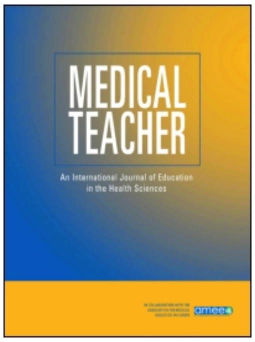

Taylor & Francis
Taylor & Francis Group

Medical Teacher

ISSN: 0142-159X (Print) 1466-187X (Online) Journal homepage: http://www.tandfonline.com/loi/imte20

Script-theory virtual case: A novel tool for education and research

Jake Hayward, Amandy Cheung, Alkarim Velji, Jenny Altarejos, Peter Gill, Andrew Scarfe & Melanie Lewis

To cite this article: Jake Hayward, Amandy Cheung, Alkarim Velji, Jenny Altarejos, Peter Gill, Andrew Scarfe & Melanie Lewis (2016): Script-theory virtual case: A novel tool for education and research, Medical Teacher, DOI: 10.3109/0142159X.2016.1170776

To link to this article: http://dx.doi.org/10.3109/0142159X.2016.1170776

MEDICAL TEACHER, 2016
http://dx.doi.org/10.3109/0142159X.2016.1170776

Script-theory virtual case: A novel tool for education and research

Jake Hayward[a], Amandy Cheung[b], Alkarim Velji[c], Jenny Altarejos[d], Peter Gill[e], Andrew Scarfe[f] and Melanie Lewis[g]

[a]Department of Emergency Medicine, University of Alberta, Edmonton, Alberta, Canada; [b]Department of Family Medicine, University of Alberta, Edmonton, Alberta, Canada; [c]Department of Emergency Medicine, University of Manitoba, Winnipeg, Manitoba, Canada; [d]Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada; [e]Department of Pediatrics, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada; [f]Department of Oncology, University of Alberta, Edmonton, Alberta, Canada; [g]Department of Pediatrics, Stollery Children's Hospital, University of Alberta, Edmonton, Alberta, Canada

ABSTRACT

Context/Setting: The script theory of diagnostic reasoning proposes that clinicians evaluate cases in the context of an "illness script," iteratively testing internal hypotheses against new information eventually reaching a diagnosis. We present a novel tool for teaching diagnostic reasoning to undergraduate medical students based on an adaptation of script theory.
Intervention: We developed a virtual patient case that used clinically authentic audio and video, interactive three-dimensional (3D) body images, and a simulated electronic medical record. Next, we used interactive slide bars to record respondents' likelihood estimates of diagnostic possibilities at various stages of the case. Responses were dynamically compared to data from expert clinicians and peers. Comparative frequency distributions were presented to the learner and final diagnostic likelihood estimates were analyzed. Detailed student feedback was collected.
Observations: Over two academic years, 322 students participated. Student diagnostic likelihood estimates were similar year to year, but were consistently different from expert clinician estimates. Student feedback was overwhelmingly positive: students found the case was novel, innovative, clinically authentic, and a valuable learning experience.
Discussion: We demonstrate the successful implementation of a novel approach to teaching diagnostic reasoning. Future study may delineate reasoning processes associated with differences between novice and expert responses.

Introduction

Diagnostic error

Effective and safe health care hinges upon accurate clinical diagnosis. Landmark studies suggest diagnostic error accounts for approximately 5–17% of preventable errors in hospitalized patients (Leape et al. 1991). A systematic review of four decades of autopsy studies found that approximately 9% of patients experienced a major premorbid diagnostic error that went undetected (Shojania et al. 2003). Common diagnoses are more frequently missed than esoteric ones (Graber et al. 2005; Singh et al. 2007; Schiff et al. 2009; Zwaan et al. 2010) and cognitive failure rather than knowledge deficit is the predominant cause of error (Campbell et al. 2007). The apparent link between diagnostic error and patient safety highlights a need to better understand diagnostic reasoning, its cognitive correlates and how best to teach diagnostic reasoning to students and trainees (Newman-Toker & Pronovost 2009).

Diagnostic reasoning, and avoidance of diagnostic error, is learned and refined as novice clinicians' progress to expert clinicians. Traditional undergraduate medical education emphasizes large group didactic teaching for knowledge transfer and small group problem-based learning for exposing knowledge processes (Nandi et al. 2000). Little time is devoted to explicit teaching of clinical reasoning, especially during the preclinical years, typically leaving this for implicit exposure during the clinical years. Sound clinical judgment relies on sound clinical reasoning, a cognitive skill in its own right that is not imparted through acquisition of factual knowledge (Elstein et al. 1990). Croskerry advocate explicit teaching of explicit strategies for reducing cognitive error (Croskerry 2002) and have demonstrated how this can be implemented in an undergraduate medical curriculum in a course teaching principles of cognitive bias (Croskerry et al. 2013).

Practice Points

- Clinical expertise and diagnostic reasoning is acquired through complex learning processes that are poorly understood.
- The best approach to teaching diagnostic reasoning remains unclear and current methods are underdeveloped.
- We present evidence that an online, group-based, reflective approach to teaching diagnostic reasoning is both feasible and valuable as an adjunct to standard undergraduate curricula.
- Our technology is grounded in the script theory of diagnostic reasoning and represents a case-based adaptation of script concordance testing methods.
- The results suggest a possible role for our technology in future research on the development of clinical expertise.

CONTACT Jake Hayward jhayward@ualberta.ca University of Alberta, Department of Emergency Medicine, Edmonton, Alberta, Canada
Supplemental data for this article can be accessed here.

Theories of diagnostic reasoning

While many theories attempt to explain the cognition of clinical reasoning, the Script Theory has gained notable popularity (Charlin et al. 2000; Bowen 2006). Script theory holds that the clinician draws upon prestored reasoning pathways in the form of illness scripts, or "profiles," when navigating new patient encounters. Acquired patient-specific data are iteratively referred to these illness profiles and the likelihood of a given hypothesis is adjusted until a sufficiently likely diagnosis is obtained.

The script theory implies many simultaneous psychological processes. As expressed by Charlin et al. (Charlin et al. 2000): "The clinician perceives or observes features of a situation and quickly accesses relevant hypotheses checking for signs and symptoms that either confirm or rule out competing hypotheses." The expert clinician carries out probability adjustments subconsciously with ease while the novice learner engages in a conscious analytic effort.

Script concordance testing

Clinical practice often involves ambiguous scenarios with imprecise or incomplete data and no single-best answer to a diagnostic or therapeutic question. The script concordance test (SCT) has been developed as a means to assess reasoning in the context of uncertainty (Charlin et al. 2000). By seeking estimates of likelihoods, it is fundamentally different from single-best answer assessments such as multiple-choice questions. By capturing some elements of thought processes, not just clinical knowledge, SCT question sequences may offer a more valid measure of clinical judgment. This potential has been demonstrated in a range of clinical assessment settings (Lubarsky et al. 2011).

A typical SCT sequence starts with a clinical vignette followed by 3–5 questions. Each question consists of three parts. First, a hypothesis is proposed (e.g. "if you were thinking of a diagnosis of pulmonary embolism…"). Second, an item of specific clinical data is offered (e.g. "And then you are informed of a positive D-dimer test result…"). Finally, the learner is asked to indicate the impact that the new data have on the likelihood of the hypothesis in part one (e.g. "This hypothesis becomes…"). By convention, a five-point Likert scale is presented and the learner is asked to record how much more or less likely the hypothesis becomes.

The SCT sequence is scored by comparing the learner's answer to the mean response of a panel of experts. Typically, a small group of experts are chosen for the topic being assessed. Experts complete the test individually and their pooled likelihood estimates are used to produce an answer key such that the learner's answer gains credit for the modal proportion of experts who also chose that answer, expressed as a percentage. Credits from each question are added to produce a total score. Although common, this scoring method is not the only method described (Gagnon et al. 2005; Charlin & van der Vleuten 2004; Lubarsky et al. 2011).

Learning theory and online cases

Adult learning theory postulates that the role of the teacher is not to transmit knowledge but to engage in a dialog that promotes critical thought, spontaneity, and mutual inquiry. Formative as opposed to summative assessment is preferred for timely and personalized feedback to encourage self-analysis and introspection. Naturally, the benefits of adult-learning theory are difficult to realize in large group settings. For these and other reasons, teaching clinical reasoning to large classes of preclinical medical students is challenging.

Online environments offer opportunities for more interactive and engaging learning and evaluation, even using multimedia simulations and virtual patients, while allowing large groups to have similar experiences (Gill et al. 2010; Kitney et al. 2014). Script theory and SCT have yet to be effectively integrated within clinically authentic online simulation environments.

Current project

We sought to develop a novel online case-based learning tool founded in script theory and to implement the tool into undergraduate medical education. Our specific objectives were to (1) evaluate the feasibility of our technology, (2) assess student-perceived educational value of the teaching tool, and (3) analyze response data for hypothesis generation and future research. We evaluate the usefulness of this teaching tool at the transition point from preclinical to clinical learning environments.

Methods

Traditional SCT question constructs were adapted for expression within an immersive virtual environment that included simulations of history taking, physical examination, laboratory test ordering and review of clinical materials in an electronic medical record (EMR). The online virtual patient survey system facilitated dynamic analysis and immediate feedback of individual results, comparisons to mean peer choices and further comparison to clinical expert choices. By allowing students to compare their answers to those of their peers, we hope to provide a benchmark for each individual's self-assessment. This is similar to the common practice of providing learners with a class average mark on an examination.

The virtual patient case was introduced as a mandatory activity in the undergraduate medical education curriculum at the Faculty of Medicine and Dentistry at the University of Alberta (Edmonton, AB, Canada). In Canada, medical school is generally divided into two preclinical years, comprised largely of didactic teaching, followed by two clinical years spent providing direct patient care. The case was introduced the end of the second preclinical year during an oncology course. Data were collected over two academic years (2012–2013 and 2013–2014) comprising two cohorts. As the case concerned diagnostic evaluation of clinical findings in a young adolescent, eight Department of Pediatrics faculty members comprised the expert panel.

Virtual patient design

Learners were instructed to log on with their usual university-provided unique identifier and could complete the case when convenient, in one or more sessions, within a preset four-day interval. Once online, a brief introduction to the process was followed by the "History" section. An audio presentation describing a clinical interview with the virtual patient (Supplementary Figure S1) was accompanied by online notepads allowing learners to record and recall observations as they progressed. Later, these notes were available when learners were asked to document their findings in an integrated EMR simulator.

After the history, the learner was asked to provide likelihood estimates for five diagnostic possibilities (Figure 2). This response format was an adaptation of the SCT question

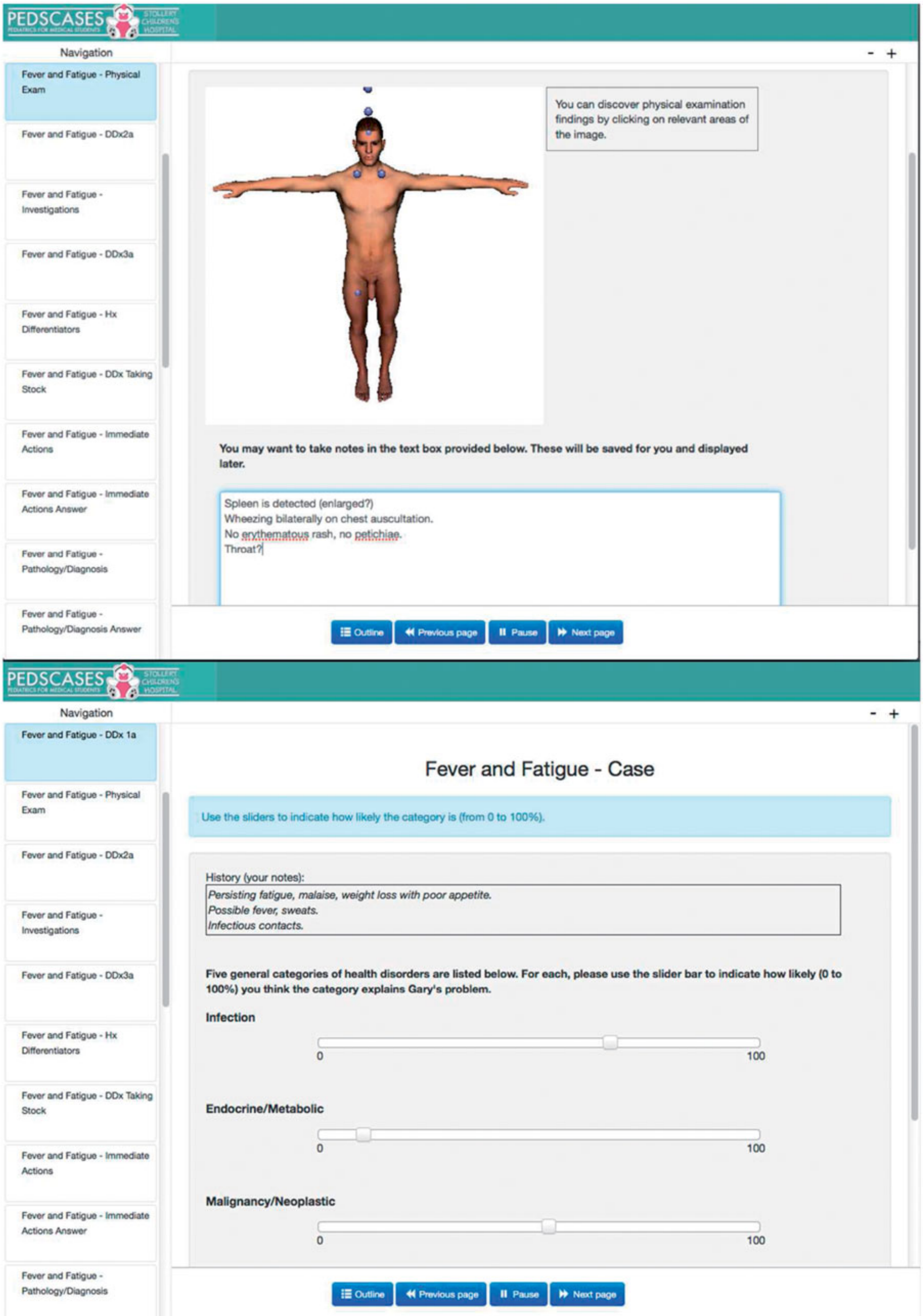

Figure 1. Screenshots from "Physical Examination" and "Sliding Scale Likelihood Estimates" sections. An interactive three-dimensional physical exam was used in conjunction with a note-pad function (upper panel). After each stage of case, virtual slide-bars were used for likelihood estimation of five diagnostic categories (lower panel).

Figure 2. Virtual patient workflow. The participant completed an immersive virtual case comprising a three stage iterative process of data acquisition and likelihood estimation for five diagnostic categories.

type in that continuously variable slide bars were used to indicate probabilities rather than a the more limited Likert scale (Figure 1).

In the next stage of the case, "Physical Examination," an interactive 3D model of the human body could be fully rotated, explored, and used to test for physical findings, with learners limited to eight priority tests from among many possibly relevant maneuvres (Figure 1). At completion of the examination, the learners were asked to re-evaluate their diagnostic likelihoods. The diagnostic response categories were kept the same and presented in the same order.

In the final stage, "Investigations," an interactive EMR simulator was used to access demographic and other information, navigate the current "Clinic Visit," edit history and physical examination notes to become part of the medical record, and order one or more investigations (Supplementary Figure S2). Results and diagnostic images were provided using the EMR, revealing clinical data in an order, style, and context typical for real-life cases in a digitally enabled workplace. The learners, once again, readjusted their diagnostic-likelihood estimates.

In effect, learners were asked to re-evaluate diagnostic likelihoods at three different stages as the clinical case unfolded. Only after all estimates at all case stages were inputted, did learners have an opportunity to review how their diagnostic impressions had evolved.

The interactive virtual patient system collected learner inputs and performed individual and group analyses in real-time. After clinical data collection, the learner was provided with feedback throughout the case. Pooled peer and expert responses could be compared to personal estimates. Interactive box plots with confidence intervals were used to estimate the variability of peer and expert estimates while transitional frequency distributions could be displayed as desired (Figure 3).

Following case completion, learner feedback and suggestions were solicited via a narrative response question. Students were asked to comment on the educational value of the interactive case, perceived weaknesses, technical difficulties, and suggestions for change.

Likelihood estimates

Learner SCT responses were pooled from both student cohorts and mean likelihood estimates were compared to those obtained from the expert clinician panel. Trends were analyzed descriptively. Likelihood estimates are reported as mean and 95% confidence intervals.

Theme analysis

Two authors (JH and AC) reviewed student feedback and identified recurring themes. Theme analysis was performed individually and then reviewed for interobserver agreement. Themes were first broadly categorized as either constructive or critical then subcategories were defined within each. NVivo qualitative data analysis software (QSR International Pty Ltd. Version 10, 2012) was used to for theme coding.

Once subthemes were delineated, both authors recorded the frequency of each comment. Codification rules were predetermined. A theme could occur no more than once per student but multiple themes might occur for any given student. Differences were flagged for discussion until consensus about theme allocation was achieved. Interobserver agreement was over 90%.

Results

Over two academic years, 301 (301/304; 99%) students participated, 144 from 2012 to 2013 and 157 from 2013 to 2014.

Likelihood estimates

Likelihood estimates for all diagnostic categories were similar across the two cohorts of students (Supplementary Figure S3). However, as illustrated in Figure 4, students' and experts' likelihood estimates differed at each stage of the virtual patient, including after history, physical examination, laboratory findings and overall assessment. For the overall final assessment, students' and experts' rated "malignancy" as the highest likelihood, with experts estimating likelihood at close to 100%, over 10% higher than the student group.

"Infection" received the second highest rating from both groups. Mean student estimates (34%; 95% confidence interval (CI): 31–37%) were nearly twice those of experts

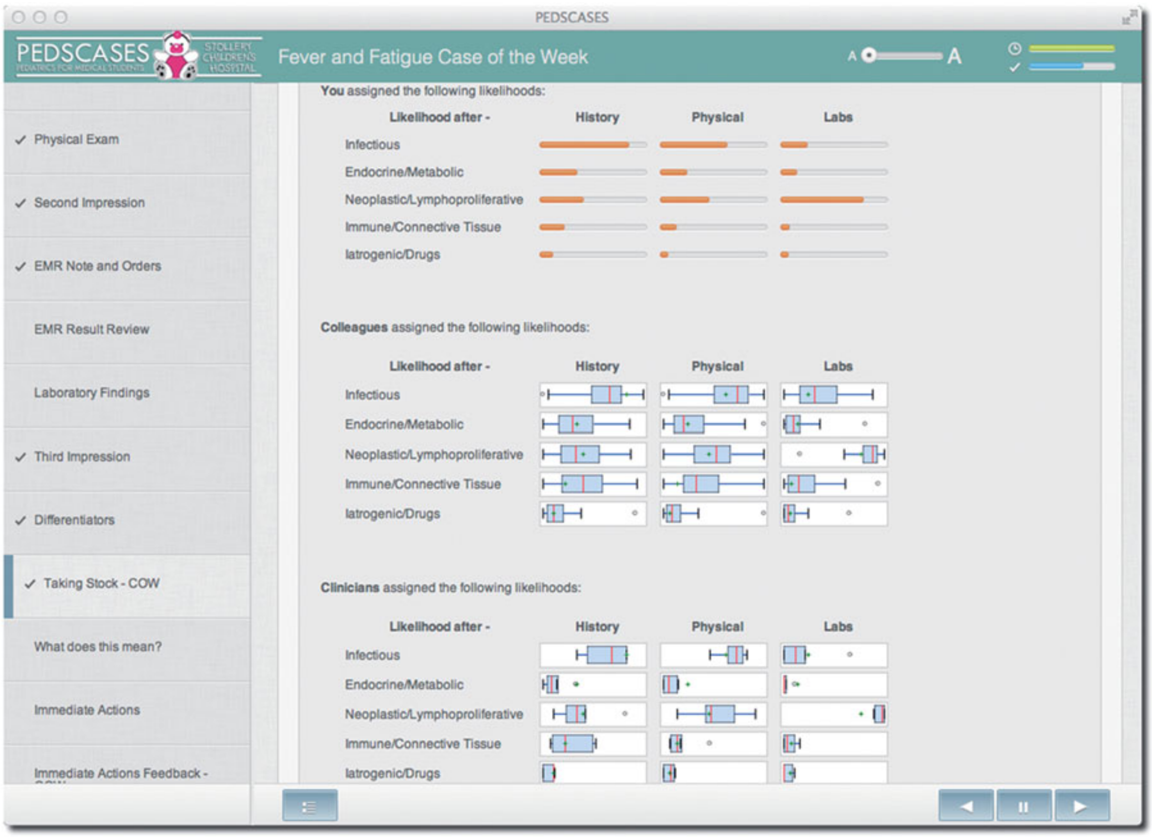

Figure 3. Screenshot from the 'Data Collection and Response Feedback' section. Individual likelihood estimates (top section) were compared to peers (middle section) and experts (bottom section) displayed with the mean, median, and interquartile ranges in box-plot format.

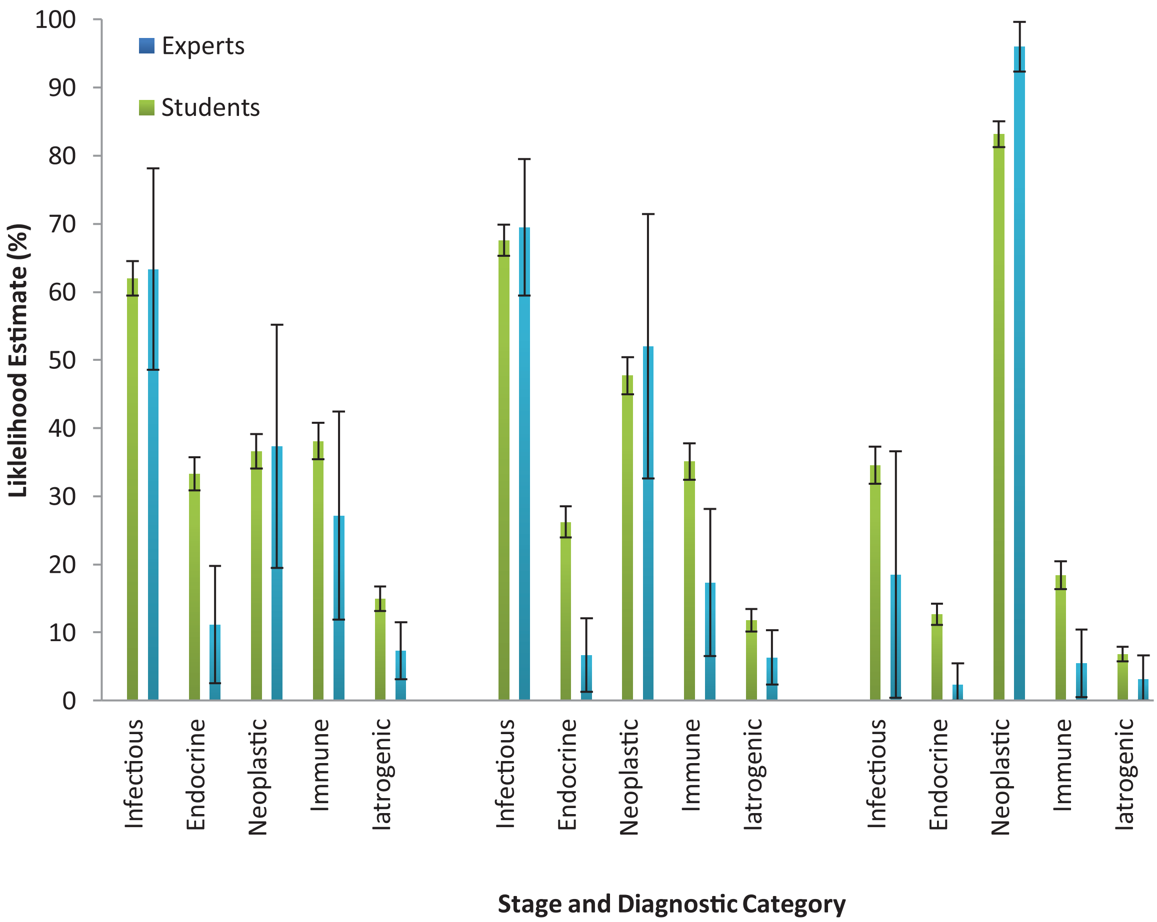

Figure 4. Likelihood estimates. Response data compared across level of expertise at three stages of estimation—after the history, physical examination, and investigations.

(19%; 95% CI: 1–36%). Expert responses were most varied for the "infection" category (interquartile range (IQR) = 2–20%), comparable to students (IQR =14–53%). Student responses were more moderate (e.g. 20–50%), while experts grouped toward the extremes of the probability scales for all categories (e.g. 0% or 100%). As more students and experts completed the case in subsequent years, we did not observe significant changes in the population means but confidence intervals narrowed.

Theme analysis

Two-thirds of medical students (225/301; 75%) provided detailed feedback on the virtual patient. Seven key themes were identified and are summarized in Table 1. The majority of respondents found the case to be a useful experience, commenting on its' novel and innovative approach (72%; 163/225). Students' commented on the clinical authenticity of the case content (32%; 71/225), the value of comparing locally credible expert case interpretations (27%; 61/225), the value of a formative learning environment (17%; 38/225), and the appeal of an immersive virtual patient (17%; 39/225).

More critical comments from students' centered largely on the technologic aspects of the case (32%; 71/225). Several students had difficulties interacting with the 3D physical exam and navigating the virtual EMR. While individual students were often critical of the virtual patient, over half (55%; 39/71) described an overall positive and educational experience. Four students felt that the case was too long, commenting that it took about an hour to complete whereas the authors had projected completion time of 30 min.

Discussion

Our study introduces the novel application of Script Theory to case-based instruction, promoting a deep level of formative learning for large groups of medical students simultaneously. The SCT question sequences encourage the learner to think in a probabilistic, Boolean fashion, while providing immediate feedback on performance, contextualized to the choices of peers and experts.

Although supportive data continues to accrue, it seems likely that complex and deep conceptual cognition forms the foundation for sound clinical reasoning. Evidence suggests that expeditious assessment methods, including multiple-choice questions, are not strongly correlated with depth and quality of learning (Trigwell & Prosser 1991; McNeil et al. 2006). Rather, it seems that the learner's approach to learning is better correlated with desired learning outcomes—a "surface level" approach leads to retention of facts but leads away from complexity, whereas a "deep approach" promotes higher complexity in knowledge and skill outcomes (Biggs & Collis 1982). Furthermore, immediate and regular formative assessment encourages students to engage sufficiently to attain a deeper understanding of the subject matter (Biggs & Collis 1982). Accordingly, our virtual case engages complex clinical reasoning processes while providing real-time expert feedback.

Previous authors have utilized reflective practice in teaching principles of clinical reasoning and cognitive bias (Croskerry et al. 2013). Such efforts are premised on the belief that clinical cognition is a skill in its own right demanding dedicated teaching strategies, a belief shared by the current authors (Elstein et al. 1990; Croskerry 2002). The most effective teaching strategies, however, remain unknown. Our virtual case explores the use of online, group-based, facilitated reflection, that is, probabilistic answers and group comparison. Our results argue for both the feasibility and value of our teaching methodology as a supplement to didactic and small group techniques.

Problem-based learning (PBL) interventions are now commonplace in preclinical medical curricula. PBL incurs heavy demands on teaching faculty and suffers from

Table 1. Theme analysis of student comments ($N = 225$).

Theme	Description	Example
Overall case design ($n = 163$)	Overall enjoyment of the case and a desire for more virtual cases of this type in the curriculum	"This was the best case study I have done in medical school so far! I think Problem Based Learning should be replaced or supplemented with this type of intelligent comparison and analysis of decision making processes!"
Clinical realism ($n = 71$)	The unique clinical realism of the case and the stimulation of clinical reasoning cognitive processes	"This case perfectly simulated the real life case scenarios. It allowed me to think like a doctor as I thought of the pertinent positives and negatives to aid my differential diagnosis, perform adequate physical exam, order appropriate investigations and re-adjusting my diagnosis".
Comparative data (Comparisons) ($n = 61$)	The value of comparative data between experts, peers, and the individual learner	"I really liked seeing what other peers and doctors put as answers it really showed how hard it is to be accurate and how your most likely diagnosis can change throughout the history, physical and labs."
Continuous response scales (Slide bars) ($n = 12$)	The unique value of continuous variable response scales as opposed to multiple choice questions (slide-bar likelihood estimates)	"This was an excellent case to work through. It is a good portrayal of how clinicians work through cases, and is a big step up from the usual multiple-choice questions that cannot realistically show a student how to fully think through a patient presentation."
Formative assessment ($n = 38$)	The value of a formative rather than summative assessment with immediate response feedback and answer explanations	"This is a great case to go through, especially when the 'stakes are low,' that is, for completion, because you are less worried about getting the right answer, and think more about what is going on."
Benefits of technology and user interface ($n = 39$)	The value of the virtual EMR and multi-media case presentation	"[the case] was straight forward and had a logical flow. I thought the application was very user friendly and was easy to figure out how to move between screens and figure out the EMR."
Challenges with technology and user interface ($n = 71$)	Difficulties with technical aspects of the software (e.g. hardware compatibility, interface design, case navigation)	"The EMR interface was glitchy for me. I couldn't access the end of the radiologic studies. I also couldn't find out how to input History and Physical Exam notes prior to finding the EMR."

variability in faculty aptitude for PBL's unique demands of learning facilitators (Norman et al. 2000; Biggs & Tang, 2011). Substituting more efficient and consistent formative learning interventions for at least some PBL sessions could improve the learners' experience without further burdening faculty. By providing a consistent virtual patient as context for SCT in an immersive online medium, we were able to manage complex response data and feedback for hundreds of students simultaneously, vastly reducing PBL facilitator demands. Our virtual case provided classwide standardization of the depth and breadth of topics covered, something difficult to achieve in small-group sessions. Interestingly, a number of students commented that the experience was the best "PBL" they had been exposed to. Many respondents suggested that more online SCT cases could replace a number of small group learning sessions.

Traditional SCT mediums are paper-based and rely on Likert scale response scales (Charlin et al. 2000; Lubarsky et al. 2011). Our SCT adaptation introduces continuous variable probability estimates and SCT integration into the more complex story of a fully realized virtual patient. The remarkably positive reception to this mix, consistent across cohort years, suggests that the script concordance model can be adapted to a wide range of clinical case contexts. It is possible that computer-assisted SCT implementations can confer advantages not previously contemplated for conventional SCT examinations. Our ability to accrue real-time data and present comparative analytics for different levels of learners proved very positive and is not possible in a traditional classroom setting. Future directions will explore the possibility of free public online access to SCT-based virtual cases through PedsCases.com, a well-established medical education website already developed and popularized by our research group (Gill et al. 2010; Kitney et al. 2014).

Our SCT adaptations generate need for new directions in clinical assessment research. We observed consistent cohort-to-cohort responses when probability sliders were used to communicate estimates of the likelihood of diagnoses. Yet at present, there is no validated SCT scoring method that can be applied to continuous response probability estimates (slider bars). Correlations between validated SCT performance measures and clinically important performance outcomes must also be further explored. In addition, longitudinal studies that track cohorts as they progress through clinical training are needed to confirm whether exposure to probabilistic learning interventions, and individual performance, correlate with superior clinical reasoning once training is complete.

The response patterns observed in our study suggest hypotheses for further testing with a larger expert cohort and cohorts of clinical learners at other stages of development (e.g. graduating students, residents, clinical fellows, generalists, and other specialists). We observed that pre-clinical students were generally more moderate (with less spread between possibilities) in their probability estimates than experts. This observation is consistent with students demonstrating a higher degree of uncertainty and lower confidence in their final diagnoses. Experts closed in on preferred diagnostic possibilities (with narrower confidence intervals) earlier than students.

An exception was rating the likelihood of "infection" as a diagnosis, the response with the widest variation for both students and experts, which may suggest that the experts themselves were in disagreement. Are experts and students uncertain for the same reasons? What does this disagreement tell us about the meaning of expert consensus and the challenges associated with teaching "xpertise"? It is possible that the variation in estimates signals ambiguity in the case data. This ambiguity may have resulted from an unrealistic virtual simulation, poor representation of case data, or other factors inherent to our particular case design. Alternatively, the virtual patient data may in fact mirror real-world clinical ambiguity. Future investigation should elucidate which of these possibilities is most likely.

Overall, the mean-likelihood estimates for experts and students exhibit unique distributions and the answer profile is reproducible across large groups of medical students. As our sample sizes for other levels of learners increase, the "clinically representative" answer for each question will become an increasingly valid reflection of the process of good clinical reasoning.

It has been observed in traditional SCT formats that responders (neither expert nor novice) are unlikely to choose answers at the extremes of the response scale, which may make the test vulnerable to exploitation by test takers (See et al. 2014). The present study replicates this finding but the effect is stronger in novices than the experts. The case design also guards against exploitation. Choosing moderate responses conferred no clear advantage—as the case progressed the diagnosis became more clear and the expert likelihood estimates grouped to the extremes of the answer scales.

Other limitations of the traditional SCT method include a lack of concurrent patient management, communication demands, and multistep problem solving. Our case addresses each of these drawbacks. The entire case represents a multistep patient management challenge including communication with an expert physician. It is, however, difficult to replicate the real-world communication nuances of clinical medicine in an online medium. This represents an area for future improvement.

Limitations

Our comparative quantitative analysis is limited by our small sample size of clinical experts. Further, most responders commented positively on the value of comparative data but did not distinguish between peer versus expert comparison. Delineating the value of peer responses represents an area for future study.

We also struggled with the uncontrolled conditions of learner-controlled, asynchronous, online sessions. Computer types, operating systems, internet browsers, and network speeds were highly variable which affected the quality of physical exam and EMR simulations for some learners. We are currently working to improve consistency amongst the previously mentioned factors.

Creating an SCT assessment embedded in an online virtual patient with physical examination and EMR simulation capabilities required significant effort. Once built, subsequent adjustments to the case were easy to deploy and similar cases could be authored at much less effort. Operational challenges included obtaining a sufficiently large panel of experts to complete the case. Other studies suggest that a panel of at least 10 experts is sufficient to achieve reliability. This number has not been widely replicated and the true minimum is uncertain (Lubarsky et al. 2011). In our study, we were only able to recruit 8 expert physicians. With increasing popularity and awareness of script-based cases, recruitment will become easier. Providing some incentive for participation, such as CME credit, might also improve participation rates.

Overall, conclusions drawn from this single-center virtual case implementation are limited and results must be replicated with other virtual cases comprising varied clinical scenarios presented to different peer and expert groups.

Conclusions

We describe a successful implementation of a novel educational tool for teaching and evaluating clinical reasoning in an online environment empowered for script concordance testing, clinical simulation and electronic medical record emulation. Our experience suggests both educational and practical benefits associated with distribution of a high-quality learning experience to large groups. Future initiatives will aim to expand the use of this tool and improve our understanding of the data it generates.

Glossary

Virtual patient: The representation of a patient in an interactive computer simulation of a real-life scenario.
Harden, R.M., and Laidlaw, J.M. (2012) Essential Skills for a Medical Teacher. Edinburgh: Churchill Livingstone.

Script theory: Script theory is a psychological theory which posits that human behaviour largely falls into patterns called "scripts" because they function analogously to the way a written script does, by providing a program for action.

Acknowledgements

Authors thank the University of Alberta Medicine Classes of 2015 and 2016 for partaking in the study, and also Drs. Maury Pinsk, Karen Forbes, Janet Ellsworth, Jennifer Walton, Samina Ali for their participation and expertise. They also thank Dr. Mara Tietzen for her research assistance.

They also extend thanks to the University of Alberta Faculty of Medicine and Dentistry for their Summer Studentship funding which helped make this project possible.

(Special thanks to Dr. Robert Hayward and Qwogo Inc. for their assistance in developing the online case and accruing data).

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

Notes on contributors

Jake Hayward, BSc, MD, is a second year emergency medicine resident at the University of Alberta.

Amandy Cheung, BSc, MD, is a second year family medicine resident at the University of Alberta.

Alkarim Velji, BSc, BEd, MD, is a first year emergency medicine resident at the University of Manitoba.

Jenny Altarejos, is the Chief Technology Officer at Qwogo Inc.

Peter Gill, BMSc, MSc, DPhil, MD, is a paediatric resident at the University of Toronto.

Andrew Scarfe, MD, FRCP, is an Associate Professor of medical oncology at the University of Alberta.

Melanie Lewis, BN, MD, MMedED, FRCP, is the Associate Dean of Learner Advocacy and Wellness at the University of Alberta.

References

Biggs JB, Tang C. 2011. Teaching for quality learning at university: what the student does. United Kingdom: McGraw-Hill Education.

Biggs JB, Collis KF. 1982. Evaluating the quality of learning: the SOLO taxonomy. New York: Academic Press.

Bowen JL. 2006. Educational strategies to promote clinical diagnostic reasoning. N Engl J Med. 355:2217–2225.

Campbell SG, Croskerry P, Bond WF. 2007. Profiles in patient safety: a "perfect storm" in the emergency department. Acad Emerg Med. 14:743–749.

Charlin B, van der Vleuten C. 2004. Standardized assessment of reasoning in contexts of uncertainty the script concordance approach. Eval Health Prof. 27:304–319.

Charlin B, Roy L, Brailovsky C, Goulet F, van der Vleuten C. 2000. The script concordance test: a tool to assess the reflective clinician. Teach Learn Med. 12:189–195.

Charlin B, Tardif J, Boshuizen HP. 2000. Scripts and medical diagnostic knowledge: theory and applications for clinical reasoning instruction and research. Academic Medicine. 75:182–190.

Croskerry P. 2002. Cognitive forcing strategies in emergency medicine. Emerg Med J. 19:A9.

Croskerry P, Singhal G, Mamede S. 2013. Cognitive debiasing 2: impediments to and strategies for change. BMJ Qual Saf. 22(Suppl 2). doi:10.1136/bmjqs-2012-001713.

Elstein AS, Shulman LS, Sprafka SA. 1990. Medical problem solving a ten-year retrospective. Eval Health Prof. 13:5–36.

Gagnon R, Charlin B, Coletti M, Sauvé E, Van der Vleuten C. 2005. Assessment in the context of uncertainty: how many members are needed on the panel of reference of a script concordance test? Med Educ. 39:284–291.

Gill P, Kitney L, Kozan D, Lewis M. 2010. Online learning in pediatrics: a student led web-based learning modality. Clin Teach. 7:53–57.

Graber ML, Franklin N, Gordon R. 2005. Diagnostic error in internal medicine. Arch Intern Med. 165:1493–1499.

Kitney L, Gill P, MacPherson P, North R, Gerdung C, Lewis M. 2014. PedsCases for medical students studying paediatrics: a prospective study. UAHS J. 10:48–52.

Leape LL, Brennan TA, Laird N, Lawthers AG, Localio AR, Barnes BA, Hebert L, Newhouse JP, Hiatt H. 1991. The nature of adverse events in hospitalized patients: results of the Harvard Medical Practice Study II. N Engl J Med. 324:377–384.

Lubarsky S, Charlin B, Cook DA, Chalk C, van der Vleuten CP. 2011. Script concordance testing: a review of published validity evidence. Med Educ. 45:329–338.

McNeil HP, Hughes CS, Toohey SM, Dowton SB. 2006. An innovative outcomes-based medical education program built on adult learning principles. Med Teach. 28:527–534.

Nandi PL, Chan JN, Chan CP, Chan P, Chan LP. 2000. Undergraduate medical education: comparison of problem-based learning and conventional teaching. Hong Kong Med J. 6:301–306.

Newman-Toker DE, Pronovost PJ. 2009. Diagnostic errors—the next frontier for patient safety. JAMA. 301:1060–1062.

Norman GR, Henk GS. 2000, Effectiveness of problem based learning curricula: theory, practice and paper darts. Med Educ. 34:721–728.

See KC, Tan KL, Lim TK. 2014. The script concordance test for clinical reasoning: re-examining its utility and potential weakness. Med Educ. 48:1069–1077.

Schiff GD, Hasan O, Kim S, Abrams R, Cosby K, Lambert BL, McNutt RA. 2009. Diagnostic error in medicine: analysis of 583 physician-reported errors. Arch Intern Med. 169:1881–1887.

Shojania KG, Burton EC, McDonald KM, Goldman L. 2003. Changes in rates of autopsy-detected diagnostic errors over time: a systematic review. JAMA. 289:2849–2856.

Singh H, Thomas EJ, Khan MM, Petersen LA. 2007. Identifying diagnostic errors in primary care using an electronic screening algorithm. Arch Intern Med. 167:302–308.

Trigwell K, Prosser M. 1991. Relating approaches to study and quality of learning outcomes at the course level. Brit J Educ Psychol. 61: 265–275.

Zwaan L, de Bruijne M, Wagner C, Thijs A, Smits M, van der Wal G, Timmermans DR. 2010. Patient record review of the incidence, consequences, and causes of diagnostic adverse events. Arch Intern Med. 170:1015–1021.