

Can artificial intelligence help Heart Teams make decisions?

Valentin Koch^{1,2*}, MSc

**Corresponding author: Helmholtz Munich, Ingolstädter Landstraße 1, 85764, Oberschleißheim, Munich, Germany.*

E-mail: valentin.koch@helmholtz-munich.de

The author's affiliations can be found at the end of this article.

In healthcare, where time is scarce and healthcare professionals have constrained schedules, the use of artificial intelligence (AI) could make healthcare professionals' lives easier. The integration of AI, in particular advanced language models like Chat Generative Pre-trained Transformer (ChatGPT), into medical decision-making processes could help streamline workflows and take some of the burden off healthcare professionals. We examine here a recent study that looks at the use of ChatGPT in augmenting the decision-making processes of Heart Teams in the treatment of severe aortic stenosis. The study suggests a high degree of alignment between ChatGPT and multidisciplinary Heart Teams in the management of severe aortic stenosis, showing how language models could impact healthcare in the future¹.

Chat-based AI tools such as ChatGPT are based on large language models (LLMs), a class of models able to process text and output a response tailored towards the request of the user. To train such a model, huge resources in terms of computational power and data are needed, which few companies can afford. As the investment in training such a model is very high, the best models are usually not publicly available and can only be accessed over an interface, often with a paid subscription only. In order not to let models output harmful content, companies instruct the models by appending custom text to the user's request that the end user cannot see and which is subject to constant optimisation. While the intention is good, this has implications: even when the underlying AI language model (e.g., GPT-4) stays the same, the behaviour of the service using this model (e.g., ChatGPT) can change without notice. Therefore, it is important to understand the difference between the former and the latter and to be aware of the implications.

The study presented in this issue of EuroIntervention by Salihu et al shows that ChatGPT aligns with Heart Team decisions in determining the treatment of a patient with

aortic stenosis, with a high agreement rate for a sample size of 150 patients¹. The authors use 14 key variables, including age, overall condition or valvular calcium score, to form a standardised report that is given to ChatGPT along with three possible treatment options it can choose from: transcatheter aortic valve implantation (TAVI), surgical aortic valve replacement (SAVR) or medical treatment. The overall agreement between the Heart Team and ChatGPT was 77%; the highest rate of agreement (90%) was achieved for suggested TAVI treatment. This is a remarkable result, even surpassing the performance of decision trees based on guidelines from the American Heart Association (AHA), with an agreement rate of 43%, or the European Society of Cardiology (ESC), with an agreement rate of 73%^{2,3}.

Article, see page e496

Patients between 70 and 80 years of age had a higher rate of misclassification by ChatGPT, which is interesting, as this is within an age range where European and US guidelines conflict: the ESC recommends surgical management for patients under the age of 75, while the AHA recommends a cutoff point of less than 65 years. It could be speculated that the behaviour of ChatGPT is a result of the training of a GPT-4 model that received mixed training signals for patients in this age group, but it could also be for different reasons that are less obvious. Furthermore, it is notable that ChatGPT did not suggest surgery for any of the patients where medical treatment was assigned by the Heart Team. However, seven of the patients were wrongfully suggested a TAVI treatment by ChatGPT. Of these seven cases, four had a high perioperative risk due to comorbidities, which was not reflected in the 14 variables that were provided to the model, showing some limitations of the presented approach. More detailed clinical data or even unstructured reports could potentially be incorporated, with hopes of improving the

performance of AI models, but this ultimately raises privacy concerns that would need to be addressed.

Despite the promising alignment with Heart Team recommendations, the reliance on external, proprietary AI models introduces uncertainties and a lack of transparency. The black box nature of these systems can lead to unexpected changes, affecting their utility and reliability in clinical settings. Ensuring reproducible and consistent results can only be guaranteed when full control over all steps – from data acquisition and processing to AI model governance – is given, which is usually not the case with an external service such as ChatGPT.

The authors rightfully conclude that “AI tools are not intended to replace clinicians but rather to support them in their decision-making process. The final clinical decision should remain in the hands of the healthcare provider, considering the patient’s unique clinical status and preferences”¹. The focus must always remain on the patient, a complex individual whose wellbeing is of the highest importance. The future may see AI becoming more powerful, specialised, and accessible, which will oblige a shift in how we trust and utilise these tools in healthcare. These opportunities are exciting but demand careful navigation to leverage the benefits of AI while prioritising patient care above all.

Author's affiliations

1. *Helmholtz Munich, Munich, Germany*; 2. *Technical University of Munich, Munich, Germany*

Conflict of interest statement

The author has no conflicts of interest to declare.

References

1. Salihu A, Meier D, Noirclerc N, Skolidis I, Mauler-Wittwer S, Recordon F, Kirsch M, Roguelov C, Berger A, Sun X, Abbe E, Marcucci C, Rancati V, Rosner L, Scala E, Rotzinger DC, Humbert M, Muller O, Lu H, Fournier S. A study of ChatGPT in facilitating Heart Team decisions on severe aortic stenosis. *EuroIntervention*. 2024;496-503.
2. Vahanian A, Beyersdorf F, Praz F, Milojevic M, Baldus S, Bauersachs J, Capodanno D, Conradi L, De Bonis M, De Paulis R, Delgado V, Freemantle N, Haugaa KH, Jeppsson A, Jüni P, Pierard L, Prendergast BD, Sádaba JR, Tribouilloy C, Wojakowski W. 2021 ESC/EACTS Guidelines for the management of valvular heart disease. *EuroIntervention*. 2022;17:e1126-96.
3. Otto CM, Nishimura RA, Bonow RO, Carabello BA, Erwin JP 3rd, Gentile F, Jneid H, Krieger EV, Mack M, McLeod C, O’Gara PT, Rigolin VH, Sundt TM 3rd, Thompson A, Toly C. 2020 ACC/AHA Guideline for the Management of Patients With Valvular Heart Disease: Executive Summary: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation*. 2021;143:e35-71.