Study Protocol

Developing a deliberate practice intervention to recalibrate physician heuristics in trauma triage

NIA R21AG072072
Clinicaltrials.gov identifier #NCT05168579
Approved by University of Pittsburgh Office of Research Protections: January 6, 2021

Table of Contents

Study Personnel

Table 1. Names of team members with role and area of expertise

Name	**Role (expertise)**
Deepika Mohan, MD	Principal Investigator (trauma surgery, physician behavior)
Douglas White, MD	Co-investigator (deliberate practice)
Jonathan Elmer, MD	Co-investigator (emergency medicine)
Raquel Forsythe, MD	Co-investigator (trauma surgery)
Robert Arnold, MD	Co-investigator (adult education)
Baruch Fischhoff, PhD	Co-investigator (behavioral science)
Kim Rak, PhD	Staff researcher (qualitative research methods)
Jaqueline Barnes, PhD	Staff researcher (qualitative research methods)
Mary Beth Ryabik, RN	Project manager

Prior Experience, Gaps in Current Knowledge, Rationale, and Significance

Diagnostic error is a particularly important problem for the 500,000 elderly patients who present to non-trauma centers every year after trauma. Emergency medicine physicians must rapidly categorize patients as having minor or severe injuries based on limited information and decide whether or not to transfer the patients to a trauma center (triage). Observations from our group demonstrate that physicians rely on heuristics (intuitive judgments) to identify severely injured patients, which results in the systematic under-triage of the elderly. Under-triaged patients experience a 10-25% increase in mortality, loss of independence, and increased pain.

The absence of an effective means of recalibrating heuristics is a critical barrier to the improvement in outcomes for elderly patients with severe injuries. Behavioral scientists agree that people develop well-calibrated heuristics through an experience-feedback loop that hones pattern recognition and ensures the recognition of relevant contextual cues. Since replicating that loop has proven challenging outside of formal training programs (e.g. residency), most initiatives to reduce diagnostic error have focused on eliminating the use of heuristics. Unfortunately, they have had limited efficacy, probably because heuristics are essential to human cognition. The National Academy of Medicine therefore recently identified the development of a method of addressing the challenge posed by heuristics to be a major priority.

In prior work, we exploited insights from the behavioral science literature to develop novel interventions to recalibrate physician heuristics. Specifically, we identified surrogates for the experience-feedback loop in other domains (e.g. threat detection), and applied them to trauma triage (NIH Stage I behavioral intervention development). We delivered these interventions as video games to increase engagement and to facilitate dissemination. In pilot trials, physicians who played the games made fewer diagnostic errors on a validated virtual simulation compared to those who completed a gold-standard, text-based educational program, an effect that persisted to six-month follow-up (Stage II development). These positive results occurred despite physicians reporting only moderate engagement with the interventions. Our overarching hypothesis is that by increasing engagement we can amplify the potency of the interventions. We therefore propose to refine the games before testing their efficacy in the real world (Stage III development).

Study Aims

Deliberate practice – goal-oriented training in the presence of a coach who can provide personalized, immediate feedback – has successfully improved performance across multiple domains, including sports, music, and combat. When used in conjunction with simulation to improve surgical skill, it has a large effect on educational outcomes. It has characteristics that make its application in this context potentially powerful (e.g., personalized feedback/relationship with coach increase engagement) but also potentially challenging (e.g. the diagnostic process does not lend itself easily to assessment). The objective of this R21 application is to test the feasibility of using deliberate practice to amplify the effect of our video game interventions. We will recruit a national sample of board-certified emergency physicians (n=30) to serve as trainees, with members of the team (n=3) serving as coaches. Trainee-coach dyads will meet for 30 minutes/week for 3 weeks, by videoconferencing, to play one of the existing video games and to use it to practice pattern recognition. We aim:

1. To assess the fidelity of intervention delivery. Approach: we will standardize coaching skill during an 'on-boarding session,' measure skill drift over the course of training sessions, and measure protocol adherence (primary outcome). Hypothesis: >90% of dyads will complete three training sessions.
2. To assess the potential effect size of the intervention. Approach: we will compare performance of trainees (n=30) with a control group of physicians (n=30) on a validated virtual simulation. Hypothesis: Trainees will make ≥25% fewer diagnostic errors than control physicians (large effect size).
3. To assess the acceptability of the intervention. Approach: we will conduct semi-structured debriefing interviews with trainees, assessing elements of the intervention that promote engagement.

This proposal will inform a future Stage III trial to compare the effect of different interventions on diagnostic error in trauma triage. If successful, this program of research will have an impact on patients by reducing the burden imposed by injury and by addressing the refractory problem of diagnostic error. It is novel conceptually in its effort to make heuristics a source of power, methodologically in its use of deliberate practice to improve diagnosis, and translationally in its use of video game technology. It is feasible because our multi-disciplinary team has clinical and behavioral science expertise, experience developing deliberate practice interventions, and a track record of successfully building video games that can transform physician behavior. It responds to two national research priorities: 1) improving the diagnostic process; 2) maintaining health and independent living among the aging.

Study Design

Overview

We will modify the interface of one of our two games to ensure its' suitability as a training task. The three team members (Mohan, Elmer, Forsythe) who will serve as coaches will undergo training with our two deliberate practice experts (White, Arnold) to standardize their coaching skills. Next, we will recruit emergency physicians working at non-trauma centers around the country (N=60): 30 trainees and 30 passive controls. We will pair trainees with a coach, and will ask coach-trainee dyads to schedule 30-minute training sessions at their convenience, once per week for the three week period. At the completion of the three weeks, we will ask trainees to complete a semi-structured, debriefing interview and a virtual simulation to assess triage performance. We will ask passive controls to complete the same simulation within 3 weeks of enrollment.

Participants

The three team members (Mohan, Elmer, Forsythe) will serve as coaches, and will undergo training with our two deliberate practice experts (White, Arnold) to standardize their coaching skills.

We will recruit emergency physicians working at non-trauma centers around the country (N=60): 30 trainees and 30 passive controls.We plan to use participants from prior behavioral trials to facilitate snowball recruitment. We will contact prior participants and will ask them to refer us to 1-2 colleagues who might be interested in participating in a research trial to assess physician decision making in trauma triage. Once we get the referrals, we will contact physicians, provide details of the trial, and then once receiving consent, will randomize them to either the intervention or the control arm of the pilot study. To supplement our numbers, we also plan to reach out to ERMI, a local professional staffing organization that provides emergency medicine physicians to non-trauma centers in the Pittsburgh region. Physicians will receive $100/hour of time spent participating in the trial, a wage-based honorarium.

Inclusion criteria

Research subjects will be board-certified physicians who work in the ED of non-trauma centers in the US, and who manage primarily adult patients.

Exclusion criteria

We will exclude physicians who work exclusively at Level I/II trauma centers, who treat only children, or who work outside of the US.

Description of *Shift with Friends*

Shift is a puzzle video game developed to recalibrate physician heuristics in trauma triage. Players engage in analogical encoding – structured case comparison – to derive their own decision principles for triage. Specifically, players review cases and then identify contextual cues associated with the presentation of severely injured patients. Next, they synthesize those cues into simple, unifying triage principles. Theoretically, the process of derivation makes the principles memorable, and therefore more likely to become part of the physicians' heuristics. The game has approximately 2 hours of content, covers 10 triage decision principles, and allows repeated play of selected sections. It should lend itself well to deliberate practice because coaches can observe the contextual cues that physicians highlight during the process of case comparison and can provide personalized feedback on how they should integrate those

cues into the pattern that they use when diagnosing trauma patients (i.e., recalibrate their heuristics).

Description of virtual simulation (used to assess efficacy)

We previously developed and validated a virtual simulation to study physician decision making in trauma triage. Physicians have to evaluate and to manage ten cases over 42 minutes, simulating a busy eight-hour ED shift. Each case includes a 2-D rendering of the patient, a chief complaint, vital signs which updates every 30 seconds, a history, and a written description of the physical exam. Physicians manage patients by selecting from a pre-specified list of 250 medications, studies, and procedures. The cases end when physicians either make a disposition decision (admit, discharge, transfer) or the patient dies. To measure diagnostic performance, we will collect information on decision making: diagnostic, therapeutic, and disposition decisions.

Research activities

At the time of enrollment, we will ask both trainees and passive control physicians to complete a questionnaire that surveys their personal characteristics (time required: 10 minutes). Trainees will be assigned a coach, and will be asked to schedule weekly 30-minute meetings for three weeks, at which time they will play a video game and will receive feedback on how to use best-practice triage decision principles (time required: 90 minutes). Trainees will also be asked to complete a short assessment of their coaches' performance after each session (time required: 5 min/session). After completion of the intervention, trainees will be scheduled to participate in a semi-structured debriefing interview (time required: 20 minutes). Both trainees and passive-control physicians will be asked to complete a virtual simulation that assesses their triage decision making (time required 42 minutes). Trainees will therefore spend 3 hours completing study tasks; passive control physicians will spend 1 hour. Trainees will receive an iPad with the video game pre-loaded at the time of enrollment. They will keep the iPad as their honorarium for participating in the study. Passive control physicians will receive a $100 gift card as a fixed, wage-based honorarium for participating in the study. We chose to provide the honorarium to minimize attrition and to ensure completion of study tasks. Based on our prior experience, the combination of framing participation as altruistic, e-mail reminders about completion of study tasks, and provision of a wage-based, fixed honorarium increases completion rates to as high as >80%. Given the costs of running a trial, and bias introduced by differential completion of study tasks, we believe that using an honorarium is warranted.

Sources of data

Physician self-report

1. Physician characteristics. We will ask all participants to respond to a questionnaire that asks about personal characteristics: demographics (e.g. age, sex, race), training (e.g. name of medical school, name of residency training program and fellowship [if applicable], year of board certification), professional experience (e.g. years in practice, number of shifts worked/month, number of patients treated/shift), and practice environment (e.g. trauma center designation of hospital; resource availability).
2. Trainee assessment of coaching performance. After each deliberate practice session, we will ask trainees to rate their coaches' skill along four domains (as in the Wisconsin Surgical Coaching Rubric), using a Likert scale.

Trainee assessment of the acceptability of the intervention. We will ask trainees to complete the User Engagement Scale (a 12-item instrument that assesses aesthetic appeal, attentional focus, perceived usability, and needs satisfaction). In addition, we will ask them to participate in a 20-minute debriefing interview in which we probe: 1) their engagement with the study tasks; 2) how the intervention affected their well-being (e.g., did they find the experience onerous or did it reinforce their intrinsic motivation); 3) barriers to implementation. Interviews will be audio-taped, transcribed, and reviewed.

Deliberate practice sessions

We will videotape all the coaching sessions between the trainee-coach dyads. We will review these video tapes to assess: 1) protocol adherence (i.e. we will calculate the proportion of dyads that complete all three training sessions); 2) coaching performance. Two independent raters will review the video tapes using the Wisconsin Surgical Coaching Rubric to evaluate coaching skill drift. The Rubric scores performance along four domains (shares responsibility, uses questions/prompts to guide trainee in self-reflection, provides constructive feedback; guides goal setting), with individual components summed together at the end for an overall assessment of skill.

Shift with Friends

The game uploads information on usage statistics (e.g. number of clicks, proportion of app used) to a database every time the iPad connects to a wireless network. We will use this added information to determine how trainees use the games, and to calculate the number who use the games outside the training sessions.

Virtual Simulation

We will ask physicians to complete a virtual simulation, and will collect information on their decisions: diagnostic (CT scan, x-rays, labs), therapeutic (medications, procedures, consults), and disposition (admit, discharge). The simulation will be available on a web-based browser, and will store responses on a secure server hosted on the University of Pittsburgh network.

Primary and secondary study endpoints

Aim 1. To assess the fidelity of intervention delivery.

Ha1: ≥90% of coach-trainee dyads will complete all three coaching sessions.

We will summarize the proportion of coach-trainee dyads that complete all the training sessions. To deem the intervention delivery strategy feasible, we will need to see ≥90% protocol adherence. We will also summarize the characteristics of game use by each trainee, including whether any of them use the game outside the deliberate practice sessions; assessments of coaching performance by independent raters and by trainees; assessments made by coaches of intervention receipt. Finally, we will perform a series of exploratory analyses using non-parametric tests (Fisher exact test, Wilcoxon-Mann Whitney, Kruskal Wallis, repeated measures regression models) to characterize the association between trainee characteristics and: a) protocol adherence; b) coaching skill (first session and overall); c) intervention receipt (first session and overall).

Aim 2. To assess the potential effect size of the intervention.

Ha2: We will detect a large difference in the diagnostic errors made by trainees and control physicians.

We will use the virtual simulation to assess physician diagnostic performance, and will use an intention-to-treat approach (i.e. all trainees regardless of whether they completed all the sessions will be asked to use the simulation). We will measure the proportion of severely injured patients appropriately triaged to a trauma center as per the American College of Surgeons guidelines. We will estimate both the group mean of diagnostic errors made by trainees and passive control physicians, and variability in the group-specific outcome. We will use a Students' t-test to compare the means of the two groups. To deem the intervention successful (and worth pursuing), we will need to detect a large difference (~25%) between the two groups. In exploratory analyses, we will also assess the effect of the intervention on heuristics by studying patterns of errors (as we have done previously). We will not test any mediators of efficacy because of the sample size.

Aim 3. To assess the acceptability of the intervention.

We will summarize responses to the User Engagement Scale-Short Form. We will also use best practice methods to report qualitative results from the semi-structured interviews, focusing on participant perceptions of the effect of the intervention on engagement.

Power analysis

We will assess the feasibility of using deliberate practice to recalibrate physician heuristics including barriers to implementation, fidelity, acceptability, the magnitude and variance of the treatment effect. We anticipate that ≥80% of physicians in both the trainee and control groups will complete the virtual simulation. Based on Cohen's power estimates for behavioral trials, with at least 20 physicians in each group, we can detect a large difference (0.80 standard deviation) in their performance, using a t-test, with power of 80% and alpha=0.1.73 Given the distribution of responses in the past, this would manifest as a 25% difference in diagnostic error.

Consent Process

Process

We will email physicians describing the study and ask if they are interested in participating. The email will include a link to a Qualtrics survey that hosts a more detailed description of the study, a consent form, and a demographics questionnaire. Once physicians have consented to participate in the trial, they will be randomized to the intervention or passive control arm.

Steps to minimize coercion

Physicians will be given three days to respond to the email, at which point they will receive a reminder. Physicians who do not respond within 1 week to the recruitment email will not be contacted further. To minimize coercion, the Qualtrics document will emphasize that participation is voluntary, and consent can be withdrawn at any point during the study.

Ongoing consent

The study will involve multiple rounds of coaching (three in total), for those assigned to the intervention arm. The study coordinator will ensure ongoing consent when scheduling followup coaching sessions.

Steps to be taken to ensure the subjects' understanding

We will ensure that all the documentation is piloted for clarity and readability prior to contacting trial participants. When recruiting physicians our email will specify that: 1) we are recruiting physicians for an NIH-funded study; 2) participation is voluntary; 3) they will receive no direct benefit from participation; 4) they will receive an honorarium for their time; 5) participation will require 3 hours of time, distributed over 3 weeks; 6) they may withdraw from the study at any time. We will also provide contact information for the PI, which they can use to gain further details about the trial. During the first coaching session and again during the debriefing interview, study personnel will confirm participant's understanding of the study and the requirements.

Electronic Data Management

Identifiers to be collected during any phase of the research including screening:

- Name
- Email address
- Phone number
- Location data (street address, city, county, zip code)
- Date information (birth date)

We will be coding participant data by removing identifiers and assigning a unique study ID/code to protect their identity. We will assign trial participants an identifier at the time of randomization. We will use the identifier when storing and analyzing all study data. The linkage file will be stored on the Pitt HSRDC virtual desktop.

No sensitive data (https://www.hrpo.pitt.edu/electronic-data-security) is collected where disclosure of identifying information could have adverse consequences for subjects or damage their financial standing, employability, insurability, educational advancement, reputation or place them at risk for criminal or civil liability.

Data Safety and Monitoring

Plan

The research team will meet weekly during data collection and then monthly for the remainder of the study period to evaluate data collection. During team meetings, we will review study protocols to ensure that no harms have occurred and that the benefits are as expected. If either study personnel or the coaches believe that an adverse event has occurred during data collection, they will report that information to the PI who will take responsibility for transmitting the information to the IRB and to the sponsor within 72 hours of notification.

Data Sharing

To support the translation of research results into policy practice, manuscripts describing research results will be drafted and submitted in a timely manner for publication in widely circulated peer-reviewed journals. Dr. Mohan will also present interim and final results at relevant academic and non-academic conferences. Raw data and derived datasets will be made available to external investigators and the public on a case-by-case basis, to be approved by the PI, Dr. Mohan, and in accordance with institutional, HIPAA, state and federal regulations. A data-sharing agreement may be instituted, depending upon the data to be shared. All data that is shared will be de-identified to protect participant privacy and confidentiality. Data and datasets will be retained and available to share for at least three years following completion of the project, in accordance with NIH regulations. The research team will track and report on the use, dissemination and sharing of all data and datasets and assist the PI with administration of data-sharing agreements as necessary.

Risks and Benefits

Foreseeable risk: breach of privacy

Risk prevention: every effort will be made to minimize risk. The investigators and study staff will achieve certification as required by the IRB at the University of Pittsburgh. All physicians will be assigned a study identifier at the time of enrollment, and all data associated with that physician will use that identifier. Information linking the data codes with subject identities will be stored separately from the recorded data. At no time will we reveal subject identities in any description or publication of the research for scientific purposes.

There are no foreseeable risks of identifying an unexpected disease since the objective of this project is to educate physicians.

Subject privacy protection: to protect participants' privacy, we will take the following steps:

1. Coaching sessions and debriefing interviews will be conducted in private rooms
2. We will not collect any sensitive information, since none is required for completion of the study aims.
3. All study subjects will be assigned unique study identifiers that will appear on all data collection instruments, documents, and files used in the statistical analysis and manuscript preparation.
4. Only limited team members will have access to personal information needed for tracking and informed consent. No personal information concerning study participants will be released without their written consent.

Foreseeable Benefits: coaching has the potential to improve physicians' performance and enhance their skill set.

Subject withdrawal from study: if physicians withdraw from the study after partial completion of the study tasks, we will analyze their residual data. In particular, understanding reasons for their withdrawal are extremely important for achieving the study objectives (assessing the feasibility of using deliberate practice interventions).

Project Timeline

Table 2. Overview of study tasks and planned timing

Task	4-7 2021	8-10 2021	11 2021	12 2021	1-4 2022	5/22-3/23
Complete regulatory tasks (e.g., set up contracts, file IRB)	x					
Modify Shift with Friends		x				
Standardize coaching skill of team members			x			
Recruit and randomize a national sample of physicians				x		
Completion of study tasks: deliberate practice sessions, virtual simulation, assessment					x	
Analysis of data					x	x