

## Comparative Evaluation of Chatbot Responses on Coronary Artery Disease

### Koroner Arter Hastalığına İlişkin Chatbot Yanıtlarının Karşılaştırmalı Değerlendirilmesi

#### ABSTRACT

**Objective:** Coronary artery disease (CAD) is the leading cause of morbidity and mortality globally. The growing interest in natural language processing chatbots (NLPCs) has driven their inevitable widespread adoption in healthcare. The purpose of this study was to evaluate the accuracy and reproducibility of responses provided by NLPCs, such as ChatGPT, Gemini, and Bing, to frequently asked questions about CAD.

**Methods:** Fifty frequently asked questions about CAD were asked twice, with a one-week interval, on ChatGPT, Gemini, and Bing. Two cardiologists independently scored the answers into four categories: comprehensive/correct (1), incomplete/partially correct (2), a mix of accurate and inaccurate/misleading (3), and completely inaccurate/irrelevant (4). The accuracy and reproducibility of each NLPC's responses were assessed.

**Results:** ChatGPT's responses were scored as 14% incomplete/partially correct and 86% comprehensive/correct. In contrast, Gemini provided 68% comprehensive/correct responses, 30% incomplete/partially correct responses, and 2% a mix of accurate and inaccurate/misleading information. Bing delivered 60% comprehensive/correct responses, 26% incomplete/partially correct responses, and 8% a mix of accurate and inaccurate/misleading information. Reproducibility scores were 88% for ChatGPT, 84% for Gemini, and 70% for Bing.

**Conclusion:** ChatGPT demonstrates significant potential to improve patient education about coronary artery disease by providing more sensitive and accurate answers compared to Bing and Gemini.

**Keywords:** Artificial intelligence, Bing Chat, ChatGPT, coronary artery disease, digital health, Gemini, natural language processing chatbots

#### ÖZET

**Amaç:** Koroner arter hastalığı (KAH), dünya çapında morbidite ve mortalitenin önde gelen nedenidir. Doğal dil işleyen sohbet robotlarına (NLPC) olan ilginin artması, bunların sağlık alanında kaçınılmaz olarak yaygın bir şekilde benimsenmesini sağlamıştır. Bu çalışmanın amacı ChatGPT, Gemini ve Bing gibi NLPC'lerin CAD ile ilgili sık sorulan sorulara verdiği yanıtların doğruluğunu ve tekrarlanabilirliğini kontrol etmektir.

**Yöntem:** KAH ile ilgili sık sorulan 50 soru ChatGPT, Gemini ve Bing üzerinden 1 hafta aryla 2 kez soruldu. İki kardiyolog bağımsız olarak cevapları 4 gruba ayırdı: kapsamlı/doğru (1), eksik/kısmen doğru (2), doğru ve yanlış/yanıltıcı karışımı (3) ve tamamen yanlış/ilgisiz (4). Her NLPC'nin yanıtlarının doğruluğu ve tekrarlanabilirliği değerlendirildi.


**Bulgular:** ChatGPT'nin puanlaması %14 eksik/kısmen doğru ve %86 kapsamlı/doğruydü. Öte yandan Gemini %68 kapsamlı/doğru yanıt, %30 eksik/kısmen doğru yanıt ve %2 karma doğru ve yanlış/yanıltıcı yanıt verdi. Son olarak Bing, %60 kapsamlı/doğru yanıtlar, %26 eksik/kısmen doğru yanıtlar ve %8 doğru ve yanlış/yanıltıcı bilgilerin bir karışımı olan yanıtlar verdi. Tekrarlanabilirlik değerleri ChatGPT için %88, Gemini için %84 ve Bing için %70 idi.

**Sonuç:** ChatGPT, Bing ve Gemini ile karşılaştırıldığında daha hassas ve doğru cevaplar sağlayarak koroner arter hastalığı hakkında hasta eğitimini artırmada önemli bir potansiyele sahiptir.

**Anahtar Kelimeler:** Yapay zeka, Bing Chat, ChatGPT, koroner arter hastalığı, dijital sağlık, gemini, doğal dil işleme sohbet robotları

#### ORIGINAL ARTICLE KLİNİK ÇALIŞMA

Levent Pay<sup>1</sup> 

Ahmet Çağdaş Yumurtaş<sup>2</sup> 

Tuğba Çetin<sup>3</sup> 

Tufan Çınar<sup>4</sup> 

Mert İlker Hayiroğlu<sup>3</sup> 

<sup>1</sup>Department of Cardiology, Istanbul Haseki Training and Research Hospital, Istanbul, Türkiye

<sup>2</sup>Department of Cardiology, Kars Harakani State Hospital, Kars, Türkiye

<sup>3</sup>Department of Cardiology, Dr Siyami Ersek Thoracic and Cardiovascular Surgery Training Hospital, Istanbul, Türkiye

<sup>4</sup>Department of Medicine, University of Maryland Medical Center Midtown Campus, Maryland, USA

#### Corresponding author:

Levent Pay

✉ leventpay@hotmail.com

Received: July 19, 2024

Accepted: October 15, 2024

**Cite this article as:** Pay L, Yumurtaş AÇ, Çetin T, Çınar T, Hayiroğlu Mİ. Comparative Evaluation of Chatbot Responses on Coronary Artery Disease. *Türk Kardiyol Dern Ars.* 2025;53(1):35-43.

DOI:10.5543/tkda.2024.78131



Available online at [archivestsc.com](http://archivestsc.com). Content of this journal is licensed under a Creative Commons Attribution - NonCommercial-NoDerivatives 4.0 International License.

Despite advancements in treatment options, coronary artery disease (CAD) is extremely common and is the leading cause of morbidity and mortality worldwide. According to recent data, cardiovascular diseases caused slightly fewer than 2.2 million deaths in women and slightly more than 1.9 million deaths in men last year.<sup>1</sup> The public availability of generative artificial intelligence (AI) has ushered in a new era in the acquisition of medical knowledge. The increasing interest in natural language processing chatbots (NLPCs) has made their widespread use in healthcare inevitable. While this transition highlights the positive role AI can play in enhancing health literacy, it has also raised concerns about the potential for misuse and misinformation.<sup>2</sup>

Released on May 13, 2024, ChatGPT version 4o offers significantly advanced features compared to its previous versions, although it requires a monthly subscription fee. The accuracy of ChatGPT's responses to frequently asked questions about atrial fibrillation, hypertension, and heart failure has been studied.<sup>3-5</sup> However, there is no research available in the literature regarding CAD. In this study, we aimed to evaluate the quality of responses provided by various NLPCs, including ChatGPT 4o, Gemini, and Bing Chat, to questions frequently asked by patients about CAD.

## Materials and Methods

### Data Collection

In this study, fifty commonly asked patient questions about CAD were identified. These questions were primarily sourced from the frequently asked questions (FAQs) sections on the websites of the Cleveland Clinic, Mayo Clinic, and the National Health Service (NHS) UK, based on the most common inquiries from patients. These questions were categorized into five subgroups: basic information, diagnosis, treatment, recovery (including operative risks/complications and follow-up), and prevention/dietary aspects. Questions are shown in Table 1, divided into categories. The questions were entered into the online interfaces of three different NLPC systems; (i) ChatGPT 4o (GPT-4, OpenAI, San Francisco, California, United States); (ii) Gemini (PaLM 2, Google, Mountain View, California, United States); (iii) Bing Chat (BC) (Microsoft, Redmond, Washington, United States). Two experienced cardiologists independently, and without knowledge of each other's assessments, evaluated the responses to each question from the different NLPC systems, focusing on various aspects. The answers were independently rated on a scale of comprehensive/correct (1), incomplete/partially correct (2), a mix of accurate and inaccurate/misleading (3), and completely inaccurate/irrelevant (4). Any discrepancies were resolved by a third reviewer.

The ChatGPT 4o, Gemini, and Bing chatbots were asked 50 questions on July 3, 2024 and again on July 10, 2024. The same questions were posed to the chatbots with a one-week

interval. The reproducibility of the responses was assessed by categorizing them into two groups based on their ratings. A response was deemed non-reproducible if the answers fell into different categories, indicating variability in the accuracy of the responses. Each question was entered separately using the "New Chat" feature. All questions were asked in English. The study was conducted in full compliance with the Declaration of Helsinki.

### Statistical Analysis

Descriptive analysis was used to present proportions as counts and percentages. Statistical analyses were conducted using Microsoft Excel (version 16.68; Microsoft Corporation, Redmond, Washington, United States) to categorize the scores of the NLPCs' responses. The answers to each question were evaluated according to current guidelines to obtain percentage-based classifications.

### Results

Overall, the scoring frequencies for ChatGPT-4 were as follows: 14% (7 answers) incomplete/partially correct and 86% (43 answers) comprehensive/correct. There were no "completely inaccurate/irrelevant" or a "mix of accurate and inaccurate/misleading" responses from ChatGPT-4. By category, the model provided "comprehensive/correct" answers to 78.5% of questions regarding "basic knowledge," 80% related to "diagnosis," 80% related to "treatment," 92.3% related to "recovery – operative risks – complications – follow-up," and 100% related to "prevention/dietary." Table 2 shows the accuracy levels of ChatGPT's responses by topic.

Gemini provided 68% (34 answers) comprehensive/correct responses, 30% (15 answers) incomplete/partially correct responses, and 2% (1 answer) mix of accurate and inaccurate/misleading responses. There were no "completely inaccurate/irrelevant" responses. According to the category breakdown, Gemini provided "comprehensive/correct" answers to 50% of questions on "basic knowledge," 60% on "diagnosis," 70% on "treatment," 84.6% on "recovery, operative risks, complications, follow-up," and 75% on "prevention/dietary." Table 3 shows the accuracy levels of Gemini's responses by topic.

Bing provided 60% (30 answers) comprehensive/correct responses, 26% (13 answers) incomplete/partially correct responses, and 8% (4 answers) mix of accurate and inaccurate/misleading responses. Additionally, Bing did not provide medical advice for 3 (6%) of the questions, instead emphasizing the importance of consulting a healthcare professional. There were no "completely inaccurate/irrelevant" responses. By category, Bing provided "comprehensive/correct" answers to 64.3% of questions on "basic knowledge," 60% on "diagnosis," 50% on "treatment," 61.5% on "recovery, operative risks, complications, follow-up," and 62.5% on "prevention/dietary." Table 4 shows the accuracy levels of Bing's responses by topic.

When assessing reproducibility, the models consistently provided reproducible answers for most questions. ChatGPT achieved a score of 88%, Gemini 84%, and Bing 70% (Table 5). Table 6 shows the chatbots' responses and accuracy for two sample questions.

## ABBREVIATIONS

AI	Artificial intelligence
CAD	Coronary artery disease
FAQs	Frequently asked questions
NHS	The National Health Service
NLPCs	Natural language processing chatbots

**Table 1. Categorization of Questions****Basic Knowledge**

1. What is coronary artery disease?
2. What are the types of coronary artery disease?
3. How common is coronary artery disease?
4. What are the symptoms of coronary artery disease?
5. What causes coronary artery disease?
6. Is coronary artery disease genetic?
7. What are the risk factors for coronary artery disease?
8. Can coronary artery disease be reversed?
9. What is a heart attack?
10. What is atherosclerosis?
11. What are stents?
12. How does smoking affect the heart and blood vessels?
13. What is angina?
14. What to do if someone has a heart attack?

**Diagnosis**

15. How is coronary artery disease diagnosed?
16. What tests will be done to diagnose coronary artery disease?
17. Why is a coronary calcium scan done?
18. I have a very high calcium score. What should I do next?
19. My father had a stent at the age of 50, and I am 45 years old. What should I do?

**Treatment**

20. How is coronary artery disease treated?
21. I suffer from chest pain after coronary angiography. What are my treatment options?
22. What is coronary artery bypass grafting?
23. Coronary artery disease: Angioplasty or bypass surgery?
24. Angina treatment: Stents, drugs, or lifestyle changes—what's best?
25. Should I take an aspirin a day?
26. I suffer from coronary artery disease (CAD). How long should I continue antiplatelet treatment?
27. What is angiography and how is it done?
28. What are the differences between femoral and radial approaches in angiography?
29. I suffer from CAD. Which is better: balloon or stent?

**Recovery – Operative Risks/Complications – Follow-up**

30. What are the complications of coronary artery disease?
31. How long does it take to recover from coronary artery disease treatment?
32. What can I expect if I have coronary artery disease?
33. How do I take care of myself?
34. When should I see my healthcare provider? When to see a doctor/How long do I need to wait to see a cardiologist?
35. When should I go to the emergency room (ER)?
36. What questions should I ask my doctor?
37. How can cardiac rehabilitation help?
38. What types of exercises are best for a patient with a history of coronary artery disease?
39. I suffer from CAD. Can I fly on a plane?
40. I suffer from CAD. Can I dive?
41. I suffer from CAD and have a stent. Can I have an MRI?
42. What are the risks of angiography?

**Prevention/Dietary**

43. Can coronary artery disease be prevented?
44. I am 40 years old. Should I screen myself for coronary artery disease?
45. I am 70 years old. Should I screen myself for coronary artery disease?
46. I suffer from CAD. Should I avoid alcohol?
47. I suffer from CAD. Should I avoid tobacco?
48. I suffer from CAD. What should I eat?
49. I suffer from CAD. What should I not eat?
50. I suffer from CAD. Can I drink coffee and how much?

**Table 2. Accuracy of ChatGPT's Responses by Topics**

	Comprehensive/Correct	Incomplete/Partially Correct
Basic Knowledge (n = 14)	11 (78.5%)	3 (21.5%)
Diagnosis (n = 5)	4 (80%)	1 (20%)
Treatment (n = 10)	8 (80%)	2 (20%)
Recovery – Operative Risks – Complications – Follow-Up (n = 13)	12 (92.3%)	1 (7.3%)
Prevention/Dietary (n = 8)	8 (100%)	0 (0%)
Total (n = 50)	43 (86%)	7 (14%)

**Table 3. Accuracy Levels of Gemini's Responses by Topics**

	Comprehensive/Correct	Incomplete/Partially Correct	Mix of Accurate and Inaccurate/Misleading Information
Basic Knowledge (n = 14)	7 (50%)	7 (50%)	0 (0%)
Diagnosis (n = 5)	3 (60%)	2 (40%)	0 (0%)
Treatment (n = 10)	7 (70%)	2 (20%)	1 (10%)
Recovery – Operative Risks – Complications – Follow-Up (n = 13)	11 (84.6%)	2 (15.4%)	0 (0%)
Prevention/Dietary (n = 8)	6 (75%)	2 (25%)	0 (0%)
Total (n = 50)	34 (68%)	15 (30%)	1 (2%)

**Table 4. Accuracy Levels of Bing's Responses by Topics**

	Comprehensive/Correct	Incomplete/Partially Correct	Mix of Accurate and Inaccurate/Misleading Information	No Suggestions
Basic Knowledge (n = 14)	9 (64.3%)	4 (28.6%)	1 (7.1%)	0 (0%)
Diagnosis (n = 5)	3 (60%)	2 (40%)	0 (0%)	0 (0%)
Treatment (n = 10)	5 (50%)	3 (30%)	1 (10%)	1 (10%)
Recovery – Operative Risks – Complications – Follow-Up (n = 13)	8 (61.5%)	1 (7.7%)	2 (15.3%)	2 (15.3%)
Prevention/Dietary (n = 8)	5 (62.5%)	3 (37.5%)	0 (0%)	0 (0%)
Total (n = 50)	30 (60%)	13 (26%)	4 (8%)	3 (6%)

**Table 5. Reproducibility of ChatGPT, Gemini, and Bing Responses to Coronary Artery Disease-Related Questions by Question Subgroup**

Question Subgroup	ChatGPT	Gemini	Bing
	Number of Responses (%)		
Basic Knowledge (n = 14)	12 (85.7%)	11 (78.5%)	11 (78.5%)
Diagnosis (n = 5)	5 (100%)	4 (80%)	4 (80%)
Treatment (n = 10)	9 (90%)	8 (80%)	6 (60%)
Recovery – Operative Risks – Complications – Follow-Up (n = 13)	11 (84.6%)	12 (92.3%)	8 (61.5%)
Prevention/Dietary (n = 8)	7 (87.5%)	7 (87.5%)	6 (75%)
Total (n = 50)	44 (88%)	42 (84%)	35 (70%)

**Table 6. Chatbots' Responses and Accuracy for Two Sample Questions****Question:** How common is coronary artery disease?**ChatGPT's Response**

Comprehensive/Correct

Coronary artery disease (CAD) is quite common and is a leading cause of morbidity and mortality worldwide. Here are some key points about its prevalence:

## 1. Global Prevalence:

- CAD is one of the most common cardiovascular diseases globally. According to the World Health Organization (WHO), cardiovascular diseases, including CAD, are the leading cause of death worldwide, accounting for approximately 17.9 million deaths each year.

## 2. Prevalence in the United States:

- In the United States, CAD is also highly prevalent. According to the American Heart Association (AHA), about 18.2 million adults aged 20 and older have CAD. It is responsible for about 1 in 4 deaths in the country, making it the leading cause of death for both men and women.

## 3. Age and Gender Factors:

- The risk of developing CAD increases with age. It is more common in men over the age of 45 and women over the age of 55. However, post-menopausal women have a higher risk compared to pre-menopausal women.

## 4. Geographical Variations:

- The prevalence of CAD varies across different regions and populations due to differences in risk factors such as diet, lifestyle, and healthcare access. Developed countries tend to have higher rates due to higher prevalence of risk factors like sedentary lifestyle and unhealthy diet, but developing countries are seeing increasing rates as well due to lifestyle changes.

## 5. Trends:

- While the incidence of CAD has been declining in many high-income countries due to better prevention and treatment strategies, it remains a significant public health issue. In contrast, low- and middle-income countries are experiencing a rise in CAD prevalence due to urbanization and lifestyle changes.

Understanding these statistics highlights the importance of prevention, early detection, and management of CAD to reduce its impact on public health.

**Gemini's Response**

Comprehensive/Correct

Coronary artery disease (CAD) is very common. In the United States alone, estimates suggest around 20.5 million adults have it, according to the Centers for Disease Control and Prevention (CDC) [National Institutes of Health (.gov)]. That's a significant number of people, roughly the size of several major cities combined. This makes it the most prevalent form of heart disease in the country.

3b. Coronary artery disease (CAD) is a very prevalent health concern around the world. Here's a look at how common it is:

## Global Burden:

- According to a 2022 study published in the Journal of the American College of Cardiology [1], there were an estimated 315 million prevalent cases of CAD globally in 2022.
- This translates to a significant portion of the world's population being affected.

## Age-Standardized Prevalence:

- The same study also reports an age-standardized prevalence of 3605 cases of CAD per 100,000 people globally in 2022.
- Age-standardized prevalence means the statistic takes into account the age distribution of the population, providing a more accurate comparison across different regions.

## Regional Variations:

- The prevalence of CAD shows some variation across different parts of the world. The 2022 study found the highest age-standardized prevalence in Central Europe, Eastern Europe, and Central Asia (8019 per 100,000), while South Asia had the lowest (2393 per 100,000).
- These variations can be attributed to factors like lifestyle habits, access to healthcare, and socioeconomic conditions.

## Trends Over Time:

- The good news is that there has been a slight decrease in the prevalence of CAD globally. The 2022 study showed an 18% decrease compared to 1990 (4390 per 100,000).
- This decrease can be partially attributed to improved public health awareness, risk factor management, and advancements in medical treatments.

## Overall:

Coronary artery disease is a widespread health problem affecting millions of people globally. However, there have been positive trends in recent years with a decrease in prevalence.

Here are some resources you can explore for further information:

- [1] Global Prevalence of Coronary Artery Disease: An Update From the Global Burden of Disease Study - Journal of the American College of Cardiology <https://www.jacc.org/doi/10.1016/S0735-1097%2824%2904310-9>
- World Health Organization: Cardiovascular diseases <https://www.who.int/health-topics/cardiovascular-diseases>

Bing Chat's Response	Comprehensive/Correct
<p>Coronary artery disease (CAD) is a prevalent condition worldwide. Here are some key statistics:</p> <ol style="list-style-type: none"> <li>Global Prevalence: <ul style="list-style-type: none"> <li>In 2021, CAD was the most common type of heart disease, causing 375,476 deaths<sup>1</sup>.</li> <li>Approximately 1 in 20 adults aged 20 and older (about 5%) have CAD<sup>1</sup>.</li> <li>About 2 in 10 CAD deaths occurred in adults under 65 years old<sup>1</sup>.</li> </ul> </li> <li>Gender Differences: <ul style="list-style-type: none"> <li>Globally, males have a higher age-standardized CVD death rate (280.8 deaths per 100,000 people in 2019) than females (204.0 deaths per 100,000 people in 2019)<sup>2</sup>.</li> </ul> </li> <li>Geographic Variation: <ul style="list-style-type: none"> <li>Around 4 in every 5 CVD deaths occur in low- and middle-income countries<sup>2</sup>.</li> <li>Progress in cardiovascular health is increasingly concentrated in high-income countries, highlighting health inequities<sup>2</sup>.</li> </ul> </li> </ol> <p>Remember that early detection and management are crucial for maintaining heart health. If you have concerns, consult a healthcare professional.</p>	
<b>Question:</b> Can coronary artery disease be reversed?	
ChatGPT's Response	Incomplete/Partially Correct
<p>While coronary artery disease (CAD) cannot be completely "cured," it can be managed effectively, and in some cases, the progression of the disease can be slowed or even partially reversed. Here are some strategies and approaches that can help in managing and potentially reversing aspects of CAD:</p> <ol style="list-style-type: none"> <li>Lifestyle Changes <ul style="list-style-type: none"> <li>Diet: <ul style="list-style-type: none"> <li>Plant-Based Diet: Diets rich in fruits, vegetables, whole grains, legumes, nuts, and seeds can help reduce cholesterol levels and improve heart health.</li> <li>Low-Fat Diet: Reducing intake of saturated fats and trans fats can lower LDL cholesterol.</li> <li>Omega-3 Fatty Acids: Consuming fish or taking fish oil supplements can have beneficial effects on heart health.</li> </ul> </li> <li>Exercise: <ul style="list-style-type: none"> <li>Regular Physical Activity: Engaging in regular aerobic exercise (such as walking, jogging, swimming, or cycling) can improve cardiovascular fitness, reduce blood pressure, and lower cholesterol levels.</li> </ul> </li> <li>Weight Management: <ul style="list-style-type: none"> <li>Healthy Weight: Achieving and maintaining a healthy weight reduces the burden on the heart and helps manage risk factors such as hypertension and diabetes.</li> </ul> </li> <li>Smoking Cessation: <ul style="list-style-type: none"> <li>Quit Smoking: Stopping smoking is crucial for improving cardiovascular health and reducing the risk of further artery damage.</li> </ul> </li> <li>Stress Management: <ul style="list-style-type: none"> <li>Stress Reduction: Techniques such as yoga, meditation, deep breathing exercises, and mindfulness can help manage stress, which is beneficial for heart health.</li> </ul> </li> </ul> </li> <li>Medications <ul style="list-style-type: none"> <li>Cholesterol-Lowering Drugs: <ul style="list-style-type: none"> <li>Statins: These medications can significantly reduce LDL cholesterol levels and stabilize plaques in the arteries, reducing the risk of heart attacks.</li> <li>Blood Pressure Medications: <ul style="list-style-type: none"> <li>ACE Inhibitors, ARBs, Beta-Blockers, and Calcium Channel Blockers: These help control blood pressure, which is vital for reducing the strain on the heart and preventing further damage to the arteries.</li> </ul> </li> <li>Antiplatelet Agents: <ul style="list-style-type: none"> <li>Aspirin or Clopidogrel: These medications help prevent blood clots, reducing the risk of heart attacks.</li> </ul> </li> </ul> </li> <li>Medical Procedures <ul style="list-style-type: none"> <li>Angioplasty and Stent Placement: <ul style="list-style-type: none"> <li>Reopening Arteries: These procedures can restore blood flow in blocked arteries, providing immediate relief from symptoms and improving heart function.</li> </ul> </li> <li>Coronary Artery Bypass Grafting (CABG): <ul style="list-style-type: none"> <li>Bypassing Blockages: This surgery can create new pathways for blood flow around blocked arteries, significantly improving blood supply to the heart.</li> </ul> </li> </ul> </li> <li>Intensive Cardiac Rehabilitation Programs <ul style="list-style-type: none"> <li>Ornish or Pritikin Programs: <ul style="list-style-type: none"> <li>These programs focus on comprehensive lifestyle changes, including diet, exercise, stress management, and group support, and have shown success in reversing some aspects of CAD.</li> </ul> </li> </ul> </li> <li>Regular Monitoring and Follow-Up <ul style="list-style-type: none"> <li>Regular Check-Ups: <ul style="list-style-type: none"> <li>Monitoring heart health regularly with your healthcare provider helps manage risk factors effectively and make timely adjustments to your treatment plan.</li> </ul> </li> </ul> </li> <li>Emerging Therapies <ul style="list-style-type: none"> <li>Research and Innovations: <ul style="list-style-type: none"> <li>Ongoing research is exploring new treatments and interventions that might offer additional ways to manage or reverse aspects of CAD in the future.</li> </ul> </li> </ul> </li> </ul></li></ol> <p>While CAD cannot be completely reversed, many patients can experience significant improvements in symptoms, a reduction in the progression of the disease, and a lower risk of heart attacks and other complications through a combination of these strategies. It is essential to work closely with your healthcare provider to develop and adhere to a personalized treatment plan.</p>	



**Gemini's Response**

Incomplete/Partially Correct

Unfortunately, there is currently no cure for coronary artery disease (CAD) that completely reverses the existing plaque buildup in the arteries. Once plaque accumulates and hardens on the artery walls, it cannot be fully removed [Mayo Clinic, National Institutes of Health (.gov)]. However, this doesn't mean there's no hope! Here's the good news:

- Management and improvement: Even though complete reversal isn't possible, CAD can be effectively managed and its progression slowed significantly. Through lifestyle changes, medications, and sometimes procedures, you can improve blood flow to the heart, prevent complications like heart attacks, and live a long and healthy life.
- Lifestyle changes: Maintaining a heart-healthy diet, regular exercise, weight management, and stress management can significantly improve your heart health and slow the progression of CAD.
- Medications: Medications like statins can help lower cholesterol levels and reduce the risk of further plaque buildup.
- Procedures: In some cases, procedures like angioplasty and stenting can open blocked arteries and improve blood flow.

Here are some studies showing progress in managing CAD:

- High-dose statins: A 2006 study called ASTEROID found that high-dose statins might be able to reverse some plaque buildup in arteries, though it highlights the need for very low LDL cholesterol levels [WebMD].
- Research on plaque reversal: There's ongoing research into new methods to potentially reverse plaque buildup, but these are still under investigation [MedicalNewsToday].

If you have CAD, the focus should be on managing the condition to prevent complications and improve your heart health. Talk to your doctor about the best course of treatment for your individual situation.

15b. Coronary artery disease (CAD) can't be completely reversed in the sense of clearing existing plaque buildup from your arteries. However, you can effectively manage CAD and even see some improvement in the function of your arteries. Here's a breakdown of what to know: Current Limitations:

- Plaque buildup in arteries is a complex process, and there's no current medical treatment that can directly reverse it.

Management Strategies:

- The focus of CAD treatment is to prevent further plaque buildup, improve blood flow to the heart, and reduce your risk of complications:
  - Lifestyle changes: A heart-healthy diet, regular exercise, weight management, smoking cessation, and stress management are crucial for slowing CAD progression.
  - Medications: Medications like statins can help lower LDL cholesterol and potentially stabilize existing plaque.
  - Procedures: In some cases, minimally invasive procedures (PCI) or bypass surgery (CABG) can open blocked arteries and improve blood flow.

Potential for Improvement:

- While existing plaque can't be fully eliminated, research is ongoing for potential future treatments that might target plaque reduction.
- Studies have shown that some medications and aggressive lifestyle changes might even lead to a slight decrease in plaque volume in certain cases. However, these reductions haven't been definitively linked to a lower risk of heart problems.

Importance of Management:

- Even though complete reversal isn't possible, effectively managing CAD can significantly improve your heart health and quality of life.
- By managing CAD, you can:
  - Reduce your risk of heart attack, stroke, and other complications.
  - Improve your symptoms like chest pain and shortness of breath.
  - Increase your exercise tolerance and overall well-being.

The takeaway:

- There's currently no way to completely reverse CAD.
- However, with proper management and lifestyle changes, you can significantly slow its progression, improve your heart health, and live a long and fulfilling life.

Here are some resources you can explore for further information:

- American Heart Association: Coronary Artery Disease <https://www.heart.org/>
- National Heart, Lung, and Blood Institute: Coronary Artery Disease <https://www.nhlbi.nih.gov/science/heart-and-vascular-diseases>

**Bing Chat's Response**

Incomplete/Partially Correct

Certainly! Coronary artery disease (CAD) can sometimes be partially reversed through a combination of lifestyle changes and medical treatments. Here's what you need to know:

1. Lifestyle Modifications:

- Healthy Diet: Adopting a heart-healthy diet low in saturated fats, trans fats, and cholesterol can slow down or even reverse the progression of CAD.
- Regular Exercise: Physical activity improves blood flow, strengthens the heart, and helps manage risk factors.
- Smoking Cessation: Quitting smoking significantly reduces CAD risk.
- Weight Management: Maintaining a healthy weight is crucial.
- Stress Reduction: Chronic stress can worsen CAD, so stress management techniques are essential.

2. Medications:

- Statins: These drugs lower cholesterol levels and stabilize plaques.
- Antiplatelet Agents: Aspirin or other medications prevent blood clots.
- Blood Pressure Medications: Controlling hypertension is vital.
- Nitroglycerin: It helps relieve chest pain (angina).

3. Medical Procedures:

- Angioplasty and Stenting: These procedures open narrowed arteries and improve blood flow.
- Coronary Artery Bypass Surgery: In severe cases, bypass grafts can restore blood flow.

4. Limitations:

- While these interventions can improve symptoms and slow disease progression, complete reversal is rare.
- Existing damage to the heart muscle may not fully recover.

Remember, consult your healthcare provider for personalized advice. Lifestyle changes play a significant role, so take proactive steps to protect your heart!

## Discussion

GPT-4o answered most questions about coronary artery disease accurately and reliably compared to Gemini and Bing. To the best of our knowledge, this is the first study in the literature to evaluate the responses of NLPCs on coronary artery disease.

Today, the internet has become a source of health-related information that is easily accessible to everyone. As more users adopt AI tools like ChatGPT, Gemini, and Bing, patients are increasingly expected to rely on these tools for information about their medical conditions. As AI capabilities advance, these tools have the potential to effectively maximize benefits and minimize risks in AI-driven patient education. However, rigorous testing and validation by regulatory medical agencies are necessary to ensure patient safety and equitable access to healthcare resources through this technology. The United States Food and Drug Administration has proposed regulating these AI technologies through continuous monitoring and periodic verification.<sup>6</sup> This regulation aims to ensure that these dynamic tools are validated for their intended purposes and undergo ongoing re-evaluation for security.

NLPCs, especially ChatGPT, have gained widespread popularity due to their advanced conversational abilities.<sup>7,8</sup> ChatGPT is an AI model that has been trained on a comprehensive dataset covering various subjects, including medicine. It provides responses in a conversational format to text-based questions posed by users.<sup>9</sup> As rapid advancements in artificial intelligence have led to the development of sophisticated large language models, their potential benefits to the scientific aspects of medicine have been highlighted in recent publications. While NLPCs can be used to collect reliable medical content, they also carry the potential risk of spreading misinformation.

Using artificial intelligence for patient education on coronary artery disease has the potential to be a valuable tool in raising patient awareness. Research is actively exploring the model's utility in cardiology, with studies showing promising results in its ability to address questions related to heart disease prevention, atrial fibrillation, and heart failure.<sup>3,4,10</sup> In this study, we evaluated the quality of ChatGPT, Gemini, and Bing's responses to questions about coronary artery disease for patient education. While 86% of the responses from ChatGPT were deemed comprehensive/correct, only 68% of Gemini's responses and 60% of Bing's responses were considered comprehensive/correct. ChatGPT consistently provided accurate and reproducible information for various medical inquiries related to coronary artery disease. Suitability varied across NLPCs; ChatGPT demonstrated the highest suitability, while Bing Chat exhibited the lowest.

Additionally, ChatGPT and Gemini showed high reproducibility across all subjects, while Bing had lower reproducibility. This may be attributed to ChatGPT's and Gemini's design, which emphasizes stability, consistency, and specific programming choices. ChatGPT typically operates independently of previous conversations or personal data, and it tends to provide very similar or identical responses to the same input. The model applies minimal randomness in its responses, which further enhances its tendency to generate consistent replies to identical inputs. On the other hand, the lower reproducibility observed in Bing

Chat may be due to its use of a broader range of data and a more dynamic inference process. Additionally, Bing Chat sometimes considers previous conversations or user history when generating responses. Personalization and context management can lead to different answers being provided for the same question.

Our study findings indicate that ChatGPT could serve as a dependable source of information for individuals seeking advice on CAD. However, patients should exercise caution when relying on artificial intelligence for medical advice and be aware of the limitations of chatbots as their integration into healthcare expands. With technological advances, multi-modal AI systems such as ChatGPT may one day become routinely used, providing patients with comprehensive information about their conditions.

Some studies in the literature have shown that ChatGPT provides less definitive answers to questions about clinical guideline strategies compared to other chatbots.<sup>11,12</sup> However, there are also studies demonstrating that ChatGPT shows higher accuracy rates compared to other chatbots in various contexts.<sup>3,5,13</sup> It is possible that lower accuracy rates observed in some studies may be attributed to the use of older versions of the technology. Our study utilized ChatGPT-4, released in May 2024, which represents a more advanced version compared to its predecessors. This advancement aligns with our findings, which indicate superior performance relative to other chatbots. In our comparison of chatbot responses to frequently asked questions about coronary artery disease, ChatGPT-4 provided more accurate results. Nonetheless, these findings should be validated through more extensive and larger studies.

Consequently, NLPCs have the potential to enhance access to medical information by offering easily understandable medical advice while considering all implications. As more patients utilize online healthcare tools, these resources can aid in patient education and increase their awareness of their conditions. However, due to the variations among different chatbots, the use of such tools should currently be considered off-label. This is because natural language processing (NLP) systems are primarily designed for general-purpose language generation rather than for providing scientific or medical advice. Given the sensitivity of health-related information and the necessity for accuracy, it is recommended that AI be supplemented by human supervision to ensure the provision of accurate and reliable information.

Artificial intelligence-based Clinical Decision Support Systems (CDSS) can enhance the efficiency of healthcare services and aid in making accurate diagnoses and treatments. It is speculated that large language models may also become part of CDSS in the future.<sup>14</sup> On the other hand, human ethics presents a significant challenge in the use of these systems.<sup>15</sup> Data privacy, bias, and model interpretability are among the primary concerns. Healthcare professionals need to follow the decisions made by CDSS, and feedback mechanisms should be established to correct potential errors.<sup>16</sup>

With further validation, NLPCs can become invaluable resources for quickly retrieving medical information in clinical settings and supporting clinicians in complex decision-making processes.



However, we are still in the early stages of AI-enabled medical services, and it is not yet recommended to rely exclusively on current versions of chatbots for medical information. The medical relevance of the responses, in terms of alignment with current guideline recommendations and good clinical practice, is limited and varies significantly across different NLPCs. Patients and caregivers should thus be aware that, at this stage, chatbots should be used cautiously for gathering medical information about CAD. As NLPCs continue to evolve rapidly, their accuracy in providing medical information needs to be continually reassessed. To effectively utilize NLPCs in patient education, it is crucial to integrate peer-reviewed content and implement a robust control mechanism.

Given the current limitations, the risk of patients and physicians over-relying on NLPCs and the potential for the spread of misinformation is a significant concern. It is important to demonstrate the accuracy and reliability of NLPCs through more comprehensive studies. Specifying the reliability rate of responses provided by chatbot applications could help mitigate this concern.

### Limitations

First, despite employing a rating system with a panel of multiple judges, bias may persist due to subjective evaluation. Second, the reviewers in the study may have had similar professional backgrounds, as they worked at the same academic institution. Third, the current version of the chatbot has an information cutoff as of April 2023, and the omission of data from the past year could potentially affect the accuracy of its answers. Fourth, since experts, not patients, were asked to evaluate the answers, the results may not fully reflect the patient's perspective. Finally, although the questions used in our study were sourced from the websites of highly respected institutions, they may not universally reflect the concerns of all patients.

### Conclusion

ChatGPT has the potential to provide important answers to frequently asked patient questions about coronary artery disease. However, ChatGPT should be used cautiously for collecting medical information about coronary artery disease, as it may occasionally omit important details in its responses. Future, more comprehensive studies, including additional NLPCs, will provide a broader perspective on the healthcare performance of these tools.

**Peer-review:** Externally peer-reviewed.

**Author Contributions:** Concept – L.P.; Design – L.P.; Supervision – T.Çınar, M.İ.H.; Resource – A.Ç.Y., T.Çetin; Materials – L.P., A.Ç.Y.; Data Collection and/or Processing – L.P., T.Çetin; Analysis and/or Interpretation – A.Ç.Y., T.Çetin; Literature Review – L.P.; Writing – L.P., M.İ.H.; Critical Review – T.Çınar, M.İ.H.

**Use of AI for Writing Assistance:** Artificial intelligence-supported technologies were not used in the writing phase of this study.

**Conflict of Interest:** The authors have no conflicts of interest to disclose.

**Funding:** This research received no external funding.

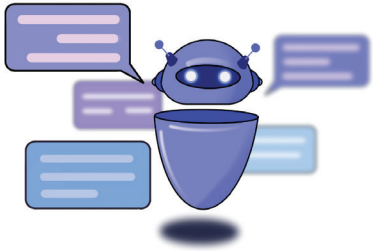
### References

1. Timmis A, Vardas P, Townsend N, et al; Atlas Writing Group, European Society of Cardiology. European Society of Cardiology: Cardiovascular disease statistics 2021. *Eur Heart J*. 2022 Feb;43(8):716-799. Erratum in: *Eur Heart J*. 2022 Feb;43(8):799. [CrossRef]
2. Hernandez CA, Vazquez Gonzalez AE, Polianovskaia A, et al. The future of patient education: AI-Driven guide for type 2 diabetes. *Cureus*. 2023;15(11):e48919. [CrossRef]
3. Hillmann HAK, Angelini E, Karfoul N, Feickert S, Mueller-Leisse J, Duncker D. Accuracy and comprehensibility of chat-based artificial intelligence for patient information on atrial fibrillation and cardiac implantable electronic devices. *Europace*. 2023;26(1):euad369. [CrossRef]
4. Almagazzachi A, Mustafa A, Eighaei Sedeh A, et al. Generative artificial intelligence in patient education: ChatGPT takes on hypertension questions. *Cureus*. 2024;16(2):e53441. [CrossRef]
5. King RC, Samaan JS, Yeo YH, Mody B, Lombardo DM, Ghashghaei R. Appropriateness of ChatGPT in answering heart failure related questions. *Heart Lung Circ*. 2024;33(9):1314-1318. [CrossRef]
6. United States Food and Drug Administration. Artificial intelligence and machine learning in software as a medical device. Accessed: Dec 2, 2024. <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>
7. Munoz-Zuluaga C, Zhao Z, Wang F, Greenblatt MB, Yang HS. Assessing the accuracy and clinical utility of chatgpt in laboratory medicine. *Clin Chem*. 2023;69(8):939-940. [CrossRef]
8. The Lancet Digital Health. ChatGPT: Friend or foe?. *Lancet Digit Health*. 2023;5(3):e102. [CrossRef]
9. ChatGPT: Optimizing language models for dialogue. Accessed Dec 2, 2024. <https://online-chatgpt.com>
10. Lee TJ, Rao AK, Campbell DJ, Radfar N, Dayal M, Khrais A. Evaluating ChatGPT-3.5 and ChatGPT-4.0 responses on hyperlipidemia for patient education. *Cureus*. 2024;16(5):e61067. [CrossRef]
11. Wang L, Wan Z, Ni C, et al. A systematic review of ChatGPT and other conversational large language models in healthcare. *medRxiv* [Preprint]. 2024:2024.04.26.24306390. Update in: *J Med Internet Res*. 2024;26:e22769. [CrossRef]
12. Kaftan AN, Hussain MK, Naser FH. Response accuracy of ChatGPT 3.5 Copilot and Gemini in interpreting biochemical laboratory data a pilot study. *Sci Rep*. 2024;14(1):8233. [CrossRef]
13. Niko MM, Karbasi Z, Kazemi M, Zahmatkeshan M. Comparing ChatGPT and Bing, in response to the Home Blood Pressure Monitoring (HBPM) knowledge checklist. *Hypertens Res*. 2024;47(5):1401-1409. [CrossRef]
14. Güneş YC, Cesur T. Large language models: Could they be the next generation of clinical decision support systems in cardiovascular diseases?. *Anatol J Cardiol*. 2024;28(7):371. [CrossRef]
15. Bozyel S, Şimşek E, Koçyiğit Burunkaya D, et al. Reply to letter to the editor: "Large language models: Could they be the next generation of clinical decision support systems in cardiovascular diseases?". *Anatol J Cardiol*. 2024;28(7):373. [CrossRef]
16. Bozyel S, Şimşek E, Koçyiğit Burunkaya D, et al. Artificial Intelligence-based clinical decision support systems in cardiovascular diseases. *Anatol J Cardiol*. 2024;28(2):74. [CrossRef]

## Comparative Evaluation of Chatbot Responses on Coronary Artery Disease



◦ Fifty frequently asked questions about CAD were asked on ChatGPT, Gemini, and Bing



	ChatGPT	Gemini	Bing Chat
Basic Knowledge	78.5%	50%	64.3%
Diagnosis	80%	60%	60%
Treatment	80%	70%	50%
Recovery/Operative Risks/Complications/Follow-up	92.3%	84.6%	61.5%
Prevention/Dietary	100%	75%	62.5%

◦ When assessing reproducibility, the models consistently provided reproducible answers for most questions.

◦ ChatGPT achieved a score of 88%, Gemini 84%, and Bing 70%



ChatGPT demonstrates significant potential to improve patient education about coronary artery disease by providing more sensitive and accurate answers compared to Bing and Gemini.

CAD\*: coronary artery disease