



Protocol Title: Artificial Intelligent Clinical Decision Support System Simulation Center Study: Trust and Usefulness of Machine Learning Risk Stratification Tool for Acute Gastrointestinal Bleeding using the Technology Acceptance Model
Principal Investigator: Dennis Shung, MD MHS PhD
Version Date: 1 January, 2023
NCT05816473

Instructions:

Certain research activities may be exempt from review, if confirmed by the IRB Chair or his/her designee and confirmed in writing to the Investigator. Research may be exempt from review when the only involvement of human subjects in the research falls into one or more of the categories noted below. The regulations allow for two additional exemption categories that are not currently implemented at Yale.

Note:

- **The IRB does not exempt studies that involve the Introductory Psychology Subject Pool.**
- **Exemption categories apply to research involving pregnant women.**
- **Exempt categories DO NOT apply to research with prisoners, except for research aimed at involving a broader subject population that only incidentally includes prisoners.**
- **Exempt categories generally apply to research with minors, except when specifically stated otherwise.**

Choose one of the following exemption categories for consideration and provide the information as requested under the corresponding category. **Delete all other categories that do not apply.** Upload the survey(s), instrument, or interview questionnaire/focus group guides to the Supporting Documents section of IRES IRB.

FOR HIPAA ONLY - The investigator assures that the protected health information for which a Waiver of Authorization has been requested will not be reused or disclosed to any person or entity other than those listed in this application, except as required by law, for authorized oversight of this research study, or as specifically approved for use in another study by an IRB.

Researchers are reminded that unauthorized disclosures of PHI to individuals outside of the Yale HIPAA-Covered entity must be accounted for in the "accounting for disclosures log", by subject name, purpose, date, recipients, and a description of information provided. Logs are to be forwarded to the Deputy HIPAA Privacy Officer.

(Category 2) 45 CFR 46.104(d)(2) Research not regulated by the FDA that only includes interactions involving educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior (including visual or auditory recording) if at least one of the following criteria is met: (Please indicate which criteria applies)

☒ (i) The information obtained is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained, directly or through identifiers linked to the subjects;

☒ (ii) Any disclosure of the human subjects' responses outside the research would not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, educational advancement, or reputation; or

☐ (iii) The information obtained is recorded by the investigator in such a manner that the identity of the human subjects can readily be ascertained, directly or through identifiers linked to the subjects, and an IRB conducts a limited IRB review to make the determination.

This exemption category applies to research with minors ONLY if the research involves educational tests or the observation of public behavior when the investigator(s) do not participate in the activities being observed.

1) Describe the purpose of the study:

The advent of new technologies, the ever-increasing use of electronic medical records, and the collection of medical data have allowed the creation of machine learning models that have the potential to revolutionize medical care. However, these novel models would only be beneficial if they are actually properly used. As such, critical to these developments is whether healthcare providers would be willing to use these tools and what can help increase the uptake of them.

The purpose of this study is to measure the effect on of a large language model interface on the usability, attitudes, and provider trust when using a machine learning algorithm-based clinical decision support system in the setting of bleeding from the upper gastrointestinal tract (upper GIB). Specifically, the investigators are looking to assess the optimal implementation of such machine learning algorithms in simulation scenarios to best engender trust and improve usability. Participants will be randomized to either machine learning algorithm alone or algorithm with a large language model interface and exposed to simulation cases of upper GIB. Simulated medical cases will be conducted at the Yale Center for Medical Simulation (YCMS). The survey does not deal in any way with health information or otherwise involve the application other than using it as a platform to assess the user's attitudes and "theory of mind" perspective about using a clinical scoring tool that is based on machine learning. Attached is a copy of the survey, which is derived from the Technology Acceptance Model (TAM).

2) Describe the location of the study.

Yale New Haven Hospital and Yale School of Medicine

a. Does this study include an international location? ☐ Yes ☒ No

If yes, specify location

3) Describe the procedures that will be used to recruit subjects.

Providers who are involved in clinical training in emergency medicine and internal medicine at Yale New Haven Hospital and the Yale School of Medicine will be recruited by the researchers.

Recruitment efforts will be through word-of-mouth, email, and communication with EM and IM leadership.

4) Describe how subjects will provide consent (and/or research authorization) to participate in the study.

Informed written and verbal consent will be gathered prior to conducting the survey. Participants will be provided with relevant information about the survey, the simulation environment, the GI bleed simulation scenarios, and educational modules about the machine learning tool & implementation of AI-CDSS in the clinical setting.

5) Describe the procedures that will be used to conduct the research. (NOTE - If using enumerators, include the name of the agency, training provided to individuals at the agency, and the specific role in this research. If using a survey platform, name the platform.)

We will recruit and provide informed consent to participants. Providers will then take the modified UTAUT survey, which will evaluate the usefulness, ease of use, attitudes towards use, and trust in AI-CDSS for evaluating risk of GIB in clinical scenarios. The initial experiment will deploy the machine learning algorithm within simulation scenarios in which a patient with acute gastrointestinal bleeding (at low, moderate, and high risk for poor outcome) is evaluated by trainees. To detect a medium effect size (Cohen's D of 0.5) with 80% statistical power in the planned study design, we would need an estimated 102 participants.

Prior to the simulation, a baseline educational module about artificial intelligence, machine learning, and clinical decision support will be provided to all participants. We will establish psychological safety by detailing what is available in the room, the opportunity to call a consultant, and availability of laboratory and radiographic studies. Each clinical scenario will run for approximately 10 minutes based on real patient cases where vital signs change over time and laboratory values are made available at specific points in the assessment. The study will evaluate the effect of a large language model-based interaction with the machine learning algorithm with interpretability dashboard compared to the machine learning algorithm with interpretability dashboard alone.

Each participant receives all three scenarios randomized order of risk. For the large language model interaction arm, participants will be provided with the computer workstation with a large language model chatbot interface of the algorithm and interpretability dashboard. For the machine learning dashboard arm, participants will be provided with the computer workstation with the algorithm and interpretability dashboard.

The study will use a common set of dependent variables to assess baseline and post-intervention attitudes towards machine learning algorithms in clinical care using an adapted Technology Acceptance Model survey assessing perceived usefulness of the system, perceived ease of use, attitudes towards using it, behavioral intentions, and trust, measured with a 5-point or 7-point Likert scale. The results will be analyzed using linear mixed-effects regression models to account for the 3x2x2 design of the study. This approach yields robust estimates and test statistics for main effects and interaction effects for each survey construct. Sensitivity analysis to evaluate errors in decision making will be performed on scenarios where there was expectation violation (i.e. the prediction of the clinician differs from the algorithmic prediction). All participants will be told prior to the simulation that the scenario will involve a patient with a GIB.

During each scenario, participants will be asked open-ended questions by simulation staff at various decision-points in the scenario regarding how they would like to proceed. At the end of each scenario, participants will be required to answer a multiple-choice question about whether to discharge for outpatient follow up, admit & observe the patient, or admit for urgent endoscopy. Each

scenario will then proceed with either patients remaining clinically stable or deteriorating and requiring further care. Debriefing about the simulation experience between researchers and participants will follow and fundamentals of GIB management will then be covered with all participants.

- 6) If subjects' identity can be readily ascertained directly or through identifiers linked to them, could any disclosure of their responses outside the research reasonably place them at risk of criminal or civil liability or be damaging to their financial standing, employability, educational advancement, or reputation? ☐ Yes ☐ No ☒ NA**

If YES –provide the list of identifiers and describe how data will be secured to protect the privacy of subjects and maintain the confidentiality of the data, and, if applicable, the coding system that will be used.:

- 7) If you are from Yale School of Medicine, School of Nursing, or another HIPAA covered entity (such as Psychology clinics) and wish to collect PHI without obtaining written HIPAA authorization, – a HIPAA waiver must be obtained. Describe why it would be impracticable to obtain the subject's authorization for use/disclosure of this data: N/A.**