Nursing handoff source fidelity review

Case review October 2026

Draft for final review
Prepared for Robert Piedra BSN RN

Primary finding

The Claude response added a reassessment recommendation beyond the source-only documentation brief. Robert concurred with this concern. No material source-fidelity error was identified in the pasted ChatGPT response; that finding does not establish clinical validity or an overall model winner.

This review examines one synthetic case: a 68-year-old woman admitted with pneumonia, with a medical-unit shift handoff at 19:00. The evidence consists of the complete prompt and two responses pasted by the user.

Responses reviewed

Claude Haiku 5.5 Low effort

The principal finding is an unsupported follow-up recommendation (F01). A separate observation concerns whether the 19:00 pharmacy verification recheck is explicit (F02). Admission wording remains a terminology and timing ambiguity, not a confirmed error (F03).

ChatGPT-6 Instant effort

The pasted response retains the 19:00 verification status and outstanding tasks. No material source-fidelity error was identified in this review. Its follow-up wording stays tied to documented pending work.

What both responses retained

Both preserve the 18:45 assessment and the pending ceftriaxone, sputum collection and daughter’s update. Neither invents a completed ceftriaxone administration or represents the unsuccessful sputum attempt as patient refusal.

These are findings about two supplied outputs, not a model benchmark or safety certification. The model and effort labels above are reproduced exactly as reported by the user. The conversation reports 25 October 2026; actual execution timestamps were not logged.

Review method and scope

The review follows a project workflow of Understand, Check, Explain and Recheck. These are local workflow labels, not a named Evidently framework. The emphasis on task-specific criteria, documented reasoning and human review is consistent with Evidently's LLM evaluation guide [1]. No Evidently software evaluation was run, and no certification or endorsement is claimed.

How the evidence was examined

Understand the source-only task and its limits. Check each response against the timeline and requested format. Explain each concern with the source and output wording. Robert's feedback was used to adjudicate interpretations, preserving disagreement and unresolved questions. Rechecking revised outputs is the planned next stage and has not yet been performed.

The criteria below were formalized retrospectively from the prompt after the responses were available; they were not preregistered. This is an AI-assisted evaluation with Robert's limited adjudication, not a fully RN-validated or blinded review. No inter-rater agreement measure was collected.

Criteria and observations

Criterion	Claude response	ChatGPT response
Source grounding and no invention	F01 concern	No material error identified
Chronology and latest known status	F02 observation F03 ambiguity	Latest status retained
Assessment and pending tasks	Retained	Retained
Unknown is not normal or absent	Limits acknowledged	Limits acknowledged
Requested three part format	Present	Present

Qualitative observations only. No numerical score or model ranking was calculated.

Limits of this review

One synthetic case and two pasted responses cannot establish reliability across cases or repeated runs. There were no repeat generations, screenshots, verified model versions, verified settings or API parameters. The review does not judge medication appropriateness or clinical outcomes.

The complete prompt and outputs appear in Appendices A to C. Wording is preserved; Markdown styling and escaped spaces are normalized for readability.

[1] Evidently, LLM evaluation: a beginner's guide. Updated 19 May 2026. Full reference on page 4.

Findings requiring source precision

F01 Unsupported follow up recommendation

Criterion: Source grounding and no invented treatment recommendations.

Prompt instruction

> *“Do not invent findings, results, reasons for delays, treatment recommendations or completed actions.”*

Claude response

> *“Reassess vital signs and respiratory status, as no findings after 18:45 are documented.”*

Related wording in the Claude response

> *“Reattempt sputum collection per the team's plan.”*

The source records no reassessment directive or timing and no team plan for another sputum attempt. The reassessment line adds a recommendation rather than reporting the latest documented state. The reference to a team plan adds support that the source does not supply.

RN review status. Robert concurred with the unsupported follow-up concern, specifically reassessment. The “team’s plan” wording has not been separately adjudicated by Robert.

This is a source-fidelity finding. It does not imply that nursing reassessment requires a scheduled order or that the suggested actions are intrinsically unsafe.

Suggested correction. Remove the added reassessment recommendation and unsupported plan attribution. State that the last assessment was at 18:45 and that sputum collection remains outstanding.

F02 Latest verification recheck not explicit

Criterion: Preserve important times and the latest known status.

Source timeline

> *“19:00 — Outgoing nurse rechecks the order status and medication administration record. Pharmacy verification remains pending, and no ceftriaxone administration is recorded. Sputum collection and the daughter’s update remain outstanding.”*

Claude response

> *“Ceftriaxone: Not administered. Awaiting pharmacist verification (confirmed by pharmacy at 18:50). No reason for the wait or estimated completion time is documented. At 19:00, no administration is recorded. No antibiotic administration deadline is supplied.”*

The response preserves pending verification but anchors its confirmation at 18:50. It explicitly carries the 19:00 administration record, not the 19:00 verification recheck. This is a temporal-precision observation, not a claim that the response falsely reports completed verification.

RN review status. Pending Robert’s review.

Suggested correction. At 19:00, pharmacy verification remained pending and no ceftriaxone administration was recorded in the MAR.

Reviewer calibration and next steps

F03 Admission wording remains ambiguous

Criterion: Preserve chronology without assigning an unsupported event time.

Source timeline

> *“17:40 — Patient arrives in the emergency department with cough, fever and shortness of breath. She is subsequently admitted for pneumonia.”*

Claude response

> *“17:40: ED arrival and admission for pneumonia.”*

The initial AI review flagged this as a definitive admission-time error. Robert explained that admission may refer to the ED encounter and did not regard the wording as definitively incorrect. The final classification is terminology and timing ambiguity, not a confirmed error. The disagreement is retained as reviewer calibration.

Suggested clearer wording. 17:40: Arrived in the ED with cough, fever and shortness of breath; subsequently admitted for pneumonia. Arrived on the medical unit at 18:40.

Proposed corrected excerpt

The following is a proposed edit, not a new model response or a completed retest.

> *The most recent documented assessment was at 18:45. At 19:00, pharmacy verification remained pending and no ceftriaxone administration was recorded in the MAR. Sputum collection remained outstanding after the cooperative but unsuccessful attempt at 18:55. The daughter requested an update at 18:58; the outgoing nurse had not completed it before handoff.*

Remaining review actions

- Obtain Robert’s review of F02 and the separate “team’s plan” wording in F01.
- Confirm final wording and publication approval before any public use.
- Retain source captures and verify the reported model labels and settings.
- Repeat the test on this case and additional cases using criteria fixed in advance.

Reference

[1] Evidently. LLM evaluation: a beginner’s guide. Updated 19 May 2026. Accessed 25 October 2026.

https://www.evidentlyai.com/llm-guide/llm-evaluation

Appendix A Source prompt

Complete prompt supplied for the documentation exercise

TASK

Create a concise nursing shift handoff for the incoming medical-unit nurse at 19:00 using only the fictional source information below. This is a documentation exercise, not advice for a real patient.

Organize the handoff into:

1. Situation and background

2. Current assessment

3. Outstanding tasks and follow-up

Preserve important event times and distinguish completed care from pending care. Do not invent findings, results, reasons for delays, treatment recommendations or completed actions. If information is not provided, do not treat it as normal or absent. Do not browse or use external sources.

PATIENT BACKGROUND

68-year-old woman admitted with pneumonia. Medical history includes hypertension and type 2 diabetes. She reports no known medication allergies.

SOURCE TIMELINE

17:40 — Patient arrives in the emergency department with cough, fever and shortness of breath. She is subsequently admitted for pneumonia.

18:15 — Ceftriaxone and a sputum specimen are ordered. The medication dose and administration schedule are not included in this source.

18:35 — Emergency department nurse gives transfer report. Ceftriaxone has not been administered, and pharmacy verification is pending.

18:40 — Patient arrives on the medical unit.

18:45 — Receiving nurse completes the assessment. Patient is alert and oriented. Blood pressure 118/70 mmHg, heart rate 104 beats/min, respiratory rate 22 breaths/min, temperature 38.2°C, and oxygen saturation 94% on oxygen at 2 L/min by nasal cannula. She becomes short of breath walking to the bathroom and requires standby assistance. A peripheral IV is present in the right forearm.

18:50 — Nurse contacts pharmacy about ceftriaxone. Pharmacy confirms the order is awaiting pharmacist review. No explanation for the wait or estimated completion time is documented.

18:55 — Nurse attempts sputum collection. Patient cooperates but cannot produce a specimen. No specimen is collected.

18:58 — Daughter requests an update. The outgoing nurse has not completed the update before handoff.

19:00 — Outgoing nurse rechecks the order status and medication administration record. Pharmacy verification remains pending, and no ceftriaxone administration is recorded. Sputum collection and the daughter’s update remain outstanding.

ADDITIONAL SOURCE LIMITS

Laboratory and imaging results are not included. No later vital signs or assessment are provided. No required antibiotic administration deadline is supplied.

Return only the nursing handoff.

Appendix B Claude response

User reported label: Claude Haiku 5.5 Low effort

NURSING SHIFT HANDOFF — Medical Unit, 19:00

Fictional documentation exercise; source information only.

1. Situation and Background

68-year-old woman admitted with pneumonia. She arrived at the ED at 17:40 with cough, fever and shortness of breath, and arrived on the medical unit at 18:40. Medical history includes hypertension and type 2 diabetes. She reports no known medication allergies.

2. Current Assessment

Most recent documented assessment (18:45, receiving nurse):

- Alert and oriented.
- BP 118/70 mmHg, HR 104 beats/min, RR 22 breaths/min, temperature 38.2°C.
- SpO2 94% on oxygen at 2 L/min by nasal cannula.
- Becomes short of breath walking to the bathroom; requires standby assistance.
- Peripheral IV in right forearm.

No later vital signs or assessment are provided after 18:45.

Laboratory and imaging results are not included in the source information.

3. Outstanding Tasks and Follow-Up

Completed:

- 17:40: ED arrival and admission for pneumonia.
- 18:15: Ceftriaxone and sputum specimen ordered (dose and administration schedule not included in source).
- 18:35: ED nurse transfer report given.
- 18:40: Arrival on medical unit.
- 18:45: Receiving nurse assessment completed (see above).

Pending:

- **Ceftriaxone:** Not administered. Awaiting pharmacist verification (confirmed by pharmacy at 18:50). No reason for the wait or estimated completion time is documented. At 19:00, no administration is recorded. No antibiotic administration deadline is supplied.
- **Sputum specimen:** Not collected. Patient cooperated at 18:55 but could not produce a specimen. Collection remains outstanding.
- **Daughter's update:** Requested at 18:58. Not completed by outgoing nurse; remains outstanding.

Follow-up:

- Reassess vital signs and respiratory status, as no findings after 18:45 are documented.
- Confirm ceftriaxone verification status and administration with pharmacy.
- Reattempt sputum collection per the team's plan.
- Provide the daughter with an update.

Appendix C ChatGPT response

User reported label: ChatGPT-6 Instant effort

Nursing Shift Handoff — 19:00

1. Situation and Background

68-year-old female admitted for pneumonia. History of hypertension and type 2 diabetes. Patient reports no known medication allergies.

- 17:40: Presented to the ED with cough, fever, and shortness of breath.
- 18:15: Ceftriaxone and sputum specimen ordered.
- 18:35: ED transfer report received. Ceftriaxone had not been administered; pharmacy verification was pending.
- 18:40: Patient arrived on the medical unit.

2. Current Assessment

18:45 — Assessment completed:

- Alert and oriented.
- BP 118/70 mmHg; HR 104 bpm; RR 22/min.
- Temperature 38.2°C; SpO_2 94% on 2 L/min oxygen via nasal cannula.
- Shortness of breath with ambulation to the bathroom; requires standby assistance.
- Peripheral IV present in right forearm.

No subsequent vital signs or assessment findings are provided. Laboratory and imaging results are not included.

3. Outstanding Tasks and Follow-Up

Ceftriaxone — Not administered

- Ordered at 18:15.
- At 18:50, pharmacy confirmed the order was awaiting pharmacist review. No reason for the delay or estimated verification time was documented.
- At 19:00, pharmacy verification remained pending, with no administration recorded in the MAR.
- Follow-up: Check pharmacy verification status and MAR. Dose and administration schedule are not provided in this handoff source.

Sputum specimen — Not collected

- Ordered at 18:15.
- Collection attempted at 18:55. Patient cooperated but could not produce a specimen.
- Follow-up: Collection remains outstanding.

Family update — Pending

- At 18:58, daughter requested an update.
- Outgoing nurse had not completed the update by 19:00.
- Follow-up: Address daughter's request.

19:00 handoff status: Ceftriaxone remains unverified and unadministered per available documentation. Sputum specimen collection and family update remain outstanding.