Chirino A, Wiemken T, Furmanek S, et al. High consistency between recommendations by a pulmonary specialist and ChatGPT for the management of a patient with non-resolving pneumonia. *Norton Healthcare Medical Journal*. 2023;1(1).
doi:10.59541/001c.75456

Case Reports/Case Series

High consistency between recommendations by a pulmonary specialist and ChatGPT for the management of a patient with non-resolving pneumonia

Alejandro Chirino, MD[1], Timothy Wiemken, PhD[2], Stephen Furmanek, MPH[3], William Mattingly, MPH[3], Thomas Chandler, MPH[3], Guillermo Cabral, MD[4], Rodrigo Cavallazzi, MD[5], Ruth Carrico, PhD, APRN[6], Julio Ramirez, MD, FACP[3]

[1] Respira Salud. Clínica Integral and Hospital Italiano de Mendoza, [2] Department of Medicine, Saint Louis University, [3] Norton Infectious Diseases Institute, Norton Healthcare, [4] Hospital Rossi, La Plata, Argentina, [5] Department of Medicine, University of Louisville, [6] Division of Infectious Diseases, University of Louisville

Keywords: ChatGpt, Artificial intelligence, Pulmonary Medicine

https://doi.org/10.59541/001c.75456

Norton Healthcare Medical Journal

Vol. 1, Issue 1, 2023

The use of artificial intelligence in healthcare is rapidly growing, with great potential to assist medical professionals in diagnosis and treatment of complex medical conditions. In this case report, the recommendations of a pulmonary consultation are compared to the recommendations from a large language model (LLM), ChatGPT, in regards to a case of non-resolving pneumonia. The recommendations by ChatGPT were found to be consistent with the pulmonary consultation, and no recommendations given were inappropriate. As shown in this case report, early generations of LLM such as ChatGPT were able to provide appropriate recommendations for the management of a patient with a complex medical condition.

INTRODUCTION

The use of Artificial Intelligence (AI) in healthcare is rapidly growing, and its applications are being utilized in various areas, including medical imaging and diagnosis.[1-3] Growing interest also is focused on the usability of large language models (LLMs),[4] and there is a need to evaluate their appropriateness for use by practicing clinicians. Such a comparison would provide a unique insight into the potential of AI in healthcare and its ability to assist medical professionals in the diagnosis and treatment of complex medical conditions. The purpose of this case report was to compare the recommendations of the pulmonologist and the LLM and to evaluate the accuracy and usefulness of AI in helping decision-making in pulmonary medicine.

In this article we present a real clinical case from daily practice. Briefly, the pulmonary service received a consultation from the primary team in regard to an admitted patient. Using this clinical case, we compared the consultation advice given by the pulmonary physician to the advice generated by a LLM, in this case, ChatGPT.[5] During the follow-up of the patient, the advice given by the pulmonary physician was also compared with advice given by ChatGPT. We structured a grid to classify the appropriateness of the AI answers into a three categories as follows: 1) Highly appropriate, when the response was both medically correct and specific to the case at hand, 2) Moderately appropriate, when the answer was medically correct but in general or not specific to the case at hand, and 3) inappropriate, when the answer was incorrect or potentially harmful.

CLINICAL CASE

An 81-year-old man with a history of ischemic stroke and subsequent hearing loss was admitted to the clinical ward with headache, asthenia and pulmonary infiltrates. Atrial fibrillation was diagnosed during evaluation in the emergency department. With a clinical diagnosis of community-acquired pneumonia, the patient was started on intravenous ampicillin/sulbactam plus azithromycin. Amiodarone was started for atrial fibrillation. The results of a viral panel and RT-PCR for SARS-CoV-2 were negative.

After hospitalization, the patient developed progressive dyspnea and fever, and leukocyte count increased from 12,500 per cubic millimeter to 15,700. On the fifth day a repeat CT scan showed progression of the infiltrates with moderate bilateral pleural and pericardial effusion. On the sixth day, an echocardiogram confirmed pericardial effusion with normal left ventricular function. On day 7, blood and sputum were re-collected for culture and the antibiotic therapy was escalated to intravenous piperacillin/tazobactam plus vancomycin. Repeat cultures were negative. On day 9, respiratory status deteriorated and the patient was admitted to the ICU for non-invasive ventilation. After 4 days in the ICU, the patient improved and returned to the medical ward but his signs and symptoms had not resolved. On the 15th day, a 3rd CT scan showed further progression of the pulmonary infiltrates. At this point, a pulmonary consultation was obtained.

ASSESSMENT AND RECOMMENDATIONS BY THE PULMONARY CONSULTANT

The pulmonary consultant assessed the patient and found shortness of breath, bilateral crackles, fever of 37.8°C, and need for oxygen via nasal cannula. The pulmonary consultant discussed the patient with the primary team and wrote a consultation with specific recommendations. After the evaluation of the patient the conclusion of the pulmonologist was the following:

The most likely clinical diagnosis was non-resolving pneumonia. The pulmonary consultant suspected a non-covered microorganism, such as *Mycobacteria* due to tree-in-bud infiltrates at CT scan in a possible immunocompromised host. Alternative clinical diagnoses were exacerbation of an undiagnosed usual interstitial pneumonia or other acute or subacute interstitial pneumonia of autoimmune or drug-related (amiodarone) origin. A consideration was given to superimposed renal and/or cardiac failure.

In regards to the management of the patient, it was recommended to perform a bronchoscopy with bronchoalveolar lavage (BAL) and transbronchial biopsy, including cell population analysis in BAL. The pulmonary consultant also recommended a full autoimmune panel and a repeat echocardiogram to evaluate possible fluid overload. If the patient continued to clinically deteriorate, it was recommended that the patient be readmitted to ICU with corticosteroid treatment and consideration for a new antibiotic treatment. The following day, the patient was afebrile. The bronchoscopy was performed, with negative gram stain analysis. Cultures, cellular population analysis in BAL and biopsy were pending. Echocardiogram showed no signs of left ventricular failure. Renal function was preserved. Two days later, the patient remained afebrile and slowly recovered with improvement in dyspnea, and reduced need for oxygen therapy. BAL cultures were negative. Autoimmune panel was negative. It was decided to discharge the patient home with a close follow-up in clinic.

ASSESSMENT AND RECOMMENDATIONS BY CHATGPT

After the consultation, the clinical data of the patient was presented to ChatGPT on the same day. ChatGPT offered the following recommendations:

1. *Re-evaluate the diagnosis and consider additional testing to rule out infections, such as fungal or mycobacterial infections.*
2. *Check for drug-resistant bacteria and consider changing the antibiotic regimen.*
3. *Assess for underlying conditions such as immune suppression, malignancy or a pulmonary embolism.*
4. *Order a bronchoalveolar lavage (BAL) or a biopsy to obtain a definitive diagnosis.*
5. *Consider corticosteroid treatment if there is evidence of an inflammatory process.*
6. *Evaluate the patient for ICU admission if his condition deteriorates.*

To refine the suggestions, ChatGPT was provided a transcript of the CT scan results showing a pattern of bilateral progressive infiltrates with ground glass opacities, reticular interstitial thickening and some traction bronchiectasis, as well as the mild bilateral pleural effusion and mild pericardial effusion.

ChatGPT offered the following updated recommendations:

1. *Consider further diagnostic tests to confirm an etiology, such as BAL or biopsy.*
2. *Rule out underlying autoimmune or interstitial lung disease.*
3. *Evaluate for possible fungal or mycobacterial infection.*
4. *Consider additional imaging studies, such as a chest MRI or a PET scan, to evaluate the extent of the disease.*
5. *Reassess the patient for ICU admission if the patient's condition continues to deteriorate.*
6. *Consider referral to a specialist for further management, such as a pulmonologist or an infectious disease specialist.*

The pulmonary consultant continued to discuss the case with ChatGPT and requested how they could rule out autoimmune diseases in a patient with no clinical signs or symptoms. ChatGPT offered the following recommendations:

1. *Serologic tests: including ANA, anti-dsDNA, anti-SSA/SSB, anti-Ro/La, anti- CCP, anti-MPO and anti-GRAN.*
2. *Lung function tests: including spirometry, diffusion capacity, and methacholine challenge test.*
3. *High-resolution CT scan of the chest to look for characteristic patterns.*
4. *Bronchoalveolar lavage (BAL) to obtain cells for analysis.*
5. *Pulmonary function tests: to evaluate for restrictive or mixed lung disease patterns.*
6. *Evaluation of serum levels of cytokines and chemokines: to assess for a systemic inflammatory response.*

Note: The presence of certain autoantibodies or characteristic imaging findings can support a diagnosis of autoimmune disease. However, further testing and specialist consultation may be necessary to make a definite diagnosis.

The following day, ChatGPT was informed that the bronchoscopy was normal and the results from the BAL and autoimmune serologic tests were pending. ChatGPT was also informed that the patient was stable and afebrile with no worsening in his pulmonary condition. It was asked for further recommendations, and offered the following:

1. *Wait for the results of the bronchoalveolar lavage (BAL) and serologic tests to determine the etiology of the patient's condition.*
2. *Consider repeating the CT scan in a few days to monitor for changes.*
3. *Reassess the patient daily for signs of worsening and consider ICU admission if necessary.*

4. *If a specific diagnosis is made, start appropriate treatment for the underlying condition.*
5. *If a definite diagnosis is not made, consider repeating the bronchoscopy with BAL or biopsy for a more definitive diagnosis.*
6. *Consider referral to a specialist for further management, such as a pulmonologist or infectious disease specialist.*

After another two days, ChatGPT was informed of the new results, that the BAL culture was negative and autoimmune tests were negative, and that the patient improved clinically with improvement in dyspnea, oxygenation, and remained afebrile. When requested for additional recommendations, ChatGPT offered the following:

1. *Discharge the patient if he is stable and has no signs of worsening, with a plan for follow-up.*
2. *Consider repeating the CT scan to monitor for further improvement.*
3. *Prescribe oral antibiotics for a few more days to ensure that the patient is fully treated.*
4. *Provide follow-up care to monitor for recurrence of symptoms.*
5. *Advise the patient to seek medical attention immediately if symptoms worsen.*
6. *Consider pulmonary rehabilitation if the patient has residual dyspnea.*

APPROPRIATENESS OF CHATGPT RECOMMENDATIONS

From a total of 31 recommendations obtained from ChatGPT, 27 were considered highly appropriate (87%) and 4 were considered to be moderately appropriate (13%). No recommendations were considered to be inappropriate. See Table 1 for categorization of each recommendation by ChatGPT.

DISCUSSION

This clinical case indicates high consistency between recommendations performed by a pulmonary specialist during initial consultation and follow-up of a patient with non-resolving pneumonia and the suggestions generated from an artificial intelligence platform.

The interface with LLMs, as ChatGPT, allows end users to participate in dialogue and provide feedback. This allows clinicians to further hone in the discussion with relevant details to the AI, mirroring the interaction that the treating physician may have with other peers when discussing a clinical case. This case suggests that future generations of AI could become a useful tool to support clinical decision-making by practicing physicians. Other possible applications in medicine may be in the area of medical education, where students can discuss cases and learn from an AI model, and in the area of clinical research.

When discussing the use of ChatGPT to provide recommendations, we expected that some recommendations may not be clinically appropriate, considering that ChatGPT has obtained medical knowledge only from open sources. We were surprised that in this particular clinical case almost all recommendations were clinically sound. A recent study evaluating recommendations for the prevention of cardiovascular disease also found that this kind of LLM provided appropriate guidance.[4]

There is great potential for AI to impact the way that medicine is currently practiced. With its ability to analyze and synthesize a large amount of data faster than a human can, AI can help in the day-to-day clinical decision-making process. Additionally, it can avoid the variability in human medical care. For example, it can provide support during shifts or days with less availability of healthcare personnel, such as nights, weekends, and holidays. All of the above may improve the quality of medical practice.[6,7]

We most likely are at the beginning of a new revolution in medical sciences.[8] As shown in this case report, early generations of LLM such as ChatGPT were able to provide appropriate recommendations for the management of a patient with a complex medical condition. As AI continues to improve, research using this novel technology will be important to define the appropriate applications of AI in the field of medicine.

Submitted: April 10, 2023 EDT, Accepted: May 17, 2023 EDT

Table 1. Categorization of each Recommendation by ChatGPT

LLM response. First row	**Appropriateness evaluation**
Re-evaluate the diagnosis and consider additional testing to rule out infections, such as fungal or mycobacterial infections	Highly Appropriate
Check for drug-resistant bacteria and consider changing the antibiotic regimen	Highly Appropriate
Assess for underlying conditions such as immune suppression, malignancy or a pulmonary embolism	Highly Appropriate
Order a bronchoalveolar lavage (BAL) or a biopsy to obtain a definitive diagnosis	Highly Appropriate
Consider corticosteroid treatment if there is evidence of an inflammatory process	Highly Appropriate
Evaluate the patient for ICU admission if his condition deteriorates.	Highly Appropriate
LLM response. Second row	**Appropriateness evaluation**
Consider further diagnostic tests to confirm an etiology, such as BAL or biopsy	Highly Appropriate
Rule out underlying autoimmune or interstitial lung disease	Highly Appropriate
Evaluate for possible fungal or mycobacterial infection	Highly Appropriate
Consider additional imaging studies, such as a chest MRI or a PET scan, to evaluate the extent of the disease	Moderately Appropriate
Reassess the patient for ICU admission if the patient's condition continues to deteriorate	Highly Appropriate
Consider referral to a specialist for further management, such as a pulmonologist or an infectious disease specialist	Highly Appropriate
LLM response. Third row	**Appropriateness evaluation**
Serologic tests: including ANA, anti-dsDNA, anti-SSA/SSB, anti-Ro/La, anti- CCP, anti-MPO and anti-GRAN	Highly Appropriate
Lung function tests: including spirometry, diffusion capacity, and methacholine challenge test	Moderately Appropriate
High-resolution CT scan of the chest to look for characteristic patterns	Moderately Appropriate
Bronchoalveolar lavage (BAL) to obtain cells for analysis	Highly Appropriate
Pulmonary function tests: to evaluate for restrictive or mixed lung disease patterns	Highly Appropriate
Evaluation of serum levels of cytokines and chemokines: to assess for a systemic inflammatory response	Highly Appropriate
Note: The presence of certain autoantibodies or characteristic imaging findings can support a diagnosis of autoimmune disease. However, further testing and specialist consultation may be necessary to make a definite diagnosis	Highly Appropriate
LLM response. Fourth row	**Appropriateness evaluation**
Wait for the results of the bronchoalveolar lavage (BAL) and serologic tests to determine the etiology of the patient's condition	Highly Appropriate
Consider repeating the CT scan in a few days to monitor for changes.	Highly Appropriate
Reassess the patient daily for signs of worsening and consider ICU admission if necessary	Highly Appropriate
If a specific diagnosis is made, start appropriate treatment for the underlying condition	Highly Appropriate
If a definite diagnosis is not made, consider repeating the bronchoscopy with BAL or biopsy for a more definitive	Highly

diagnosis	Appropriate
Consider referral to a specialist for further management, such as a pulmonologist or infectious disease specialist	Highly Appropriate
LLM response. Fifth row	**Appropriateness evaluation**
Discharge the patient if he is stable and has no signs of worsening, with a plan for follow-up	Highly Appropriate
Consider repeating the CT scan to monitor for further improvement	Highly Appropriate
Prescribe oral antibiotics for a few more days to ensure that the patient is fully treated.	Highly Appropriate
Provide follow-up care to monitor for recurrence of symptoms	Highly Appropriate
Advise the patient to seek medical attention immediately if symptoms worsen	Highly Appropriate
Consider pulmonary rehabilitation if the patient has residual dyspnea	Moderately Appropriate
Highly appropriate: 87% Moderately appropriate: 13% Inappropriate: 0	

REFERENCES

1. Obermeyer Z, Emanuel EJ. Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. *N Engl J Med*. 2016;375(13):1216-1219. doi:10.1056/nejmp1606181

2. Chen M, Decary M. Artificial intelligence in healthcare: An essential guide for health leaders. *Healthc Manage Forum*. 2020;33(1):10-18. doi:10.1177/0840470419873123

3. Khemasuwan D, Sorensen JS, Colt HG. Artificial intelligence in pulmonary medicine: computer vision, predictive model and COVID-19. *Eur Respir Rev*. 2020;29(157):200181. doi:10.1183/16000617.0181-2020

4. Sarraju A, Bruemmer D, Van Iterson E, Cho L, Rodriguez F, Laffin L. Appropriateness of Cardiovascular Disease Prevention Recommendations Obtained From a Popular Online Chat-Based Artificial Intelligence Model. *JAMA*. 2023;329(10):842-844. doi:10.1001/jama.2023.1044

5. OpenAI. ChatGPT. Accessed January 15, 2023. https://chat.openai.com/

6. Garcia-Vidal C, Sanjuan G, Puerta-Alcalde P, Moreno-García E, Soriano A. Artificial intelligence to support clinical decision-making processes. *EBioMedicine*. 2019;46:27-29. doi:10.1016/j.ebiom.2019.07.019

7. Chen YM, Kao Y, Hsu CC, et al. Real-time interactive artificial intelligence of things–based prediction for adverse outcomes in adult patients with pneumonia in the emergency department. *Acad Emerg Med*. 2021;28(11):1277-1285. doi:10.1111/acem.14339

8. Beam AL, Drazen JM, Kohane IS, Leong TY, Manrai AK, Rubin EJ. Artificial Intelligence in Medicine. *N Engl J Med*. 2023;388(13):1220-1221. doi:10.1056/nejme2206291