Clinical decision-making augmented by simulation training: neural correlates demonstrated by functional imaging: a pilot study


1 University Division of Anaesthesia, Addenbrooke’s Hospital, Hills Road, Cambridge CB2 0QQ, UK
2 Queens’ College, University of Cambridge, Cambridge CB3 9ET, UK
3 School of Clinical Medicine, University of Cambridge, Hills Road, Cambridge CB2 0SP, UK
* Corresponding author. E-mail: sg613@cam.ac.uk

Editor’s key points

- The neurocognitive mechanisms underlying the beneficial effects of simulation training on performance are poorly understood.
- Functional magnetic resonance imaging was used to monitor brain activation patterns in parallel with physiological indices of stress in subjects answering multiple choice questions after undergoing simulation or online-based training.
- This pilot proof-of-concept study demonstrated differences in cerebral activation patterns between the two groups in areas activated in response to stress consistent with simulation-induced stress attenuation.

Background. Investigation of the neuroanatomical basis of clinical decision-making, and whether this differs when students are trained via online training or simulation training, could provide valuable insight into the means by which simulation training might be beneficial.

Methods. The aim of this pilot prospective parallel group cohort study was to investigate the neural correlates of clinical decision-making, and to determine if simulation as opposed to online training influences these neural correlates. Twelve third-year medical students were randomized into two groups and received simulation-based or online-based training on anaphylaxis. This was followed by functional magnetic resonance imaging scanning to detect brain activation patterns while answering multiple choice questions (MCQs) related to anaphylaxis, and unrelated non-clinical (control) questions. Performance in the MCQs, salivary cortisol levels, heart rate, and arterial pressure were also measured.

Results. Comparing neural responses to clinical and non-clinical questions (in all participants), significant areas of activation were seen in the ventral anterior cingulate cortex and medial prefrontal cortex. These areas were activated in the online group when answering action-based clinical questions related to their training, but not in the simulation group. The simulation group tended to react more quickly and accurately to clinical MCQs than the online group, but statistical significance was not reached.

Conclusions. The activation areas seen could indicate increased stress when answering clinical questions compared with general non-clinical questions, and in the online group when answering action-based clinical questions. These findings suggest simulation training attenuates neural responses related to stress when making clinical decisions.

Keywords: computer simulation; magnetic resonance imaging, functional; stress, psychological

Accepted for publication: 12 July 2013

Simulation has been increasingly used in anaesthesia training. Previous studies indicate that students taught using high-fidelity simulation performed better in simulated and clinical settings. Reasons cited for simulation as a tool for effective learning have included the provision of feedback and evaluation, and the facilitation of experiential learning. Simulation training may also facilitate stress reduction in the workplace, which can adversely affect human and team performance, and impair non-technical skills. Research is still required to demonstrate the cognitive processes which may be responsible for the perceived benefits of simulation training.

We used functional magnetic resonance imaging (fMRI) to examine the neural correlates of clinical decision-making. Analysis of fMRI images allows identification of brain regions involved in the function of interest, referred to as activated regions. fMRI has been used, for example, to compare situation-based and text-based learning of language, but has not yet been used to investigate clinical methods of learning. This pilot exploratory study was used to demonstrate proof of concept, using a small number of subjects. While it would not be possible to draw firm conclusions from the study, proof of concept should aid the justification of a larger scale study with more subjects.
This was a pilot prospective parallel group cohort study. We hypothesized that by using fMRI, we could explore the cognitive processes behind clinical decision-making in medical students. Further, that we might be able to identify neural processing differences, if any, between learners trained by simulation vs online training.

It remains unclear what neural differences exist between simulation-based or online-based learning. Therefore, the imaging analysis was run in an exploratory fashion in order to avoid missing important areas of neural activation. Given the possibility raised for stress reduction after simulation training, we measured behavioural data to correlate with stress, that is, salivary cortisol levels, heart rate, arterial pressure, and visual analogue scales. The control of human emotions involves various limbic and paralimbic structures. Specifically, the anterior cingulate cortex and medial prefrontal cortex have been shown to respond to stress. Therefore, changes in activation patterns in any of these areas could be an indication for the induction of stress during the testing session.

The main aim of this study was to investigate the neural correlates of clinical decision-making, that is, the identification of cognitive processes behind clinical decision-making. The secondary aim was to determine if simulation training as opposed to online training influences these neural correlates.

**Methods**

Ethical approval was granted by Cambridgeshire 2 Research Ethics Committee. Written informed consent was obtained from participants.

**Participants**

We recruited medical students from Cambridge University (aged 20–21) in the academic year 2010–11. As this was a pilot study, a power calculation was not performed. We recruited 12 participants (6 in each group). Inclusion criteria were that participants were in their third undergraduate year, and spoke English as their first language. These students had no prior experience in the clinical diagnosis and management of anaphylaxis.

Exclusion criteria were refusal to participate in the study, contraindications to MRI scanning, claustrophobia, pregnancy, medication such as antidepressants, or medical conditions which may affect cognitive performance, such as history of significant traumatic brain injury or of significant psychiatric illnesses. Since the experiment had a strong language element, and because of language lateralization in the brain with handedness, we excluded left-handed volunteers.

Volunteers were randomly assigned into two groups using computer-generated random numbers. One group was assigned to simulation-based training and the other to online-based training.

**The teaching session**

Volunteers were taught using simulation-based or online-based training, for a total of 30 min each (identical factual content), on anaphylaxis. The online-based session was a supervised session, and the student could ask questions freely. The simulation-based session was structured into an introductory 5 min of factual slides with the same slides from the online session on features, diagnosis, treatment, and aftermath, and a shorter version of the pathophysiology. Information from the online session describing the airway, breathing, circulation, disability, and exposure approach to managing a critically ill patient were instead taught for 10 min using the simulator (METI Human Patient Simulator, Sarasota, FL, USA). The student then participated in a scenario on anaphylaxis for 10 min, followed by 5 min for debriefing and summary.

The same researcher structured the teaching sessions, facilitated the simulation-based sessions, and supervised the e-learning sessions for all volunteers in order to reduce bias. The researcher was not actively teaching during the online session, so the researcher was not blinded to the nature of the study. In the information sheet, it was stated that high-fidelity simulation may lead to improved teaching and learning, and that the aim of the study was to determine the effect of simulation-based learning on brain networks which mediate clinical decision-making by comparing brain activations between the group taught online and the group taught via simulation. An interval of 3–7 weeks elapsed before the scanning session. The difference in the intervals between the teaching and scanning sessions in the two groups was not statistically significant ($P=0.62$).

**The scanning session**

At the beginning of the scanning session, the volunteers underwent computerized neuropsychological tests taken from the Cambridge Neuropsychological Test Automated Battery (CANTAB, Cambridge Cognition Ltd; www.cantab.com) to test for equivalent cognitive abilities between the groups. The following tests were performed in a quiet room: simple reaction time, paired associates learning, big/little circle, intra/extradimensional shift, and information sampling task. The tests measured speed of response, episodic memory, comprehension and learning, rule acquisition and impulsivity, and decision-making, respectively.

Scanning involved answering questions to a multiple choice question (MCQ) task. Volunteers familiarized themselves with the format of the task before entering the scanner by answering 10 practice questions. They also completed a visual analogue scale (VAS) to measure anxiety, and a salivary cortisol sample was taken. Heart rate and non-invasive arterial pressure was measured every 5 min during the scanning session (Precess, InVivo Corp., Orlando, FL, USA). Immediately after the scan, a second salivary cortisol sample was taken, and a repeat VAS for anxiety completed.

**The fMRI MCQ task**

An event-related fMRI paradigm was used. The participants answered 80 MCQs presented in random order using a box with four buttons. These were equally divided into clinical and non-clinical (control) questions. All clinical questions were based on anaphylaxis, testing only what was taught in...
the teaching session. In each of these two categories, questions were further equally divided into those based on knowledge, or performing an action. In other words, there were 20 ‘action clinical’, 20 ‘action control’, 20 ‘knowledge clinical’, and 20 ‘knowledge control’ questions. An example of a knowledge question would be ‘The drugs that most commonly trigger anaphylaxis are ….’ An action question would, for example, be ‘When assessing a patient with suspected anaphylaxis begin by ….’

Each MCQ was displayed for 10 s. Questions were of similar length, taking ~8.9 s to read. This was determined beforehand by recording the time taken for the same researcher to read each question three times. No statistically significant difference in sentence length was found (P > 0.05). The time to answer the question was kept short in order to maintain the attention of the volunteers over the 80 MCQs. The MCQs were tested on volunteers unrelated to the study, who felt that the 10 s for reading and answering the question were sufficient. The pace of the paradigm could perhaps have induced stress, but both groups would have experienced the same effect.

A fixation cross was displayed for 3 s between MCQs, during which no cognitive task was performed, that is, MCQ, fixation, MCQ, fixation for the entire experiment, which was carried out in one continuous scan. The timing between stimuli was not varied, as there was natural variation in time intervals between the onset of the MCQ and the response. The fixation blocks were used as a coarse baseline to ensure we achieved activations in places we expected them to be in, and to provide adequate separation between haemodynamic peaks of interest. E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA, USA) was used for stimulus presentation and response recording. The fMRI paradigm took 1040 s to complete.

Image acquisition
A Siemens Trio 3 T scanner (Wolfson Brain Imaging Centre, Cambridge, UK) was used for the MRI scanning.

Statistical analysis

Behavioural data analysis
Statistical analysis was performed using PASW Statistics 18 Release Version 18.0.0 (SPSS Inc., Chicago, IL, USA). Standard parametric tests were applied to analyse the behavioural data, using two-tailed unpaired t-tests. Some of the MCQ responses were not logged (on average <5%). This could have been for reasons such as uncertainty about the answer, or not pressing the button sufficiently. Naturally, these responses were not included in the analyses.

Imaging analysis
Statistical Parametric Mapping (SPM8, Wellcome Trust Centre for Neuroimaging, London, UK) was used for pre-processing and modelling of the fMRI (BOLD) images implemented in MATLAB version 7.5 (Mathworks, Natick, MA, USA).

Analysis was carried out in an exploratory fashion, in order to assess possible neural differences between simulation and online-based learners. A fixed-effects analysis was used due to the small number of subjects. Reported areas of activation survived a voxel threshold of P ≤ 0.005 uncorrected and a random field cluster threshold of P ≤ 0.05 corrected for multiple comparisons for the entire brain.

Supplementary methods

Image acquisition
Each functional BOLD volume consisted of 33 interleaved, descending, oblique axial slices, 3 mm thick with interslice gap of 0.75 mm and in-plane resolution of 3 mm, field of view=192 × 192 mm, repetition time (TR)=2 s, acquisition time=2 s, time echo (TE)=30 ms, and flip angle=78°. We also acquired T1-weighted structural images at 1 mm isotropic resolution in the sagittal plane, using a magnetization prepared rapid gradient echo (MPRAGE) sequence with TR=2250 ms, inversion time (TI)=900 ms, TE=2.99 ms, and flip angle=9°, for localization purposes.

Imaging analysis
Pre-processing of the fMRI data involved slice-timing correction to account for time differences in slice acquisition time and within-subject realignment to the first fMRI image acquired to account for head motion. fMRI images were then spatially normalized to the Montreal Neurological Institute (MNI) template and finally were spatially smoothed using an isotropic (6 mm³) full-width half-maximal Gaussian kernel. The structural T1-weighted images were spatially normalized to the T1-weighted MNI template and used to superimpose the functional results from individuals to ascertain anatomical correspondence.

We used a fixed-effects general linear model to model the five conditions of interest namely: fixation, action-based clinical questions, knowledge-based clinical questions, action-based non-clinical (control) questions, and knowledge-based non-clinical (control) questions. The model included all 12 volunteers. Gender and co-variates of no interest were not included in the model. For some of the analyses described below, we obtained weighted parameter estimate images (contrasts) for each volunteer and combined them into group-level analyses.

For the contrast ‘action clinical vs action control’, we performed a correlational analysis between extracted βs (proxy for activation) in the anterior cingulate cluster, and response accuracy to the action clinical questions. This was to determine if this cluster of activation could have arisen as a result of poorer performance in the MCQs. This contrast was chosen due to poorer performance in the action clinical MCQ questions being found in the online group as opposed to the simulation group.

To establish the relationship between the change in salivary cortisol levels, and brain activation patterns common to both groups, we used a linear regression model with cortisol levels (µg dl⁻¹) as a regressor, combining simulation and online groups.

We also performed a correlational analysis between cortisol levels and activity in the anterior cingulate cluster (as detected
for the action clinical vs action control questions). To achieve this, we extracted region of interest (ROI) data from the anterior cingulate cluster, that is, we extracted mean weighted $\beta$s (proxy for activation).

Given the role of the ventral anterior cingulate, medial prefrontal cortex, and amygdala in emotional processing, we also calculated functional connectivity of the ventral anterior cingulate cortex with the amygdala for activity attributed to the action clinical vs action control questions using a correlational ROI approach. MarsBaR (MARSeille Boîte À Région d’Intérêt, a statistical parametric mapping toolbox used for ROI analysis) was used to extract ROI data from the weighted parameter estimate images for the anterior cingulate cortex and amygdala.

We report MNI coordinates. They were converted to Talairach space in order to assign Talairach and Tournoux (1988) nomenclature to peak activation voxels using the Talairach and Tournoux atlas. Voxels are volume elements, or cubes of brain tissue, several thousand of these making up each fMRI scan.

### Results

Twenty-seven volunteers were examined for eligibility aiming to recruit 12 volunteers. Fifteen were eligible. Twelve were initially recruited. However, due to technical problems with data logging during the fMRI scanning of two volunteers, behavioural data was lost and therefore the volunteers could not be included in the study. One other volunteer became claustrophobic during scanning, and could not complete the procedure. Hence, three further volunteers were recruited. They were randomized into the simulation-based training group (four females, two males) and the online-based training group (three females, three males).

#### Behavioural data

There was no statistically significant difference ($P>0.05$) between cognitive abilities of participants in the groups, as measured using the Cambridge Neuropsychological Test Automated Battery.

There was a trend suggesting that the mean reaction times from responses to the MCQs of the simulation group were shorter than those of the online group, but the differences were not statistically significant. There was no statistically significant difference between the two groups for the MCQ scores (Table 1).

The difference in salivary cortisol levels of the volunteers in each group before the scan was not statistically significant ($P=0.35$), nor was the change in levels during the scan. There was a mean reduction of 0.12 (SEM 0.09) $\mu$g $dl^{-1}$ in the simulation group compared with 0.03 (SEM 0.05) $\mu$g $dl^{-1}$ in the online group ($P=0.44$). The mean difference in the reduction in salivary cortisol in the simulation minus online group was $-0.08$ (95% confidence interval $-0.31$ to $0.15$).

There were no statistically significant differences between the two groups in the anxiety VAS scores before the scan or after the scan. There was no statistically significant difference in the change in VAS scores from before to after the scan. There were also no statistically significant differences between the two groups in mean heart rate, maximum heart rate, mean ‘mean arterial pressure’, or maximum ‘mean arterial pressure’ (Table 2).

#### Imaging data

While performing the task, that is, when comparing activation patterns resulting from answering all clinical and non-clinical MCQs with baseline (fixation), participants activated the cerebellum, primary somatosensory cortex, primary motor cortex, premotor cortex, supplementary motor area, dorsolateral and dorsomedial prefrontal cortex, primary visual cortex, Broca’s area, primary auditory cortex, orbitofrontal cortex, insula, thalamus, caudate, supramarginal gyrus, anterior cingulate, ventrolateral prefrontal cortex, putamen, hippocampus, posterior parietal cortex, angular gyrus, superior temporal gyrus, temporopolar area, and fusiform gyrus bilaterally. Given the complexity of the task, we expected widespread brain areas to

### Table 1

The mean accuracy in percentages; and response times (in seconds) are shown, with the standard error of the mean (SEM) in parentheses. The mean difference in percentage accuracy and reaction times are shown with 95% confidence intervals of the difference in parentheses.

<table>
<thead>
<tr>
<th>Multiple choice question category</th>
<th>Action-based clinical questions</th>
<th>Knowledge-based clinical questions</th>
<th>Action-based non-clinical (control) questions</th>
<th>Knowledge-based non-clinical (control) questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of correct questions (simulation group)</td>
<td>60 (8.7)</td>
<td>58 (5.3)</td>
<td>96 (4.2)</td>
<td>82 (4.8)</td>
</tr>
<tr>
<td>Percentage of correct questions (online group)</td>
<td>48 (3.6)</td>
<td>64 (5.1)</td>
<td>93 (1.7)</td>
<td>85 (2.9)</td>
</tr>
<tr>
<td>Mean difference in percentage accuracy (simulation minus online group)</td>
<td>11.7 ($-9.2$ to $32.5$)</td>
<td>$-6.7$ ($-23.0$ to $9.6$)</td>
<td>$2.5$ ($-7.5$ to $12.5$)</td>
<td>$-3.3$ ($-15.8$ to $9.1$)</td>
</tr>
<tr>
<td>Significance level</td>
<td>0.24</td>
<td>0.38</td>
<td>0.59</td>
<td>0.56</td>
</tr>
<tr>
<td>Reaction times (simulation group)</td>
<td>6.1 (0.5)</td>
<td>5.7 (0.4)</td>
<td>4.9 (0.4)</td>
<td>4.6 (0.3)</td>
</tr>
<tr>
<td>Reaction times (online group)</td>
<td>6.8 (0.1)</td>
<td>6.0 (0.2)</td>
<td>5.2 (0.3)</td>
<td>4.8 (0.2)</td>
</tr>
<tr>
<td>Mean difference in reaction times (simulation minus online group)</td>
<td>$-0.6$ ($-1.7$ to $0.4$)</td>
<td>$-0.3$ ($-1.2$ to $0.6$)</td>
<td>$-0.3$ ($-1.5$ to $0.9$)</td>
<td>$-0.2$ ($-1.1$ to $0.7$)</td>
</tr>
<tr>
<td>Significance level</td>
<td>0.21</td>
<td>0.50</td>
<td>0.58</td>
<td>0.65</td>
</tr>
</tbody>
</table>
Next, we examined neural responses when answering all clinical questions, over and above neural responses to all control questions (i.e. action control questions alone). We found that answering clinical questions over and above control questions in both groups were significant for the same contrast. This difference between the two groups did not reach statistical significance.

These activations were largely overlapping with those identified when comparing responses to all clinical questions with those elicited by all control questions in all 12 participants (Table 3, Fig. 1). The simulation group did not yield any significant activations for the same contrast. This difference between the two groups did not reach statistical significance.

There were no significant differences in activation when contrasting responses to knowledge-based clinical questions with knowledge-based control questions in either group (simulation group knowledge clinical vs knowledge control, or online group knowledge clinical vs knowledge control).

Next, we tried to establish whether response accuracy related to our cluster of interest in the anterior cingulate cluster activated in the action clinical vs action control contrast. We found no correlation ($P=0.20$, $r=0.40$) between response accuracy and the activation in the anterior cingulate.

Since this is an exploratory study, we also attempted to establish whether emotional processing carried out by amygdalo-limbic circuits plays any role in the processing of action clinical questions. We calculated functional connectivity between the ventral anterior cingulate cortex and the amygdala (for action clinical questions) by correlating $\beta$s (effects
sizes/strength of activation) from each participant, extracted from both areas, and found a significant correlation. Anterior cingulate activity correlated with activity in the right \( P=0.002, r=0.80 \) and left \( P=0.01, r=0.70 \) amygdala, suggesting an emotional element to the processing of action clinical questions.

Finally, \( \beta \)-values of the ventral anterior cingulate from the contrast of the action clinical vs action control questions (for all 12 participants) correlated positively with the change in salivary cortisol levels (salivary cortisol after the MCQ task minus salivary cortisol before) \( P=0.007, r=0.73 \), Fig. 3).

**Discussion**

We found differences in cerebral activation patterns between the two groups: those who did not have teaching in the simulator, when answering questions based on performing an action, displayed activation of the ventral anterior cingulate cortex and medial prefrontal cortex, areas critical in emotional processing\(^\text{11 13 14}\) and known to be activated in response to stress.\(^\text{12}\) These same areas were activated when all the students answered clinical compared with control questions. There was also a significant correlation between the change in cortisol levels and the effect size of activations in the ventral anterior cingulate cortex.

The areas of activation observed when answering MCQs overall involved areas of the brain responsible for planning movements, visuomotor coordination, processing language, reading, and association areas. This is unsurprising considering the participants were engaged in a complex task that involved a lot of cognitive components. We also observed activation in decision-making areas such as the lateral prefrontal cortex,\(^\text{18 19}\) anterior and ventral cingulate cortex,\(^\text{18}\) and posterior parietal cortex.\(^\text{19}\) Hence, activation patterns upon answering the MCQs were as widespread as expected.
The activation patterns we observed in the medial prefrontal cortex and ventral anterior cingulate cortex in response to answering clinical questions, and in the online group when answering action clinical questions, may suggest that stress and anxiety are evoked by the prospect of facing an emergency clinical scenario. Various emotions, including anxiety, have been shown to increase cerebral blood flow in the pregenual portion of the ventral anterior cingulate cortex. The anterior cingulate cortex has been reported to be consistently activated in response to stress. The ventral anterior cingulate cortex and medial prefrontal cortex may be important in the extinction of fear, and the rostral anterior cingulate cortex has been activated in response to the resolution of emotional conflict. It has been suggested that the medial prefrontal cortex responds to a variety of emotions, and has a general role in emotional processing. Activations in the ventral-rostral medial prefrontal cortex caused by emotional responses extend into the ventral anterior cingulate cortex. It is postulated that ventral-rostral areas of the medial prefrontal cortex interact with the ventral anterior cingulate cortex in regulating interconnected cognitive and emotional tasks.

Other roles of the medial prefrontal cortex include reward processing, retrieval of memory, and decision-making. Therefore, the activation we saw in the medial prefrontal cortex could have been due to these other cognitive processes. However, concurrent activation of the ventral anterior cingulate cortex and medial prefrontal cortex suggests that the activation patterns were more likely due to emotional processing, such as anxiety and stress. The possibility that simulation-based training may reduce stress in clinical situations fits with earlier findings, suggesting that learners benefit from simulation-based training, in that it provides deliberate practice.

As there was no correlation found in the action clinical vs action control contrast between the response accuracy of the MCQs and the activation in the ventral anterior cingulate cortex and medial prefrontal cortex, this would indicate that the activations were unlikely to be a result of poorer performance in the action clinical questions seen in the online group.

In the connectivity analysis between the ventral anterior cingulate cortex and amygdala, a correlation was found between activation in the amygdala and the ventral anterior cingulate cortex. The amygdala, as part of the limbic system, also has an established role in the control of emotion. This suggests a link between the ventral anterior cingulate cortex and emotion. The medial prefrontal cortex and anterior cingulate cortex may be top-down modulators of emotional responses and mediate responses in the amygdala. Since the amygdala did not respond to experimental stimuli (activation did not survive conservative statistical thresholds), we perhaps witnessed a ‘top-down’ relationship of activation between the amygdala and ventral anterior cingulate cortex.

Overall, there were no statistically significant differences in salivary cortisol levels between the groups. The use of salivary cortisol levels has limitations since adrenocortical activity can be influenced by other factors such as context and type of task. Cortisol may decrease after a stressful event if the stress produced before the task elicited a cortisol rise greater than that produced by the task itself. Alternatively, the time course of sampling may have missed the peak, which can be variable, or levels may have declined over time with circadian changes. There was, however, a significant correlation between the change in cortisol levels, and the effect size of the activations in the ventral anterior cingulate cortex.

Existing studies suggesting that simulation training may reduce stress are sparse. Previous research has suggested a reduced stress response to test scenarios after training in a simulation setting. However, the opportunity for the participants to become familiar with the simulator may have been a confounding factor. In addition, given the difficulties involved in objectively measuring stress, the combination of techniques used have highlighted the possibility that simulation-based training may reduce stress in clinical scenarios. This is important given the effects of stress on performance of tasks, such as those requiring divided attention, and decision-making.

Our study was limited by the sample size. This might explain why there were trends suggesting that those trained in the simulator answered more quickly and accurately, but these findings did not reach statistical significance. As this was a pilot study, a power calculation was not performed.

Statistical significance was not reached with the behavioural data collected to correlate with the induction of stress, that is, salivary cortisol levels, heart rate and arterial pressure measurements, and VAS scores. This would have been difficult to achieve with the small sample size. Therefore, the behavioural data could not be used to confirm results from the imaging analysis. We can only postulate the reasoning behind the neural activation patterns seen with the knowledge of what these patterns have been demonstrated in other studies to indicate. Our findings do not allow firm conclusions to be drawn, but can be used to inform hypotheses which can be tested in further studies.
Although the fixed-effects model used in the imaging analysis allowed us to postulate that there is a difference between the students studied, it does not allow generalization of results to the population. Moreover, this study can only be interpreted in the context of teaching medical students the subject of anaphylaxis. To draw more generalizable conclusions, the study would need to include more subjects and would need to be repeated with more experienced subjects such as trained anaesthetists and with different clinical scenarios. MCQs were chosen as a method of testing, since they can be responded to via a button box, which is routinely used in the scanner environment and does not result in movement artifacts. Despite the problems associated with the use of MCQs for testing clinical decision-making, MCQ scores have previously been reported to correlate with clinical reasoning test scores.

Evidence exists that simulation can improve some learning outcomes and clinical performance. However, further research for the cognitive processes which may take place in simulation learners, and the possible role of simulation in reducing stress is still required, perhaps using the methods we have described. fMRI has previously been used to study different methods of learning, in the field of languages, and arithmetic. fMRI has also recently been used to investigate clinical reasoning, using MCQs. However, to our knowledge, fMRI has so far not been used for investigating methods of learning in the field of clinical medicine.

In conclusion, this proof-of-concept study suggests that there may be a difference in the induction of stress in medical students taught by simulation vs online teaching. The results are consistent with the possibility that teaching medical students the subject of anaphylaxis using simulation may result in less stress when tested on this subject, in comparison to online teaching. The techniques we have described could be used in a larger scale study in order to further explore this subject and to confirm these findings.

**Authors’ contributions**

S.S.H.G.: design of study, study co-ordination, data collection, data analysis, writing up, revising, and final approval of the paper. E.A.S.: design of study, data analysis, writing up, revising, and final approval of the paper. R.M.A.: design of study, data analysis, writing up, revising, and final approval of the paper. M.K.: design of study, writing up, revising, and final approval of the paper. S.B.: design of study, writing up, revising, and final approval of the paper. D.F.W.: design of study, writing up, revising, and final approval of the paper. D.W.W.: design of study, writing up, revising, and final approval of the paper. D.K.M.: design of study, writing up, revising, and final approval of the paper. A.K.G.: design of study, writing up, revising, and final approval of the paper.

**Acknowledgements**

Salivary Cortisol samples were analysed by the NIHR Cambridge Biomedical Research Centre, Core Biochemical Assay Laboratory, Cambridge, Cambridgeshire, UK.

**Declaration of interest**

None declared.

**Funding**

This work was supported by Cambridge University Hospitals NHS Foundation Trust Postgraduate Medical Education Centre, Cambridge, UK. The simulator used in the study is part of the Postgraduate Medical Education Centre, but otherwise had no role in the study. E.A.S. is funded by the Stephen Erskine Fellowship, Queens’ College, Cambridge, UK. R.M.A. was funded by a Wellcome Trust Clinical Research Training Fellowship, Raymond and Beverly Sackler Studentship and an honorary Cambridge Commonwealth Scholarship, UK.

**References**


Krawczyk DC. Contributions of the prefrontal cortex to the neural basis of human decision making. *Neurosci Biobehav Rev* 2002; 26: 631–64


Euston DR, Gruber AJ, McNaughton BL. The role of medial prefrontal cortex in memory and decision making. *Neuron* 2012; 76: 1057–70


van der Vleuten CP, Newble DI. How can we test clinical reasoning? *Lancet* 1995; 345: 1032–4


*Handling editor: H. C. Hemmings*