic Study

Chih-Chi Chen¹, MD; Cheng-Ta Wu², MD; Carl P C Chen¹, MD, PhD; Chia-Ying Chung¹, MD; Shann-Ching Chen³, PhD; Mel S Lee⁴, MD, PhD; Chi-Tung Cheng^{5*}, MD, PhD; Chien-Hung Liao^{5*}, MD

¹Department of Physical Medicine and Rehabilitation, Chang Gung Memorial Hospital, Taoyuan, Taiwan

²Department of Orthopaedic Surgery, Chang Gung Memorial Hospital, Kaohsiung, Taiwan

³Compal Electronics, Smart Device Business Group, Taipei, Taiwan

⁴Department of Orthopaedic Surgery, Pao-Chien Hospital, Pingtung, Taiwan

⁵Department of Trauma and Emergency Surgery, Chang Gung Memorial Hospital, Taoyuan City, Taoyuan, Taiwan

*these authors contributed equally

Corresponding Author:

Chi-Tung Cheng, MD, PhD

Department of Trauma and Emergency Surgery

Chang Gung Memorial Hospital

Trauma Center, 5, Fuxin Street, Kweishiang District

Taoyuan City, Taoyuan, 333

Taiwan

Phone: 886 3 3281200 ext 3651

Email: atong89130@gmail.com

Abstract

Background: Total hip replacement (THR) is considered the gold standard of treatment for refractory degenerative hip disorders. Identifying patients who should receive THR in the short term is important. Some conservative treatments, such as intra-articular injection administered a few months before THR, may result in higher odds of arthroplasty infection. Delayed THR after functional deterioration may result in poorer outcomes and longer waiting times for those who have been flagged as needing THR. Deep learning (DL) in medical imaging applications has recently obtained significant breakthroughs. However, the use of DL in practical wayfinding, such as short-term THR prediction, is still lacking.

Objective: In this study, we will propose a DL-based assistant system for patients with pelvic radiographs to identify the need for THR within 3 months.

Methods: We developed a convolutional neural network-based DL algorithm to analyze pelvic radiographs, predict the hip region of interest (ROI), and determine whether or not THR is required. The data set was collected from August 2008 to December 2017. The images included 3013 surgical hip ROIs that had undergone THR and 1630 nonsurgical hip ROIs. The images were split, using split-sample validation, into training (n=3903, 80%), validation (n=476, 10%), and testing (n=475, 10%) sets to evaluate the algorithm performance.

Results: The algorithm, called SurgHipNet, yielded an area under the receiver operating characteristic curve of 0.994 (95% CI 0.990-0.998). The accuracy, sensitivity, specificity, and F_1 -score of the model were 0.977, 0.920, 0.932, and 0.944, respectively.

Conclusions: The proposed approach has demonstrated that SurgHipNet shows the ability and potential to provide efficient support in clinical decision-making; it can assist physicians in promptly determining the optimal timing for THR.

(JMIR Form Res 2023;7:e42788) doi: [10.2196/42788](https://doi.org/10.2196/42788)

KEYWORDS

osteoarthritis; orthopedic procedure; artificial intelligence; AI; deep learning; machine learning; orthopedic; pelvic; radiograph; predict; hip replacement; surgery; convolutional neural network; CNN; algorithm; surgical; medical image; medical imaging

Introduction

Deteriorating hip joint disorder causes disability worldwide, and hip osteoarthritis alone ranks as the 11th highest contributor [1-3]. Total hip replacement (THR) [4] is considered the international gold standard for treating refractory degenerative and rheumatologic hip disorders [5]. The indication criteria for THR in guidelines still appear unclear, mentioning domains like pain, function, radiologic change, and the limited effectiveness of nonoperative therapy, but they lack specific cut-off values or ranges [6,7]. THR has been considered one of the most cost-effective and successful orthopedic interventions currently available [8,9], which can improve pain and function in patients for whom it is indicated [10]. More than 1 million THRs are performed each year [11]. Identifying patients who are more likely to receive THR in the short term is of interest, as pain and self-reported functional status in patients might not deteriorate with a shorter wait time for surgery [12].

The timing of the decision to refer patients for surgical consultation is vital for primary health providers. Better preoperative hip function in patients is indicative of better postoperative functional outcomes [13], and a delayed operation could result in poorer outcomes [14]. The decision and timing of referring a patient to a hip specialist are important for the general practitioner. Referring patients too early is not recommended for those patients who respond well to conservative treatment, as it could result in longer waiting times for the patients who truly require THR referrals. A late referral might also delay the timing of the operation, affecting the surgical outcome and patient prognosis, and it could decrease patients' quality of life [15,16]. Additionally, conservative treatments like intra-articular corticosteroid injection may result in higher odds of prosthetic joint infection for patients who are scheduled to receive THR within 3 months or even longer [17,18]. Therefore, identifying patients who should receive operations in the short term is crucial for primary health providers [19].

Anteroposterior pelvic radiography is the most commonly used primary diagnostic tool for assessing hip joint conditions, such as fractures, osteonecrosis, and degenerative disorders [20,21]. These assessments provide critical information to guide and offer a prognosis as well as treatment options [22]. A recent systematic review [23] suggested that information obtained from PXR, including higher Kellgren and Lawrence grades, superolateral femoral head migration, and subchondral sclerosis, were predictive of a faster progression to THR [23]. However, accurate recognition of these features needs expertise and is time-consuming. An automated detection model for predicting THR within a short-term period has potential benefits, including increasing efficiency, reducing delayed referrals, and improving patient outcomes.

Significant breakthroughs in medical imaging applications have been achieved in the field of deep learning (DL), improving the speed of diagnosis and decision-making with comparable efficacy to that of professional clinicians [24-28]. DL has been used to detect lesions in pathologic images, electrocardiography,

and retinography [29-32]. In the field of hip studies, DL techniques have been applied to detecting hip fractures [33], diagnosing osteoarthritis [34], grading the severity of osteoarthritis [35], and predicting patient-specific payment models [36,37].

Artificial intelligence (AI) has been suggested not only for aiding in diagnosis prediction but also for practical wayfinding in clinical decision-making [38]. Although the advancement of medical AI in the orthopedic field is on the rise, DL applications in practical wayfinding, such as short-term THR prediction, are still lacking. In this study, we use a DL-based framework to develop SurgHipNet, a fully automated diagnosis assistance system, using PXR to obtain THR prediction within 3 months. The goal of this study was to develop and validate a DL model that automatically extracts radiographic features of the hips and assists primary health providers in identifying patients who need THR in the short term for early referral.

Methods

Data Acquisition

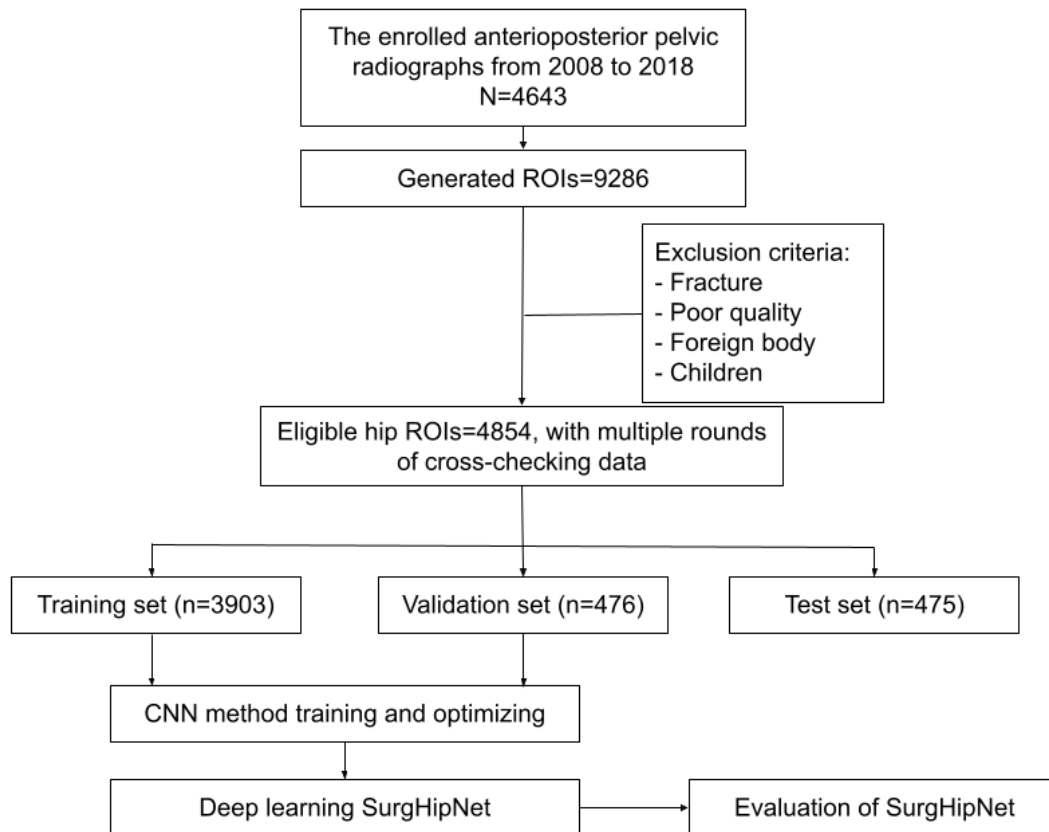
We extracted the PXR and surgical reports of all patients who underwent THR in Linkou Chang Gung Memorial Hospital (CGMH) from August 2008 to December 2017. This data collection consisted of 3013 PXR within 3 months before THR surgery (THR group). To identify the non-THR group, we used the trauma data bank in CGMH to search for patients who had PXR during the same time period as the THR group. From these records, we identified 1630 individuals who did not receive any hip surgery at our hospital within the following 3 months, as confirmed by their medical records. These individuals were included in our study as the non-THR group. We excluded images depicting previous hip surgery with foreign bodies, hip fractures, and pediatric skeletal images, as well as those with poor image quality.

Ethical Considerations

For deidentification purposes, each image file name was converted to a serial number, and the upper and lower sections of the images containing patient privacy information, such as names, medical record numbers, and date of birth, were cropped out. For the fully anonymized reports and radiographs used in this study, the requirement to obtain informed consent was waived. The Internal Review Board of CGMH approved this study (201801784B0).

Data Annotation

The final data set consisted of 4854 hip joints, using 4643 weight-bearing anterior-posterior PXR from 4643 participants (Figure 1). Hip regions of interest (ROI) were split into training (n=3903, 80%), validation (n=476, 10%), and independent test (n=475, 10%) data sets. We used the training data set for model training and performed 20-fold cross-validation. Then we used the validation set to adjust the hyperparameters to elevate the performance and prevent overfitting. Lastly, we used the independent test data set (n=475) to validate the final performance.

Figure 1. Flowchart showing participant selection from the database. CNN: convolutional neural network; ROI: region of interest.

The images were initially labeled as THR or non-THR according to the surgical reports. For the THR group, we reviewed all surgical reports to ensure that patients had undergone THR and then acquired the corresponding images for training. For the non-THR group, we could only confirm that patients had not received THR within the subsequent 3 months at our hospital. For joint localization training, we developed a labelless practical framework to automatically detect hips. Bounding boxes were placed at the center of the femoral head. To annotate the hip ROI, we used 3 annotators trained to place square bounding boxes approximately centered at the femoral head or the artificial hip joint with customized graphical user interface software. All the labeled ROIs in the data set were visually reviewed by 1 physician with 15 years of clinical experience, and the ROI annotators used the same rules to annotate the data sets. The hip ROIs were cropped from PXR by single-shot detector-based algorithms. A preprocessing procedure including cropping of square regions and mirroring of the left hip ROI (a process of flipping the left hip ROI into mirrored left hip ROIs) was needed. The square hip ROI was resized to 224 224 pixels with an 8-bit grayscale color format to reduce complexity and computation. For nonsquare input radiographs, the image was padded to achieve a square size, with zero values added to the width or height to ensure that the convolution operation preserves the aspect ratio of the hip and pelvis shape in the radiograph. The detailed framework of the development method has been described in [Multimedia Appendix 1](#) and a previous study [39]. The hip ROIs were then inputted into a further network for classification.

SurgHipNet Architecture, Data Preprocessing, and Model Training

The SurgHipNet was designed based on ResNet 101 and pretrained using ImageNet. It is used with PyTorch v0.4 and fastai API 2020 implementation and runs on CUDA 9.0 within the Ubuntu 16.04 operating system, powered by 1 Nvidia Tesla V100 graphic processing unit. For model training, we used Adam optimizer and focal loss, with $\alpha=.5$ and $\gamma=5$. The training schedule used a cyclical learning rate and the 1-cycle policy, where 2 cycles were trained with the first 10 epochs on the last layer and 10 additional epochs on the last 2 layers. We used 1 Nvidia 1080Ti graphic processing unit on the Ubuntu 16.04 operating system.

A 20-fold cross-validation approach was used on the data set. We used the classification error rate as the loss function and Adam optimizer ($\beta_1=.9$ and $\beta_2=.99$) and added a final BatchNorm layer. We trained the model with minibatches of size 16 and used a cyclical learning rate strategy and 1-cycle policy, where we first trained the last layer with differential learning rates, using $\text{max_lr}=0.003$ for 10 cycles; we then trained all layers using the same learning rate setting for 10 cycles. The details of the comparative experiments are listed in [Multimedia Appendix 2](#). We augmented the data during training with the following fastai augmentation settings:

```

tfms = get_transforms(do_flip=False, flip_vert=False,
max_rotate=15, max_zoom=1.1, p_affine=0.5, max_lighting=0.5,
p_lighting=0.1, max_warp=0)
  
```

Algorithm Prediction Visualization

For the SurgHipNet models that predict the need for THR, we also used Gradient-weighted Class Activation Mapping (Grad-CAM) to generate a heatmap activated by the model for the pathologic areas of the hip joint space or femoral head. This heatmap provided evidence that the model accurately identified the potential pathologic sites. This method ensures that the algorithm detects and classifies based on the area around the femoral head, rather than other segmentations of the image, as previous experiences have shown [40].

SurgHipNet Evaluation and Statistical Analysis

All statistical analyses were carried out using R 3.6.3 (R Core Team) with the packages “pROC,” “tableone,” “caret,” and “ggplot2.” We reported overall accuracy, sensitivity, specificity,

positive and negative predictive value, F_1 -score, and the area under the receiver operating characteristic curve (AUC) along with the 95% CIs, which were estimated using bootstrapping with 2000 replicates to evaluate the performance of SurgHipNet.

Results

Data Characteristics

In this study, we used a total of 4854 hip ROIs, including 1138 ROIs with THR and 3716 ROIs without THR, to develop our algorithm. The characteristics and epidemiology of the images are presented in Table 1. We included all the ROIs and split the data set into training (n=3903, 80%), validation (n=476, 10%), and testing sets (n=475, 10%).

Table 1. Characteristics of etiology and grading of the data sets.

Group and data set etiology and grades	Training data set (n=3903), n (%)	Validation data set (n=476), n (%)	Testing data set (n=475), n (%)
THR^a group (n=1383)			
Avascular necrosis-1	16 (0.4)	3 (0.6)	4 (0.8)
Avascular necrosis-2	189 (4.8)	17 (3.6)	23 (4.8)
Avascular necrosis-3	232 (5.9)	36 (7.6)	26 (5.4)
Avascular necrosis-4	282 (7.2)	28 (5.9)	32 (6.7)
Avascular necrosis-5	2 (0.05)	N/A ^b	N/A
Osteoarthritis-1	8 (0.2)	7 (1.5)	1 (0.2)
Osteoarthritis-2	23 (0.6)	7 (1.5)	4 (0.8)
Osteoarthritis-3	70 (1.8)	6 (1.3)	4 (0.8)
Osteoarthritis-4	82 (2.1)	2 (0.4)	10 (2.1)
Osteoarthritis-5	3 (0.08)	1 (0.2)	1 (0.2)
Other etiologies	17 (0.4)	1 (0.2)	1 (0.2)
Non-THR group (n=3716)			
Normal	2246 (57.5)	283 (59.4)	269 (56.6)
Osteoarthritis-1	733 (18.8)	85 (17.9)	100 (21.1)

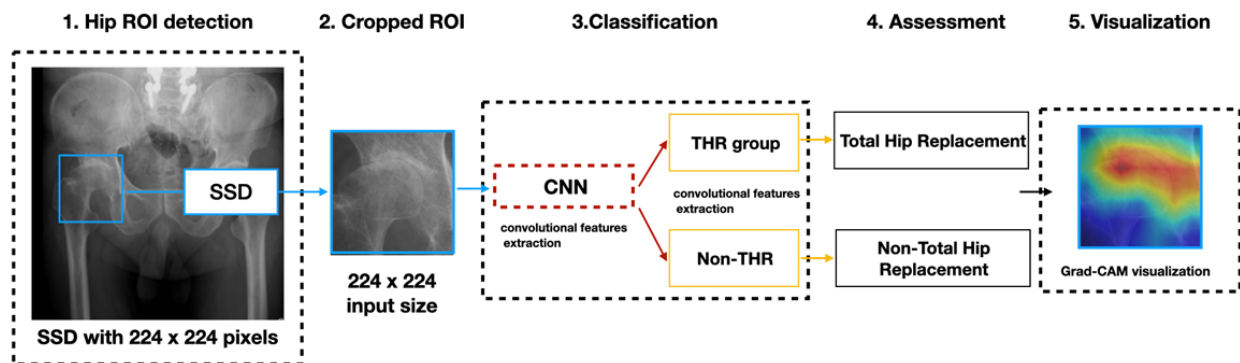
^aTHR: total hip replacement.

^bN/A: not applicable.

In the beginning, each hip region in a radiograph was labeled by an autonomic hip detection system [39]. The cropped ROI was used with a convolutional neural network–based algorithm for classification into the THR or non-THR groups. Furthermore,

the Grad-CAM technique was adapted for visualization. The deep neural network learning architecture is presented in Figure 2.

Figure 2. The SurgHipNet pipeline. The input images detected the hip region of interest (ROI) first, and then, we cropped the hip area as our target ROI. Then the cropped images were inputted into a convolutional neural network (CNN) to be further classified into total hip replacement (THR) or non-THR groups. Finally, a Gradient-weighted Class Activation Mapping (Grad-CAM) was applied for the final visualization. SSD: single-shot detector.



SurgHipNet Performance on the Independent Test Data Set

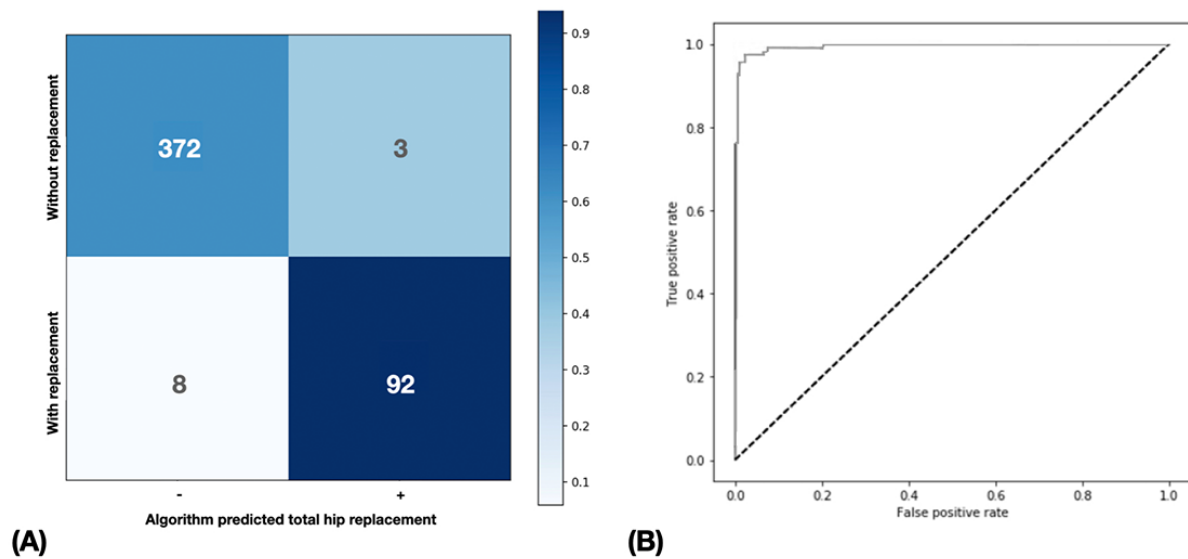
We used the isolated testing data set to evaluate the performance of SurgHipNet. It predicted the THR with an acceptable

performance; the overall accuracy, sensitivity, specificity, and AUC were 0.977, 0.9200, 0.992, and 0.994 (95% CI 0.990-0.998), respectively (Table 2; Figure 3).

Table 2. The performance of SurgHipNet classification results in the testing data set.

Features	Values
True positive	92
True negative	372
False positive	3
False negative	8
Accuracy	0.977
Prevalence	0.211
Sensitivity	0.920
Specificity	0.992
Positive predictive value	0.968
Negative predictive value	0.979
F_1 -score	0.944
Area under the receiver operating characteristic curve (95% CI)	0.994 (0.990-0.998)

Figure 3. (A) The confusion matrix plot of the SurgHipNet algorithm. (B) The receiver operator characteristic curve and the summary of classification results on other independent test data sets comprising 475 ROIs. The area under the receiver operator characteristic curve was 0.9947.



Model Interpretation Using Grad-CAM Visualization

Grad-CAM was applied after the last convolutional layer of the model and then overlaid with the radiograph to show the relevance of specific areas for the model's classification process. Figure 4 represents 3 ROIs through visualizations, each of which was classified using the THR prediction model. The heatmap

identified either a narrowing of the hip joint space or the presence of a destructive femoral head or acetabular rims. The last image in Figure 4 is the ROI of the hip, classified with a normal prediction.

After all the processing steps, the ROIs were integrated into the previous pelvic film to create a predictive image as shown in Figure 5 for clinical physicians to review.

Figure 4. Representative Gradient-weighted Class Activation Mapping (Grad-CAM) visualization examples (the total hip replacement group). (A) The hip region of interest shows the collapse of the femoral head and the narrowing of the joint space. (B) The heatmap highlights the collapsed femoral head with the destruction of the hip space. (C) The heatmap highlights the destruction of the femoral head. (D) Normal hip region of interest visualization.

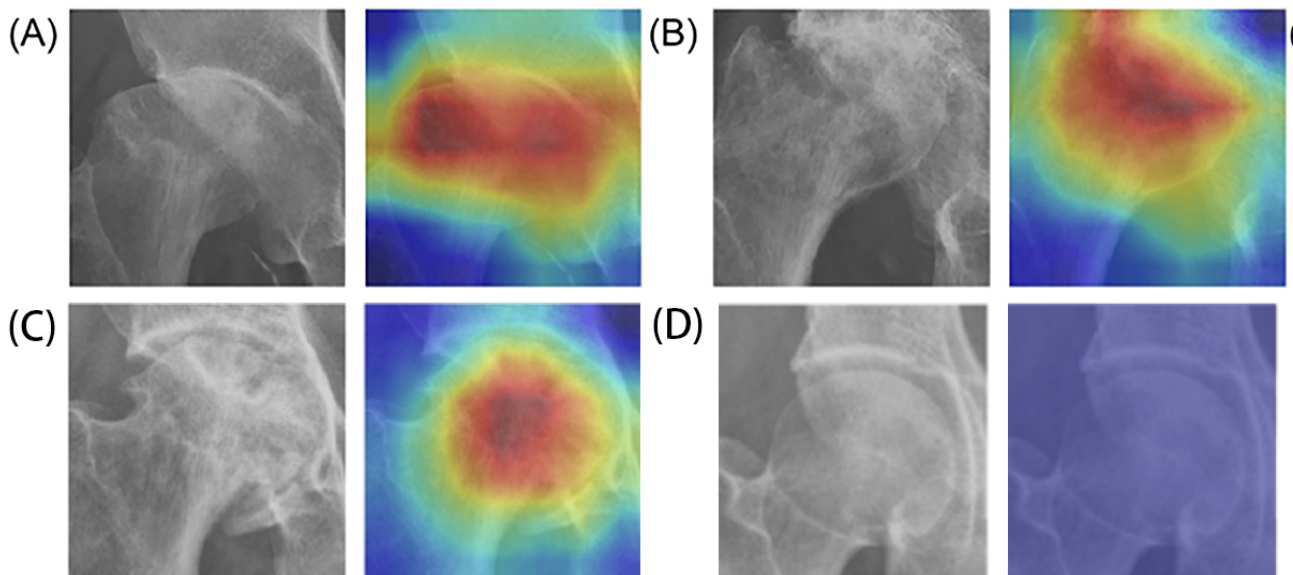
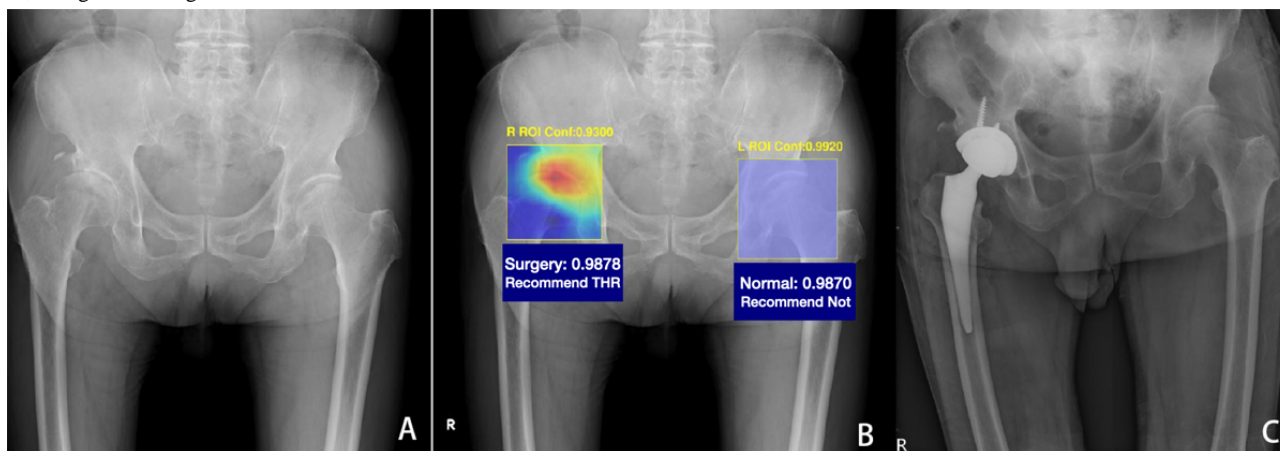


Figure 5. The demonstration of SurgHipNet (the total hip replacement group). (A) The input image of plain pelvic film of a patient with the right hip narrowing of the joint space and femoral head collapse. (B) SurgHipNet detects the hip region of interest and presents the predictive value for a hip replacement recommendation. The right hip joint was predicted to need surgery, and the left one was predicted to be a normal hip joint. The heatmap showed the right femoral head collapse with the joint space narrowing. (C) 3 months later, the post-THR plain film of this patient. Conf: configuration; L: left; R: right. ROI: region of interest.



Discussion

Principal Findings

In this study, it has been demonstrated that the SurgHipNet system can assist frontline doctors in identifying patients who have a high possibility of needing to receive THR within 3 months. In the testing set, the accuracy, sensitivity, specificity, and AUC for the prediction of THR were 0.977, 0.920, 0.992, and 0.994, respectively. To mitigate the “black box” effect, we applied Grad-CAM to generate a heatmap that shows the affected areas of the joint, further drawing the user's attention to potential lesions detected by SurgHipNet. To the best of our knowledge, this is the first study that provides THR prediction. Frontline physicians can timely refer selected patients to orthopedic experts by recommendations from the SurgHipNet model. It has the potential to be used in a wide range of medical imaging applications.

Musculoskeletal disorders, including osteoarthritis, are the most significant contributors to disability worldwide. Hip osteoarthritis is one of the most commonly affected large joints, and its global prevalence continues to rise due to factors such as aging and obesity [3,16]. Conservative treatments, including pharmacologic treatment, exercise, physical therapy, weight reduction, proper footwear, and assistive devices, are the primary therapeutic options [41-43]. It is still unclear who should be referred to THR. Guidelines only suggest that THR is suitable for patients with intractable pain, failed nonsurgical treatments, and functional limitations, and patients often receive the operation with delay [43,44]. However, more established evidence suggests that early surgical intervention reduces associated complications and improves the quality of life for patients experiencing hip discomfort [45]. Higher postoperative levels of range of motion and muscle strength were also observed in patients who had THR surgery earlier, compared to patients who had the surgery later in their disease progression [46]. THR has proven to be cost-effective in the management of hip osteoarthritis, especially for those aged 30-80 years [16].

Medical AI is changing the health care system, and DLs have been used to detect lesions in pathologic images, electrocardiography, and radiography [29]. These algorithms have demonstrated remarkable achievements in disease detection and prediction, achieving a similar level of accuracy to that of experienced physicians. AI has also developed rapidly in the orthopedic field, with applications in image diagnostics, prediction of surgical risk, clinical decision-making, and outcome prediction [27,33,47]. DL has shown its potential in investigating hip arthropathy. Xue et al [34] used a VCG-16 model for hip osteoarthritis classification. Their model showed high sensitivity, specificity, and accuracy compared to chief physicians with 10 years of experience. von Schacky et al [35] recently developed a multitasking DL model for the grading of hip osteoarthritis severity, which also showed a comparable performance with that of expert radiologists. Zhang et al [48] developed a DL model for the diagnosis of developmental dysplasia of the hip, which was highly consistent and more effective in comparison with clinician-led diagnoses.

Recently, medical AI has been suggested to be not only a diagnosis prediction but also a wayfinding tool in clinical decision-making [38]. Leung et al [49] developed a DL prediction model for the prediction of total knee replacement for osteoarthritis progression evaluation. Our study used DL to predict the timing of THR in months, providing practical clinical utility. To our knowledge, this is the first proposed DL model that predicts THR. With this developed model, frontline physicians can effectively refer selected patients to experts in hip replacement, which is a wayfinding support in clinical decision-making.

Study Limitations

The proposed algorithm's use can improve diagnostic accuracy for patients who may need THR, and it can support decision-making in surgical consultations. However, there are still some limitations. First, the algorithm currently does not integrate clinical information, which differs from clinical practice considerations. Decisions regarding arthroplasty might be affected by clinical presentations, such as rest pain, range of movement, underlying conditions [50], as well as the economic

and environmental conditions of the patient [51]. In these situations, patients would not necessarily undergo surgery even if surgery is recommended by SurgHipNet's review of the PXR's. On the other hand, there are occasions where SurgHipNet will advise against performing THR, but the patient will still need surgical intervention; for example, if they need to partake in more activity or if their pain and restricted movement prevent them from being able to access an ambulance. As we suggested, the goal of this developed algorithm is to assist the decision-making process for referring patients to THR surgeons without delay, but it is not intended to be a substitution for a clinician's judgment [52]. Using their expertise, clinicians will make a final suggestion after integrating image findings, the patient's clinical condition, the patient's willingness, and the algorithm results. Future studies could attempt to integrate patients' clinical data into the system, thereby enhancing the algorithm's comprehensiveness. The second limitation is that the limited number of physicians participating in this evaluation could result in an underpowered study. Third, the indication for performing a PXR among individuals in the non-THR group was not clear, which could potentially introduce confounding results. Fourth, because the algorithm focused on proposing the need for THR rather than detecting the severity of patients' conditions, the training data, labeling, and design were limited to using DL to determine the severity of the degenerative joint

disease. Unlike previous studies [35,53] that focused on providing high-resolution heatmaps, further processing is not recommended. Through our study, the physician's attention can be directed toward the ROI within the hip joint; this ensures that the algorithm makes assessments and predictions based on the features of the hip lesion site. However, the resolution might be relatively coarse, leading to possible false negative predictions (Multimedia Appendix 3). In the future, adding labels around the joints to create better heatmaps could increase the confidence of physicians using this model. Further, any hip surgeries that patients in the non-THR group might have undergone in other hospitals, which cannot be detected through our system's medical records, were also likely to potentially contribute to confounding results. Finally, the independent test data set images were only from one institution; even though we already performed validation, the overfitting could not be prevented. A prospective multicenter study should be completed to investigate the system's function in the real world.

Conclusions

In conclusion, we developed SurgHipNet to predict the probability of THR within 3 months. This is the first algorithm designed to suggest THR referrals. Future prospective studies are required to validate whether applying SurgHipNet as a computer-aided tool in a clinical environment leads to more precise diagnoses and facilitates patient management.

Acknowledgments

The authors thank Mr Weichung Wang for his critical reading of the manuscript. This publication was supported by the National Science and Technology Council of Taiwan (NSTC; 2314-B-182A-063; grants NMRPG3M6111 and NMRPG3M6112) and a Chang Gung Memorial Hospital grant (CMRPG1K0191).

Authors' Contributions

CCC, SCC, and CHL designed the experiments; CCC, CTC, and CHL acquired the radiographs for use in the study and provided strategic support; SCC wrote the codes to achieve different tasks; CTC implemented the annotation tools for data annotation; CTW, CYC, and CHL provided labels for use in measuring the algorithm performance; CCC, CTC, and CHL reviewed the computational results; CCC and SCC drafted the manuscript; CHL and CPCC supervised the project; MSL, CCC, and CTC revised the manuscript critically. All authors read and approved the final manuscript.

Conflicts of Interest

SCC is an employee of Compal Electronics. CCC, CTC, and CHL declare no relationships with any companies whose products or services may be related to the subject matter of this paper.

Multimedia Appendix 1

The hip region of interest detection algorithm description.
[\[DOCX File , 16 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

The comparison experiments for the SurgHipNet development.
[\[DOCX File , 15 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

An example of incorrect classification of SurgHipNet.
[\[DOCX File , 849 KB-Multimedia Appendix 3\]](#)

References

1. O'Neill TW, McCabe PS, McBeth J. Update on the epidemiology, risk factors and disease outcomes of osteoarthritis. *Best Pract Res Clin Rheumatol* 2018 Apr;32(2):312-326 [doi: [10.1016/j.berh.2018.10.007](https://doi.org/10.1016/j.berh.2018.10.007)] [Medline: [30527434](https://pubmed.ncbi.nlm.nih.gov/30527434/)]
2. Peat G, Thomas MJ. Osteoarthritis Cartilage 2021 Feb;29(2):180-189 [FREE Full text] [doi: [10.1016/j.joca.2020.10.007](https://doi.org/10.1016/j.joca.2020.10.007)] [Medline: [33242603](https://pubmed.ncbi.nlm.nih.gov/33242603/)]
3. Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, et al. The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study. *Ann Rheum Dis* 2014 Jul;73(7):1323-1330 [doi: [10.1136/annrheumdis-2013-204763](https://doi.org/10.1136/annrheumdis-2013-204763)] [Medline: [24553908](https://pubmed.ncbi.nlm.nih.gov/24553908/)]
4. Trouvin A, Perrot S. Pain in osteoarthritis. Implications for optimal management. *Joint Bone Spine* 2018 Jul;85(4):429-434 [doi: [10.1016/j.jbspin.2017.08.002](https://doi.org/10.1016/j.jbspin.2017.08.002)] [Medline: [28889010](https://pubmed.ncbi.nlm.nih.gov/28889010/)]
5. Abdelaal MS, Restrepo C, Sharkey PF. Global perspectives on arthroplasty of hip and knee joints. *Orthop Clin North Am* 2020 Apr;51(2):169-176 [doi: [10.1016/j.ocl.2019.11.003](https://doi.org/10.1016/j.ocl.2019.11.003)] [Medline: [32138855](https://pubmed.ncbi.nlm.nih.gov/32138855/)]
6. Gademan MGJ, Putter H, Van Den Hout WB, Kloppenburg M, Hofstede SN, Cannegieter SC, et al. The course of pain and function in osteoarthritis and timing of arthroplasty: the CHECK cohort. *Acta Orthop* 2018 Oct;89(5):528-534 [FREE Full text] [doi: [10.1080/17453674.2018.1502533](https://doi.org/10.1080/17453674.2018.1502533)] [Medline: [30350757](https://pubmed.ncbi.nlm.nih.gov/30350757/)]
7. Gademan MGJ, Hofstede SN, Vliet Vlieland TPM, Nelissen RGHH, Marang-van de Mheen PJ. Indication criteria for total hip or knee arthroplasty in osteoarthritis: a state-of-the-science overview. *BMC Musculoskelet Disord* 2016 Nov 09;17(1):463 [FREE Full text] [doi: [10.1186/s12891-016-1325-z](https://doi.org/10.1186/s12891-016-1325-z)] [Medline: [27829422](https://pubmed.ncbi.nlm.nih.gov/27829422/)]
8. Varacallo M, Luo TD, Johanson NA. Total hip arthroplasty techniques. *StatPearls* 2023 [Medline: [29939641](https://pubmed.ncbi.nlm.nih.gov/29939641/)]
9. Knight SR, Aujla R, Biswas SP. Total hip arthroplasty - over 100 years of operative history. *Orthop Rev (Pavia)* 2011 Sep 06;3(2):e16 [FREE Full text] [doi: [10.4081/or.2011.e16](https://doi.org/10.4081/or.2011.e16)] [Medline: [22355482](https://pubmed.ncbi.nlm.nih.gov/22355482/)]
10. Jeldi AJ, Deakin AH, Allen DJ, Granat MH, Grant M, Stansfield BW. Total hip arthroplasty improves pain and function but not physical activity. *J Arthroplasty* 2017 Jul;32(7):2191-2198 [doi: [10.1016/j.arth.2017.02.002](https://doi.org/10.1016/j.arth.2017.02.002)] [Medline: [28285898](https://pubmed.ncbi.nlm.nih.gov/28285898/)]
11. Pivec R, Johnson AJ, Mears SC, Mont MA. Hip arthroplasty. *Lancet* 2012 Nov 17;380(9855):1768-1777 [doi: [10.1016/S0140-6736\(12\)60607-2](https://doi.org/10.1016/S0140-6736(12)60607-2)] [Medline: [23021846](https://pubmed.ncbi.nlm.nih.gov/23021846/)]
12. Hoogboom TJ, van den Ende CHM, van der Sluis G, Elings J, Dronkers JJ, Aiken AB, et al. The impact of waiting for total joint replacement on pain and functional status: a systematic review. *Osteoarthr Cartil* 2009 Nov;17(11):1420-1427 [FREE Full text] [doi: [10.1016/j.joca.2009.05.008](https://doi.org/10.1016/j.joca.2009.05.008)] [Medline: [19500526](https://pubmed.ncbi.nlm.nih.gov/19500526/)]
13. Smith GH, Johnson S, Ballantyne JA, Dunstan E, Brenkel IJ. Predictors of excellent early outcome after total hip arthroplasty. *J Orthop Surg Res* 2012 Mar 25;7:13 [FREE Full text] [doi: [10.1186/1749-799X-7-13](https://doi.org/10.1186/1749-799X-7-13)] [Medline: [22445184](https://pubmed.ncbi.nlm.nih.gov/22445184/)]
14. Jo W, Lee Y, Ha Y, Kim T, Koo K. Delay of total hip arthroplasty to advanced stage worsens post-operative hip motion in patients with femoral head osteonecrosis. *Int Orthop* 2018 Jul;42(7):1599-1603 [doi: [10.1007/s00264-018-3952-5](https://doi.org/10.1007/s00264-018-3952-5)] [Medline: [29700582](https://pubmed.ncbi.nlm.nih.gov/29700582/)]
15. Aresti N, Kassam J, Nicholas N, Achan P. Hip osteoarthritis. *BMJ* 2016 Jul 06;354:i3405 [doi: [10.1136/bmj.i3405](https://doi.org/10.1136/bmj.i3405)] [Medline: [27383835](https://pubmed.ncbi.nlm.nih.gov/27383835/)]
16. Agarwal N, To K, Khan W. Cost effectiveness analyses of total hip arthroplasty for hip osteoarthritis: a PRISMA systematic review. *Int J Clin Pract* 2021 Feb;75(2):e13806 [doi: [10.1111/ijcp.13806](https://doi.org/10.1111/ijcp.13806)] [Medline: [33128841](https://pubmed.ncbi.nlm.nih.gov/33128841/)]
17. Werner BC, Cancienne JM, Browne JA. The timing of total hip arthroplasty after intraarticular hip injection affects postoperative infection risk. *J Arthroplasty* 2016 Apr;31(4):820-823 [doi: [10.1016/j.arth.2015.08.032](https://doi.org/10.1016/j.arth.2015.08.032)] [Medline: [26803711](https://pubmed.ncbi.nlm.nih.gov/26803711/)]
18. Forlenza EM, Burnett RA, Korrapati A, Yang J, Forsythe B, Della Valle CJ. Preoperative corticosteroid injections demonstrate a temporal and dose-dependent relationship with the rate of postoperative infection following total hip arthroplasty. *J Arthroplasty* 2021 Jun;36(6):2033-2037.e1 [doi: [10.1016/j.arth.2021.01.076](https://doi.org/10.1016/j.arth.2021.01.076)] [Medline: [33618958](https://pubmed.ncbi.nlm.nih.gov/33618958/)]
19. Bannuru RR, Osani MC, Vaysbrot EE, Arden NK, Bennell K, Bierma-Zeinstra SMA, et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. *Osteoarthr Cartil* 2019 Nov;27(11):1578-1589 [doi: [10.1016/j.joca.2019.06.011](https://doi.org/10.1016/j.joca.2019.06.011)] [Medline: [31278997](https://pubmed.ncbi.nlm.nih.gov/31278997/)]
20. Kim C, Linsenmeyer KD, Vlad SC, Guermazi A, Clancy MM, Niu J, et al. Prevalence of radiographic and symptomatic hip osteoarthritis in an urban United States community: the Framingham osteoarthritis study. *Arthritis Rheumatol* 2014 Nov 26;66(11):3013-3017 [FREE Full text] [doi: [10.1002/art.38795](https://doi.org/10.1002/art.38795)] [Medline: [25103598](https://pubmed.ncbi.nlm.nih.gov/25103598/)]
21. Altman R, Alarcón G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum* 1991 May;34(5):505-514 [doi: [10.1002/art.1780340502](https://doi.org/10.1002/art.1780340502)] [Medline: [2025304](https://pubmed.ncbi.nlm.nih.gov/2025304/)]
22. Moltu C, Stefansen J, Svisdahl M, Veseth M. Negotiating the coresearcher mandate - service users' experiences of doing collaborative research on mental health. *Disabil Rehabil* 2012;34(19):1608-1616 [doi: [10.3109/09638288.2012.656792](https://doi.org/10.3109/09638288.2012.656792)] [Medline: [22489612](https://pubmed.ncbi.nlm.nih.gov/22489612/)]
23. Teirlinck CH, Dorleijn DMJ, Bos PK, Rijkels-Otters JBM, Bierma-Zeinstra SMA, Luijsterburg PAJ. Prognostic factors for progression of osteoarthritis of the hip: a systematic review. *Arthritis Res Ther* 2019 Aug 23;21(1):192 [FREE Full text] [doi: [10.1186/s13075-019-1969-9](https://doi.org/10.1186/s13075-019-1969-9)] [Medline: [31443685](https://pubmed.ncbi.nlm.nih.gov/31443685/)]
24. Ebrahimighahnavieh MA, Luo S, Chiong R. Deep learning to detect Alzheimer's disease from neuroimaging: a systematic literature review. *Comput Methods Programs Biomed* 2020 Apr;187:105242 [doi: [10.1016/j.cmpb.2019.105242](https://doi.org/10.1016/j.cmpb.2019.105242)] [Medline: [31837630](https://pubmed.ncbi.nlm.nih.gov/31837630/)]

25. Majkowska A, Mittal S, Steiner DF, Reicher JJ, McKinney SM, Duggan GE, et al. Chest radiograph interpretation with deep learning models: assessment with radiologist-adjudicated reference standards and population-adjusted evaluation. *Radiology* 2020 Feb;294(2):421-431 [doi: [10.1148/radiol.2019191293](https://doi.org/10.1148/radiol.2019191293)] [Medline: [31793848](https://pubmed.ncbi.nlm.nih.gov/31793848/)]
26. Asiri N, Hussain M, Al Adel F, Alzaidi N. Deep learning based computer-aided diagnosis systems for diabetic retinopathy: a survey. *Artif Intell Med* 2019 Aug;99:101701 [doi: [10.1016/j.artmed.2019.07.009](https://doi.org/10.1016/j.artmed.2019.07.009)] [Medline: [31606116](https://pubmed.ncbi.nlm.nih.gov/31606116/)]
27. Cheng C, Wang Y, Chen H, Hsiao P, Yeh C, Hsieh C, et al. A scalable physician-level deep learning algorithm detects universal trauma on pelvic radiographs. *Nat Commun* 2021 Feb 16;12(1):1066 [FREE Full text] [doi: [10.1038/s41467-021-21311-3](https://doi.org/10.1038/s41467-021-21311-3)] [Medline: [33594071](https://pubmed.ncbi.nlm.nih.gov/33594071/)]
28. Yao L, Zhang J, Liu J, Zhu L, Ding X, Chen D, et al. A deep learning-based system for bile duct annotation and station recognition in linear endoscopic ultrasound. *EBioMedicine* 2021 Mar;65:103238 [FREE Full text] [doi: [10.1016/j.ebiom.2021.103238](https://doi.org/10.1016/j.ebiom.2021.103238)] [Medline: [33639404](https://pubmed.ncbi.nlm.nih.gov/33639404/)]
29. Esteva A, Chou K, Yeung S, Naik N, Madani A, Mottaghi A, et al. Deep learning-enabled medical computer vision. *NPJ Digit Med* 2021 Jan 08;4(1):5 [FREE Full text] [doi: [10.1038/s41746-020-00376-2](https://doi.org/10.1038/s41746-020-00376-2)] [Medline: [33420381](https://pubmed.ncbi.nlm.nih.gov/33420381/)]
30. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016 Dec 13;316(22):2402-2410 [doi: [10.1001/jama.2016.17216](https://doi.org/10.1001/jama.2016.17216)] [Medline: [27898976](https://pubmed.ncbi.nlm.nih.gov/27898976/)]
31. Sbröllini A, De Jongh MC, Ter Haar CC, Treskes RW, Man S, Burattini L, et al. Serial electrocardiography to detect newly emerging or aggravating cardiac pathology: a deep-learning approach. *Biomed Eng Online* 2019 Feb 12;18(1):15 [FREE Full text] [doi: [10.1186/s12938-019-0630-9](https://doi.org/10.1186/s12938-019-0630-9)] [Medline: [30755195](https://pubmed.ncbi.nlm.nih.gov/30755195/)]
32. Huang B, Tian S, Zhan N, Ma J, Huang Z, Zhang C, et al. Accurate diagnosis and prognosis prediction of gastric cancer using deep learning on digital pathological images: a retrospective multicentre study. *EBioMedicine* 2021 Nov;73:103631 [FREE Full text] [doi: [10.1016/j.ebiom.2021.103631](https://doi.org/10.1016/j.ebiom.2021.103631)] [Medline: [34678610](https://pubmed.ncbi.nlm.nih.gov/34678610/)]
33. Cheng C, Ho T, Lee T, Chang C, Chou C, Chen C, et al. Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs. *Eur Radiol* 2019 Oct;29(10):5469-5477 [FREE Full text] [doi: [10.1007/s00330-019-06167-y](https://doi.org/10.1007/s00330-019-06167-y)] [Medline: [30937588](https://pubmed.ncbi.nlm.nih.gov/30937588/)]
34. 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(6):e0178992 [FREE Full text] [doi: [10.1371/journal.pone.0178992](https://doi.org/10.1371/journal.pone.0178992)] [Medline: [28575070](https://pubmed.ncbi.nlm.nih.gov/28575070/)]
35. von Schacky CE, Sohn JH, Liu F, Ozhinsky E, Jungmann PM, Nardo L, et al. Development and validation of a multitask deep learning model for severity grading of hip osteoarthritis features on radiographs. *Radiology* 2020 Apr;295(1):136-145 [FREE Full text] [doi: [10.1148/radiol.2020190925](https://doi.org/10.1148/radiol.2020190925)] [Medline: [32013791](https://pubmed.ncbi.nlm.nih.gov/32013791/)]
36. Karnuta JM, Navarro SM, Haeberle HS, Helm JM, Kamath AF, Schaffer JL, et al. Predicting inpatient payments prior to lower extremity arthroplasty using deep learning: which model architecture is best? *J Arthroplasty* 2019 Oct;34(10):2235-2241.e1 [doi: [10.1016/j.arth.2019.05.048](https://doi.org/10.1016/j.arth.2019.05.048)] [Medline: [31230954](https://pubmed.ncbi.nlm.nih.gov/31230954/)]
37. Ramkumar PN, Karnuta JM, Navarro SM, Haeberle HS, Iorio R, Mont MA, et al. Preoperative prediction of value metrics and a patient-specific payment model for primary total hip arthroplasty: development and validation of a deep learning model. *J Arthroplasty* 2019 Oct;34(10):2228-2234.e1 [doi: [10.1016/j.arth.2019.04.055](https://doi.org/10.1016/j.arth.2019.04.055)] [Medline: [31122849](https://pubmed.ncbi.nlm.nih.gov/31122849/)]
38. Adler-Milstein J, Chen JH, Dhaliwal G. Next-generation artificial intelligence for diagnosis: from predicting diagnostic labels to "wayfinding". *JAMA* 2021 Dec 28;326(24):2467-2468 [doi: [10.1001/jama.2021.22396](https://doi.org/10.1001/jama.2021.22396)] [Medline: [34882190](https://pubmed.ncbi.nlm.nih.gov/34882190/)]
39. Liu F, Chen C, Cheng C, Wu C, Hsu C, Fu C, et al. Automatic hip detection in anteroposterior pelvic radiographs—a labelless practical framework. *J Pers Med* 2021 Jun 07;11(6):522 [doi: [10.3390/jpm11060522](https://doi.org/10.3390/jpm11060522)] [Medline: [34200151](https://pubmed.ncbi.nlm.nih.gov/34200151/)]
40. Albahli S, Yar GNAH. Fast and accurate detection of COVID-19 along with 14 other chest pathologies using a multi-level classification: algorithm development and validation study. *J Med Internet Res* 2021 Feb 10;23(2):e23693 [FREE Full text] [doi: [10.2196/23693](https://doi.org/10.2196/23693)] [Medline: [33529154](https://pubmed.ncbi.nlm.nih.gov/33529154/)]
41. Katz JN, Arant KR, Loeser RF. Diagnosis and treatment of hip and knee osteoarthritis: a review. *JAMA* 2021 Feb 09;325(6):568-578 [FREE Full text] [doi: [10.1001/jama.2020.22171](https://doi.org/10.1001/jama.2020.22171)] [Medline: [33560326](https://pubmed.ncbi.nlm.nih.gov/33560326/)]
42. Lane NE. Osteoarthritis of the Hip. *N Engl J Med* 2007 Oct 04;357(14):1413-1421 [doi: [10.1056/nejmcp071112](https://doi.org/10.1056/nejmcp071112)]
43. Lespasio MJ, Sultan AA, Piuzzi NS, Khlopas A, Husni ME, Muschler GF, et al. Hip osteoarthritis: a primer. *Perm J* 2018;22:17-084 [FREE Full text] [doi: [10.7812/TPP/17-084](https://doi.org/10.7812/TPP/17-084)] [Medline: [29309269](https://pubmed.ncbi.nlm.nih.gov/29309269/)]
44. Mujica-Mota RE, Watson LK, Tarricone R, Jäger M. Cost-effectiveness of timely delayed primary total hip replacement in Germany: a social health insurance perspective. *Orthop Rev (Pavia)* 2017 Sep 30;9(3):7161 [FREE Full text] [doi: [10.4081/or.2017.7161](https://doi.org/10.4081/or.2017.7161)] [Medline: [29071040](https://pubmed.ncbi.nlm.nih.gov/29071040/)]
45. Quintana JM, Escobar A, Aguirre U, Lafuente I, Arenaza JC. Predictors of health-related quality-of-life change after total hip arthroplasty. *Clin Orthop Relat Res* 2009 Nov;467(11):2886-2894 [FREE Full text] [doi: [10.1007/s11999-009-0868-9](https://doi.org/10.1007/s11999-009-0868-9)] [Medline: [19412646](https://pubmed.ncbi.nlm.nih.gov/19412646/)]
46. Montin L, Leino-Kilpi H, Suominen T, Lepistö J. A systematic review of empirical studies between 1966 and 2005 of patient outcomes of total hip arthroplasty and related factors. *J Clin Nurs* 2008 Jan;17(1):40-45 [doi: [10.1111/j.1365-2702.2007.01944.x](https://doi.org/10.1111/j.1365-2702.2007.01944.x)] [Medline: [18088258](https://pubmed.ncbi.nlm.nih.gov/18088258/)]

47. Myers TG, Ramkumar PN, Ricciardi BF, Urish KL, Kipper J, Ketonis C. Artificial intelligence and orthopaedics: an introduction for clinicians. *J Bone Joint Surg Am* 2020 May 06;102(9):830-840 [FREE Full text] [doi: [10.2106/JBJS.19.01128](https://doi.org/10.2106/JBJS.19.01128)] [Medline: [32379124](https://pubmed.ncbi.nlm.nih.gov/32379124/)]
48. Zhang S, Sun J, Liu C, Fang J, Xie H, Ning B. Clinical application of artificial intelligence-assisted diagnosis using anteroposterior pelvic radiographs in children with developmental dysplasia of the hip. *Bone Joint J* 2020 Nov;102-B(11):1574-1581 [doi: [10.1302/0301-620X.102B11.BJJ-2020-0712.R2](https://doi.org/10.1302/0301-620X.102B11.BJJ-2020-0712.R2)] [Medline: [33135455](https://pubmed.ncbi.nlm.nih.gov/33135455/)]
49. Leung K, Zhang B, Tan J, Shen Y, Geras KJ, Babb JS, et al. Prediction of total knee replacement and diagnosis of osteoarthritis by using deep learning on knee radiographs: data from the osteoarthritis initiative. *Radiology* 2020 Sep;296(3):584-593 [FREE Full text] [doi: [10.1148/radiol.2020192091](https://doi.org/10.1148/radiol.2020192091)] [Medline: [32573386](https://pubmed.ncbi.nlm.nih.gov/32573386/)]
50. Dreinhöfer KE, Dieppe P, Stürmer T, Gröber-Grätz D, Flören M, Günther K, et al. Indications for total hip replacement: comparison of assessments of orthopaedic surgeons and referring physicians. *Ann Rheum Dis* 2006 Oct;65(10):1346-1350 [FREE Full text] [doi: [10.1136/ard.2005.047811](https://doi.org/10.1136/ard.2005.047811)] [Medline: [16439438](https://pubmed.ncbi.nlm.nih.gov/16439438/)]
51. Youm J, Chan V, Belkora J, Bozic KJ. Impact of socioeconomic factors on informed decision making and treatment choice in patients with hip and knee OA. *J Arthroplasty* 2015 Feb;30(2):171-175 [doi: [10.1016/j.arth.2014.09.006](https://doi.org/10.1016/j.arth.2014.09.006)] [Medline: [25301018](https://pubmed.ncbi.nlm.nih.gov/25301018/)]
52. Murphy MP, Brown NM. CORR synthesis: when should the orthopaedic surgeon use artificial intelligence, machine learning, and deep learning? *Clin Orthop Relat Res* 2021 Jul 01;479(7):1497-1505 [FREE Full text] [doi: [10.1097/CORR.0000000000001679](https://doi.org/10.1097/CORR.0000000000001679)] [Medline: [33595930](https://pubmed.ncbi.nlm.nih.gov/33595930/)]
53. Tiulpin A, Thevenot J, Rahtu E, Lehenkari P, Saarakkala S. Automatic knee osteoarthritis diagnosis from plain radiographs: a deep learning-based approach. *Sci Rep* 2018 Jan 29;8(1):1727 [FREE Full text] [doi: [10.1038/s41598-018-20132-7](https://doi.org/10.1038/s41598-018-20132-7)] [Medline: [29379060](https://pubmed.ncbi.nlm.nih.gov/29379060/)]

Abbreviations

- AI:** artificial intelligence
AUC: area under the receiver operating characteristic curve
CGMH: Chang Gung Memorial Hospital
DL: deep learning
Grad-CAM: Gradient-weighted Class Activation Mapping
PXR: pelvic radiograph
ROI: region of interest
THR: total hip replacement

Edited by A Mavragani; submitted 19.09.22; peer-reviewed by M Farad, B Qian, HW Chiu; comments to author 20.12.22; revised version received 27.01.23; accepted 04.08.23; published 20.10.23

Please cite as:

Chen CC, Wu CT, Chen CPC, Chung CY, Chen SC, Lee MS, Cheng CT, Liao CH

Predicting the Risk of Total Hip Replacement by Using A Deep Learning Algorithm on Plain Pelvic Radiographs: Diagnostic Study
JMIR Form Res 2023;7:e42788

URL: <https://formative.jmir.org/2023/1/e42788>

doi: [10.2196/42788](https://doi.org/10.2196/42788)

PMID: [37862084](https://pubmed.ncbi.nlm.nih.gov/37862084/)

©Chih-Chi Chen, Cheng-Ta Wu, Carl P C Chen, Chia-Ying Chung, Shann-Ching Chen, Mel S Lee, Chi-Tung Cheng, Chien-Hung Liao. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 20.10.2023. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.