Acquiring Clinical Reasoning Competency: Group versus Individual Practice of Virtual Patient Simulations

DANE M. CHAPMAN, MD, PhD[1,2,3]
JUDITH G. CALHOUN, PhD, MBA[1]
ADRIAN P. VAN MONDFRANS, PhD[2]
WAYNE K. DAVIS, PhD[1]

Institutional Affiliations:

1. *University of Michigan Medical School, Department of Medical Education, Ann Arbor, MI*
2. *Brigham Young University, McKay School of Education, Department of Instructional Psychology and Technology, Provo, UT*
3. *University of Missouri-Columbia, School of Medicine, Center for Clinical Reasoning and Procedural Competency, Department of Emergency Medicine, Columbia, MO*

Abstract

Background:
It is unknown whether group or individual practice using free-inquiry virtual patient simulations would most facilitate acquisition of clinical reasoning skills required of competent physicians.

Objective:
To determine the effect of virtual patient simulation group practice on clinical reasoning competency.

Methods:
We used an experimental, pretest-posttest, control group design. Ninety-seven of 191 post-second year medical students were randomly assigned to group practice or individual practice and solved six virtual patient simulations: one pretest (individual), three practice (group or individual), and two posttest (individual) simulations. Multivariate ANOVA and univariate ANCOVA statistics were used to compare groups.

Results:
Computer transcripts (N=486) were generated by 81 post-second year medical students with complete data. Compared to individual-practice students (n=41), group-practice students (n=40) performed as well as or better on 18 of 19 clinical reasoning proficiency measures, demonstrated greater overall clinical reasoning proficiency, indicated more diagnostic hypotheses and used more focused inquiries. Individual-practice students were more efficient in making critical physical examination inquiries.

Conclusions:
Instructional effectiveness was established for both individual and group virtual patient simulation practice, with a combined group practice and virtual patient simulation effect in promoting clinical reasoning competency.

Keywords:
virtual patient simulation, group practice, clinical reasoning, clinical decision-making, medical students

Introduction

In our current healthcare climate, physicians are rewarded for being sufficiently thorough, yet efficient in data gathering and for using problem-solving strategies that limit diagnostic tests and commensurate costs, but still promote maximal diagnostic proficiency. These critical clinical reasoning (CR) skills (proficiency, efficiency, thoroughness, and strategy) are neither adequately mastered nor measured in medical schools and residency programs.[1-3] CR involves both data gathering (i.e. history taking, physical examination, and the selection and interpretation of appropriate diagnostic tests) and diagnostic and therapeutic decision-making skills.[3-4] Cognitive models are needed to teach and assess data-gathering and decision-making competencies,[5-6] preferably earlier in medical school since the organization of clinical knowledge and the directionality of CR acquired during medical school carries over into subsequent resident and physician performance.[5-9] Knowledge organization and schema acquisition seem to be more important for CR expertise than the use of problem-solving methods, and their development over many years of experience requires exposure to many patient cases while emphasizing the association of disease-specific features, signs and symptoms.[7,9]

High-fidelity, free-inquiry virtual patient simulations (VPS) provide increasingly sophisticated opportunities to engage in virtual patient encounters, and have been implicated in teaching and testing cognitive CR skills.[10-13] Virtual patient simulations are considered "high-fidelity" if they closely simulate an actual physician-patient encounter, and are "free-inquiry" if they allow free questioning without menu or branching limitations. It seems likely that such simulations could provide trainees with the equivalent of "many years of patient care experience," and facilitate their knowledge organization and schema acquisition.[3,11-12]

While the cost of high-fidelity VPSs has limited their widespread use in training,[10] group study could facilitate their integration into medical school and residency curricula by requiring fewer start-up multimedia configurations. Small group study using the problem-based learning (PBL) format has been shown to improve CR competency as measured by United States Medical Licensing Examination (USMLE) and residency program director evaluations, but with greater financial and faculty resource costs.[8] Once developed, VPSs could markedly decrease dependance upon overburdened faculty and limited training resources.[13] However, it is unknown whether GP would facilitate or hinder CR acquisition using an interactive, free-inquiry VPS model. It is conceivable that VPS GP might dilute the frequency of individual student-simulation interactions resulting in an antagonistic rather than a synergistic effect. Accordingly, we conducted this study to determine whether free-inquiry VPS group (GP) versus individual (IP) practice would impact the development of CR competency. We hypothesized that: (1) VPS practice would enhance medical student CR competency, (2) CR competency would vary by VPS case content, and (3) GP students would do as well as or better than IP students on measures of CR competency.

Methods

Study Design:

Using an experimental, pretest-posttest, control group design, we assessed the effect of three hours of VPS GP versus IP practice (independent variable) upon nineteen previously established CR dependent outcome measures.[3]

Study Setting and Population:

The study qualified for internal review board (IRB) exemption as a curriculum innovation project and was conducted at the Taubman Health Sciences Library Learning Resource Center at the University of Michigan Medical School. Ninety-seven of 191 pre-clinical medical students participated in a one-week VPS elective during a required four-week PBL curriculum, with 23-25 participants being randomly assigned to each of the four weeks between the preclinical and clerkship years.

Virtual Patient Simulation Elective:

The 6.5 hour VPS elective included two morning sessions (3.0 and 3.5 hours) during which students worked through six free-inquiry virtual patient simulations: one 60-minute pretest as individuals (Case 1: Cardiology), three 60-minute GP or IP simulations with corrective feedback (Case 2: Pediatric Endocrinology, Case 3: Infectious Disease, and Case 4: Pulmonary), and two 45-minute posttests as individuals (Case 5: Pulmonary and Case 6: Cardiology).

Group Versus Individual Practice:

GP students worked in groups of three and were assigned to one of three roles which changed until each group member had experienced each role: (1) typist: typed group inquiries at the keyboard; (2) recorder: recorded times, type of inquiries, diagnostic hypotheses and likelihood rankings on a VPS log; and (3) chairperson: insured that all group members participated in making decisions, and cast the deciding vote if group members were indecisive. IP students worked through the three practice VPS alone and were responsible for typing inquiries at the keyboard and maintaining their own VPS log.

Virtual Patient Simulations:

Multi-problem, text-driven, network-based virtual patient simulations were selected for practice and assessment as they were the most sophisticated high-fidelity, free-inquiry simulations available at the time of the study.[3] Following an "opening scene," users assumed the role of physicians and moved to and from history, physical examination, diagnostic study, diagnosis and treatment sections without cueing or branching limitations. The VPS responded to user inquiries with questions or feedback in predictable ways using artificial intelligence.

Assessments and Procedure:

VPS transcripts (N=486) were generated by 81 medical students with complete data and documented student-simulation interactions for 243 hours of medical student practice, and 202 hours of assessment. Outcome performance scores along the nineteen predetermined CR indices were derived from 243 hard-copy VPS assessment transcripts (one pretest and two posttests) and were classified into one of four previously validated[3] clinical reasoning constructs: proficiency, efficiency, thoroughness, and strategy (See Table 1).

Clinical reasoning proficiency refers to how effectively critical data were gathered and correct diagnoses made. The dependent CR proficiency variables were: percent of critical data gathering inquiries obtained (history, exam, and diagnostic test proficiencies), percent of correct diagnoses made (diagnosis

TABLE 1: Mathematical Descriptions of Nineteen Clinical Reasoning Performance Indices Derived for Use in Multi-Problem Virtual Patient Simulations

Index	Abbreviation	Description[a]
Proficiency		
History Taking	HTP	(Obtained CHT/Total CHT) X 100
Physical Examination	PEP	(Obtained CPE/Total CPE) X 100
Diagnostic Tests	DTP	(Obtained CDT/Total CDT) X 100
Correct Diagnoses	DP	(Obtained CD/Total CD) X 100
Program Solving Index	PSI	(HTP + PEP + DTP + DP) / 4
Proficiency Index	PI	(Obtained CHT + CPE + CDT) X 100 / (Total CHT + CPE + CDT)
Efficiency		
History Taking	HTE	(CHT Obtained/HTT) X 100
Physical Examination	PEE	(CPE Obtained/PET) X 100
Diagnostic Tests	DTE	(CDT Obtained/DTT) X 100
Thoroughness		
History Taking	HTT	Total HT
Physical Examination	PET	Total PE
Diagnostic Tests	DTT	Total DT
Total Data-Gathering	TDG	(HTT + PET + DTT)
Diagnosis	DT	Total D
Strategy		
History Taking	HTS	[HTT/(HTT+PET + DTT)] X 100
Physical Examination	PES	[PET/(HTT+PET + DTT)] X 100
Diagnostic Tests	DTS	[DTT/(HTT+PET + DTT)] X 100
Focused Strategy Index	FSI	(HH + PP+ DD +1) / (HP + HD + PH + PD + DH + DP + 1)
Invasiveness/Cost Index	ICI	[DTT/(HTT + PET)] X 100

[a]*Symbol Key: HT*= history taking inquiries, *PE*= physical examination inquiries, *DT*= diagnostic test inquiries, *D*= diagnoses indicated, *C*= critical inquiry or diagnosis (e.g. CHT=critical history taking inquiries), *HH*= history to history transition, *PP*= physical exam to physical exam transition, *DD*= diagnostic test to diagnostic test, *HP*= history to physical exam, HD= history to diagnostic test, *PH*= physical exam to history, *PD*= physical exam to diagnostic test, *DH*= diagnostic test to history, and *DP*= diagnostic test to physical exam transition.

proficiency), Problem Solving Index–an average of data gathering and decision making proficiencies, and Proficiency Index–the percent of data-gathering critical information obtained.

Clinical reasoning efficiency was defined as the percentage of data gathering inquiries that were critical in making the diagnosis of a patient's problem(s). Higher scores represent greater efficiency in making medical inquiries. Clinical reasoning efficiency variables included history, physical examination and diagnostic test efficiencies.

Clinical reasoning thoroughness reflects the frequency of data gathering inquiries made or diagnoses indicated. Clinical reasoning thoroughness variables included: total number of history inquires (history thoroughness), physical examination inquiries (physical examination thoroughness), and diagnostic test inquiries (diagnostic test thoroughness); total number of history, physical examination and diagnostic test inquiries combined (total data gathering thoroughness); and total number of diagnoses hypothesized at the completion of each simulated case (diagnosis thoroughness).

Clinical reasoning strategy refers to the cognitive strategies used to arrive at correct diagnoses; and, reflects individual preference for certain data gathering techniques (e.g. to use either a focused inquiry approach or a "shot gun" or haphazard approach). CR strategy indices included: percent of total data-gathering inquiries that relate to history taking (history strategy), physical examination (physical examination strategy), or diagnostic test (diagnostic test strategy); Focused Strategy Index–the standardized proportion of data gathering inquiry transitions of similar type (e.g. history to history) to all other combinations of possible inquiry transitions from one type of inquiry to another (e.g. history to physical examination, diagnostic test to history, etc), where high scores reflect a more focused and systematic data-gathering approach; and Invasiveness/Cost Index–the standardized proportion of diagnostic test inquiries (relatively invasive and costly) to the sum of history taking and physical examination inquiries (relatively non-invasive and less costly), where higher scores reflect a more invasive and costly data-gathering approach.

Data Analysis:

BMDP[14] multivariate factorial, repeated-measures ANOVA statistics were used to determine any overall effect of GP/IP practice (independent variable) upon the four CR performance constructs. When pretest differences were found between treatment groups, then analyses of covariance (ANCOVA) were utilized to adjust for pretreatment differences with the pretest treated as the covariate. If a significant multivariate effect was observed foir a CR construct, then univariate ANOVA or ANCOVA statistics were used to test for CR index differences among treatment groups.

Results

The study sample (N=97) appeared to be representative of the entire medical school class (N=191) as VPS students did not differ significantly from other class members on ethnicity, sex, prior clinical experience on the hospital wards, or on PBL CR assessments. Complete hard-copy VPS assessment transcripts (n=243) were obtained for 81 study participants: GP(n=40) and IP(n=41).

VPS Practice. Significant pretest-posttest differences were found for both treatment groups (GP/IP) and suggest the utility of three hours of high-fidelity VPS practice in teaching selective CR skills (See Table 2).

VPS Case Content. Multivariate analyses of posttests (Case 5

TABLE 2: Group Practice (GP) and Individual Practice (IP) Pretest and Posttest Mean Scores (SD) for Nineteen Clinical Reasoning Performance Measures[a]

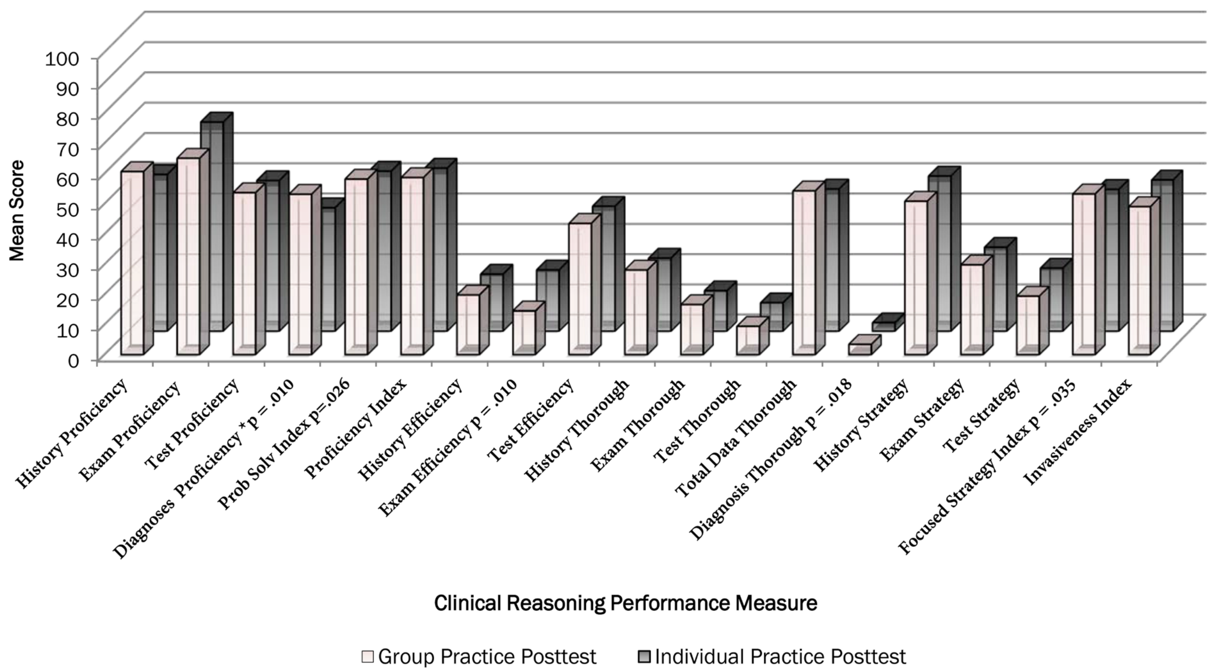
Index	Case 1 Pretest (Cardiology)		Case 5 Posttest (Pulmonary)		Case 6 Posttest (Cardiology)	
	GP	IP	GP	IP	GP	IP
Proficiency						
History Taking	50.0 (24.9)	40.2 (23.0)	60.6 (15.6)[c]	51.8 (21.2)[c]	51.0 (16.5)	47.1 (13.5)
Physical Examination	57.5 (22.1)	44.5 (22.0)	65.0 (18.4)[c]	69.1 (17.3)[c]	62.0 (19.1)[c]	62.4 (20.0)[c]
Diagnostic Tests	62.5 (24.7)	54.9 (26.9)	53.6 (13.2)[c]	49.8 (11.1)[c]	53.6 (15.5)[c]	51.5 (17.9)[c]
Correct Diagnoses	38.3 (28.8)	35.8 (24.0)	53.1 (22.8)[b,c]	40.8 (18.3)[b,c]	32.5 (18.0)[c]	28.0 (17.0)[c]
Program Solving Index	52.1 (15.0)	43.9 (13.9)	58.1 (9.3)[b,c]	52.9 (10.9)[b,c]	49.8 (9.4)	47.3 (9.9)
Proficiency Index	56.7 (15.8)	46.5 (15.0)	58.6 (9.2)[c]	53.9 (11.8)[c]	54.3 (10.1)	51.9 (10.4)
Efficiency						
History Taking	7.8 (4.0)	8.5 (5.8)	19.8 (8.1)[c]	18.9 (8.9)[c]	24.5 (11.1)[c]	25.7 (10.7)[c]
Physical Examination	17.9 (9.6)	17.5 (7.4)	14.5 (7.9)[b]	20.3 (11.4)[b]	21.0 (8.8)[c]	22.8 (10.3)[c]
Diagnostic Tests	44.8 (25.4)	26.3 (15.0)	43.5 (13.8)[c]	41.3 (18.2)[c]	36.5 (9.5)	35.6 (16.7)
Thoroughness						
History Taking	26.4 (10.2)	21.0 (8.5)	28.1 (12.4)[c]	24.2 (9.1)[c]	23.6 (11.1)	21.2 (9.0)
Physical Examination	13.6 (5.3)	10.9 (5.4)	16.6 (8.0)[c]	13.4 (6.6)[c]	16.3 (6.1)[c]	15.4 (6.4)[c]
Diagnostic Tests	7.0 (3.8)	9.8 (4.9)	9.4 (3.4)[c]	9.4 (4.0)[c]	13.8 (4.5)[c]	15.2 (7.8)[c]
Total Data-Gathering	47.1 (12.7)	41.7 (12.4)	54.2 (16.6)[c]	47.0 (15.2)[c]	53.8 (12.7)[c]	51.8 (15.6)[c]
Diagnosis	3.0 (1.2)	3.1 (1.1)	3.4 (1.2)[b]	2.8 (1.0)[b]	4.6 (1.5)[c]	5.1 (2.1)[c]
Strategy						
History Taking	55.7 (12.4)	49.7 (11.8)	50.8 (10.6)	51.3 (9.3)	42.0 (13.9)[c]	40.6 (13.3)[c]
Physical Examination	28.9 (8.2)	26.1 (9.3)	29.8 (8.3)	27.8 (8.8)	30.4 (9.4)[c]	30.1 (10.0)[c]
Diagnostic Tests	15.4 (8.9)	24.2 (11.2)	19.4 (9.5)	20.9 (7.7)	27.5 (12.1)[c]	29.3 (11.4)[c]
Focused Strategy Index	53.3 (12.2)	48.3 (7.6)	53.2 (12.3)[b]	46.9 (6.2)[b]	51.4 (9.0)	48.3 (10.7)
Invasiveness Index	46.0 (6.5)	53.1 (10.8)	49.0 (9.3)	50.0 (7.4)	49.6 (10.2)	50.7 (9.9)

[a] Repeated-Measures, Factorial ANCOVA for Pretest (Case 1) and Posttest (Case 5 and Case 6) comparisons for 81 medical students (GP, n=40; IP, n=41) with complete data.

[b] ANCOVA, F-Test, $p < .050$, for significant GP (n=40) versus IP (n=41) differences.

[c] ANCOVA, F-Test, $p < .050$, for significant pretest-posttest differences (N=81).

FIGURE 1: Mean Scores on Nineteen Case 5 Clinical Reasoning Measures after Three Hours of Group or Individual Practice Using Virtual Patient Simulations in Post-Second Year Medical Students (N=81) (*ANCOVA, F test)



and Case 6) revealed significant GP/IP differences only for the posttest with content similar to that encountered during a practice simulation (Case 5), and were related to CR thoroughness (Wilks' λ=12.1, $F(4, 62)=2.89$; $p =.03$), and strategy constructs (Wilks' λ=14.2, $F(5, 61)=2.67$; $p = .03$), although differences between treatment groups approached significance along the efficiency construct as well (Wilks' λ=8.40, $F(3, 63)=2.71$; $p = .052$) (See Table 2). There were no multivariate GP/IP differences on the posttest (Case 6) with content not previously encountered on a practice simulation.

Group Versus Individual Practice. GP students performed as well as or better than IP students on 18 of 19 CR competency measures (See Table 2; Figure 1). Despite randomization into GP/IP treatment groups, multivariate analyses revealed pretest GP/IP differences across clinical reasoning proficiency (Wilks' λ=15.0, $F(6, 60)=2.31$; $p = .04$), efficiency (Wilks' λ=14.6, $F(3, 63)=4.71$; $p = .00$), thoroughness (Wilks' λ=16.5, $F(4, 62)=3.93$; $p = .01$), and strategy measures (Wilks' λ=16.4, $F(5, 61)=3.08$; $p = .02$).

When covarying on the pretest, univariate ANCOVA of the Case 5 posttest detected significantly higher GP versus IP proficiency scores along Diagnosis Proficiency ($F(1, 76)=7.06$, $p = .010$) and the Problem-solving Index ($F(1, 76)=5.17$, $p = .026$), but not along any of the indices measuring only data gathering proficiency. Univariate ANCOVA of Case 5 efficiency scores detected a significant decrease in GP physical examination efficiency compared to IP students ($F(1, 76)=7.03$, $p = .010$). Significant differences in GP versus IP Case 5 thoroughness scores were found only for Diagnosis Thoroughness ($F(1, 76)=5.85$, $p = .018$), but not for any of the data gathering thoroughness measures. Univariate ANCOVA of Case 5 strategy measures revealed a significant tendency for GP students to use a more focused, and less haphazard data gathering strategy compared to IP students ($F(1, 76)=4.59$, $p = .035$) (See Figure 1). When covarying on the pretest, no significant GP/IP treatment effects were observed for the second posttest (Case 6) with content not previously encountered on a practice simulation (See Table 2).

Discussion

The results of this study confirm the efficacy of both group and individual practice in teaching CR skills and demonstrate that free-inquiry VPSs can be successfully implemented into the medical school curriculum. Both GP and IP students acquired selective CR skills after only three hours of free-inquiry VPS practice. That GP is at least as effective as IP in teaching all but physical exam efficiency, and superior to IP in teaching more focused data gathering and more elaborate hypothesis generation are important findings. As expected, GP/IP treatment differences were found only with case content encountered during VPS practice.

GP students acquired CR skills beyond that expected from interacting with the VPS alone. GP students were able to use their collective knowledge on a practice case of similar content, which carried over to their more focused data gathering and to greater diagnostic hypotheses generated as individuals on the assessment case of similar content (Case 5 Posttest: Pulmonary). Presumably as a result of generating more hypotheses,[15] GP students were also more likely to make correct diagnoses; and this may be the single greatest impact of VPS GP upon acquisition of CR competency.

Group use of VPSs would greatly reduce initial VPS purchasing and upkeep costs when implementing VPSs into the medical school curriculum. The challenge of doing more with less, with greater training expectations and reduced training resources, has greatly impacted health professions education.[13,16] Combining virtual patients and small group study using computer-based clinical scenarios, web-based and otherwise may help to reconcile this seeming paradox of better training with fewer resources.[13,17,18]

Since students were randomly assigned to treatment groups, it is unlikely that GP/IP pretest differences along all four CR constructs were due to chance alone. What, then could account for these differences? One tenable explanation is that a GP/IP treatment effect occurred prior to the pretest. An indirect GP treatment could have occurred as students were informed of their GP/IP treatment assignments several days before the pretest assessment. They were not, however, given the identity of the other group members. It is possible that students assigned to the GP treatment were more motivated to learn CR skills in anticipation of performing with peers.

This study has limitations. It was conducted nearly three decdes ago as part of a PhD dissertation,[19] and was never formally published. While study findings have become relevant with developments in VPS training and assessment, VPSs have changed in some ways that might impact study results. Each generation of VPSs have become more sophisticated with patient scenario video clips, actual EKGs and radiographs requiring user interpretation, and more advanced scoring, data storage, retrieval and web-based capabilities. In a study of web-based VPSs, users found demonstrations of physical exam abnormalities in heart or lung sounds, skin lesions, and neurological findings quite helpful.[20] Still, similar or more concerning limitations exist in OSCE-based and USMLE® step 3® computer-based exams and in live simulated patients.[2] The essential VPS components required for teaching and assessing clinical reasoning have remained the same: free inquiry, high fidelity, no cueing or branching limitations, and artificial intelligence interactive capabilities that require users to indicate history, exam and laboratory inquiries. Finally, the study was conducted at a single traditional medical school with post-second-year, pre-clinical medical students. It is uncertain how the results would generalize to clinical medical students, residents, or physicians in continuing medical education. In this regard, web-based VPSs have been found to have greater acceptance among pre-clinical second-year medical students compared to clinical fourth-year students.[20]

In summary, pre-clinical medical students assigned to VPS GP performed as well as or better than those assigned to VPS IP on 18 of 19 CR competency measures. This is an important finding since GP requires fewer VPS training resources, and would thereby facilitate implementing virtual patient simulations into the medical school curriculum. Moreover, these results suggest that VPS and GP had a combined or additive facilitating effect upon student acquisition of CR skills. The indirect treatment effect of prior knowledge of being assigned to work with peers is a new finding and has many implications for motivating students to acquire CR competency.

Received for publication August 13, 2012. Revisions received January 15, 2013. Accepted January 19, 2013. Available online February 1, 2013.

Presented in part as an "Innovations in Medical Education Exhibit" at the 108th annual meeting of the Association of American Medical Colleges (AAMC), Washington, DC, November 2-4, 1997; and at the 27th annual meeting of the Society for Academic Emergency Medicine (SAEM), Washington DC, May 22, 1997.

Contributions of Authors: Conception (DMC, JGC, APV, WKD), design (DMC, JGC), analysis and interpretation (DMC, JGC), drafting (DMC, JGC, WKD), and revising (DMC) the article. DMC takes responsibility for the paper as a whole.

Acknowledgements: The authors are indebted to Robert H. Bartlett, MD and Richard D. Judge, MD, course directors of the Problem-Based Learning Curriculum, for their enthusiastic endorsement of the medical problem solving computer simulation elective; the ninety-seven University of Michigan medical students who enrolled in the elective; the Taubman Medical Library Learning Resource Center staff for their essential role in making the computer simulation elective an overwhelming success; Fredric M. Wolf, PhD; Elaine M. Hockman, PhD; Nancy P. Allen, PhD; C. Victor Bunderson, PhD; Russell T. Osgethorpe, PhD; and Sally H. Cavanaugh, PhD for their insightful review and recommendations to improve this work; and to Becky Bluett and Christine Downs for their assistance in preparing the manuscript.

Funding: In part by a University of Michigan curriculum innovation grant.

Conflicts of Interest: No personal, commercial, political, academic or financial conflict of interest was reported by any of the authors. DMC is editor-in-chief of JCRPC.

Ethical Approval: Authors report the study qualified for IRB exemption.

Correspondence: Dane M. Chapman, MD, PhD, Director, Center for Clinical Reasoning and Procedural Competency, Department of Emergency Medicine, University of Missouri-Columbia School of Medicine, 1 Hospital Drive, Columbia MO, 65201, e-mail: <chapmandan@health.missouri.edu>

References

1. AAMC Report of the Ad Hoc Committee of Deans. Educating doctors to provide high quality medical care. 2004: 1-12. Available from: https://members.aamc.org/eweb/upload/Educating%20Doctors%20to%20Provide%20July%202004.pdf.
2. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA*. 2002;287(2):226-35.
3. Chapman DM, Calhoun JG, Van Mondfrans AP, Davis WK. Measuring clinical reasoning competency using a virtual patient model. *JCRPC*. 2013;1(1):1-9
4. Chapman DM, Char DM, Aubin C. Clinical decision making. Chapter 10 in: Marx JA et al., editors. *Rosen's Emergency Medicine, Concepts and Clinical Practice* 6th Edition. Philadelphia, PA: Mosby/Elsevier; 2006:125-33.
5. Norman GR. The epistemology of clinical reasoning: Perspectives from philosophy, psychology, and neuroscience. *Acad Med.* 2000;75(10):S127-33.
6. Elstein AS. Clinical problem solving and decision psychology: Comment on "The epistemology of clinical reasoning." *Acad Med*. 2002;75(10):S134-36.
7. Elstein AS, Schwarz A. Clinical problem solving and diagnostic decision making: Selective review of the cognitive literature. *BMJ* 2002;324:729-32.
8. Hoffman K, Hosokawa M, Blake R, Headrick L, Johnson G. Problem-based learning outcomes: Ten years of experience at the University of Missouri-Columbia School of Medicine. *Acad Med*. 2006;81:617-625.
9. Patel VL, Arocha JF, Leccisi, MS. Impact of undergraduate medical training on housestaff problem-solving performance: Implications for problem-based curricula. *J Dent Educ.* 2001;65(11):1199-1218.
10. Friedman CP. The marvelous medical education machine or how medical education can be unstuck in time. *Acad Med*. 2000;75(10):S137-42.
11. Stevens SM, Goldsmith TE, Summers KL et al. Virtual reality training improves students' knowledge structures of medical concepts. In: Westwood JD et al., editors. *Medicine Meets Virtual Reality* 13. Amsterdam: IOS Press; 2005:519-25.
12. Jarrell BE. Simulation for teaching decision making in medicine: The next step. Presented at Medicine Meets Virtual Reality 13, Long Beach, CA, Jan 28, 2005.
13. Cook DA, Triola MM. Virtual patients: A critical literature review and proposed next steps. *Med Educ*. 2009;43(4):303-11.
14. BMDP Statistical Software, Inc. 1964 Westwood Blvd, Suit 202, Los Angeles, CA. Available from: http://www.statistical-solutions-software.com/bmdp-statistical-software/bmdp/.
15. Gruppen LD, Palchik NS, Wolf FM, et al. Medical student use of history and physical information in diagnostic reasoning. *Arthritis Care Res*. 1993;6(2):64-70.
16. Chapman DM, Hayden S, Sanders AB, Binder LS, Chinnis A, Corrigan K, LaDuca T, Dyne P, Perina DG, Smith-Coggins R, Sulton L, Swing S. Integrating the Accreditation Council for Graduate Medical Education core competencies into the Model of the Clinical Practice of Emergency Medicine. *Ann Emerg Med.* 2004;33(6):756-769.
17. Eshach H, Bitterman H. From case-based reasoning to problem-based learning. *Acad Med.* 2003;78(5):491-96.
18. Tärnvik A. Revival of the case method: A way to retain student-centred learning in a post-PBL era. *Med Teach*. 2007;29:32–36.
19. Chapman DM. Teaching and evaluating clinical reasoning through computer-based patient management simulations. *Dissertation Abstracts International.* 1985;46:784-B.
20. Gesundheit N, Brutlag P, Youngblood P, et al. The use of virtual patients to assess the clinical skills and reasoning of medical students: Initial insights on student acceptance. *Med Teach.* 2009;31(8): 739-42.