Date: 15/10/2020

Validating virtual world simulation to assess decision-making of managing an open lower limb fracture: a prospective cohort study

Study Protocol and Statistical Analysis Plan

Participants and study design

Postgraduate (PG) doctors of varying levels were voluntarily recruited and randomly divided into three cohorts on the basis of training grades (and number of years of postgraduate training): Novice (PG 1-2 years, n=10), Intermediate (PG 3-4 years, n=10), and Experts (PG 5-10 years, n=10). This was a London-wide multi-centre prospective cohort study conducted at Imperial College London, an academic and public university. An invitation to participate was extended to medical personnel through the hospitals and medical students through their universities. Those who replied were consented and included in the study. There was no selection bias since the authors did not personally select the participants. There were strict criteria set for the eligibility of participation in the study. The participants represent a heterogenous mix which should not affect the overall outcome. Every participant was tested individually under the same supervision, on the same laptop, using the same online VW program between March to June 2020.

Selection criteria

Inclusion criteria: PG orthopaedic trainees and naïve to surgical simulation.

Exclusion criteria: undergraduates and previous exposure to surgical simulation involving cognitive task analysis.

Simulation design

An immersive clinical scenario was created, allowing interaction with both the virtual environment and patient to augment the learning experience and improve decision-making skills. The self-directed learning experience was enhanced by using real-time interactive multimedia, through external links to official peer-reviewed guidelines, and instructional videos by international professional trauma, orthopaedic and emergency medicine bodies.

Modified Delphi technique for expert approval

In order to ensure that the content of VW is accurate and represents current clinical practice, the schema and questions were sent to three independent experts (Trauma and Orthopaedic Consultants/Attendings) until a consensus was reached. There were ten rounds of drafts generated until the VW was enrolled out to participants. The agreement was on the content and the core knowledge expected of trainees with respect to ATLS and BOAST guidelines.

Tasks

Each participant was taken through the same scenario by answering a series of multiple choice and single best answer questions regarding each management step. Participants interacted with the on-screen patient while gathering information on the history, examination findings and investigation results. Optimal treatment options were also presented from resuscitation to surgical intervention and post-operative complications (venous thromboembolic disease and compartment syndrome). Correct answers, each equalling a score of 1, to questions were displayed on-screen immediately after an answer was submitted, with explanations and external links provided.

Outcome measures

(i) ***Face and content validity***

The face and content validity were assessed through the use of five questions which looked at how realistic the scenario seemed, and if the participants felt that it reflected real-life practice. The participants scored their answers using a 7-point adapted Likert scale which was chosen as research confirms significantly less accurate data when the number of scale points drops below five or above seven [19].

(ii) ***Construct validity***

This was determined through the objective score of the participants' overall performance using the programme. The programme recorded the selected answers, if they were correct and the time taken to complete the exercise.

(iii) ***Comparison with Need for Cognition Scale and Mental Rotation Test scores***

The objective scores were correlated with validated educational theory tests consisting of the Need for Cognition Scale (NFC) [20] and the Mental Rotation Test (MRT) [21]. The NFC scale, acts as a measure of how much the participant enjoys thoughtful processes, whilst the MRT looks at the participants' inherent visuo-spatial ability.

(iv) ***Semi-structured interviews***

Interviews formed the subjective assessment for thematic analysis. The participants' opinions were explored further through six open-ended questions each lasting for one minute. Participants expressed their thoughts about the programme qualitatively. The questions looked into what they felt could be improved, their likes and dislikes, and if they thought it was useful as an assessment and training tool.

Data analysis

Statistical analysis

Shapiro-Wilk normality test indicates that the data followed normal distribution (mean=191.7, median 192, SD=9.99, W=0.95, $p>0.93$ if p-value is set as 0.05). Hence the sample population represents both the sample and the wider population. The median ± median absolute deviation with 95% Bonett-Price confidence intervals were calculated for both objective and subjective data. The data were treated uniformly, and parametric testing through 95% Bonett-Price confidence intervals between two medians, 2-tail unpaired t-test between two cohorts and a one-way ANOVA test between all three cohorts were used to determine significant differences ($p<0.05$) in the overall scores.

Thematic analysis

Thematic analysis was based on the answers from the semi-structured interviews to identify key words and phrases, using the protocol set out by Braun and Clarke [22].

Ethics

The study did not require ethical approval from NHS REC. The study was conducted on university and academic premises outside of clinical commitments. All participants volunteered and offered their informed consent with the option of withdrawing from the study at any point during its course. All data was anonymised and destroyed after the completion of the study.