OPEN

PRS GLOBAL OPEN

LETTER TO THE EDITOR

# Developing Machine Learning Algorithms to Support Patient-centered, Value-based Carpal Tunnel Decompression Surgery

Angelos Mantelakis, BSc (Hons), MBBS (Lond), MRCS (Eng)*; Ankur Khajuria, MRCS (Eng), MSc (Oxon), PhD†‡

Sir:

We read with great interest the article entitled "Developing Machine Learning Algorithms to Support Patient-centered, Value-based Carpal Tunnel Decompression Surgery."[1] The authors describe the application of machine learning (ML) models in the prediction of functional and symptomatic improvement following carpal tunnel decompression (CTD) surgery. We commend the authors on the originality of this article, as this is one of the first ever reported applications of this technology in hand surgery.[2] However, the methodology of creation of the machine learning algorithm, and the clinical applicability presented in the discussion is questionable.

From a technological standpoint, it is unclear why some of the predictors were chosen to train the described ML algorithms, leading to "black-box" medical decisions. To-date, there is no evidence that factors such as anemia, stomach ulcers, lung disease, or backpain affect the postoperative outcome of CTD. However, these are utilized as predictors in the described ML models and are given significant weight to predict CTD surgery outcomes. This contributes to what is known as "black-box" medical decision-making, meaning there is no justification or explanation as to why a set of seemingly unrelated factors influence the algorithms decision-making. In healthcare, there is very low acceptance to utilizing algorithms with "black-box" decisions to treat patients, due to a lack of trust in the absence of utilization evidence-based practices.[3,4] To overcome barriers of introduction of this technology to clinical practice, it is pivotal to make ML interpretable, by choosing to incorporate key predictors that relate to the outcome in hand and allowing the algorithm to provide justification as to why the outcome is chosen.

During the creation of the novel ML algorithms proposed in this article, the authors describe the use of training set (80%) and test set (20%). This fundamentally lacks the utilization of a validation set.[2] When creating an ML model, the authors must split the total dataset into three subsets: training, validation, and testing sets. After training, algorithms need to be validated on a set of different data than what it was trained on, called a validation set, aiming to select the best model architecture and perform tuning of the hyperparameters. Following this process, the final model can then be finally examined on the test set. The authors have not utilized a validation set and as such the proposed ML models are likely overfitted to the test set and report a misleading higher area under the curve performance.[5]

Besides these methodological limitations, the proposed utility of these novel models is contentious. The authors suggest that the utilization of the described ML algorithm can reduce costs and risks associated with unbeneficial surgery. With an area under the curve of 0.759, this statement is inaccurate, as such models would refuse surgery to approximately one in four patients that would have benefited from CTD. Although this is a promising technology, it is still in its infancy and its potential at present has been overstated.

*Angelos Mantelakis, BSc (Hons), MBBS (Lond), MRCS (Eng)*
Guys and St Thomas's NHS Foundation Trust
Westminster Bridge Road, SE1 7EH
London, United Kingdom
E-mail: aggelosmantelakis@yahoo.com

## DISCLOSURE

*The authors have no financial interest to declare in relation to the content of this article.*

## REFERENCES

1. Harrison CJ, Geoghegan L, Sidey-Gibbons CJ, et al. Developing machine learning algorithms to support patient-centered, value-based carpal tunnel decompression surgery. *Plast Reconstr Surg Glob Open*. 2022;10:e4279.
2. Mantelakis A, Assael Y, Sorooshian P, et al. Machine learning demonstrates high accuracy for disease diagnosis and prognosis in plastic surgery. *Plast Reconstr Surg Glob Open*. 2021;9:e3638.
3. Poon AIF, Sung JJY. Opening the black box of AI-Medicine. *J Gastroenterol Hepatol*. 2021;36:581–584.
4. Kanevsky J, Corban J, Gaster R, et al. Big data and machine learning in plastic surgery: a new frontier in surgical innovation. *Plastic Reconstr Surg*. 2016;137:890e–897e.
5. An C, Park YW, Ahn SS, et al. Radiomics machine learning study with a small sample size: single random training-test set split may lead to unreliable results. *PLoS One*. 2021;16:e0256152.

*From the *Guys and St Thomas's NHS Foundation Trust, London, UK; †Kellogg College, University of Oxford, UK; ‡Department of Surgery and Cancer, Imperial College London, London, UK.*



*Plast Reconstr Surg Glob Open 2022;10:e4494; doi: 10.1097/GOX.0000000000004494; Published online 25 August 2022.*