

Pre-Survey

- **[0-120]** Age
- **[Female/Male/Other]** Gender
- **[M1-M4]** Which year of medical school?
- **[Yes/No]** USMLE Step 1 passed?
- **[0-50]** How many previous studies/tasks have they done with Centaur Labs?
- **[Short Answer]** Which medical school do they attend?
- **[1: Not familiar, 2: Some familiarity, 3: High familiarity]** Are they familiar with any of the following clinical challenges? JAMA Clinical Challenge, NEJM Image Challenge, NEJM Resident 360
 - **[1. not at all; 2. occasionally; 3. every one that is released]** If so, how regularly do they follow these challenges?
- **[1: Not familiar, 2: Some familiarity, 3: High familiarity]** Are they familiar with the MedBullets website?
 - **[1. not at all; 2. occasionally; 3. every one that is released]** If so, how regularly do they follow the clinical challenges released there?
- **[1-5]** Which of the following best describes your experience with large language models (LLMs) in clinical or healthcare-related contexts? (Select the option that most closely matches your experience.)
 1. I have a technical understanding of how LLMs work (.g., architecture, training, fine-tuning).
 2. I have used LLMs in clinical workflows or research (e.g., for documentation, decision support, or data analysis).
 3. I have interacted with LLMs for general or clinical questions (e.g., using ChatGPT to look up clinical facts or guidelines).
 4. I am aware of LLMs but have not used them in any clinical or technical setting.
 5. I am not familiar with LLMs.
- **[1-5]** Which of the following best reflects your view on the readiness of LLMs for clinical deployment? (Select the option that most closely matches your opinion.)
 1. Not ready at all – LLMs are unreliable and unsafe for any clinical use.
 2. Ready for limited, low-risk tasks – e.g., documentation assistance, summarization.
 3. Ready for clinician-in-the-loop tasks – e.g., decision support with human oversight.
 4. Ready for autonomous use in select areas – e.g., triage or patient education.
 5. Fully ready for broad clinical deployment with minimal oversight.