

OBJECTIVE STRUCTURED CLINICAL EVALUATION (OSCE) OF AN ARTIFICIAL INTELLIGENCE (AI) CLINICAL DECISION SUPPORT SYSTEM (CDSS) TOOL



Human Research Program
Systems Integration Office

June 2025

Ariana Nelson¹, Ali Al², Arian Anderson³, Jonathan Steller¹, Truong Le²,
Lynn Boley², Kurt L. Berens², Jay Lemery⁴, David Hilmers^{5,6}

¹ University of Texas Medical Branch, Galveston, TX.

² KBR, Inc., Houston, TX.

³ University of Colorado School of Medicine, Boulder CO.

⁴ IPA, NASA Johnson Space Center, Houston, TX.

⁵ Baylor College of Medicine, Houston, TX

⁶ TRISH

“Expanding the Boundaries of Space Medicine and Technology”



Disclosure Information

2025 AsMA-UHMS Annual Scientific Meeting

I will not discuss off-label use and/or investigational use in my presentation

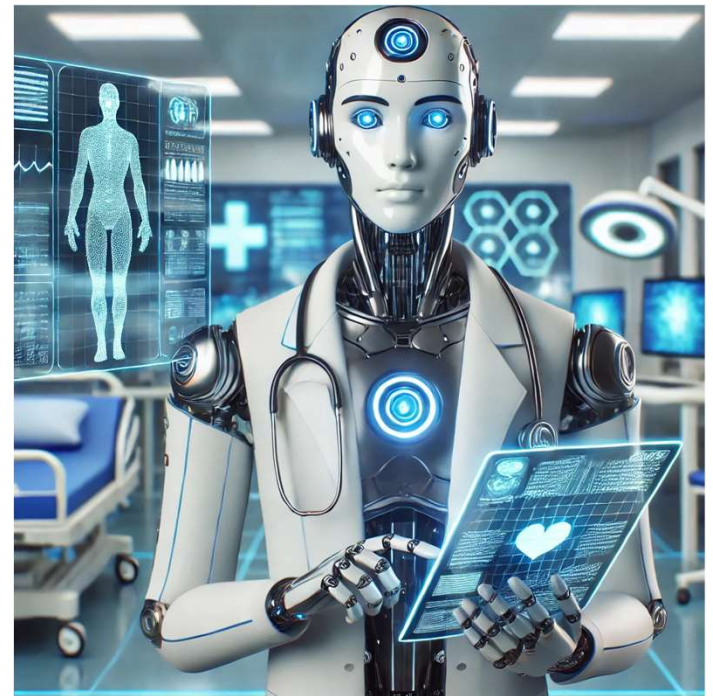
I was employed by NASA/KBR/UTMB at the time this work was done but I am not a representative of these entities, and I am not presenting on their behalf.

I am employed by Vast Space. I have no other relevant financial relationships to disclose.

Agenda



- **Background**
- **Objective**
- **Approach**
- **Challenges and Limitations**
- **Lessons Learned**

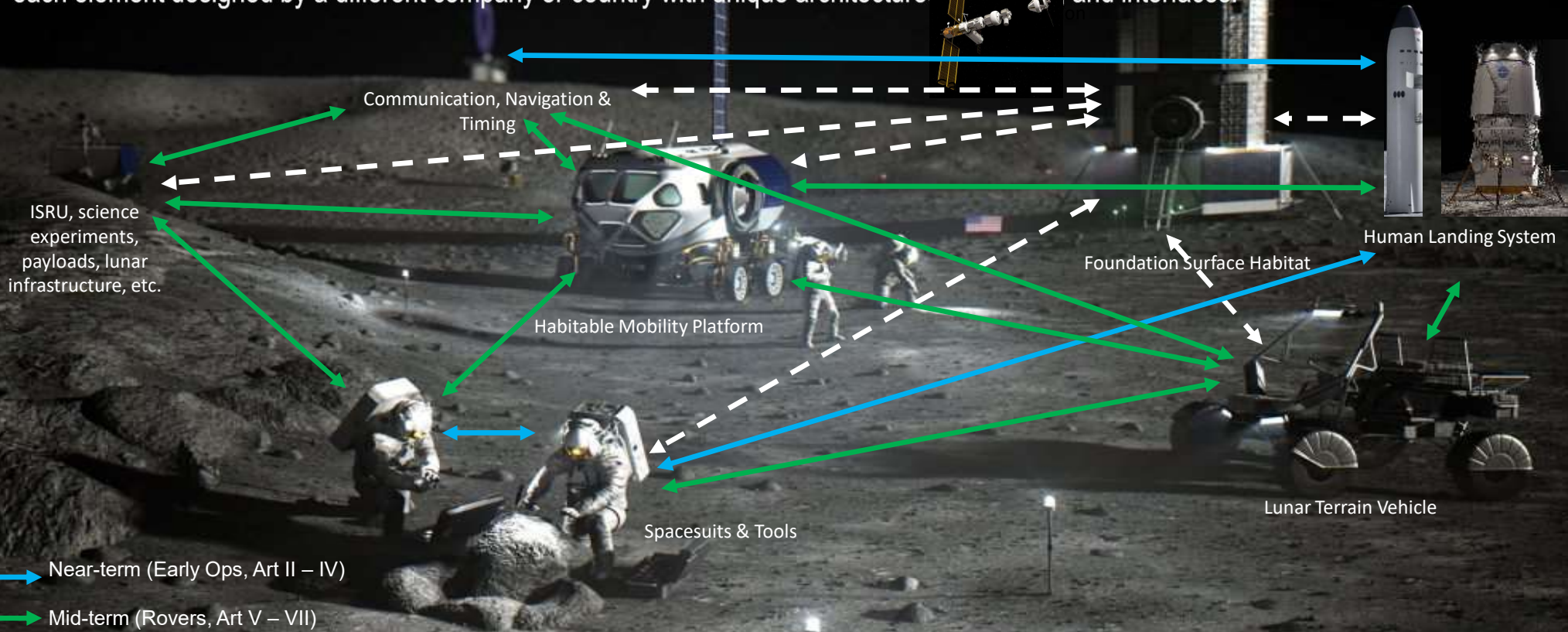


• Image courtesy of ChatGPT

Lunar Operations



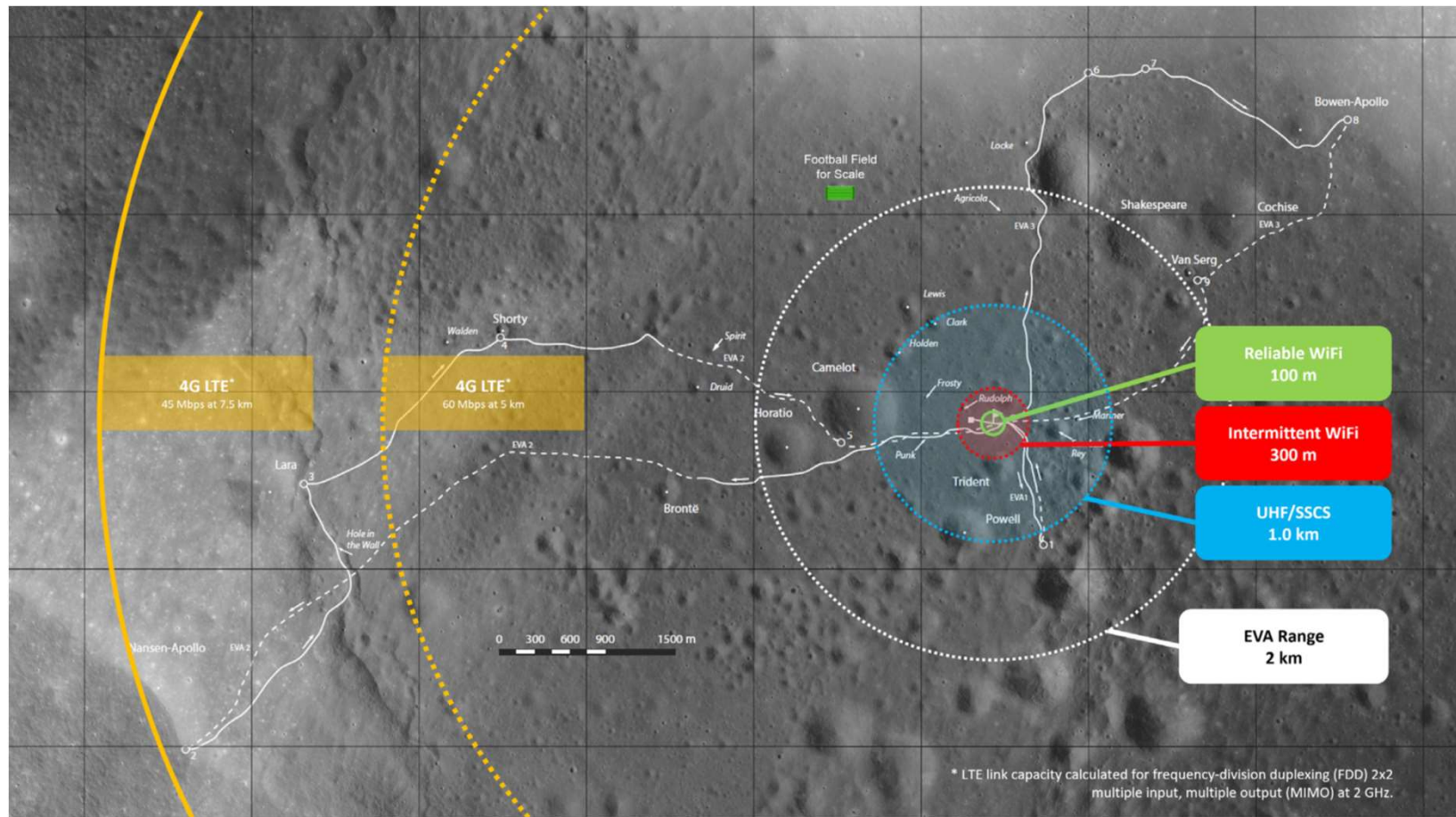
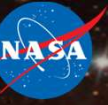
Artemis crews and autonomous systems will work in tandem to operate multiple Lunar surface elements on a single mission with each element designed by a different company or country with unique architecture and interfaces.



- ↔ Near-term (Early Ops, Art II – IV)
- ↔ Mid-term (Rovers, Art V – VII)
- ↔ Long-term (Base Camp, Hab, Art VIII+)

[https://ntrs.nasa.gov/citations/20240003016#:~:text=Lunar%20Command%20and%20Control%20Interoperability%20\(LuCCI\)%20Project%20Overview%20The%20Lunar,commercial%20partners%2C%20and%20international%20partners.](https://ntrs.nasa.gov/citations/20240003016#:~:text=Lunar%20Command%20and%20Control%20Interoperability%20(LuCCI)%20Project%20Overview%20The%20Lunar,commercial%20partners%2C%20and%20international%20partners.)

Lunar Command and Control Interoperability



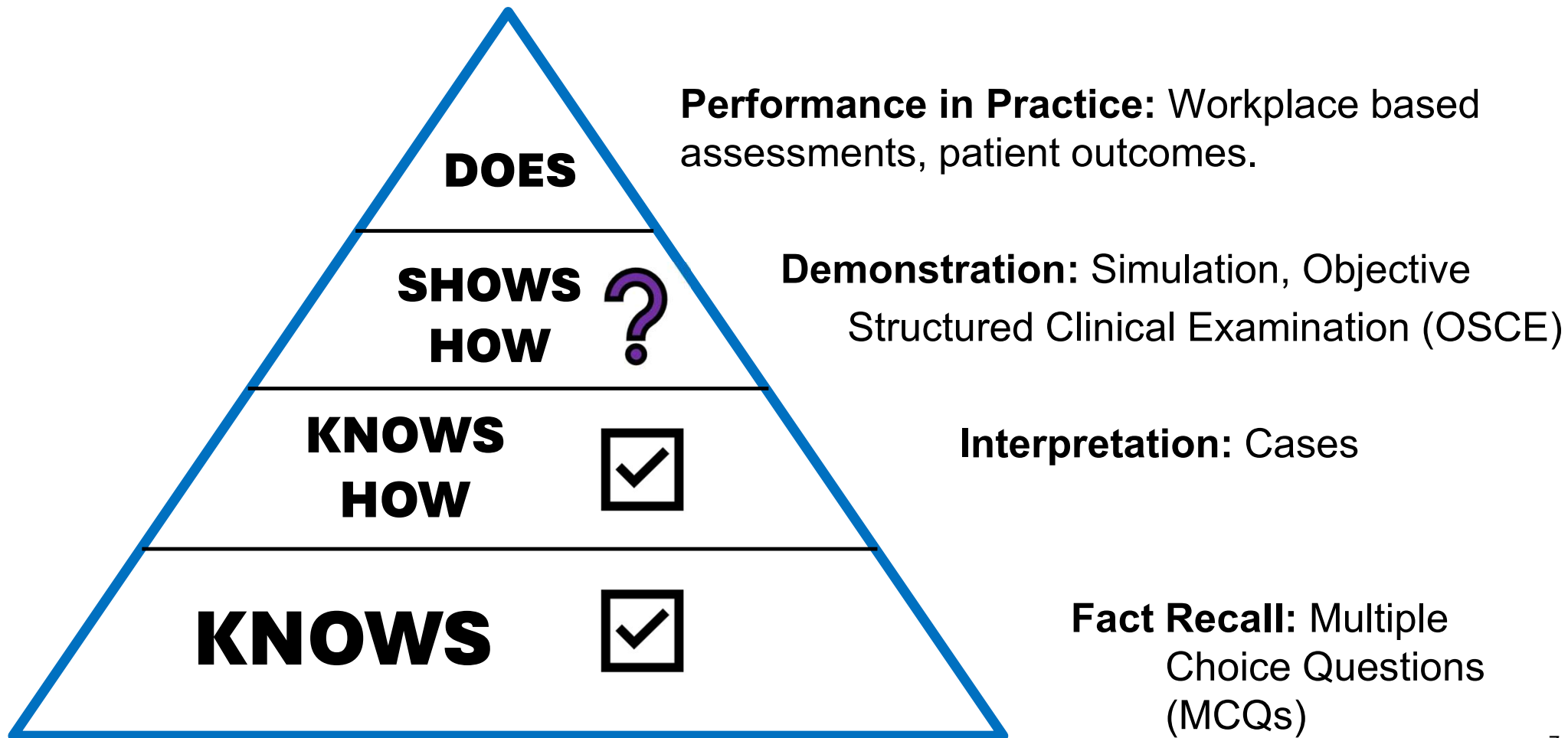
- NASA's Lunar Communications and Navigation Architecture. 2023 Moon to Mars Architecture Concept Review. <https://www.nasa.gov/wp-content/uploads/2024/01/lunar-communications-and-navigationarchitecture.pdf?emrc=f1a91a>

Objective



- Explore the use of AI as a novel Clinical Decision Support System (CDSS) tool.
- Develop a natural language model-based AI tool to streamline clinical evaluation of astronaut crewmembers in support of Earth Independent Medical Operations (EIMO).
- Evaluate this “Doc in a Box” via an Objective Structured Clinical Evaluation (OSCE).

Miller's Pyramid of Mastery



The Cases



- **3 cases based on “chief complaint”**
 - From EIMO ConOps
- **Llama 3.1 model**
 - Optimized for multilingual dialogue
 - 8B, **70B** and 405B

Example: "Has the patient experienced any previous episodes of similar symptoms, such as severe back pain or vomiting?"

Next: symptoms_collector

Next: __end__

Next: symptoms_collector

Supervisor: Re-ask for clarification on the location of the pain, specifically if it's limited to the left flank or if it radiates to other areas, and also ask about the duration of the pain, using the HPI sequence.

Next: __end__

Next: symptoms_collector

Supervisor: Re-ask about the duration of the pain, as the HPI sequence was not fully explored, and also verify if the pain is constant or intermittent.

Next: __end__

Next: symptoms_collector

Supervisor: Re-ask about the patient's vital signs, specifically temperature, and also explore the relationship between the pain and vomiting, using the HPI sequence. Additionally, verify if the patient has experienced any similar episodes in the past.

Next: __end__

Next: symptoms_collector

Supervisor: Re-ask about the patient's medications, social history, family history, and trauma history, as these areas have not been explored yet. Also, verify if the patient has experienced any recent changes in appetite, urine output, or bowel movements.

Next: __end__

Next: symptoms_collector

Supervisor: Re-ask about the patient's social history, specifically inquiring about food and exercise habits, as this area has not been fully explored yet.

Next: __end__

Next: symptoms_collector

Supervisor: Re-ask about the patient's past medical history, specifically inquiring about any previous episodes of similar symptoms or conditions that may be relevant to their current symptoms, and also verify if the patient has experienced any changes in urine output or bowel movements recently.

Supervisor: Re-ask one question at a time, and also verify if the patient has experienced any changes in urine output or bowel movements recently, as this was not fully explored yet.

Example: "Has the patient experienced any previous episodes of similar symptoms, such as severe back pain or vomiting?"

Next: __end__

Next: symptoms_collector

Next: __end__

Next: symptoms_collector

Supervisor: Re-ask about the relationship between the pain and vomiting, and also verify if the patient has experienced any changes in the severity or frequency of the pain, using the HPI sequence. Additionally, remind the user to answer the question about the relationship between the pain and vomiting, and then explore the urinary urgency symptom further.

Example: "Can you please tell me more about the relationship between the pain and vomiting, or do they occur independently of each other? Additionally, has the patient noticed any changes in the severity or frequency of the pain over time?"

Next: __end__

Next: symptoms_collector

Next: __end__

Next: symptoms_collector

Supervisor: Re-ask about the relationship between the pain and vomiting, and also verify if the patient has experienced any changes in the severity or frequency of the pain over time?"

Agentic Approach



- **History**
- **Physical Exam**
- **Differential Diagnosis**
 - **Diagnostics**
- **Plan of Care**



- Image courtesy of ChatGPT

Agentic AI = Chatbot



Next Steps:

I will consult with the physician to determine the next course of action.

Please let me know if this summary is accurate and if you have any questions.

 this is good. we can switch now.. thanks

 ||Switch||

 hi

 I'm here to assist you with the physical examination process for Alex.

Let's begin with the **Respiratory Examination**.

Purpose: To assess Alex's respiratory system and identify any abnormalities.

Step-by-Step Instructions:

1. **Positioning:** Have Alex sit comfortably on an exam table or lie down, depending on the procedure.
2. **Observation:** Observe Alex's breathing pattern, looking for any signs of distress or abnormality.
 - o Increased respiratory rate
 - o Use of accessory muscles (e.g., neck and shoulder muscles)
 - o Nasal flaring

 heart sounds are normal and pulses are normal.

 Thank you for those findings.

Based on the information gathered so far, we have:

- **Respiratory Examination:** Normal respiratory effort, decreased breath sounds on the right side, no chest tenderness.
- **Cardiovascular Examination:** Normal heart sounds, regular rhythm, normal pulses.

||Switch||

 hi

 Based on Alex's symptoms and examination findings, a chest X-ray would be the most appropriate initial test. This will help us visualize his lungs for pneumonia or collapsed lung.

Here are the instructions for performing a chest X-ray:

1. **Preparation:** Alex will need to remove any jewelry or metal objects that could interfere with the X-ray image. He should also wear a hospital gown.
2. **Positioning:** Alex will be asked to stand or lie down on an X-ray table. The technician will guide him into the correct position, ensuring his arms are raised and his back is flat against the table.

Evaluators



Who scored the OSCEs?

01

Anesthesiologist
Pain Medicine

02

Obstetric
Gynecologist

03

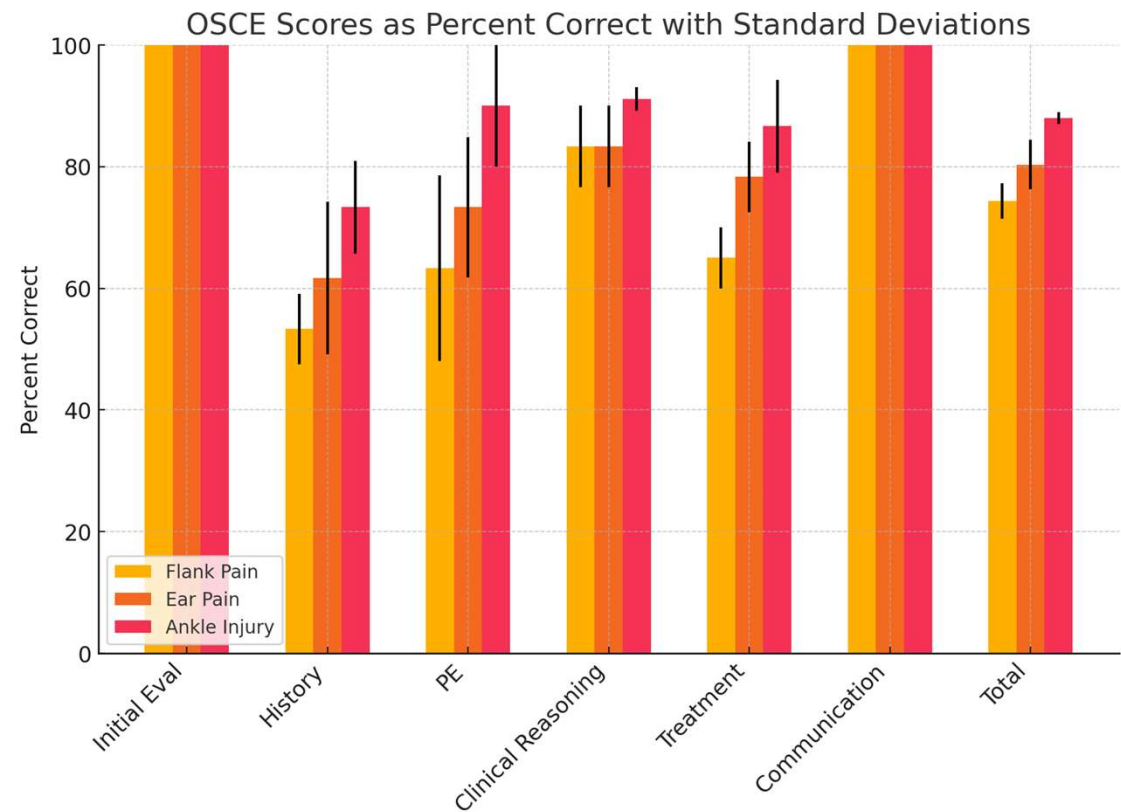
Internal
Medicine

Astronaut

Scores



- Flank Pain 74%
- Ear Pain 80%
- Ankle Injury 88%



Evaluator Correlation



- **Pearson correlation matrix shows high correlations among raters**
- **0.74 to 0.99**
- **Strong agreement in scoring.**

	Evaluator 1	Evaluator 2	Evaluator 3
Evaluator 1	1.00	0.74	0.76
Evaluator 2	0.74	1.00	0.9996
Evaluator 3	0.76	0.9996	1.00

Lessons Learned



- **Lobotomizing consequences**
- **Prompt is EVERYthing**
- **EIMO must evolve in parallel with Moon to Mars**
- **Bias**



• Image courtesy of ChatGPT



Questions?