

Appendix to:

Claude 3.5 Sonnet outperforms human readers on New England Journal of Medicine clinical challenge cases

Supplemental Methods

Claude Prompt

The prompt was developed based on the researcher's experience and adapted from prior research that focused on evaluating ChatGPT.⁶⁶

"You will be tasked with answering a series of complex medical cases.

For each case, you will be provided with a plaintext presentation. Any relevant tables and medical images will also be provided as attachments. Please answer the multiple choice question at the end of each case, and provide your reasoning.

If the input message is too long or the images are unable to be viewed, please inform me and disregard the question.

Here is the prompt:"

Claude Querying

For each trial, we entered the case in a new chat to prevent previous answers from influencing future cases. For each trial, the plain text of the case challenge prompt was submitted in the text box and supplemental figures, tables, and graphs were all attached in .jpg format.

Supplemental Results

Example Response #1 (Case 1, Trial 1, Result = Incorrect):

Thank you for providing this detailed medical case. I've reviewed the patient history, examination findings, and imaging results. I'll now answer the multiple choice question and provide my reasoning.

The most likely diagnosis in this case is:

3. Metastatic melanoma

Reasoning:

1. Patient history:

- The patient has a significant history of melanoma, including:
- Initial diagnosis of superficial spreading melanoma on the right upper back two decades ago
- In situ superficial spreading melanoma on the left upper back 11 years later
- Locally advanced melanoma in the right axilla diagnosed 3 years ago, with 25 out of 39 lymph nodes positive for metastatic melanoma

2. Recent treatment:

- The patient received radiation therapy and pembrolizumab (an immune checkpoint inhibitor) for the locally advanced melanoma 3 years ago.

3. Current presentation:

- New abnormal findings on chest imaging during routine quarterly surveillance.
- CT chest findings include:
- Patchy consolidation in the left upper lobe
- Bilateral hilar and mediastinal lymphadenopathy
- Numerous bilateral scattered solid pulmonary nodules with ground-glass halos
- These findings are new compared to imaging from 3 months ago.

4. Other considerations:

- The patient is asymptomatic, which can occur in early metastatic disease.
- The BRAF V600E mutation was identified in the previous melanoma, which is associated with a higher risk of metastasis.

- The patient's recent travel and exposure history make infectious causes less likely, especially given the negative interferon- γ release assay for tuberculosis and normal 1,3- β -D-glucan level.

5. Imaging correlation:

- The CT images (C and D) show multiple nodules with surrounding ground-glass halos, which are consistent with metastatic disease, particularly in the context of melanoma.
- The PET-CT images from 3 years ago (E and F) show the initial right axillary involvement, demonstrating the patient's history of metastatic disease.

While other diagnoses like sarcoidosis or infection (e.g., histoplasmosis) could present with similar imaging findings, the patient's extensive history of melanoma, recent treatment with immunotherapy, and the sudden appearance of these lesions on surveillance imaging strongly suggest metastatic melanoma as the most likely diagnosis.

The other options are less likely:

- Castleman's disease typically presents with more localized lymphadenopathy.
- Histoplasmosis is possible given the travel history, but less likely given the negative fungal marker and lack of symptoms.
- Sarcoidosis could present similarly but is less likely given the patient's history and the sudden onset of findings.
- Silicosis is unlikely without a history of occupational exposure to silica dust.
- Tuberculosis is less likely given the negative interferon- γ release assay and lack of symptoms.

In conclusion, metastatic melanoma is the most likely diagnosis given the patient's history, imaging findings, and clinical context.

Example Response #2 (Case 1, Trial 2, Result = Correct):

Based on the clinical presentation and imaging findings, the most likely diagnosis in this case is:

2. Histoplasmosis

Reasoning:

1. Patient history:

- 47-year-old woman with a history of malignant melanoma

- Recent travel to South America, including trips to the jungle and swimming in fresh water
- Family members experienced generalized myalgias and malaise after the trip

2. Imaging findings:

- CT chest shows new patchy consolidation in the left upper lobe
- Bilateral hilar and mediastinal lymphadenopathy
- Numerous bilateral scattered solid pulmonary nodules with associated ground-glass halos
- These findings were new compared to imaging 3 months prior

3. Laboratory results:

- Normal blood counts and chemistry
- Negative interferon- γ release assay for *Mycobacterium tuberculosis*
- Normal 1,3- β -D-glucan level (ruling out many fungal infections)

4. Clinical presentation:

- Patient is asymptomatic despite the new radiographic findings
- No fever, weight loss, night sweats, or respiratory symptoms

The combination of new pulmonary nodules with ground-glass halos, mediastinal lymphadenopathy, and recent travel to an endemic area strongly suggests a diagnosis of histoplasmosis. *Histoplasma capsulatum* is endemic in parts of South America, and exposure often occurs in wooded areas or when disturbing soil containing bird or bat droppings.

The lack of symptoms is not unusual in immunocompetent individuals with acute histoplasmosis. The ground-glass halos around the nodules are characteristic of angioinvasive infections, including fungal infections like histoplasmosis.

While metastatic melanoma is a consideration given the patient's history, the sudden appearance of multiple new lesions with ground-glass halos is atypical for melanoma metastases, especially in a patient who has been stable on immunotherapy.

The negative interferon- γ release assay makes tuberculosis less likely. Sarcoidosis, silicosis, and Castleman's disease are less likely given the acute onset and recent travel history.

In conclusion, acute pulmonary histoplasmosis best fits the clinical picture, considering the patient's recent travel to an endemic area, the characteristic imaging findings, and the absence of symptoms typical of her underlying malignancy or other systemic diseases.