

Study 4 - Diagnostic Uncertainty and Information Seeking in Virtual Reality Paediatric Scenarios

Introduction

In this study, we aim to extend our previous findings using a virtual reality (VR) experimental paradigm that is more naturalistic to real medical practice. Participants were unable to see the patient in the vignette task, which is important given that the visual state (or distress) of a patient can be informative for a doctor in diagnosing the patient. By simulating this, we aim to investigate the link between information seeking and confidence in a more open-ended clinical situation that has a wider range of possible options for history taking, physical examination, testing and treatment options when compared to our vignette paradigm (which constrained the amount of information available on each case for usability). Given this increased flexibility, we can look at more fine-grained aspects of information seeking, as well as the effect of ongoing treatment of patients on confidence. Our vignette task was static in time, in that the patient does not change over the course of a case (i.e. improving or deteriorating over time). This VR paradigm then allows for doctors to start managing the patient's symptoms and even using reactions to their treatment plan in order to change their understanding of the patient.

Methods

Participants

Materials

We used VR scenarios implemented by Oxford Medical Simulation (OMS), a company that uses VR for medical education and simulation, in their bespoke software. Participants in this study were medical students based in Oxford who were at the time taking part in VR-based teaching sessions as part of their medical degrees. Students performed the scenarios using Oculus Quest 2 VR headsets. Scenarios were based in paediatrics, meaning that the patients in the scenario were children who were attending the hospital with their legal guardian. Each scenario features a visual 3D implementation of a basic ward room in a hospital. Participants are shown a (child) patient, their guardian and a nurse who can help with certain treatment and testing. All of the ‘avatars’ in the scenario can be questioned by the participant using a predefined set of requests/actions (e.g. asking the nurse to check blood pressure, asking the patient/child about if they are in pain). The scenarios have full sound (e.g. being able to hear the patient’s lung auscultation) and the avatars are voiced.

Each participant completed two scenarios over two separate VR sessions. The sessions were held around one month apart. During each session, the participants each performed one scenario in VR and observed their partner during their scenario. Participants also engaged in peer-to-peer feedback discussions as part of their education. The scenarios presented in each sessions are described below (students are split into two groups, shown below as groups A and B, each performing a different pair of scenarios in a fixed order):

* **Session One:** + * *Group A:* patient/child is a 6-year-old-girl presenting with a 1 day history of central abdominal pain and thirst. She was generally unwell for

2 days prior, with reduced appetite and a sore throat. Collateral history reveals Type 1 Diabetes and erratic blood sugars. (**Underlying Condition: Diabetic Ketoacidosis**) + * *Group B*: patient/child is a 5-year-old boy presenting with worsening shortness of breath, wheeze, and signs of respiratory distress, on the background of 2 days of likely viral illness. He has a medical history of asthma and has had similar exacerbations in the past. (**Underlying Condition: Acute Severe Exacerbation of Asthma**)

* **Session Two:** + * *Group A*: patient/child is a 5-year-old boy presenting with shortness of breath and drowsiness (**Underlying Condition: Chest Sepsis/Pneumonia**) + * *Group B*: patient/child is a 5-year-old girl with a 1 day history of sore throat and fever. She starts having a generalised tonic clonic seizure during the scenario. (**Underlying Condition: Febrile seizure on background of tonsillitis**)

Procedure

The aim for students in the scenarios was to diagnose the patient, begin treatment and hand over the case to a senior with appropriate understanding of the patient (handovers were conducted using a standardised framework known as SBAR, meaning that clinicians have to brief the senior on the Situation, Background, Assessment and Recommendation for the patient). Whilst in the scenario, participants can learn about the patient's medical history, check key parameters (such as temperature, pulse, blood pressure, respiratory rate etc), perform physical exams/tests and begin certain treatment actions (such as administering oxygen or prescribing medication). Compared to the previous studies, participants have a much wider array in terms of the information and tests they can request, as well as being able to begin a treatment plan.

After 5 minutes in the scenario (by which point it is expected that participants would have a history of the patient and have started some early assessment of

the patient), participants are asked to pause the scenario (taking off their VR headset) and fill in a brief questionnaire on paper. Multiple VR participants were performing the scenario simultaneously and were paired with another student who would watch their performance. This other student would aid with administering the questionnaire, with the student subsequently switching roles for the other scenario. The VR participant was asked in the questionnaire to answer the follow (this is considered time point 1):

- “Please say all the conditions that you are currently considering or are concerned about for this patient. Include any/all common, rare or contributing conditions you are considering. For each, please rate how likely you think they are on a scale of 1 (low) to 5 (high).”
- “On a scale of 1-10, how confident are you that you understand the patient’s condition?”
- “How severe do you think the patient’s condition is on a scale of 1 to 10?” (Each point of the scale represented a different clinical action/course, with 1 representing “Discharge in <4 hours, no follow up” and 10 representing “Requires arrest/peri arrest team.”)

The questionnaire was kept relatively short to minimise disruption to the scenario. This was due to the extra time that could be expended by asking participants to take off and put on the headset again to readjust to VR. Participants were given 20 minutes to complete the scenario, but could end the scenario early if they feel that they have completed the necessary care and tests for the patient. After completing the scenario, participants completed a second questionnaire on a separate sheet (this is considered time point 2). The second questionnaire featured the same three questions as the first questionnaire (see above), as well as the following questions:

- “To what extent would you be prepared to leave the patient prior to a senior review” (this question was answered using a visual analogue scale)
- “Did you complete all the history, examinations and investigations necessary? If not, what else would you do if given more time?”

Data Analysis

The dependent variables that we derive are as follows:

- **Performance:** The OMS software implements a series of objectives for each scenario, which are tasks or actions that the participant is expected to have completed within the allotted time. This can include administering oxygen, prescribing a particular medication or calculating the Patient Early Warning Score (PEWS). The proportion of completed objectives is used as a score of the participant's performance during the scenario.
- **Confidence Change:** the participants' confidence in their understanding of the patient's condition is recorded at two time points, with the first being after 5 minutes (out of the 20 minute time limit) and the second being after the participant has finished the scenario. Confidence at each stage is recorded on a 10 point scale (1-10). The difference between the second and the first confidence rating is taken, such that a positive value indicates that the participant has increased their confidence over the course of the scenario.
- **Number of Differentials:** participants are asked to record all the diagnostic differentials that they are considering at the two aforementioned time points. Hence, the total number of differentials is recorded at each stage. The Initial Number of differentials is the number of diagnoses provided at the pause point.
- **Diagnostic Appropriateness:** each participant's set of differentials are assessed for how appropriate they are for the scenario. Each scenario has a set of differentials that are considered most likely, probable and improbable (with any others considered incorrect). To calculate a score for how appropriate the diagnoses are, we sum the likelihood values provided for all differentials that were marked as most likely or probable. We then add these to the sum of likelihood values for improbable differentials divided by two. This sum is divided by the total sum of all differentials. This overall measure then

measures what proportion of the participants' likelihoods are dedicated to probable differentials. However, we also penalised participants for providing few differentials, such that high scoring sets of differentials are larger sets of likely or probably differentials.

We also derived measures of information seeking similar to previous studies. The VR scenarios are far richer in terms of the available set of information for participants when compared to the vignette paradigm. For our analysis, we record all actions (or 'clicks') made by participants whilst in the scenario. Actions are categorised into a number of groups. The main categories are labelled as History, Examination or Testing, similar to in the vignette study. This set of information is mostly similar across scenarios though there are minor differences especially in the History category. Across scenarios, there are 35 possible History actions, 29 Examination actions and 18 Testing actions. This especially means that in comparison to the vignette paradigm, participants can take more detailed patient histories and can receive very different pieces of information depending on what they request from patient documentation and from asking the patient/guardian in the scenario. Outside of these categories, there are other actions available to participants, such as administering medication for the patient, calling for help or providing reassurance to the patient/guardian, but these are not used for our analysis. After categorising the participants' actions, we define a number information seeking measures:

- History Taking: this is the number of History actions for a given scenario that take place before the pause point.
- Total Information Seeking: this is the number of actions classified under History, Examination or Testing across the scenario.
- Information Value: to calculate the value of each information sought across these categories, we calculate the difference in OMS performance score for

participants with or without that information. We then sum all sought information values for each participant within each of the information categories (History, Examination, Testing).

- Amount of Treatment: this is the number of actions classified as treatment of the patient across the scenario.

As all actions are recorded with timestamps in the output dataset, we categorise whether actions occurred before or after the pause point (5 minutes in). Hence, we can investigate information seeking before and after the pause point where participants record their initial diagnoses and confidence.

Results

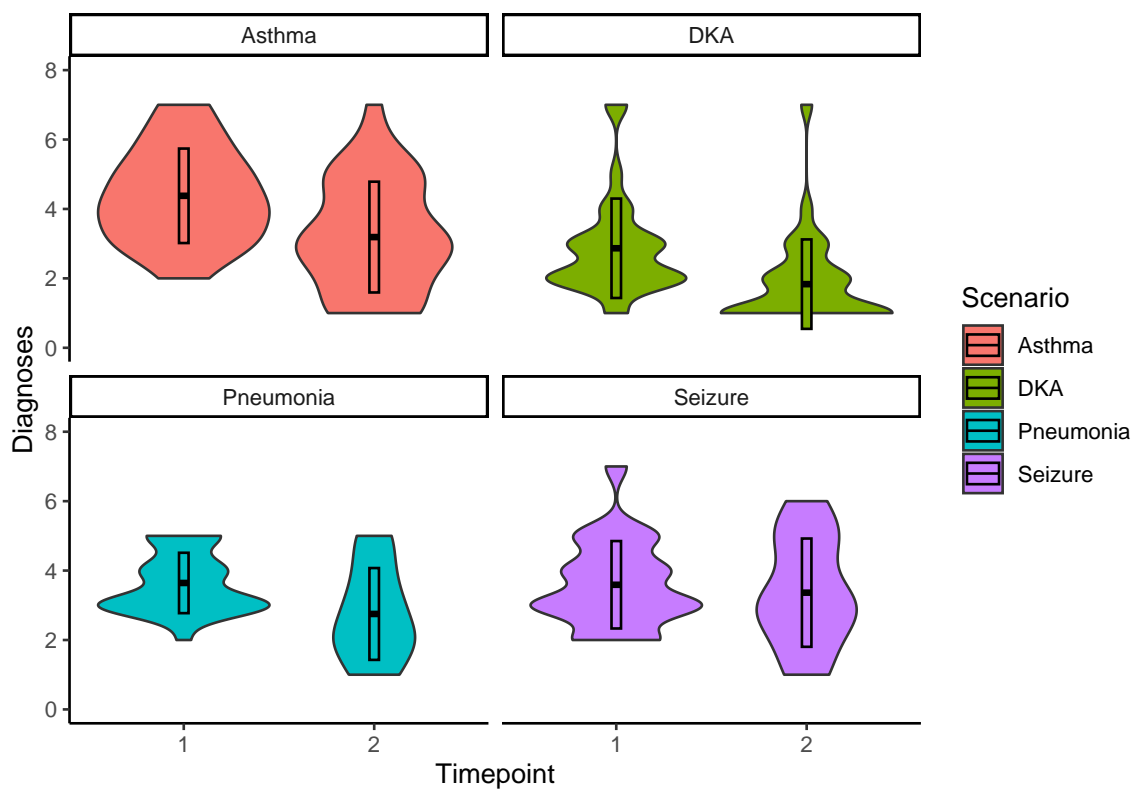
Overall Performance

Say how many participants completed each scenario

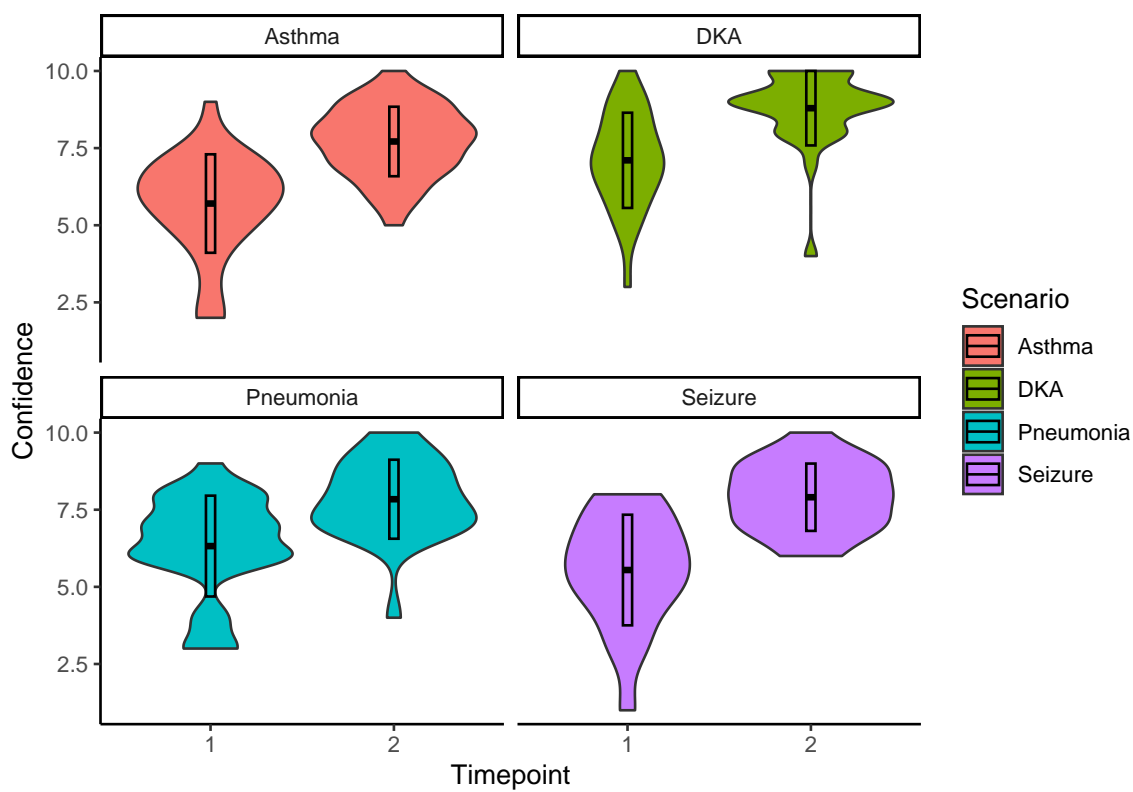
Table 1 below shows average values for our key dependent variables by scenario.

Scenario	n	MeanPerformance	MeanInformationSeeking	MeanInitialConfidence	MeanC
Asthma	37	68.14	19.73	5.70	
DKA	30	57.30	19.47	7.10	
Pneumonia	28	63.31	24.46	6.32	
Seizure	22	63.81	24.09	5.55	

Diagnoses by timepoint and scenario



Confidence by timepoint and scenario



Initial Diagnostic Breadth

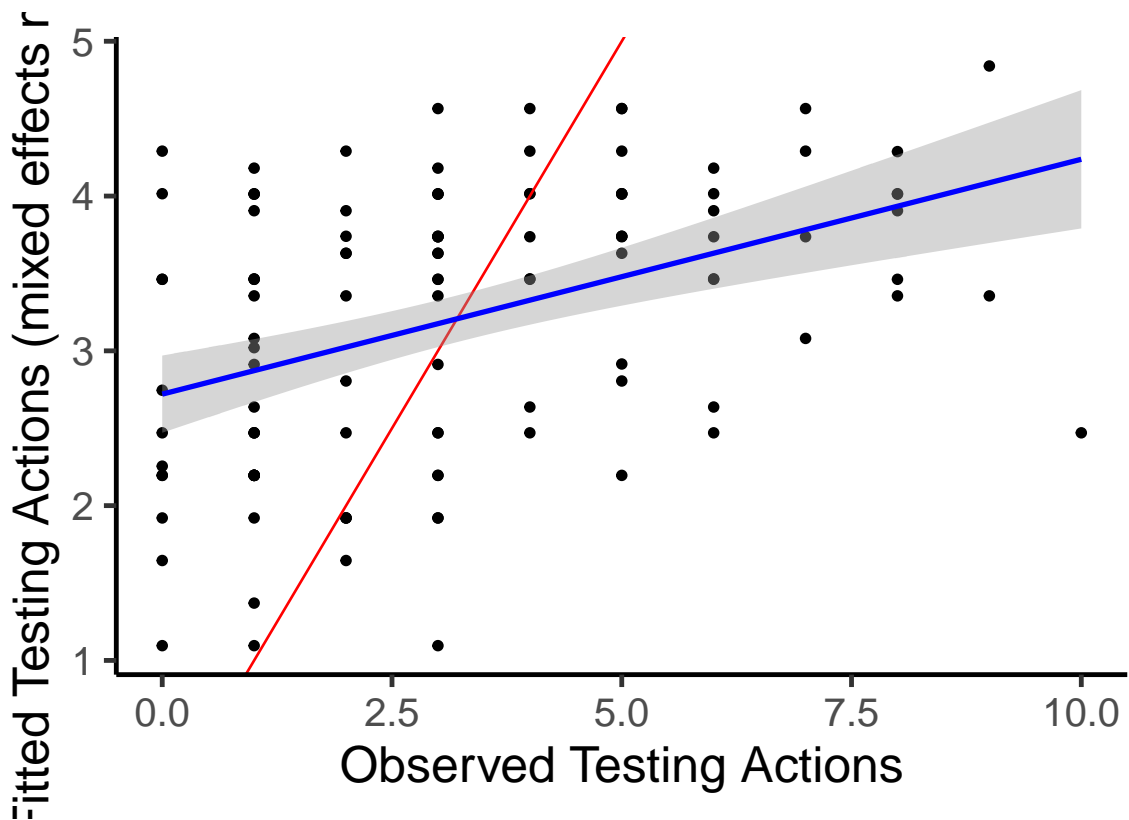
ConfidenceChange \sim InitialDiagnoses + (1|Scenario) + (1|Participant)

InformationSeeking \sim InitialDiagnoses + (1|Scenario) + (1|Participant)

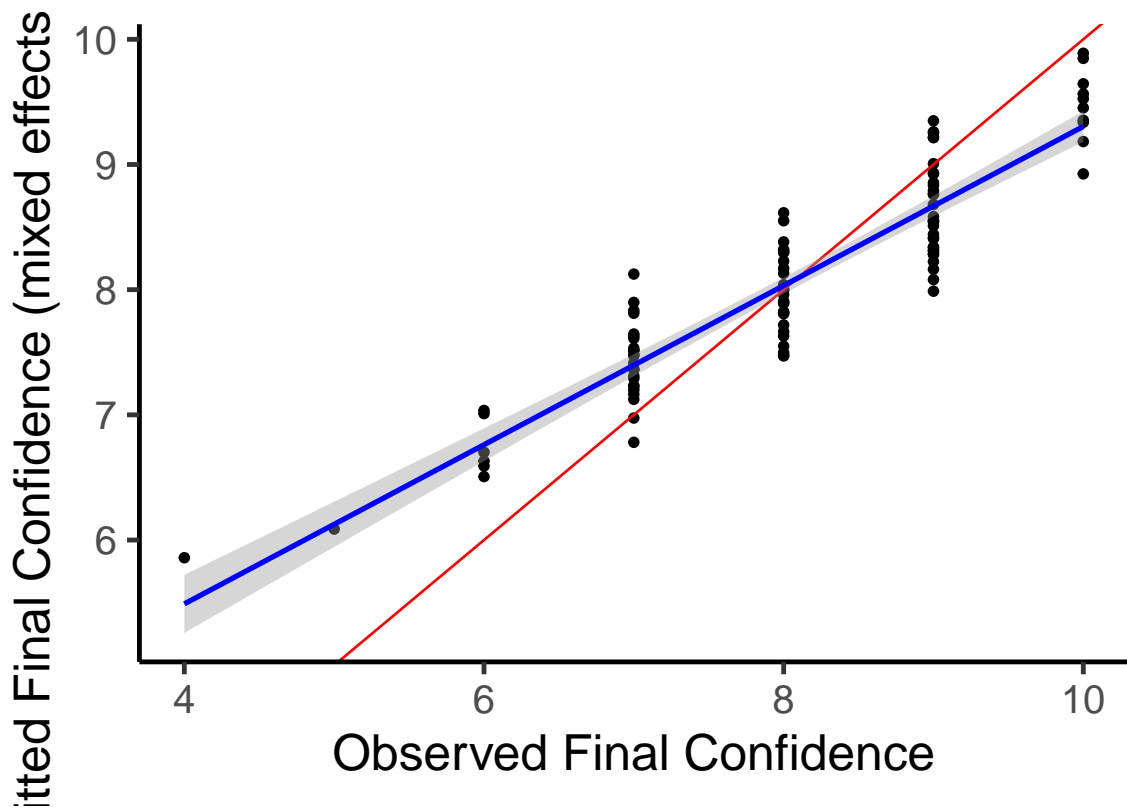
Information Seeking and Confidence

Initial Confidence \sim Information Seeking Before

Information Seeking After \sim Initial Confidence



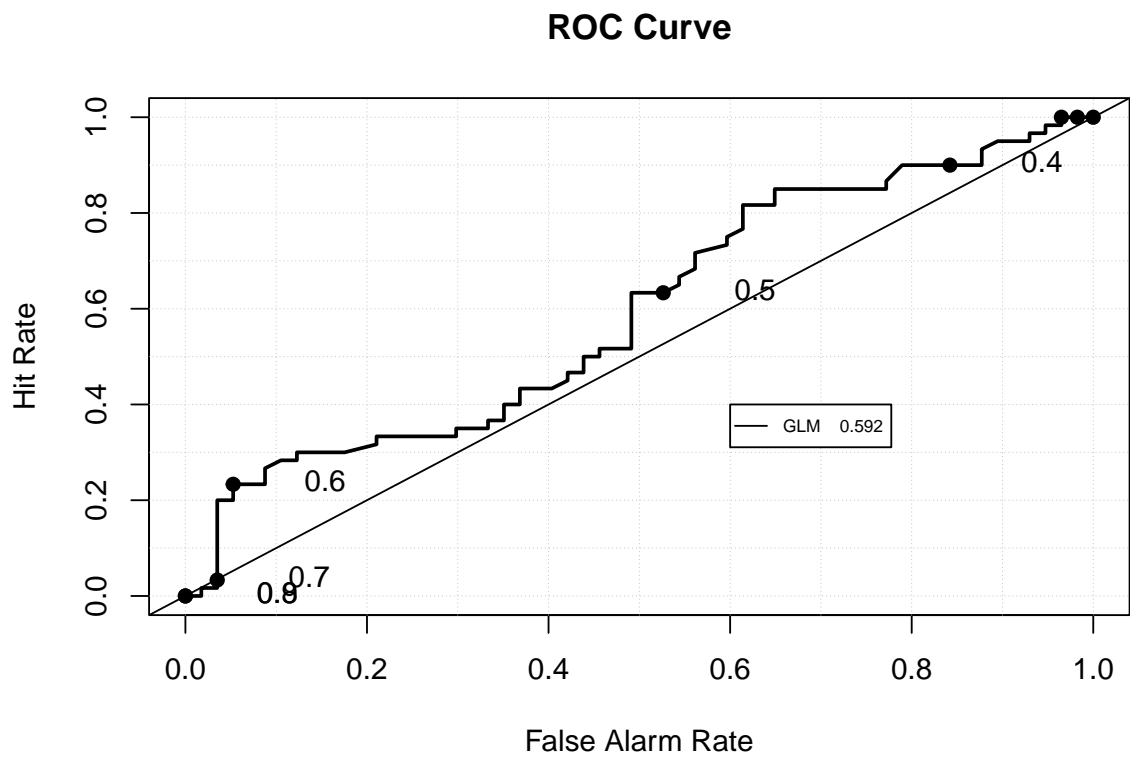
Final Confidence \sim Treatment + Time to Ask for Help



Diagnostic Appropriateness

Diagnostic Appropriateness ~ Value Diagnostic Appropriateness split with binary

classifier



##	Model	Area	p.value	binorm.area
## 1	Model 1	0.5918129	0.04367647	NA