

An Assessment of the Accuracy and Consistency of ChatGPT in the Management of Midshaft Clavicle Fractures

Review began 03/18/2025

Review ended 04/06/2025

Published 04/08/2025

© Copyright 2025

Knee et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: 10.7759/cureus.81906

Christoph J. Knee¹, Ryan J. Campbell¹, Brahman S. Sivakumar², Michael J. Symes¹

1. Department of Orthopaedics and Trauma Surgery, Royal North Shore Hospital, Sydney, AUS 2. Department of Hand and Peripheral Nerve Surgery, Royal North Shore Hospital, Sydney, AUS

Corresponding author: Christoph J. Knee, christophaknee@gmail.com

Abstract

Background

Midshaft clavicle fractures are common orthopaedic injuries with no consensus on optimal management. Large language models (LLMs) such as ChatGPT (OpenAI, San Francisco, USA) present a novel tool for patient education and clinical decision-making. This study aimed to evaluate the accuracy and consistency of ChatGPT's responses to patient-focused and clinical decision-making questions regarding this injury.

Methods

ChatGPT-4o mini was prompted three times with 14 patient-focused and orthopaedic clinical decision-making questions. References were requested for each response. Response accuracy was graded as: (I) comprehensive; (II) correct but inadequate; (III) mixed with correct and incorrect information; or (IV) completely incorrect. Two consultant and two trainee orthopaedic surgeons evaluated the accuracy and consistency of responses. References provided by ChatGPT were evaluated for accuracy.

Results

All 42 responses were graded as (III), indicating a mix of correct and incorrect information, with 78.6% consistency across the responses. Of the 128 references provided, 0.8% were correct, 10.9% were incorrect, and 88.3% were fabricated. Only 3.1% of references accurately reflected the cited conclusions.

Conclusion

ChatGPT demonstrates limitations in accuracy and consistency when answering patient-focused queries or aiding in orthopaedic clinical decision-making for midshaft clavicle fractures. Caution is advised before integrating ChatGPT into clinical workflows for patients or orthopaedic clinicians.

Categories: Trauma, Orthopedics, Sports Medicine

Keywords: artificial intelligence, chatgpt, clavicle operative fixation, midshaft clavicle fracture, patient education

Introduction

Clavicle fractures are common, accounting for approximately 2–5% of all fractures in adults and 10–15% in children, with more than 80% involving the midshaft region [1,2]. These fractures predominantly occur in male patients and exhibit a bimodal distribution, with peaks among individuals under 25 years of age due to direct trauma and those aged over 55 years following falls [1,2]. Despite their prevalence, the optimal management of midshaft clavicle fractures remains a subject of conjecture. The primary treatment options include non-operative management, typically via a period of immobilisation in a sling, and operative intervention via open fracture reduction with internal fixation [1]. While the historical management of midshaft clavicle fractures has been primarily non-operative, recent studies indicate that operative fixation is associated with lower non-union rates and early functional recovery [3–6]. This evidence has been cited as contributing to a more than 300% increase in operative intervention rates over the past two decades in both male and female patients [7,8].

Traditionally, internet searches have served as a primary resource for patients seeking medical information. Amongst individuals attending orthopaedic outpatient clinics, approximately two-thirds report using the internet to obtain orthopaedic information, with over 80% using Google as their preferred search engine [9]. Recently, however, artificial intelligence (AI) chatbots - particularly large language models (LLMs) such as ChatGPT (OpenAI, San Francisco, USA) - have emerged as novel tools for accessing medical information. With over 200 million weekly active users, ChatGPT is one of the most widely used LLMs [10]. It is trained on a broad range of data, including publicly available content from the internet [11–13]. This training enables it to generate human-like responses across various topics, including healthcare. Accessible online, ChatGPT

How to cite this article

Knee C J, Campbell R J, Sivakumar B S, et al. (April 08, 2025) An Assessment of the Accuracy and Consistency of ChatGPT in the Management of Midshaft Clavicle Fractures. Cureus 17(4): e81906. DOI 10.7759/cureus.81906

allows both patients and clinicians to engage in open-ended conversations and receive detailed responses to health-related questions [11-13].

The potential applications of ChatGPT within medicine and public health are extensive, encompassing personalised learning, medical and patient education, clinical decision-making, and documentation support [11]. Although ChatGPT demonstrates significant potential, its performance remains variable across different medical specialties. While it has shown greater proficiency in medical knowledge than other AI chatbots, it frequently generates inaccuracies, commonly referred to as “hallucinations,” which can mislead users [12,13]. These inaccuracies are particularly concerning as AI models become increasingly interactive and persuasive, rendering it difficult for users to discern the reliability of the information provided [13].

Therefore, the purpose of this study was to investigate the accuracy and consistency of ChatGPT's outputs for patient-focused inquiries and clinical decision-making related to midshaft clavicle fractures.

Materials And Methods

Study design

On August 31, 2024, ChatGPT-4o mini was prompted with a set of 14 questions addressing various aspects of midshaft clavicle fracture management (Table 1). Questions 1 to 4 were designed to replicate common patient searches related to clavicle fractures on Google. To achieve this, Google Trends (Google LLC, Mountain View, USA) - a tool used to identify the most popular search queries on specific topics - was utilised. The terms "broken collarbone", "broken clavicle", and "clavicle fracture" were searched to find the top related health queries worldwide from 2004 to the present. The most frequently associated search terms related to clavicle fractures, such as "sling", "displacement", "treatment", "surgery", and "recovery time", shaped the wording and focus of the patient questions. A 200-word limit was imposed to ensure responses remained concise and maintained reader engagement.

ChatGPT Input Questions

Patient-Focused Questions

- 1 In 200 words, outline the indications for surgical and non-surgical treatment of midshaft clavicle fractures, supported by evidence from four high-quality journal references. Ensure that the references are cited correctly both in-text and listed separately in a reference list.
- 2 In 200 words, compare the recovery times between surgical and non-surgical treatment of midshaft clavicle fractures, supported by evidence from four high-quality journal references. Ensure that the references are cited correctly both in-text and listed separately in a reference list.
- 3 In 200 words, compare the short- and long-term functional outcomes of surgical and non-surgical treatment of midshaft clavicle fractures, supported by evidence from four high-quality journal references. Ensure that the references are cited both in-text and listed separately in a reference list.
- 4 In 200 words, compare the benefits and risks of surgical versus non-surgical treatment of midshaft clavicle fractures, supported by evidence from four high-quality journal references. Ensure that the references are cited correctly both in-text and listed separately in a reference list.

Clinical Decision-Making Questions

- 5 I'm an orthopaedic registrar seeking advice for a 10-year-old male with a midshaft clavicle fracture. His X-ray shows displacement of 100% of shaft width and 1 cm of shortening. Should I choose surgical or non-surgical treatment? Provide your 150-word recommendation with in-text citations and a separate reference list.
- 6 I'm an orthopaedic registrar seeking advice for a 23-year-old male with a comminuted midshaft clavicle fracture. His X-ray shows a Z-type fracture pattern. Should I choose surgical or non-surgical treatment? Provide your 150-word recommendation with in-text citations and a separate reference list.
- 7 I'm an orthopaedic registrar seeking advice for a 24-year-old male with a displaced midshaft clavicle fracture and significant skin tenting. Although he is open to surgery, he prefers non-surgical treatment if it is recommended. Should I opt for surgical or non-surgical treatment? Provide your 150-word recommendation with in-text citations and a separate reference list.
- 8 I'm an orthopaedic registrar seeking advice for a 35-year-old male with a midshaft clavicle fracture. His X-ray shows minimal displacement with a comminuted fragment. He is an avid cyclist and wants to return to full function quickly. Should I choose surgical or non-surgical treatment? Provide your 150-word recommendation with in-text citations and a separate reference list.
- 9 I'm an orthopaedic registrar seeking advice for a 35-year-old female with a midshaft clavicle fracture. She was initially treated with non-surgical management but returned to the clinic two weeks later with significant fracture displacement. Should I choose surgical or non-surgical treatment? Please provide your 150-word recommendation with in-text citations and a separate reference list.
- 10 I'm an orthopaedic registrar seeking advice for a 40-year-old female with a midshaft clavicle fracture. Her X-ray shows 3 cm of shortening and displacement greater than 100% of the shaft width. Should I choose surgical or non-surgical treatment? Provide your 150-word recommendation with in-text citations and a separate reference list.
- 11 I'm an orthopaedic registrar seeking advice for a 40-year-old male with a displaced midshaft clavicle fracture and a suspected subclavian vascular injury. He has poorly controlled diabetes and leads a sedentary lifestyle. Should I choose surgical or non-surgical treatment? Provide your 150-word recommendation with in-text citations and a separate reference list.
- 12 I'm an orthopaedic registrar seeking advice for a 60-year-old male with a displaced open midshaft clavicle fracture. He has pulmonary fibrosis, chronic kidney disease, and lives in a nursing home. Should I choose surgical or non-surgical treatment? Provide your 150-word recommendation with in-text citations and a separate reference list.
- 13 I'm an orthopaedic registrar seeking advice for a 60-year-old male with a displaced open midshaft clavicle fracture and a floating shoulder. He has pulmonary fibrosis, chronic kidney disease, and lives in a nursing home. Should I choose surgical or non-surgical treatment? Provide your 150-word recommendation with in-text citations and a separate reference list.
- 14 I'm an orthopaedic registrar seeking advice for a 70-year-old male with a midshaft clavicle fracture. His X-ray shows displacement of 100% of shaft width and 1 cm of shortening. He lives in a nursing home and has low functional demands. Should I choose surgical or non-surgical treatment? Provide your 150-word recommendation with in-text citations and a separate reference list.

TABLE 1: ChatGPT input questions on the management of midshaft clavicle fractures

Questions 5 to 14 were designed to simulate clinical decision-making scenarios that an orthopaedic clinician might present to ChatGPT for assistance with treatment decisions. The spectrum of management topics covered included clavicle fractures in paediatric and geriatric patients, comminuted fractures (both simple and complex), significant displacement and shortening, failed non-operative management, absolute surgical indications, and patient preferences. References were requested for all responses to assess the accuracy of

the information and recommendations provided. For patient-focused questions, four references were requested to allow for sufficient analysis across responses.

Each question was submitted three times on the same day to assess response consistency. All entries were conducted in separate chat sessions to prevent prior interactions from influencing subsequent responses. Screenshots were taken to document the responses, with examples provided in Appendices (Figures 3-7).

Assessment of outputs

The responses were independently evaluated by two orthopaedic registrars (CK and RC) and two senior orthopaedic surgeons (MS and BS). An established accuracy grading scale was used to assess each response: (I) comprehensive; (II) correct but inadequate; (III) mixed with correct and incorrect information; and (IV) completely incorrect [14]. Each reviewer assessed consistency, defined by similar accuracy grading, content, and recommendations across the three responses for each of the 14 questions. Consistency was assigned a dichotomous grade of "yes" or "no," and any discrepancies in grading were resolved through discussion until a consensus was reached.

The accuracy of the references provided in the responses was verified through independent searches conducted in MEDLINE, Embase, and Scopus, including searches of article titles, abstracts, keywords, and authors. References were classified as incorrect if they could be located but contained inaccuracies (such as errors in the title, authors, publication date, or journal), or fabricated if they could not be located at all. Both correct and incorrect references were further analysed to determine whether the response accurately reflected the cited material.

Data analysis

Descriptive statistics, including frequencies, counts, percentages, and the standard error of the mean (SEM) for relevant variables, were calculated using Microsoft Excel v16.77 (Microsoft Corporation, Redmond, USA).

Results

ChatGPT response characteristics

ChatGPT was presented with 14 questions, each repeated three times, generating 42 responses. The mean response length was 135.9 ± 3.9 words (SEM) (range: 108-197 words). A total of 128 references were provided with a mean of 3.0 references per response. ChatGPT consistently included in-text citations and reference lists as requested; however, in one response to Question 1, it provided only three references, despite being asked for four.

For patient-focused questions (Questions 1 to 4), the mean response length was 168.9 ± 6.6 words (SEM) (range: 123-197 words), with a mean of 3.9 references per response (range: 3-4). For clinical decision-making questions (Questions 5 to 14), responses averaged 122.7 ± 1.54 words (SEM) (range: 108-140 words), with 2.7 references per response (range: 2-4). Examples of ChatGPT responses to a patient-focused question and a clinical decision-making question are provided in Figures 1-2, respectively.

In 200 words, outline the indications for surgical and non-surgical treatment of midshaft clavicle fractures, supported by evidence from four high-quality journal references. Ensure that the references are cited correctly both in-text and listed separately in a reference list.

Midshaft clavicle fractures are commonly treated based on displacement, symptoms, and functional needs. Non-surgical management is often recommended for undisplaced or minimally displaced fractures. Treatment typically involves a sling or brace for immobilization, analgesics, and physiotherapy. A study by Nove-Josserand et al. (2004) supports this, showing that non-surgical management yields satisfactory results in terms of functional recovery and pain relief for non-displaced fractures (Nove-Josserand, J., et al., 2004).

In contrast, surgical intervention is indicated for displaced fractures, particularly those with significant shortening or skin tenting. A systematic review by Roberson et al. (2018) found that surgical fixation often leads to better functional outcomes and reduced rates of nonunion compared to conservative treatment in displaced fractures (Roberson, J., et al., 2018). Surgical options include plate fixation or intramedullary fixation, which are effective in stabilizing the fracture and promoting early return to function.

References:

1. Nove-Josserand, J., et al. (2004). Nonoperative treatment of mid-shaft clavicle fractures: A randomized trial. *Journal of Orthopaedic Trauma*, 18(7), 406-411.
2. Roberson, J., et al. (2018). Operative vs. nonoperative treatment of midshaft clavicle fractures: A meta-analysis. *Journal of Shoulder and Elbow Surgery*, 27(2), 314-322.
3. Neer, C. S. (1960). Fractures of the clavicle. *Journal of the American Medical Association*, 172(4), 283-289.
4. Handoll, H. H., et al. (2008). Surgical versus non-surgical treatment for acute mid-shaft clavicular fractures. *Cochrane Database of Systematic Reviews*, 4, CD007428.



Message ChatGPT



ChatGPT can make mistakes. Check important info.



FIGURE 1: ChatGPT response to the indications for surgical and non-surgical treatment of midshaft clavicle fractures (Question 1, response 3)

I'm an orthopaedic registrar seeking advice for a 10-year-old male with a midshaft clavicle fracture. His X-ray shows displacement of 100% of shaft width and 1 cm of shortening. Should I choose surgical or non-surgical treatment? Provide your 150-word recommendation with in-text citations and a separate reference list.

For a 10-year-old male with a midshaft clavicle fracture showing 100% displacement and 1 cm shortening, surgical intervention is generally recommended. In children, displaced midshaft clavicle fractures can be treated non-operatively with good results in many cases; however, significant displacement and shortening may increase the risk of nonunion or malunion, potentially leading to functional impairment or cosmetic concerns (Wipperman & Goerl, 2016). Surgical treatment, typically via open reduction and internal fixation, can restore proper alignment and function, especially in cases with substantial displacement and shortening (Hsu et al., 2011). Although non-surgical management is often effective, the degree of displacement in this case suggests that surgery may provide a more reliable outcome and prevent potential long-term complications.

References

Hsu, J. R., Koulouvaris, P., & Jones, K. (2011). *Surgical vs. Non-Surgical Management of Midshaft Clavicle Fractures: A Review*. *Journal of Orthopaedic Trauma*, 25(9), 608-614.

Wipperman, J., & Goerl, K. (2016). *Clavicle Fractures: Diagnosis and Management*. *American Family Physician*, 94(12), 964-970.



Message ChatGPT

ChatGPT can make mistakes. Check important info.

FIGURE 2: ChatGPT management recommendation for paediatric midshaft clavicle fracture (Question 5, response 2)

Information accuracy and consistency

The accuracy of all 42 responses was rated as Grade III, indicating a mix of correct and incorrect information. Consistency was observed in 11 of the 14 questions (78.6%) when comparing the three responses provided for each question (Table 2). There was agreement between graders regarding accuracy and consistency across all responses.

Question	R1 Grade (I-IV)	R1 Ref (n/nt)	R2 Grade (I-IV)	R2 Ref (n/nt)	R3 Grade (I-IV)	R3 Ref (n/nt)	Consistency (Y/N)
1	III	0/4	III	0/3	III	0/4	Y
2	III	0/4	III	1/4	III	0/4	Y
3	III	0/4	III	0/4	III	0/4	Y
4	III	0/4	III	0/4	III	0/4	Y
5	III	0/2	III	0/2	III	0/3	N
6	III	0/3	III	0/3	III	0/3	Y
7	III	0/2	III	0/3	III	0/3	Y
8	III	0/3	III	0/3	III	0/2	N
9	III	0/3	III	0/3	III	0/2	Y
10	III	0/3	III	0/4	III	0/2	Y
11	III	0/3	III	0/3	III	0/3	Y
12	III	0/3	III	0/3	III	0/3	N
13	III	0/2	III	0/2	III	0/2	Y
14	III	0/2	III	0/4	III	0/2	Y

TABLE 2: Evaluation of the accuracy and consistency of ChatGPT's responses
'R1', first response; 'R2', second response; 'R3', third response; Ref, reference(s); 'n', number; 'nt', total number; 'Y', yes; 'N', no
Grades I-IV: I. Comprehensive, II. Correct but inadequate, III. Mixed with correct and incorrect information, IV. Completely incorrect

Reference analysis

Of the 128 references presented by ChatGPT, one was correct (0.8%), 14 contained incorrect details (10.9%), and 113 appeared fabricated (88.3%). Of the 15 references that could be verified (both correct and incorrect references), four (3.1%) accurately represented the research conclusions from the cited sources.

Discussion

This study demonstrates the limitations of ChatGPT in providing accurate and consistent responses to both patient-focused inquiries and clinical recommendations for midshaft clavicle fractures. Notably, each response contained incorrect information, as well as fabricated references.

ChatGPT continues to see widespread adoption, with usage expanding rapidly across various sectors, including healthcare [10,12,13]. Its responses are becoming more sophisticated, making it increasingly difficult to distinguish AI-generated content from that of medical professionals [15]. This trend highlights the need for ongoing evaluation of ChatGPT's output accuracy and the credibility of its cited sources.

Kung et al. [16] reported that ChatGPT provided verifiable references 87.9% of the time when answering Orthopaedic In-Training Examination questions. In contrast, our study found that 88.3% of ChatGPT's references were fabricated, with only 11.7% being verifiable. A possible explanation for this discrepancy is that ChatGPT's reference accuracy may vary depending on the query type. Nonetheless, these findings emphasise the need for significant improvements in its ability to deliver consistent evidence-based information prior to incorporation in patient or clinician decision-making algorithms.

A recent systematic review and meta-analysis revealed that most evaluations of LLMs focus on medical knowledge tasks like answering licensing exam questions (44.5%) and making diagnoses (19.5%), while only 3.3% and 0.4% of these studies focused on orthopaedics and physical medicine, respectively. The authors recommended further studies that evaluate the use of clinical data, as well as expanding research into underrepresented specialties [17].

Evaluations of ChatGPT's performance in orthopaedics are mixed. Cuthbert and Simpson found that ChatGPT lacked accuracy and higher-order judgment when answering orthopaedic surgery and trauma-related questions, with 35.8% of answers being correct [18]. In contrast, Christy et al. [19] reported that 78%

of ChatGPT's responses to common distal radius fracture questions were deemed appropriate by a consensus of three fellowship-trained hand and upper-extremity surgeons. Truhn et al. [20] found that ChatGPT can provide largely accurate and clinically relevant treatment recommendations for common knee and shoulder conditions and is not prone to hallucinations. Seth et al. [21] assessed ChatGPT's effectiveness in generating information on hip osteoarthritis and reported that, although the responses were generally accurate, they were often superficial, lacked judgment, and omitted important research. Knee et al. [22] reported that early models of ChatGPT demonstrated significant limitations in accuracy and consistency when providing information on distal radius fractures. Other studies on carpal tunnel syndrome and scaphoid fractures have acknowledged ChatGPT's ability to provide validated information but highlighted inconsistencies in diagnostic pathways and clinical guidance [23,24].

In response to patient-focused questions in this study, ChatGPT generally recommended non-surgical management for non-displaced or minimally displaced midshaft clavicle fractures. Surgical management was consistently reported to offer benefits such as faster recovery, reduced non-union rates, and improved long-term functional outcomes. Whilst the literature supports these conclusions, the vast majority of the references provided by ChatGPT were fabricated. Additionally, ChatGPT failed to present a balanced discussion of evidence showing no significant long-term functional differences between surgical and non-surgical management [3,4,25]. Given that ChatGPT's knowledge extends to October 2023, it is unlikely that outdated input data accounts for this oversight [26]. Conflicting sources of information with varying weighting in ChatGPT's algorithms may be a contributory factor. Nevertheless, providing balanced information is crucial for facilitating collaborative, shared decision-making with patients, especially given the controversies surrounding the optimal management of clavicle fractures, and is a major limitation of the LLM tested.

The evaluation of ChatGPT's clinical decision-making on midshaft clavicle fractures revealed mixed results in terms of both accuracy and consistency. ChatGPT consistently recommended surgical management for Z-type fracture patterns, cases with skin tenting, failed non-surgical management, gross displacement, shortening, and subclavian vascular injury. It also recommended non-surgical management for geriatric patients with low functional demands. While these recommendations were generally appropriate, the reasoning provided often lacked depth and expert knowledge and frequently diverged from the evidence cited in its sources.

ChatGPT did not consistently recognise open fractures as an absolute indication for surgery and provided inconsistent advice for paediatric patients and simple comminuted fractures (see Figures 3-7 within Appendices). It made decisions without clarifying important information about patient comorbidities. For instance, in both Questions 12 and 13 (involving open fractures in patients with chronic kidney disease and pulmonary fibrosis), ChatGPT failed to assess the severity of these conditions before recommending non-surgical management and did not suggest a multidisciplinary approach to preoperative planning. These findings align with those of Sparks et al. [27] who observed that ChatGPT often omitted critical information, contained inaccuracies, and lacked thoroughness in addressing treatment options and relevant risk factors in orthopaedic cases.

A strength of this study was the use of Google Trends to formulate patient-focused questions, incorporating the most commonly searched terms related to clavicle fractures on Google and enhancing relevance to typical patient queries. Including both patient- and clinician-focused questions enabled a comprehensive evaluation of ChatGPT's utility across different user groups in orthopaedic care. Independent assessment by four orthopaedic reviewers reduced subjectivity in evaluating responses. Additionally, analysing the references provided by ChatGPT allowed for direct assessment of the accuracy of its responses and recommendations.

However, several limitations should be acknowledged. Although questions were based on commonly searched terms, they may not fully reflect how patients naturally phrase their queries when interacting with ChatGPT. Only single, isolated prompts were tested, which may not capture the more conversational or interactive nature of user engagement. Imposed word and reference limits may have restricted the depth of ChatGPT's responses. Notably, when no reference limit was specified, ChatGPT often provided fewer citations. Finally, this study used ChatGPT-4o mini - the most accessible and widely available version - though results may vary with other versions of the model.

Conclusions

This study found that ChatGPT has notable limitations in accuracy and consistency when providing patient information or recommending clinical decisions for midshaft clavicle fractures. Every response from ChatGPT contained incorrect or fabricated content. These findings highlight the need for patients and orthopaedic clinicians to exercise caution before integrating ChatGPT into clinical workflows.

Appendices

I'm an orthopaedic registrar seeking advice for a 10-year-old male with a midshaft clavicle fracture. His X-ray shows displacement of 100% of shaft width and 1 cm of shortening. Should I choose surgical or non-surgical treatment? Provide your 150-word recommendation with in-text citations and a separate reference list.

For a 10-year-old male with a midshaft clavicle fracture showing 100% displacement and 1 cm shortening, non-surgical treatment is generally recommended. Studies indicate that non-operative management, including the use of a figure-of-eight harness or sling, is often effective in children with displaced clavicle fractures, leading to satisfactory outcomes in most cases (Mehlman et al., 2005). Surgical intervention, such as open reduction and internal fixation, is typically reserved for severe cases with significant displacement, neurovascular compromise, or non-union risk (Wipperman & Goerl, 2016). In this case, given the patient's age and the nature of the displacement, non-surgical treatment should be the initial approach, with close follow-up to monitor healing and functional recovery.

References:

- Mehlman, C. T., & Wipperman, J. (2005). Clavicle fractures in children. *Journal of the American Academy of Orthopaedic Surgeons*, 13(4), 244-255.
- Wipperman, J., & Goerl, K. (2016). Clavicle fractures. *American Family Physician*, 94(12), 982-989.



Message ChatGPT



ChatGPT can make mistakes. Check important info.



FIGURE 3: ChatGPT management recommendation for paediatric midshaft clavicle fracture (Question 5, response 1)

I'm an orthopaedic registrar seeking advice for a 35-year-old male with a midshaft clavicle fracture. His X-ray shows minimal displacement with a comminuted fragment. He is an avid cyclist and wants to return to full function quickly. Should I choose surgical or non-surgical treatment? Provide your 150-word recommendation with in-text citations and a separate reference list.

For a 35-year-old male with a minimally displaced, comminuted midshaft clavicle fracture, non-surgical treatment is typically recommended unless there are specific functional concerns or associated injuries. Studies show that non-surgical management, including immobilization with a sling or figure-of-eight brace, results in satisfactory outcomes in terms of healing and functional recovery for most midshaft clavicle fractures (O'Brien et al., 2014; Neer, 1960). The majority of these fractures heal well without surgery, with good functional outcomes and a low rate of complications. Given that the fracture is minimally displaced and comminuted but not severely malaligned, non-surgical management should enable a return to full function, especially with adherence to rehabilitation exercises (Jensen et al., 2012). However, if the patient desires a quicker return to cycling or if there are any complications or specific concerns, surgical intervention might be considered.

References:

Jensen, M. B., Tuxøe, J., Hølmkjær, M. F., & Jensen, J. L. (2012). Non-operative versus operative treatment of mid-shaft clavicle fractures. *Acta Orthopaedica*, 83(2), 141-146.

Neer, C. S. (1960). Fractures of the Clavicle. *Journal of the American Medical Association*, 172(1), 29-35.

O'Brien, P. H., & Saed, M. (2014). Non-operative management of clavicle fractures: a systematic review. *Journal of Shoulder and Elbow Surgery*, 23(12), 1882-1889.



Message ChatGPT



ChatGPT can make mistakes. Check important info.



FIGURE 4: ChatGPT management recommendation for comminuted midshaft clavicle fracture (Question 8, response 1)

I'm an orthopaedic registrar seeking advice for a 35-year-old male with a midshaft clavicle fracture. His X-ray shows minimal displacement with a comminuted fragment. He is an avid cyclist and wants to return to full function quickly. Should I choose surgical or non-surgical treatment? Provide your 150-word recommendation with in-text citations and a separate reference list.

For a 35-year-old active cyclist with a minimally displaced midshaft clavicle fracture and a comminuted fragment, surgical intervention is often recommended. Non-surgical treatment, such as conservative management with a sling, is typically effective for minimally displaced fractures but might not suffice for fractures with comminuted fragments due to potential instability and prolonged recovery (Altamimi et al., 2018). Surgical options, including plate fixation, can provide better alignment and stability, which is crucial for an individual who seeks to return to high-function activities like cycling (Miller et al., 2020). Early surgical intervention can also minimize the risk of non-union or malunion, which might otherwise delay recovery and affect functional outcomes (Neer et al., 2007). Thus, considering the patient's activity level and fracture characteristics, surgical treatment might offer a more favorable and quicker return to full function.

References:

Altamimi, S., & Bhandari, M. (2018). *Clavicle fractures: A review of current management strategies*. *Journal of Orthopaedic Trauma*, 32(5), 254-262.

Miller, A. N., & Tashjian, R. Z. (2020). *Surgical versus nonsurgical treatment of midshaft clavicle fractures*. *Journal of Shoulder and Elbow Surgery*, 29(7), 1369-1377.

Neer, C. S., 2nd, & Rockwood, C. A. (2007). *Displaced midshaft clavicular fractures*. *Journal of Bone and Joint Surgery*, 89(1), 1-12.



Message ChatGPT

ChatGPT can make mistakes. Check important info.

FIGURE 5: ChatGPT management recommendation for comminuted midshaft clavicle fracture (Question 8, response 2)

I'm an orthopaedic registrar seeking advice for a 60-year-old male with a displaced open midshaft clavicle fracture. He has pulmonary fibrosis, chronic kidney disease, and lives in a nursing home. Should I choose surgical or non-surgical treatment? Provide your 150-word recommendation with in-text citations and a separate reference list.

Given the patient's age, comorbidities, and living situation, non-surgical management of the displaced open midshaft clavicle fracture is advisable. Surgical intervention, such as open reduction and internal fixation (ORIF), presents higher risks for individuals with pulmonary fibrosis and chronic kidney disease due to potential complications from anesthesia and impaired healing (Jung et al., 2021). Furthermore, this patient's nursing home residency suggests limited postoperative rehabilitation support, which may impact the effectiveness of surgical treatment (Ranger et al., 2023). Non-surgical treatment, including a sling and pain management, can be effective, with studies indicating satisfactory outcomes in similar cases (Robinson et al., 2004). However, close follow-up is essential to monitor for complications like nonunion or malunion, given the patient's comorbid conditions.

References

Jung, S. W., Lee, J. H., & Kim, H. K. (2021). *Surgical management of clavicle fractures in patients with comorbidities*. *Journal of Orthopaedic Trauma*, 35(2), 85-91.

Robinson, C. M., Goudie, E. B., & Murray, I. R. (2004). *Open reduction and internal fixation of displaced midshaft clavicular fractures: a randomized controlled trial*. *Journal of Bone and Joint Surgery*, 86(9), 1415-1421.

Ranger, T. A., Hamilton, J. A., & Mills, L. A. (2023). *Impact of post-acute care settings on surgical recovery and outcomes*. *Clinical Orthopaedics and Related Research*, 481(2), 456-463.



Message ChatGPT

ChatGPT can make mistakes. Check important info.

FIGURE 6: ChatGPT management recommendation for open midshaft clavicle fracture (Question 12, response 1)

I'm an orthopaedic registrar seeking advice for a 60-year-old male with a displaced open midshaft clavicle fracture. He has pulmonary fibrosis, chronic kidney disease, and lives in a nursing home. Should I choose surgical or non-surgical treatment? Provide your 150-word recommendation with in-text citations and a separate reference list.

For a 60-year-old male with a displaced open midshaft clavicle fracture, considering his comorbidities—pulmonary fibrosis, chronic kidney disease, and nursing home residence—surgical intervention might be preferred over non-surgical treatment. Open fractures typically require surgical management to prevent infection and ensure proper alignment, especially given his high risk of complications due to pulmonary fibrosis and chronic kidney disease (Graham et al., 2021). Surgical fixation can facilitate better fracture stabilization and potentially improve functional outcomes, which is crucial for patients with limited mobility and existing health challenges (Postacchini & Gumina, 2020). Non-surgical management could lead to prolonged immobilization, increasing the risk of complications such as pneumonia and muscle atrophy, particularly in elderly patients (Hsu et al., 2019). Therefore, surgical treatment is recommended, provided the patient can tolerate the procedure and post-operative care.

References

- Graham, J., Schaefer, R., & Reddy, S. (2021). Management of open clavicle fractures: A review of treatment options and outcomes. *Journal of Orthopaedic Trauma*, 35(7), 420-428.
- Hsu, J. R., & Mauffrey, C. (2019). Management of clavicle fractures in the elderly. *Orthopedic Clinics of North America*, 50(1), 51-59.
- Postacchini, F., & Gumina, S. (2020). The treatment of midshaft clavicle fractures: Surgical versus conservative management. *Clinical Orthopaedics and Related Research*, 478(3), 620-628.



Message ChatGPT

ChatGPT can make mistakes. Check important info.

FIGURE 7: ChatGPT management recommendation for open midshaft clavicle fracture (Question 12, response 3)

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Christopha J. Knee, Ryan J. Campbell, Brahman S. Sivakumar, Michael J. Symes

Acquisition, analysis, or interpretation of data: Christopha J. Knee, Ryan J. Campbell, Brahman S. Sivakumar, Michael J. Symes

Drafting of the manuscript: Christopha J. Knee, Ryan J. Campbell

Critical review of the manuscript for important intellectual content: Brahman S. Sivakumar, Michael J. Symes

Supervision: Brahman S. Sivakumar, Michael J. Symes

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue.

Animal subjects: All authors have confirmed that this study did not involve animal subjects or tissue.

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

- Morgan C, Bennett-Brown K, Stebbings A, Li L, Dattani R: Clavicle fractures. *Br J Hosp Med (Lond)*. 2020, 81:1-7. [10.12968/hmed.2020.0158](https://doi.org/10.12968/hmed.2020.0158)
- Postacchini F, Gumina S, De Santis P, Albo F: Epidemiology of clavicle fractures. *J Shoulder Elbow Surg*. 2002, 11:452-6. [10.1067/mse.2002.126613](https://doi.org/10.1067/mse.2002.126613)
- Qvist AH, Jensen SL: Minimal early functional gains after operative treatment of midshaft clavicular

- fractures: a meta-analysis of 10 randomized controlled trials including 1333 patients. *JSES Int.* 2024, 8:400-6. [10.1016/j.jseint.2023.12.011](#)
4. Martin JR, Saunders PE, Phillips M, Mitchell SM, Mckee MD, Schemitsch EH, Dehghan N: Comparative effectiveness of treatment options for displaced midshaft clavicle fractures: a systematic review and network meta-analysis. *Bone Jt Open.* 2021, 2:646-54. [10.1302/2633-1462.28.BJO-2021-0112.R1](#)
5. Smeeing DP, van der Ven DJ, Hietbrink F, et al.: Surgical versus nonsurgical treatment for midshaft clavicle fractures in patients aged 16 years and older: a systematic review, meta-analysis, and comparison of randomized controlled trials and observational studies. *Am J Sports Med.* 2017, 45:1937-45. [10.1177/0363546516673615](#)
6. Nonoperative treatment compared with plate fixation of displaced midshaft clavicular fractures. A multicenter, randomized clinical trial. *J Bone Joint Surg Am.* 2007, 89:1-10. [10.2106/JBJS.F.00020](#)
7. Sepehri A, Guy P, Roffey DM, O'Brien PJ, Broekhuysen HM, Lefaivre KA: Assessing the change in operative treatment rates for acute midshaft clavicle fractures: incorporation of evidence-based surgery results in orthopaedic practice. *JB JS Open Access.* 2023, 8: [10.2106/JBJS.OA.22.00096](#)
8. Campbell RJ, Handford C, Donaldson MJ, Sivakumar BS, Jiang E, Symes M: Surgical management of clavicle fractures in Australia: an analysis of Australian Medicare Benefits Schedule database from 2001 to 2020. *ANZ J Surg.* 2023, 93:656-62. [10.1111/ans.18312](#)
9. Burrus MT, Werner BC, Starman JS, Kurkis GM, Pierre JM, Diduch DR, Hart JM: Patient perceptions and current trends in internet use by orthopedic outpatients. *HSS J.* 2017, 13:271-5. [10.1007/s11420-017-9568-2](#)
10. ChatGPT statistics. (2024). Accessed: October 19, 2024: <https://www.demandsage.com/chatgpt-statistics/>.
11. Shahriar S, Hayawi K: Let's have a chat! A conversation with ChatGPT: technology, applications, and limitations [PREPRINT]. *ArXiv.* 2023, [10.48550/arXiv.2302.13817](#)
12. Rahsepar AA, Tavakoli N, Kim GH, Hassani C, Abtin F, Bedayat A: How AI responds to common lung cancer questions: ChatGPT vs Google Bard. *Radiology.* 2023, 307: [10.1148/radiol.230922](#)
13. OpenAI, Achiam J, Adler S, et al.: GPT-4 technical report [PREPRINT]. *ArXiv.* 2024, [10.48550/arXiv.2303.08774](#)
14. Yeo YH, Samaan JS, Ng WH, et al.: Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. *Clin Mol Hepatol.* 2023, 29:721-32. [10.3350/cmh.2023.0089](#)
15. Thirunavukarasu AJ, Ting DS, Elangovan K, Gutierrez L, Tan TF, Ting DS: Large language models in medicine. *Nat Med.* 2023, 29:1930-40. [10.1038/s41591-023-02448-8](#)
16. Kung JE, Marshall C, Gauthier C, Gonzalez TA, Jackson JB 3rd: Evaluating ChatGPT performance on the Orthopaedic In-Training Examination. *JB JS Open Access.* 2023, 8: [10.2106/JBJS.OA.23.00056](#)
17. Bedi S, Liu Y, Orr-Ewing L, et al.: Testing and evaluation of health care applications of large language models: a systematic review. *JAMA.* 2025, 333:319-28. [10.1001/jama.2024.21700](#)
18. Cuthbert R, Simpson AI: Artificial intelligence in orthopaedics: can chat generative pre-trained transformer (ChatGPT) pass section 1 of the fellowship of the Royal College of Surgeons (Trauma & Orthopaedics) Examination?. *Postgrad Med J.* 2023, 99:1110-14. [10.1093/postmj/qgad053](#)
19. Christy M, Morris MT, Goldfarb CA, Dy CJ: Appropriateness and reliability of an online artificial intelligence platform's responses to common questions regarding distal radius fractures. *J Hand Surg Am.* 2024, 49:91-8. [10.1016/j.jhsa.2023.10.019](#)
20. Truhn D, Weber CD, Braun BJ, Bressem K, Kather JN, Kuhl C, Nebelung S: A pilot study on the efficacy of GPT-4 in providing orthopedic treatment recommendations from MRI reports. *Sci Rep.* 2023, 13: [10.1038/s41598-023-47500-2](#)
21. Seth I, Rodwell A, Tso R, Valles J, Bulloch G, Seth N: A conversation with an open artificial intelligence platform on osteoarthritis of the hip and treatment. *J Orthop Sports Med.* 2023, 112-20. [10.26502/josm.511500088](#)
22. Knee CJ, Campbell RJ, Graham DJ, Handford C, Symes M, Sivakumar BS: Examining the role of ChatGPT in the management of distal radius fractures: insights into its accuracy and consistency. *ANZ J Surg.* 2024, 94:1391-6. [10.1111/ans.19143](#)
23. Seth I, Xie Y, Rodwell A, Gracias D, Bulloch G, Hunter-Smith DJ, Rozen WM: Exploring the role of a large language model on carpal tunnel syndrome management: an observation study of ChatGPT. *J Hand Surg Am.* 2023, 48:1025-33. [10.1016/j.jhsa.2023.07.003](#)
24. Seth I, Lim B, Xie Y, Hunter-Smith DJ, Rozen WM: Exploring the role of artificial intelligence chatbot on the management of scaphoid fractures. *J Hand Surg Eur Vol.* 2023, 48:814-18. [10.1177/17531934231169858](#)
25. Woltz S, Krijnen P, Schipper IB: Plate fixation versus nonoperative treatment for displaced midshaft clavicular fractures: a meta-analysis of randomized controlled trials. *J Bone Joint Surg Am.* 2017, 99:1051-7. [10.2106/JBJS.16.01068](#)
26. GPT-4o mini: advancing cost-efficient intelligence. (2024). Accessed: October 22, 2024: <https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/>.
27. Sparks CA, Fasulo SM, Windsor JT, Bankauskas V, Contrada EV, Kraeutler MJ, Scillia AJ: ChatGPT is moderately accurate in providing a general overview of orthopaedic conditions. *JB JS Open Access.* 2024, 9: [10.2106/JBJS.OA.23.00129](#)