

90357_Auto_Edited.docx

Deep learning automation of radiographic patterns for hallux valgus diagnosis

Hussain A *et al.* AI automation of radiographic patterns

Abstract

Artificial intelligence (AI) and deep learning are becoming increasingly powerful tools in diagnostic and radiographic medicine. Deep learning has already been utilized for automated detection of pneumonia from chest radiographs, diabetic retinopathy, breast cancer, skin carcinoma classification, and metastatic lymphadenopathy detection, with diagnostic reliability akin to medical experts. In this article of the *World Journal of Orthopedics*, the authors apply an automated and AI-assisted technique to determine the hallux valgus angle (HVA) for assessing HV foot deformity. With the U-net neural network, the authors constructed an algorithm for pattern recognition of HV foot deformity from anteroposterior high-resolution radiographs. The performance of the deep learning algorithm was compared to expert clinician manual performance and assessed alongside clinician-clinician variability. The authors found that the AI tool was sufficient in assessing HVA and proposed the system as an instrument to augment clinical efficiency. Though further sophistication is needed to establish automated algorithms for more complicated foot pathologies, this work adds to the growing evidence supporting AI as a powerful diagnostic tool.

Key Words: Artificial intelligence; Hallux valgus; Deep learning; Automated radiography

Hussain A, Lee C, Hu E, Amirouche F. Deep learning automation of radiographic patterns for hallux valgus diagnosis. *World J Orthop* 2023; In press

Core Tip: This editorial summarizes and outlines the original paper “Automated decision support for hallux valgus treatment options using anteroposterior foot radiographs”. We summarize the scope of the deep learning process and compare it to existing artificial intelligence studies used in clinical diagnostic studies. We additionally describe its limitations and impact in the field of automated diagnostic tools.

INTRODUCTION

Artificial Intelligence (AI) and deep learning have emerged as potential assets in the realm of diagnostic and radiographic medicine, significantly impacting the quality and efficiency of healthcare^[1]. Automated systems have demonstrated their value in various medical domains, such as AI-derived pneumonia detection from chest radiographs, diabetic retinopathy assessment, breast cancer diagnosis, skin carcinoma classification, and metastatic lymphadenopathy detection^[2]. The diagnostic reliability of some AI systems has been found to parallel that of medical experts, marking a transformative era in medical imaging and analysis^[3].

This study by Kwolek *et al*^[4], featured in the *World Journal of Orthopedics*, represents a supportive contribution to the application of AI in orthopedic diagnostics. The authors employed an automated and AI-assisted technique to determine the hallux valgus angle (HVA), a critical parameter in assessing HV foot deformity. Leveraging the capabilities of the U-net neural network, the authors constructed an algorithm tailored for pattern recognition of HV foot deformity from anteroposterior high-resolution radiographs. The authors evaluated the performance of their deep learning algorithm by comparing it to expert clinician manual assessments and also considered clinician-clinician variability. The study found that the AI tool demonstrated sufficient accuracy in assessing the HVA when compared to clinician readings and circumvented the issue of clinician-clinician variability. Since the proposed system was effective, the authors suggest that it may have an eventual role in supporting clinical efficiency or efficiency in the evaluation of HV deformities.

While acknowledging the success achieved in this study, the authors emphasize the need for further sophistication in automated algorithms to address more intricate foot pathologies. Automated algorithms like the one presented here are not well-adapted to perform analysis for patients outside of the narrow scope of a very specific pathology. In this case, exclusion of patients with pathologies such as osteoarthritis, pes cavus, Charcot foot, or other joint deformities limits the clinical application of the proposed algorithm. Especially considering the simple diagnostic nature of assessing HVA, clinician discretion is still important for evaluating concomitant foot deformities or pathologies.

As a general commentary, the integration of AI in medical imaging has significant implications for patients and physicians. The use of AI-assisted diagnostics may lead to enhanced accuracy and efficiency in medical practice, potentially leading to quicker and more precise treatment decisions^[5]. Patients may benefit from improved diagnostic capabilities, resulting in more timely interventions and potentially more accurate diagnoses. Physicians may experience a shift in their roles, with AI serving as a valuable supportive tool in diagnostic processes. There is a growing fear that AI may eventually have the power to overtake the clinician's role in diagnosing certain pathologies, but at this stage, AI simply serves as a tool for augmenting efficient delivery of care^[6]. The proposed AI system in this study does not replace clinical expertise but rather complements it, offering a tool that can handle repetitive tasks efficiently and contribute to diagnostic accuracy^[3]. This study reinforces the growing evidence supporting AI as a powerful diagnostic tool in orthopedics. As AI continues to evolve, its impact on the quality of healthcare, patient outcomes, and the workflow of medical practitioners will transform a new era in the integration of technology with traditional medical practices.

The HV is one of the most common forefoot conditions and is defined as angular deviation of more than 15 degrees of the hallux with respect to the first metatarsal bone^[7,8]. Risk factors for the development of HV include female sex, increased age, body mass index, pes planus, hammertoe, and ill-fitting footwear^[8,9]. Though the exact biomechanical etiology is unclear, recent literature suggests that HV is likely due to soft

tissue contracture and attenuation leading to malalignment at the bone articulations^[7]. When evaluating HVA, it is important to evaluate the osseous tissue within the foot. The first ray osseous components are composed of the first metatarsal and medial cuneiform with stability dependent on several static and dynamic structures at the first metatarsophalangeal (MTP) and tarsometatarsal joints, making the first ray intrinsically unstable^[10]. In the early stages of HV deformity, the prevailing theory posits a weakening of medial support structures of the first toe. This weakening manifests as a medial displacement of the first metatarsal accompanied by a lateral deviation and pronation of the big toe. Consequently, a gradual varus deformity takes shape at the first MTP joint. As the metatarsal head undergoes medial shift and rotation in the frontal plane, its position relative to the sesamoid apparatus is altered. Consequently, the first metatarsal head rests on the medial sesamoid while the lateral sesamoid is in the first intermetatarsal (IMA) space. Concurrently the developed deformity at the MTP joint permits the hallux flexor and extensor tendons to bow laterally, causing further deformity. Simultaneously, the displaced abductor hallucis flexes and pronates the phalanx contributing to the distortion. The increased prominence at the first MTP joint stems from the increased prominence of the first metatarsal head^[9,10]. The HVA is commonly measured by plain radiograph. The widely-accepted method to measure HVA consists of an angle constructed between the center longitudinal axis of the first metatarsal and the axis of the hallux. The angle has typically been determined through use of a protractor by use of the physician but more recently has been determined with assistance from technology^[11].

In this article of *World Journal of Orthopedics*, Kwolek *et al*^[4] demonstrated a novel approach to HVA and IMA measurements by utilizing deep-learning computer automated methods. By comparing the measurements obtained by their trained U-Net neural network to measurements performed manually by clinicians, they sought to demonstrate the algorithm's clinical efficacy. The study was conducted with a cohort of 133 patients, comprising 265 preoperative radiographic images, with the sole inclusion criteria being weight-bearing symptomatic HV. The authors excluded radiographs with

other underlying pathologies that could complicate a read (*i.e.*, severe osteoarthritis, joint deformation, Charcot foot). The U-Net neural network was first trained using anteroposterior foot radiographs with labeled bones and segmental separation. Rather than utilizing binary segmentation of bones with bone extraction, the authors opted for a multi-class segmentation that only selects three bones (1MT, 2MT, and hallucial PP) to calculate HVA/IMA ultimately. This was done to limit the difficulties with reliably training the model despite radiographic noise/artifacts and complex bone structure in the anteroposterior view. Following training and validation of the model, the algorithm was used to automate the measurement of the HVA/IMA in 84 radiographs. In those same radiographs, HVA/IMA were measured manually multiple times by clinicians who were blinded to clinical outcomes. The measurements between AI clinicians and clinician clinicians were compared.

Kwolek *et al*^[4] found that there was a large correlation (HVA: 0.96-0.99; IMA: 0.78-0.95) between the AI-generated measurements and clinician measurements of HVA/IMA. They found that the ratio of operative decisions made based on AI recommendations compared to clinician decisions was almost 0.80, which is equal if not higher than the ratio among different clinicians. The authors state that these results are strongly suggestive of a successful achievement that would ultimately save time for the radiologist and orthopedic surgeon, while producing clinically actionable results.

A limitation of this study includes the standardization of initial training of the U-Net neural network, as the dataset was largely relying on measurements of three segmented bones, which may not be reliable in patients with varying anatomy. In addition, radiographs were selected on the basis that there was no secondary bone pathology present. These limited data sets may not be reflective of the general patient population suffering from HV. These concerns largely coincide with the literature regarding the implementation of AI in clinical practice. These include the amount of dependence that is put on high-quality training data which requires large and clinically relevant datasets^[12]. In addition, there are ethical concerns with making clinical decisions based on deep learning technology when not all of the logical bases of the system are

understood. Finally, concerns have been raised regarding the clinician's lack of acceptance and trust in AI innovations, as well as their lack of awareness and familiarity with the technology to be used readily in clinical practice^[13].

Despite this, the study does well to demonstrate an efficient metric that has produced promising results. The study design is strong, and the authors were able to create a standardized neural network that is able to reliably make clinically meaningful decisions at the same rate, if not better, than clinicians. Despite the potential limitations of AI implications, it is hard to ignore the potential benefits of the clinical implications that the authors posit.

CONCLUSION

The featured study by Kwolek *et al*^[4] supports the integration of AI into orthopedic diagnostics, specifically in the assessment of HV foot deformity. The authors utilized a U-net neural network to determine the HVA. This has implications on the diagnosis of foot pathologies, but also the use of automated tools in healthcare in general.

However, the study also acknowledges the need for further sophistication in automated algorithms to address complex foot pathologies. Considering the rate of concomitant procedures performed during hallux alignment surgery, it is important to note that HV does not always occur in isolation^[14]. The authors of this study needed to exclude concomitant pathologies such as osteoarthritis, pes cavus, Charcot foot, and joint deformities, which does not yield a clinically-accurate patient population. The proposed automated system was effective in determining the HVA, a somewhat simple measurement, but only for foot radiographs where no other foot pathologies were noted. Further optimization of the algorithm is needed for this tool to become clinically relevant.

Additionally, while we have previously discussed the potential for AI to increase clinical efficiency, the value of reduced diagnostic time may be marginal in the case of measuring HVA. The conventional method used to acquire HVA takes $12.3 \text{ s} \pm 0.6 \text{ s}$ to measure, and even as short as $5.9 \text{ s} \pm 0.2 \text{ s}$ with the point-connection method^[15]. While

automation of this measurement may be convenient, the marginal time saved by the algorithm may not outweigh the limitations that currently exist with this model. The automated process may save an expertly trained clinician a number of seconds or minutes, but that time would be needed to screen the radiographs for exclusion criteria previously mentioned.

Finally, as with any deep-learning algorithm, the outputs are subject to bias in training data and patient demographic inputs^[16]. The study would be strengthened with the use of a larger, more diverse dataset, potentially utilizing a multi-center approach, to ensure that the automated system was able to process patient radiographs from a wide range of demographics including age, sex, race, and lifestyle.

In summary, while the study gives support to the exciting topic of AI as a powerful diagnostic tool in medicine, there are many limitations that need to be addressed before the algorithm presented here reaches clinical relevance. The deep-learning system would benefit from training with a wider variety of pathologies and a larger sample size of more diverse patient demographics. Nevertheless, this work by Kwolek *et al*^[4] is an important reminder of the potential impacts of automated systems in medicine.

ACKNOWLEDGEMENTS

¹ The views expressed in this article are solely those of the authors and do not necessarily reflect the position or policy of the affiliated institutions.

REFERENCES

- 1 **Malamateniou C**, Knapp KM, Pergola M, Woznitza N, Hardy M. Artificial intelligence in radiography: Where are we now and what does the future hold? *Radiography (Lond)* 2021; **27** Suppl 1: S58-S62 [PMID: 34380589 DOI: 10.1016/j.radi.2021.07.015]
- 2 **Patel BN**, Rosenberg L, Willcox G, Baltaxe D, Lyons M, Irvin J, Rajpurkar P, Amrhein T, Gupta R, Halabi S, Langlotz C, Lo E, Mammarappallil J, Mariano AJ, Riley G, Seekins J, Shen L, Zucker E, Lungren M. Human-machine partnership with artificial intelligence

for chest radiograph diagnosis. *NPJ Digit Med* 2019; **2**: 111 [PMID: 31754637 DOI: 10.1038/s41746-019-0189-7]

3 **Currie G**, Hawk KE, Rohren E, Vial A, Klein R. Machine Learning and Deep Learning in Medical Imaging: Intelligent Imaging. *J Med Imaging Radiat Sci* 2019; **50**: 477-487 [PMID: 31601480 DOI: 10.1016/j.jmir.2019.09.005]

4 **Kwolek K**, Gądek A, Kwolek K, Kolecki R, Liszka H. Automated decision support for Hallux Valgus treatment options using anteroposterior foot radiographs. *World J Orthop* 2023; **14**: 800-812 [PMID: 38075473 DOI: 10.5312/wjo.v14.i11.800]

5 **Koski E**, Murphy J. AI in Healthcare. *Stud Health Technol Inform* 2021; **284**: 295-299 [PMID: 34920529 DOI: 10.3233/SHTI210726]

6 **Sauerbrei A**, Kerasidou A, Lucivero F, Hallowell N. The impact of artificial intelligence on the person-centred, doctor-patient relationship: some problems and solutions. *BMC Med Inform Decis Mak* 2023; **23**: 73 [PMID: 37081503 DOI: 10.1186/s12911-023-02162-y]

7 **Gonzalez TA**, Smith JT, Bluman EM, Ready LV, Ciurylo W, Chiodo CP. Republication of "Treatment of Hallux Valgus Deformity Using a Suture Button Device: A Preliminary Report". *Foot Ankle Orthop* 2023; **8**: 24730114231195342 [PMID: 37655934 DOI: 10.1177/24730114231195342]

8 **Nguyen US**, Hillstrom HJ, Li W, Dufour AB, Kiel DP, Procter-Gray E, Gagnon MM, Hannan MT. Factors associated with hallux valgus in a population-based study of older women and men: the MOBILIZE Boston Study. *Osteoarthritis Cartilage* 2010; **18**: 41-46 [PMID: 19747997 DOI: 10.1016/j.joca.2009.07.008]

9 **Ray JJ**, Friedmann AJ, Hanselman AE, Vaida J, Dayton PD, Hatch DJ, Smith B, Santrock RD. Hallux Valgus. *Foot Ankle Orthop* 2019; **4**: 2473011419838500 [PMID: 35097321 DOI: 10.1177/2473011419838500]

10 **Perera AM**, Mason L, Stephens MM. The pathogenesis of hallux valgus. *J Bone Joint Surg Am* 2011; **93**: 1650-1661 [PMID: 21915581 DOI: 10.2106/JBJS.H.01630]

- 11 **Meng HZ**, Zhang WL, Li XC, Yang MW. Radiographic angles in hallux valgus: Comparison between protractor and iPhone measurements. *J Orthop Res* 2015; **33**: 1250-1254 [PMID: 25763918 DOI: 10.1002/jor.22872]
- 12 **Lee JG**, Jun S, Cho YW, Lee H, Kim GB, Seo JB, Kim N. Deep Learning in Medical Imaging: General Overview. *Korean J Radiol* 2017; **18**: 570-584 [PMID: 28670152 DOI: 10.3348/kjr.2017.18.4.570]
- 13 **Eltawil FA**, Atalla M, Boulos E, Amirabadi A, Tyrrell PN. Analyzing Barriers and Enablers for the Acceptance of Artificial Intelligence Innovations into Radiology Practice: A Scoping Review. *Tomography* 2023; **9**: 1443-1455 [PMID: 37624108 DOI: 10.3390/tomography9040115]
- 14 **Chodaba Y**, Andrews NA, Halstrom J, Sankey T, Sanchez T, Harrelson WM, Agarwal A, Shah A. First Metatarsophalangeal Arthrodesis for Hallux Rigidus and Hallux Valgus: A Comparison of Patient Reported and Clinical Outcomes. *J Foot Ankle Surg* 2023; **62**: 683-688 [PMID: 36964117 DOI: 10.1053/j.jfas.2023.02.008]
- 15 **Seo JH**, Ahn JY, Boedijono D. Point-Connecting Measurements of the Hallux Valgus Deformity: A New Measurement and Its Clinical Application. *Yonsei Med J* 2016; **57**: 741-747 [PMID: 26996576 DOI: 10.3349/ymj.2016.57.3.741]
- 16 **Gianfrancesco MA**, Tamang S, Yazdany J, Schmajuk G. Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data. *JAMA Intern Med* 2018; **178**: 1544-1547 [PMID: 30128552 DOI: 10.1001/jamainternmed.2018.3763]

ORIGINALITY REPORT

1%

SIMILARITY INDEX

PRIMARY SOURCES

1

pubmed.ncbi.nlm.nih.gov

Internet

24 words — 1%

- EXCLUDE QUOTES

ON

EXCLUDE SOURCES

< 15 WORDS
- EXCLUDE BIBLIOGRAPHY

ON

EXCLUDE MATCHES

< 10 WORDS