## **Original Article**

# Content analysis of Al-generated (ChatGPT) responses concerning orthodontic clear aligners

## Sarah Abu Arqub<sup>a</sup>; Dalya Al-Moghrabi<sup>b</sup>; Veerasathpurush Allareddy<sup>c</sup>; Madhur Upadhyay<sup>d</sup>; Nikhilesh Vaid<sup>e</sup>; Sumit Yadav<sup>f</sup>

## ABSTRACT

**Objectives:** To assess the accuracy of ChatGPT answers concerning orthodontic clear aligners. **Materials and Methods:** A cross-sectional content analysis of ChatGPT generated responses to queries related to clear aligner treatment (CAT) was undertaken. A total of 111 questions were generated by three orthodontists based on a set of predefined domains and subdomains. The artificial intelligence (AI)-generated (ChatGPT) answers were extracted and their accuracy was determined independently by five orthodontists. The accuracy of answers was assessed using a prepiloted four-point scale scoring rubric. Descriptive statistics were performed.

**Results:** The total mean accuracy score for the entire set was  $2.6 \pm 1.1$ . It was noted that 58% of the AI-generated answers were scored as objectively true, 18% were selected facts, 9% were minimal facts, and 15% were false. False claims included the ability of CAT to reduce the need for orthognathic surgery ( $4.0 \pm 0.0$ ), improve airway function ( $3.8 \pm 0.5$ ), achieve root parallelism ( $3.6 \pm 0.5$ ), alleviate sleep apnea ( $3.8 \pm 0.5$ ), and produce more stable results compared to fixed appliances ( $3.8 \pm 0.5$ ).

**Conclusions:** The overall level of accuracy of ChatGPT responses to questions concerning CAT was suboptimal and lacked citations to relevant literature. Ability of the software to offer current and precise information was limited. Therefore, clinicians and patients must be mindful of false claims and relevant facts omitted in the answers generated by ChatGPT. (*Angle Orthod.* 2024;94:263–272.)

KEY WORDS: ChatGPT; Content validity; Clear aligners; Artificial intelligence

<sup>b</sup> Assistant Professor, Department of Preventive Dental Sciences, College of Dentistry, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

<sup>c</sup> Professor and Head, Brodie Craniofacial Endowed Chair, Department of Orthodontics, University of Illinois, Chicago, IL, USA.

<sup>d</sup> Associate Clinical Professor, Division of Orthodontics, UConn Health, Farmington, CT, USA.

<sup>e</sup> Adjunct Professor, Department of Orthodontics, Saveetha Dental College, Saveetha Institute of Medical and Technical Sciences, Chennai, India.

<sup>f</sup> Professor and Chair, Department of Growth and Development, University of Nebraska Medical Center College of Dentistry; and Children's Hospital, Omaha, NE, USA.

Corresponding author: Sarah Abu Arqub, clinical assistant professor, Department of Orthodontics, University of Florida, Gainesville, Florida, USA

(e-mail: sabuarqub@ufl.edu)

Accepted: November 2023. Submitted: July 2023. Published Online: January 10, 2024 © 2024 by The EH Angle Education and Research Foundation, Inc.

## INTRODUCTION

With advancement in the field of artificial intelligence (AI), ChatGPT (Chat Generative Pre-trained Transformer) was launched in November 2022 by Open AI (OpenAI LLC, San Francisco, CA, USA), as a large, AI-based language model (LLA) with capability to create human-like responses to any question input.<sup>1</sup> It has surprised scientists and the general public with its comprehensive and detailed responses to various text input.<sup>1</sup> ChatGPT is based on a generative pretrained transformer GPT structure that processes natural tests with the use of neural networks. Thus, the generated responses are often based on the context of the input wording.<sup>1</sup> Its superiority lies in its ability to create accurate and highly refined answers due to its wide range of training data from internet sources.<sup>2</sup>

There are several features that make ChatGPT a powerful tool, including its ability to generate responses in different styles and languages in addition to understanding the context of the question. Its benefits can be generally summarized in its increased efficacy in automating conversations and answering questions, which

<sup>&</sup>lt;sup>a</sup> Clinical Assistant Professor, Department of Orthodontics, University of Florida, Gainesville, FL, USA.

Category (Score)	Definition
Objectively true (1)	A claim that is based on available scientific evidence and presented all relevant information whether positive or negative
Selected facts (2)	A claim that presented true selected facts based on available scientific evidence, but omitted important information related to clear aligners
Minimal facts (3)	A claim that exaggerated or overemphasized benefits of clear aligners beyond available evidence
False (4)	A claim that is objectively false either due to lack of supporting evidence or contradicting available evidence

Table 1. Definition of the Various Categories Used to Determine Accuracy of the Al-Generated Responses

saves time and resources.<sup>2</sup> Due to the fact that it is trained on large data sets of information and conversations, it is characterized by improved accuracy compared to manual responses, outperforming other natural language processing systems (NLP).<sup>2</sup> Despite those benefits, there are challenges associated with ChatGPT related to potential security concerns associated with risk of adversarial attacks that aim to manipulate the model and result in incorrect output. Additionally, it has limited capabilities to provide up-to-date and accurate information and/or answers to complex guestions for a wide variety of topics due to its hindered ability to browse the internet and inability to access external information from the internet.<sup>2</sup> Therefore, uncertainty still remains with regard to accessing high-guality and reliable information via ChatGPT.

Technological advancements have led healthcare professionals and patients to increasingly refer to artificial intelligence (AI) chatbots and online search engines as a convenient source for medical and dental information.<sup>3</sup> The conversational interactions and the seemingly correct responses to many medical and dental inquiries have increased reliance on these platforms to answer many questions and reduce reliance on other professional and accurate resources for medical and dental information.<sup>3</sup> Risks of this phenomenon have been well illustrated in the healthcare literature comprehensively, and precautionary warnings are thereby warranted.<sup>4</sup>

Clear aligner treatment (CAT) is one of the most popular and debated orthodontic development among contemporary appliances due to increased demand for esthetic orthodontic treatment.<sup>4</sup> In a recent cross-sectional study concerning marketing claims related to clear aligners (Invisalign) made by professional accounts on social media, a total of 92 claims were identified from 50 Instagram posts.<sup>5</sup> These claims were mainly related to acceptability of these appliances and less associated pain; only 5% of these claims were objectively true.<sup>5</sup> ChatGPT is a convenient source for patients to access information with regard to CAT. Its accuracy, reliability, and content validity with regard to information associated with CAT in orthodontics has not been evaluated, particularly in connection with questions that dentists, orthodontists, and patients are likely to ask. Therefore, this study was designed to give content analysis of ChatGPT in providing comprehensive information about outcomes related to CAT. The hypothesis was that ChatGPT Al-generated responses about CAT would be accurate and reliable.

#### MATERIALS AND METHODS

A cross-sectional content analysis of ChatGPT generated responses to queries related to CAT was undertaken. An initial dataset of 150 questions was generated by two authors (AAS, AD) and reviewed by a third author (YS). These questions were initially categorized based on 14 key treatment outcome domains and their associated 54 subdomains, as described previously by Tsichlaki et al. using standardized methodology that focused on encompassing clinician- and patient-focused outcomes.<sup>6</sup> The initial dataset of questions was refined through joint discussion, resulting in a final selection of 111 questions. Additional questions that did not fit into any of the key domains were categorized into "other."

To assure consistency, one author (YS) collected the AI-generated answers for the 111 questions (chat. openai.com). A customized Excel sheet was created for data collection and scoring. The accuracy of the collected answers was scored independently by five orthodontists and CAT experts (AAS, AD, UM, AV, and VN). The accuracy of the answers was assessed based on the best available evidence and their clinical expertise in CAT using a modified four-point scale as follows: 1: Objectively true; 2: Selected facts; 3: Minimal Facts; 4: False (Table 1).<sup>7</sup> Prior to scoring, a meeting was held to ensure common understanding of the scoring system among raters. Accuracy assessment was piloted by the five raters on a subset of answers (n = 10).

## **Statistical Analysis**

Simple descriptive statistics were used to summarize the data. Outcomes for each question and each domain were examined using GraphPad prism version 10 (GraphPad Software Inc, La Jolla, CA). Average accuracy score results were summarized as mean and standard deviation for each question separately, per domain and subdomain, and for the entire dataset. Distribution of median accuracy scores was also calculated.

## RESULTS

Among the 111 scored questions, the total mean accuracy score for the entire set was (2.6  $\pm$  1.1)

er Domain, Subdomain, and Question for Domain 1

			Average Score/Question		Avera Score/Sub	ige Idomain	Avera Score/D	age omain
Domain	Subdomains	Questions	Mean	SD	Mean	SD	Mean	SD
1. Harms	1. Pain	1. Are clear aligners associated with more or less pain compared to fixed appliance?	3.4	1.3	2.3	1.3	2.2	0.5
		2. Do clear aligners' attachments cause discomfort/ irritation?	1.2	0.5				
	2. Enamel demineralization	3. Are clear aligners associated with greater white spot lesions/demineralization than fixed appliance?	2.4	1.5	-			
	3. Patient reported adverse events	4. What are the most common reported adverse events with clear aligners?	1.8	0.8	-			
	4. Root resorption	5. Are subjects treated with fixed appliance more prone to root resorption than those treated with clear aligners?	2.2	1.6	-			
	5. Enamel roughness post debond	6. Which bur is the safest to enamel during removal of aligners' attachments?	2.8	1.6	2.7	0.7		
		7. Do clear aligners adversely affect the enamel?	2.6	1.1				
	6. Caries	8. Are clear aligners associated with more caries than fixed appliance?	2.8	1.3	-			
	7. Halitosis and tongue coating	9. Does clear aligner treatment adversely affect the breath?	1.4	0.6	-			
	8. Mucosal ulceration	10. Are clear aligners associated with irritation to lips and cheeks?	1.6	0.6	1.9	0.8		
		11. Are clear aligners associated with mucosal ulceration?	2.2	1.1				
	9. Gingival irritation	12. Are clear aligners associated with gingival irritation?	1.4	0.5	-			

<sup>a</sup> SD indicates standard deviation.

(Tables 2–5). Of note, 58% of the generated answers fell into the objectively true category, 18% were scored as selected facts, 9% as minimal facts, and 15% were considered false (Figure 1).

#### **Inaccurate Information**

The mean accuracy score concerning harms associated with CAT (domain 1) was 2.2  $\pm$  0.5 (Table 2). Inaccurate information was provided by ChatGPT concerning pain associated with CAT compared to fixed appliances (3.4  $\pm$  1.3) (Table 2).

Claims related to the effect of CAT on jaw movements  $(3.2 \pm 1.3)$  and associated condylar changes were inaccurate  $(3.0 \pm 0.0)$  (Table 3). Additionally, there was inaccurate information identified in soft tissue changes associated with CAT  $(2.6 \pm 1.5)$ , as well as occlusal and alignment changes  $(2.9 \pm 0.3)$ , especially those concerning the effectiveness of CAT in achieving root parallelism  $(3.6 \pm 0.9)$  and torque control  $(3.4 \pm 0.9)$  (Table 3). Inaccurate information was noted in ChatGPT responses to questions related to the ability of CAT to reduce need for orthognathic surgery  $(4.0 \pm 0.0)$ , affect airway volume  $(3.8 \pm 0.5)$ , alleviate sleep apnea  $(3.8 \pm 0.5)$ , and provide more stable outcomes compared to fixed appliances  $(3.8 \pm 0.5)$  (Table 5).

#### **Accurate Information**

Responses associated with periodontal health (1.3  $\pm$  1.6), and microbial and physiological changes (1.3  $\pm$  0.5)

in CAT were considered accurate (Table 3). Additionally, ChatGPT provided accurate answers concerning patient satisfaction with CAT (1.0  $\pm$  1.0) and its effect on sleeping, eating, and oral health-related quality of life (1.1  $\pm$  3.1) (Table 4), aligner breakage and loss, hygiene and wear requirements, and frequency and number of appointments (1.4  $\pm$  0.5) (Table 4). Accurate information was also noted regarding treatment duration (1.0  $\pm$  1.2), rate of tooth movement (1.0  $\pm$  2.0), and costeffectiveness (1.0  $\pm$  1.8) (Table 4).

#### DISCUSSION

Based on the results of this cross-sectional study conducted 5 months into the existence of ChatGPT, the null hypothesis was rejected. Responses provided concerning CAT were short of being completely reliable. The ability to generate evidence-based answers about CAT were suboptimal, with only 58% of the answers being objectively true. False claims primarily included the ability of CAT to reduce the need for orthognathic surgery, improve airway function, achieve root parallelism, alleviate sleep apnea, and produce more stable results compared to fixed appliances. Therefore, the adequacy of ChatGPT in its current form and its use in academia and research is disputable. The lack of expert level and evidence-based opinions in the composition of many of the generated answers to the questions was often the trend, and the language used in the answers was simple (Table 6). Overall, the answers for all guestions were

				score/ ions	Average Subdo	Score/ main	Average Dom	Score
Domain	Subdomain	Question	Mean	SD	Mean	SD	Mean	SD
2. Function	10. Speech	13. Is there any effect of clear aligners on speech?	1.6	1.3	-		2.5	0.1
	11. Mandibular excursion	14. Do clear aligners affect jaw movements?	3.2	1.3	-			
	12. Other (function)	15. Are clear aligners the appliance of choice for musicians and singers?	1.6	1.3	2	0.3		
		16. Do clear aligners have an effect on the chewing efficacy?	2.4	1.1				
	13. Condylar changes	17. Does clear aligner treatment change the condylar adaptation?	3	1.4	3	0		
		18. Is there any relationship between temporomandibular jaw joint health and clear aligners?	3	1.4				
		19. Does clear aligner wear reduce the tendency of clenching and bruxism for TMD subjects?	3	1.4				
3. Soft tissues	<ol> <li>Soft tissue profile changes</li> </ol>	20. What are the effects of clear aligners on the lip position and thickness?	2.6	1.5	-		-	
4. Occlusal/	15. Tooth angulation/	21. Can root parallelism be achieved with clear aligners?	3.6	0.9	3.3	0.1	2.9	0.3
alignment	inclination	22. Can the roots be torqued with clear aligners?	3.4	0.9				
change		23. Can teeth rotations be corrected with clear aligners?	2.8	0.8				
-	16. Alignment relapse	24. Is alignment stable with clear aligners?	2	1	-			
	17. Occlusal outcome	25. Do clear aligners cause posterior open bites?	2.6	1.5	-			
<ol> <li>Periodontal outcomes</li> </ol>	18. Enamel reduction	26. What are the adverse effects of interproximal stripping as part of clear aligner treatment?	2	1	-		1.4	0.7
	19. Periodontal Health	27. Are clear aligners associated with more gingival recession compared to fixed appliance?	1	0	1.3	1.6		
		28. Are clear aligners associated with more bone loss compared to fixed appliance?	1	0				
		29. Are clear aligners associated with more gingival inflammation compared to fixed appliance?	1.6	1.3				
		30. Are clear aligners associated with more bleeding on probing compared to fixed appliances?	1.6	1.3				
<ol> <li>Microbiological/ physiological</li> </ol>	20. Microbial composition/count	31. Do clear aligners alter the microbial components of oral cavity?	1.2	0.4	1.4	1.3	1.3	0.5
, ,		32. Do clear aligners promote more plaque formation compared to fixed appliances?	1.6	1.3				
	21. Salivary pH	33. Do clear aligners alter the level of salivary pH?	1.2	0.4	-			

Table 3. Average Scores per Domains, Subdomains, and Questions for Domains 2–6<sup>a</sup>

<sup>a</sup> SD indicates standard deviation; TMD, temporomandibular joint disorder.

generally lengthy and referred to trials and reviews without in-text citations (Supplementary 1). However, some ChatGPT answers to CAT queries were useful in areas of periodontal health (1.3  $\pm$  1.6), patient satisfaction (1.0  $\pm$  1.0), impact on daily activities (1.1  $\pm$  3.1), in addition to oral/aligner hygiene (1.4  $\pm$  1.3) and wear (1.3  $\pm$  2) requirements.

ChatGPT indicated that CAT was associated with overall less pain compared to fixed appliances. However, a recent systematic review indicated that patients treated with clear aligners reported less pain in the first couple of days only and there was no difference after a week; therefore, these findings were controversial among the included studies and the certainty of evidence was low.<sup>8</sup> In terms of function, ChatGPT generated answers indicating that CAT could affect jaw movement and cause jaw pain and difficulty in movement. Likewise, with regard to responses related to condylar changes, ChatGPT indicated the presence of a potential relationship between jaw joint health, condylar adaptation, and use of clear aligners. This was in disagreement with the available evidence that orthodontic treatment might not increase or be related to the prevalence of temporomandibular disorders.9

The AI-generated responses indicated that root parallelism and torque can be achieved with aligners. Likewise, ChatGPT stated the high efficacy in correcting rotations with CAT. Unfortunately, achieving pure root movement and complete derotation on teeth with clear aligners are onerous tasks and attempts to achieve these have been unsuccessful.<sup>10</sup> Evidence has also shown that torquing or root movement is the least predictable of all movements with CAT.<sup>11</sup> Large language-based algorithms perform relatively well on knowledge-based tests, but their performance is subpar on medical/dental concepts and literature. To deliver excellent performance, AI-based large language models require high-guality data but, currently, they are trained on biased data sets, which may be the reason for inaccurate answers to gueries in this specific research field.

The responses related to the effects of CAT on facial and smile esthetics were flawed. The ChatGPT response explained that, if too much space was created between the upper and lower teeth, it could result in a "gummy" smile or an overbite that may detract from the overall appearance of the face. A recent study showed that smile esthetics after the use of Invisalign and fixed appliances was superior (buccal corridors, smile cant, gingival

Table 4	Average Score per	Question	Subdomain	and Domain for	Domains 7 8	and 9 <sup>a</sup>
	Average Score per	Question.	Subuullalli.		$D_{0}$	anus

				Average Score/Question		Average Score/Subdomain		Average Score/Domai	
Domain	Subdomains	Questions	Mean	SD	Mean	SD	Mean	SD	
7. Knowledge and satisfaction	22. Satisfaction	34. Are patients treated with clear aligners more satisfied than those treated with fixed appliances?	1	1	-		1.6	0.5	
	23. Information comprehension	35. Are patients' instructions regarding wear of clear aligners easy to comprehend by patients?	2	1.4	-				
	24. Acceptability of	36. Are the clear aligner attachments esthetic?	1	1.2	2.1	0.8			
	appliance	37. Does clear aligner treatment alter facial esthetics?	3	1					
	25. Self-esteem	38. Do clear aligners result in better confidence when	1	1.8	2	2.2			
		compared to fixed appliances?	_						
		39. Do clear aligners result in better self-esteem when compared to fixed appliances?	2	2.2					
	26 Aligner impact	40 Do clear aligners affect eating?	1	12	11	31			
	Lot / ingrior impact	41 Do clear aligners affect sleeping?	1	1		0			
		42. Is the oral health-related quality of life better with clear	1	1.2					
		aligner treatment when compared to fixed appliances?							
8. Compliance	27. Aligner breakage/	43. How frequent are the reported clear aligners breakages?	1	2	1.8	4.6	1.4	0.5	
F	loss	44. What will happen if I lose my clear aligner?	1	1.6					
	28. Oral and aligner	45. Do patients with clear aligners have superior oral hygiene compared to those with fixed appliances?	1	1.2	1.4	1.3			
	nygione	46. What are the methods of cleaning and disinfecting of	1	1.8					
		47. Do the clear aligners' attachments make toothbrushing difficult?	1	1					
		48. How frequent is the reported loss of clear aligners?	1	1.6					
	29. Duration of aligner wear/day	49. What is the common duration of retainer wear with clear aligners?	1	1.8	1.3	1.2			
		50. Is the "blue indicator" accurate in measuring compliance with clear aligners?	1	1					
		51. What methods can be used to measure compliance of patients wearing clear aligners?	1	1					
		52 What would happen if Lskip wearing some aligners?	1	12					
	30. Attendance	53. How frequently do I have to visit my orthodontist during	1	1.2	1.2	1.6			
		54. Is there a difference in number of appointments needed	1	1.2					
9. Efficiency/cost	31. Treatment duration	55. Is there a difference in treatment duration between clear	1	1.2	1.1	2.7	2	0.5	
effectiveness		56. Is there a difference in number of emergency appoint-	1	1					
	20 Data of tooth	ments between clear aligners and fixed appliances?	4	0					
	movement	aligners?	I	2	-				
	33. Treatment success	58. What are the possible variables that affect treatment success with clear aligners rather than fixed appliances (age, sex, root length, bone quality, and bone levels)?	2	3	2.2	3.3			
		59. Is there a difference in the outcomes between adults and teens treated with clear aligners?	1	1.4					
		60. What is the mean improvement of deep bite with clear aligners?	1	2.2					
	34. Cost	61. Are clear aligners cost-effective to the patient? 62. Are clear aligners cost-effective to the orthodontist/dentist?	2 1	3.2 1.8	2.5	0.8			

<sup>a</sup> SD indicates standard deviation.

display, maxillary midline, and smile index) to fixed appliances and there was no effect on lip position.<sup>12</sup>

ChatGPT indicated the ability to reduce need for orthognathic surgery with use of clear aligners, thus failing to provide clinician/orthodontist level insight. Also, the ChatGPT generated response lacked the outcomes achieved with orthognathic surgery treated with clear aligners or fixed appliances.<sup>13</sup> Additionally, ChatGPT lacked sources (references) for its answers. In its current format, an Al-generated large languagebased model needs further algorithmic training to implement current dental/orthodontic knowledge, principles, and concepts in real-world settings.

ChatGPT seemed to be equally unreliable when it was questioned on the effects of CAT on airway changes and sleep apnea, suggesting positive effects on airway volume and relief of sleep apnea. This was erroneous and overlooked the available evidence on the lack of association between airway improvement and use of CAT in treating any malocclusion.<sup>14</sup>

ChatGPT used in this research was not a useful tool for generating answers to scientific queries. ChatGPT lacked the knowledge and expertise necessary to accurately convey simple and complex orthodontics concepts adequately. Another major problem with ChatGPT response was redundancy and plausible-sounding false information. Therefore, despite the fact that it might provide patients with some useful insights and general information with regard to their treatment, its use for teaching and research purposes is currently limited and should be avoided. Orthodontic professionals should also be aware of how patients may use these tools and

			Average Score/ Question		Average Score/ Average Sco Question Subdomain		Average Score/ Average S Question Subdom		rage Score/ Average Score		Average Score	
Domain	Subdomain	Question	Mean	SD	Mean	SD	Mean	SD				
10. Other	35. Orthognathic	63. Can clear aligner treatment reduce the need for	4	0	-		2.1	0.5				
	36. Airway volume/	64. Are there airway changes associated with clear aligner treatment?	3.8	0.5	3.8	1.5						
	sleep upried	65. Is there any evidence of relief of sleep apnea and clear aligner therapy?	3.8	0.5								
	37 Aligner/attach-	66 What gender is mostly treated with aligners?	22	11	19	14						
	ment indication	67. What age group is mostly treated with clear aligners?	1.8	0.8								
	and uses	68. Can clear aligners be used for extraction treatment with predictable outcomes?	3	1.4								
		69. Can clear aligners be used for cleft lip and palate patients?	2.2	1.3								
		70. Can clear aligners be used for amelogenesis imperfecta?	1.4	0.6								
		71. Can patients with impacted teeth be treated with clear aligners?	2.8	1.3								
		72. Is distalization effective with clear aligners?	2.2	1.3								
		73. Can clear aligners be used during sports?	1.6	1.3								
		74. Can clear aligners be used as nightguards?	1	0								
		75. Can clear aligners be used as sports mouthguard?	1.8	1.3								
		76. Can I use the clear aligner as a bleaching tray?	1.2	0.5								
		77. Can the clear aligner tray serve as a retainer?	1.8	1.3								
		78. Can clear aligners be used for osteogenesis imperfecta?	1.8	1.3								
		79. Can clear aligners be used for treating patients with craniofacial syndromes?	1.4	0.9								
		80. Are clear aligners a common treatment option for orthognathic surgical cases?	1.2	0.5								
		81. Are clear aligners effective to be used with temporary anchorage devices?	2	1.2								
		82. Is there a difference between conventional and optimized attachments in outcomes obtained with clear aligners?	2.8	1.1								
		83. What are the types of malocclusions best treated with clear aligners?	1.6	0.9								
		84. Are clear aligners common for early treatment/phase one (mixed dentition)?	1.6	1.3								
		85. What malocclusions can be treated in phase one with clear aligners?	2	1.4								
	38 Commercially	86 What are clear aligners made of?	16	0 0	1.8	1						
	available aligners/	87. Which commercially available clear aligner are most	2.2	1.3	1.0	'						
	material	<ul> <li>88. Are there any environmental concerns with the polymer used in the clear aligner therapy? Are they environmentally friendly?</li> </ul>	1.8	1.1								
		89. Are clear aligners made of the same material as clear retainers?	1.4	0.6								
		90. Are 3D printed (in-house aligners) clear aligners better than commercially available aligners?	1.6	0.9								
		91. What clear aligner material is best in terms of wear resistance?	2	1.4								
	39. Aligner mecha-	92. What is the main mechanism of open bite correction with	3.2	1.3	3	0.3						
	TISH OF ACTION	93. What is the main mechanism of overjet correction with	2.8	1.3								
	40. Patient-clinician	clear aligners? 94. How can the orthodontist know if I wear my aligners?	1.2	0.5	-							
	relationship 41. Aligners vs other	95. Are clear aligners better than twinblock for Class II early	2.2	1.3	2.5	0.3						
	appliances	treatment?										
		96. In which way is fixed appliances better than clear aligners?	3.2	1.3								
		97. In which cases are aligners better than fixed appliances?	1.8	0.8								
		98. Is patient cooperation better with clear aligner treatment	2.8	1.1								
		When compared to conventional braces?	2.0	0 5								
		99. IS Stability of treatment better with clear aligners when	3.8	0.5								
		100 Are clear aligners better for the treatment of periodentally	0	1 /								
		compromised patients compared to fixed appliances?	ے 1 ہ	1.4								
	40. Definement store	appliances?	1.0	1.0	17	0						
	42. Heinement stage	refinements?	1.8	1.3	1./	U						
	10. 0	<ul> <li>Now many numbers or reinnements would patients treated with clear aligners need?</li> <li>Not which is the state state of the state state of the state state</li></ul>	1.6	1.3	10	0.0						
	43. Scanners	<ul><li>104. What is the most accurate scanner available for aligners?</li><li>105. Do the outcomes differ whether a scanner or impression</li></ul>	1.6 1	0.9 0	1.3	0.6						
		was used to take the records for clear aligner treatment?		_								
	44. Retention phase	106. What are the common retainers used with clear aligners?	1	0	1	0						

 Table 5.
 Average Score/Question, Subdomain, and Domain for Additional Questions<sup>a</sup>

	Table 5.	Continued.
--	----------	------------

			Average Score Question		Average Subdor	Score/ main	Average Dom	Score/ ain
Domain	Subdomain	Question	Mean	SD	Mean	SD	Mean	SD
	45. Accuracy of online information	<ul><li>107. Are retainers required after clear aligner treatment?</li><li>108. Is the content of social media platforms about clear aligners accurate in terms of reliability of the information?</li></ul>	1 1.2	0 0.5	-			
	46. Evidence-based	109. Are there RCTs proving the effectiveness of clear aligners?	2.2	1.6	-			
	47. Additional questions	<ul><li>110. What are the myths about clear aligners?</li><li>111. What type of emergencies are there with clear aligners?</li></ul>	3 1.8	1.2 0.8	2.4	0.3		

<sup>a</sup> RCT indicates randomized controlled trial; SD, standard deviation.

provide them with appropriate guidance. An additional concern was that ChatGPT provided varied responses for the same questions when assessed at different timepoints, which might raise ethical concerns and lead to discrepancies in accessing the desired information. It is possible that, with enough repetition, more accurate answers could be generated due to the machine learning capabilities of ChatGPT. Additionally, it seems like the algorithms need to be trained and tested using published journal articles of high quality as evidence to upgrade the accuracy and capabilities of ChatGPT in answering questions from an evidence-based perspective. Relying exclusively on the currently available version of ChatGPT as a source for valid and reliable information with regard to CAT is not recommended. And false information delivery poses a risk to the profession and patients. Societies need to debunk misleading information and increase awareness among the public and professionals while using this Al tool.

On the other hand, acceptable accuracy levels were observed for answers to questions concerning knowledge, satisfaction, compliance, and cost-effectiveness. Therefore, depending on regular advances in the model's algorithm and the influences of reinforced learning from human feedback on the system, in the future, ChatGPT might be a useful source for orthodontic patients seeking to comprehend aspects related to CAT since it provides an interactive interface for treatment-related information. However, at the moment, use of these advances cannot be an alternative to conventional means of communicating information related to orthodontic treatment to the patient.



Figure 1. Distribution of median accuracy scores.

Table 6. Examples of Answers Scored as False (4), Minimal Facts (3), Selected Facts (2), and Objectively True (1)

Question	ChatGPTAnswer	Score
Does clear aligner therapy	"Clear aligner treatment can potentially have positive effects on the airway in certain	False (4)
	<ul> <li>Research has shown that certain types of malocclusion, such as a deep overbite or a retruded mandible, can contribute to airway issues such as sleep-disordered breathing and obstructive sleep apnea. By correcting these types of malocclusions with clear aligners or other orthodontic treatments, it may be possible to improve the airway and reduce the severity of sleep-disordered breathing.</li> <li>In addition, clear aligner treatment can potentially improve nasal breathing by expanding the arches of the teeth and improving the position of the tongue. This can help to increase the space in the nasal passages and improve airflow.</li> <li>However, it is important to note that not all cases of malocclusion are related to airway issues, and not all patients with airway issues will benefit from orthodontic treatment. It is important for a qualified orthodontist or oral surgeon to evaluate the case and determine the most appropriate treatment plan.</li> <li>Additionally, it is important for patients to seek treatment for airway issues from a qualified medical professional, such as a sleep medicine specialist or an otolaryngologist, in addition to seeking orthodontic treatment. Treatment for airway issues may include a combination of medical and dental interventions, depending on the specific case."</li> </ul>	
Do clear aligners cause posterior open bites?	<ul> <li>"Clear aligners do not typically cause posterior open bites, but there are some cases where this can occur.</li> <li>Posterior open bite is a type of malocclusion where there is a gap between the upper and lower teeth when the back teeth are together. It can be caused by various factors, including skeletal discrepancies, habits such as thumb-sucking or tongue-thrusting, and certain types of orthodontic treatment.</li> <li>Clear aligner treatment is designed to gradually move the teeth into their desired position, which can sometimes lead to changes in the bite. However, clear aligners are typically customized to apply pressure to specific areas of the teeth to achieve the desired tooth movement without compromising the bite or causing any malocclusions.</li> <li>In some cases, however, clear aligner treatment can exacerbate an existing posterior open bite or create a new one, especially if the malocclusion is caused by skeletal discrepancies or habits. This is why it is important for patients to receive a thorough evaluation and treatment plan from a qualified orthodontist or dental provider to ensure that clear aligners are an appropriate treatment option for their specific needs.</li> <li>If a patient experiences a posterior open bite during or after clear aligner treatment, the orthodontist may need to adjust the treatment plan or use additional orthodontic appliances to correct the bite. The patient may also need to wear a retainer or other device after treatment to help maintain the alignment of their teeth and prevent any relapse or malocclusion."</li> </ul>	Minimal facts (3)
How frequent is the reported appliance breakages?	<ul> <li>Clear aligners are a type of dental appliance that are designed to straighten teeth. These appliances are typically made of a clear plastic material and are custom-fitted to the patient's teeth. They are often used as an alternative to traditional metal braces.</li> <li>In general, clear aligners are designed to be durable and long-lasting, and they are not prone to breaking or cracking. However, like any dental appliance, they can occasionally experience issues such as cracking, warping, or breaking.</li> <li>The frequency of reported appliance breakages with clear aligners can vary depending on a number of factors, including the specific brand of aligners, the patient's usage patterns, and how well the aligners are designed to be worn for 20 to 22 hours per day, and they should be removed before eating or drinking anything other than water. If patients do not follow these guidelines or if they are particularly rough with their aligners, they may experience more frequent breakages.</li> <li>Overall, the frequency of reported appliances. However, it is still possible for these appliances to experience issues from time to time, and patients should follow their orthodontist's instructions for care and usage to minimize the risk of breakages.</li> </ul>	Selected facts (2)
Are the clear aligner attachments esthetic?	<ul> <li>"Clear aligner attachments are small, tooth-colored shapes that are attached to the teeth during clear aligner treatment. They serve as anchors or handles for the clear aligners to exert force and achieve tooth movements. The color and size of the attachments can vary depending on the type and location of the tooth movement required.</li> <li>While clear aligner attachments are designed to be as inconspicuous as possible, they may still be visible up close. The exact level of esthetics of the attachments can vary depending on the patient's individual situation and the specific type of attachment used.</li> </ul>	True (1)

Question	ChatGPT Answer	Score
	In general, clear aligner attachments are made from a tooth-colored composite material that closely matches the shade of the teeth. This material is designed to be resistant to staining and wear, ensuring that the attachments remain as esthetic as possible through- out the treatment process. Overall, while clear aligner attachments are not completely invisible, they are designed to be as esthetic as possible to ensure that patients can maintain their confidence and com- fort throughout the treatment process."	

#### Strengths, Limitations, and Future Directions

A strength of the study was the inclusion of 111 questions developed by three orthodontists based on outcomes relevant to clinicians and patients. However, the selection of a cohort of five orthodontists from the academic sector to rank the generated answers may have introduced response bias. It is important to note that these orthodontists are well-versed in CAT and evidence-based orthodontics and have published extensively. One limitation was that the validation of ChatGPT may not necessarily apply to other AI models. Additionally, ethical and privacy concerns related to the quality of information shared by ChatGPT should be taken into consideration. It is highly recommended to integrate high-guality orthodontic information from sources such as peer-reviewed journals. Additionally, incorporation of in-text citations into these Al algorithms can help ensure the delivery of high-quality information. Establishing a regulatory framework for orthodontic websites is an area on which the specialty should consider focusing.<sup>15,16</sup> Future research should be done to perform content analysis of different large language-based AI models.

## CONCLUSIONS

- The Al-generated answers to questions related to CAT in orthodontics displayed suboptimal accuracy and lacked reference to the current evidence.
- False claims were identified, especially on topics related to orthognathic surgery, airway, efficacy of tooth movement, and root control. In the future, ChatGPT could be a useful adjunct tool to improve knowledge and answer questions regarding orthodontic treatment.
- However, patients and orthodontists should be aware of the limitations and ethical concerns associated with ChatGPT and actively check available evidence from trusted sources.
- Attempts should be made to improve the robustness of these AI models prior to their integration in the healthcare profession.

## ACKNOWLEDGMENTS

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R88), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors report no conflicts of interest.

## SUPPLEMENTAL DATA

Supplemental Table #1 is available online.

## REFERENCES

- Brown T, Mann B, Ryder N, et al. Language models are few-shot learners. *Adv Neural Inf Process Syst.* 2020;33: 1877–1901.
- 2. Deng J, Lin Y. The benefits and challenges of ChatGPT: an overview. *Front Comput Intell Sys.* 2022;2:81–83.
- 3. Shen Y, Heacock L, Elias J, et al. ChatGPT and other large language models are double-edged swords. Radiological Society of North America; 2023; p. 230163.
- Vaid NR. Artificial intelligence (AI) driven orthodontic care: a quest toward utopia? Seminars in Orthodontics; Elsevier; 2021:57–61.
- Alkadhimi A, Al-Moghrabi D, Fleming PS. The nature and accuracy of Instagram posts concerning marketed orthodontic products: a cross-sectional analysis. *Angle Orthod*. 2022; 92:247–254.
- Tsichlaki A, O'Brien K, Johal A, Fleming PS. Orthodontic trial outcomes: plentiful, inconsistent, and in need of uniformity? A scoping review. *Am J Orthod Dentofacial Orthop*. 2018;153:797–807.
- Alkadhimi A, Al-Moghrabi D, Fleming PS. The nature and accuracy of Instagram posts concerning marketed orthodontic products. *Angle Orthod.* 2022;92:247–254.
- Cardoso PC, Espinosa DG, Mecenas P, Flores-Mir C, Normando D. Pain level between clear aligners and fixed appliances: a systematic review. *Prog Orthod.* 2020;21: 1–17.
- Kim M-R, Graber TM, Viana MA. Orthodontics and temporomandibular disorder: a meta-analysis. *Am J Orthod Dentofacial Orthop.* 2002;121:438–446.
- Upadhyay M, Arqub SA. Biomechanics of clear aligners: hidden truths & first principles. J World Fed Orthod. 2022; 11:12–21.
- Dai F-F, Xu T-M, Shu G. Comparison of achieved and predicted tooth movement of maxillary first molars and central incisors: first premolar extraction treatment with Invisalign. *Angle Orthod*. 2019;89:679–687.

- Christou T, Abarca R, Christou V, Kau CH. Smile outcome comparison of Invisalign and traditional fixed-appliance treatment: a case-control study. *Am J Orthod Dentofacial Orthop*. 2020;157:357–364.
- Kwon T, Alshehri A, Palo L, et al. Assessment of the occlusal outcomes in patients treated with orthognathic surgery and clear aligners. Orthod Craniofac Res. 2023;26:371–377.
- Al-Jewair T, Kurtzner K, Giangreco T, Warunek S, Lagravère-Vich M. Effects of clear aligner therapy for Class II malocclusion on upper airway morphology and

daytime sleepiness in adults: a case series. *Int Orthod*. 2020;18:154–164.

- Gandedkar NH, Vaid NR, Darendeliler MA, Premjani P, Ferguson DJ. The last decade in orthodontics: a scoping review of the hits, misses and the near misses! *Semin Orthod*. 2019:339–355.
- Gandedkar NH, Wong MT, Darendeliler MA. Role of virtual reality (VR), augmented reality (AR) and artificial intelligence (AI) in tertiary education and research of orthodontics: an insight. *Semin Orthod*. 2021;69–77.