ChatGPT takes on the european exam in core cardiology: an AI success story

I. Skalidis¹, A. Cagnina¹, O. Luangphiphat¹, O. Muller¹, E. Abbe², S. Fournier¹

¹University Hospital Centre Vaudois (CHUV), Lausanne, Switzerland
²Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland

Funding Acknowledgements: None.

Background: ChatGPT, the trending novel artificial intelligence has triggered ongoing debate regarding its capabilities. Recently, preliminary reports showed that answered correctly in the majority of questions of the United States Medical License Examinations (USMLE). However, its ability to succeed a more precise, challenging and high-stakes post-graduate test, such as the final exam for the completion of medical residency, like the European Exam in Core Cardiology (EECC), is not known yet.

Purpose: We sought to evaluate the performance of ChatGPT on EECC, to test its capability on a more demanding, high-stakes post-graduate exam in Cardiology training.

Methods: A total of 488 publicly-available single-answer multiple choice questions (MCQs) were randomly obtained from three different MCQs sources that are traditionally used for the preparation for the EECC: 88 from the sample exam questions released since 2018 from the official ESC website, 200 from the 2022 edition of StudyPRN and 200 from the Braunwald’s Heart Disease Review and Assessment (BHDRA). Questions containing audio or visual assets were excluded. After filtering, 362 MCQ items (ESC sample: 68, BHDRA:150, StudyPRN: 144) were included and considered as input source. False responses and indeterminate responses were considered as not correct.

Results: ChatGPT answered to 340 questions out of 362, with 22 indeterminate answers in total. The overall accuracy was 58.8% across all the question sources. More specifically, it demonstrated an accuracy for the ESC sample, BHDRA and StudyPRN of 61.7%, 52.6%, 63.8% respectively. It answered correctly 42/68 (4 indeterminate) of ESC sample questions, 79/150 (11 indeterminate) of the BHDRA and 92/144 (7 indeterminate) of the StudyPRN.

Conclusion: ChatGPT manages to correctly answer the majority of EECC’s questions and perform within the passing threshold range. Although it cannot yet process visual content, it can provide rational and correct answers to text-based inputs in most scenarios. The model may be able to efficiently handle a massive amount of acquired medical knowledge, but the current approach may not substitute for critical thinking, innovation and creativity; some of the key attributes that doctors are expected to showcase.

Performance of ChatGPT on EECC
You review a 45-year-old gentleman in the cardiology outpatient clinic who is under surveillance follow-up for a dilated thoracic aorta. He works as a teacher. He exercises regularly and is entirely asymptomatic. His father died aged 68 years old from an acute aortic dissection. His latest echo shows a bicuspid aortic valve with moderate aortic regurgitation. The left ventricle is not dilated and left ventricular ejection fraction is 65%. Magnetic resonance (MR) angiogram of the aorta shows effacement of the sino-tubular junction with maximum aortic root dimensions of $52 \times 54$ mm at the proximal ascending aorta. According to European Society of Cardiology guidelines, what is the next most appropriate course of action?

a. Echo and MR angiogram in 12 months
b. Exercise echocardiography
c. Initiate beta-blocker therapy
d. Refer to cardiothoracic for consideration of surgery
e. Transoesophageal echocardiography

Based on the information provided, the patient has a bicuspid aortic valve with moderate aortic regurgitation, a dilated thoracic aorta, and a family history of aortic dissection. The maximum aortic root dimension is greater than 50 mm, which exceeds the threshold for elective surgery according to European Society of Cardiology guidelines.

Therefore, the next most appropriate course of action would be to refer the patient to cardiothoracic surgery for consideration of surgery (option d). The surgery will aim to prevent progression to aortic dissection, which is associated with significant morbidity and mortality. In addition to surgery, the patient should receive pharmacologic therapy, such as beta-blockers, to slow down the progression of the aortic dilatation.