Original Publication

OPEN ACCESS

Clinical Reasoning Workshop: Cervical Spine and Shoulder Disorders

Alex Moroz, MD*

*Corresponding author: alex.moroz@nyumc.org

Citation: Moroz A. Clinical reasoning workshop: cervical spine and shoulder disorders. *MedEdPORTAL*. 2017;13:10560.
https://doi.org/10.15766/mep_2374-8265.10560

Copyright: © 2017 Moroz. This is an open-access publication distributed under the terms of the Creative Commons Attribution-NonCommercial-Share Alike license.

Abstract

Introduction: Information processing and cognitive factors may be a cause of physician diagnostic errors. While the conceptual framework of dual processing in clinical reasoning is widely accepted, how can residents be taught to switch between automatic and reflective modes, and will doing so improve their decision making? Developing effective clinical reasoning habits while in training may facilitate progression to expertise, reduce diagnostic errors, and improve patient safety. **Methods:** This workshop allows residents to practice engaging in and toggling between both modes of information processing using exemplar musculoskeletal vignettes. Originally implemented with a group of 26 physical medicine and rehabilitation residents, the workshop includes pre- and posttests, small-group learning, and a small-group competition. **Results:** Posttest scores improved on pretest scores. In an online session evaluation, residents indicated they liked the workshop and thought it improved their diagnostic ability. **Discussion:** This workshop, which includes team- and case-based learning, key features assessment, dual processing theory, and gamification, was effective in engaging residents and resulted in high resident satisfaction and perception of increased ability to tackle clinical problems. Faculty time required was moderate after the initial setup, which in our case primarily involved uploading content into an online learning management system.

Keywords

Clinical Reasoning, Musculoskeletal, Dual Processing Theory

Appendices

- A. Introduction Presentation and Schedule.pptx
- B. Pre- and Posttest.docx
- C. Exemplar Cases.pptx
- D. Session Evaluation Questions.docx

All appendices are peer reviewed as integral parts of the Original Publication.

Educational Objectives

By the end of this session, learners will be able to:

1. Identify the correct diagnosis at least 80% of the time when presented with perfect cases of common painful cervical spine and shoulder conditions.
2. Identify at least one history key feature and at least one physical exam key feature for common cervical spine diagnoses and shoulder diagnoses.

Introduction

Gaps in physician data collection and medical knowledge cannot account for the majority of diagnostic errors; information processing and related cognitive factors may be the root cause. While the conceptual framework of dual processing in clinical reasoning is widely accepted, it is not clear how medical students and residents can be taught to switch between automatic and reflective modes and whether doing so improves their decision making. Developing effective clinical reasoning habits while in training may facilitate progression to expertise, reduce diagnostic errors, and improve patient safety.

It is widely accepted that most medical errors involve both system and physician factors.¹ Physician cognitive factors (information processing), rather than knowledge or data collection, are thought to account for a majority of diagnostic errors.²

The conceptual framework of dual process of reasoning originated in cognitive psychology and has been widely accepted to explain and study physician clinical reasoning.^{3,4} Reflexive mode, also known as System 1, is a rapid, intuitive pattern-recognition process largely outside of conscious control. In contrast, reflective mode, also known as System 2, is slower and involves deliberate and intentional systematic analysis of available data.

Mamede and colleagues found that structured reflection appears to result in the acquisition of clinical knowledge more effectively than the generation of immediate or differential diagnoses.⁵ Forty-six senior medical students first engaged in diagnosing six clinical cases under three experimental conditions: providing an immediate diagnosis, generating a differential diagnosis, and structured reflection. Mamede and colleagues asked the subjects to diagnose four of the six cases immediately and then 1 week later. While students under immediate and differential diagnosis conditions did better during immediate testing, structured reflection resulted in significantly better diagnostic performance 1 week later.

Supporting the effectiveness of deliberate reflection, the same research group exposed 38 internal medicine residents to a media article about two diseases.⁶ Six hours later, participating residents were asked to diagnose eight seemingly unrelated cases, two of which had features resembling diseases described in the media article. The researchers went to great lengths to mask any connection between the first and the second parts of the study in a systematic and thoughtful way. When subjects were tested, diagnostic accuracy was significantly lower on the cases with expected (availability) bias. Repeated diagnosis using deliberate reflection improved diagnostic accuracy, restoring it to the prebias level.

Sibbald and de Bruin echoed these findings and discovered that although residents were able to reflect on and accurately recognize their mode of reasoning, using balancing strategies (either opposing or similar) did not reduce diagnostic error.⁷ Studying six medical residents' electrocardiogram interpretation skills, Sibbald and de Bruin found that regardless of strategy used initially, instructions to systematically analyze the electrocardiogram (rather than look for patterns) during reinterpretation resulted in improved diagnostic accuracy.

In contrast, Ilgen and colleagues found that instructions to trust one's first impressions resulted in similar diagnostic performance compared with instructions to engage in systematic analysis.⁸ In a large multicenter randomized study that involved students, residents, and faculty, the researchers used clinical vignettes developed and tested by Mamede and colleagues.⁵ Three hundred and ninety-three participants were randomized into either a group instructed to use first impressions or a group encouraged to perform a directed (systematic) search. Instructions to use first impressions not only resulted in similar diagnostic performance and shorter testing time but also showed stronger correlation with participants' United States Medical Licensing Examination scores.

Similarly, Norman and colleagues demonstrated that encouraging residents to slow down and engage in analytical reasoning did not result in increased diagnostic accuracy.⁴ In a cohort study of 204 second-year medical residents at three medical schools, these researchers instead found that incorrectly diagnosed cases took longer regardless of the reasoning strategy used.

The ability to enhance residents' diagnostic accuracy using directed training designed to foster switching between reflexive and reflective modes of processing is still a matter of debate within the educational research community. Additionally, to our knowledge, comparative effectiveness and feasibility of available instructional methods have not been evaluated.

However, developing effective clinical reasoning skills while in training may facilitate progression to expertise, reduce diagnostic errors, and improve patient safety. This brief workshop was designed to provide resident physicians with practice in engaging in and toggling between both modes of information processing using exemplar case vignettes.

The fact that residents in our program invariably selected "improving clinical reasoning and developing a differential diagnosis" as one of the goals for their musculoskeletal and sports medicine rotation served as

a local impetus for developing this workshop. Additionally, program evaluation results suggested that first-year residents were not prepared for outpatient musculoskeletal rotations and that a traditional lecture curriculum might not adequately prepare them for the challenges encountered in clinics.

An introductory clinical reasoning curriculum has previously been published on MedEdPORTAL.⁹ However, it did not specifically address any particular content domain, nor has there been conclusive evidence that learning about clinical reasoning may lead one to being better at clinical reasoning.¹⁰ A series of neck and shoulder OSCE cases created for use with internal medicine residents has also been published on MedEdPORTAL, but its instructional methodology was not underpinned by the latest theory of clinical reasoning.¹¹ The unique contribution of our workshop is that it utilizes both systems of clinical reasoning, offers learners practice in toggling from one to another (a skill that is necessary for real-world diagnostic reasoning), and at the same time applies this broad skill to a concrete domain of medical knowledge, musculoskeletal neck and shoulder conditions.

Methods

Our program is an ACGME-accredited physical medicine and rehabilitation (PM&R) residency program with 12 residents in each of the 3 years (residents enter at a PGY-2 level). In addition to workplace base learning¹² during monthly clinical rotations (inpatient, outpatient, consultation, procedures), all residents participate in a weekly core curriculum delivered by faculty. At the beginning of the year, residents are divided (by the program director) into three smaller learning groups, with 12 residents in each (four residents in each year of training). These smaller groups are used for all small-group learning activities, such as case-based learning, throughout the year.

This workshop was given once midyear; the assessment data thus are from one iteration. First-, second-, and third-year PM&R residents who participated had, respectively, approximately 6 months, 18 months, and 30 months of the didactic core curriculum and clinical rotations, roughly a third of which cover musculoskeletal and sports medicine.

The Workshop

This workshop was conducted by the faculty during the core curriculum time, when all residents physically come to a single location. Four classrooms were utilized: a large classroom for the entire class and three small-group rooms. A single faculty member conducted all segments of the workshop.

Introduction and learning objectives: The workshop began with a slide-show presentation by the faculty (Appendix A). After a brief review of the process of clinical reasoning and dual reasoning theory,¹⁰ the slide show reviewed (1) the role of semantic qualifiers,¹³ (2) the workshop structure of toggling between recognizing the diagnosis from an exemplar case vignette and identifying key features given a common diagnosis,¹⁴ and (3) the learning objectives. The slide show also covered the workshop schedule.

Pretest: In the next segment, the entire class took an 18-question online pretest (Appendix B) utilizing an institutional learning management system. Pretest questions were multiple choice¹⁵ with nine answer options and were written by me using my 16 years of clinical and teaching experience in the musculoskeletal content domain. The first nine questions were based on vignettes of patients with common painful neck conditions, and questions 10-18 were based on vignettes of patients with common painful shoulder conditions. The exemplar case diagnoses were selected based on faculty's listing of typical and common outpatient problems seen by PM&R residents over the course of their teaching career. Ten minutes were allocated for the test, and residents were encouraged to go with their first impression or gut feeling (System 1). Appendix B lists the test questions and answer options in a .docx format that can be easily imported into a local learning management system.

Small-group preparation: After the pretest, residents separated into their individual learning groups in separate, smaller classrooms. Residents were provided with the list of 18 diagnoses (Appendix C), and the senior year 2 and year 3 residents in each group were instructed to train the year 1 residents and prepare

them for the team competition in the next segment. Residents were encouraged to use any resources and references they needed to accomplish that task.

Game time: Residents returned to the larger classroom for the next segment of the workshop. Year 1 residents from each learning group sat at the front of the room and participated in the competition. Year 2 and year 3 residents were encouraged to cheer for their group but not allowed to help with answers or hints. The faculty facilitated the competition. One of the chief residents kept the score, added up the points, and announced the winning group and the runner-up at the end of the segment. As a prize, the winning group was invited to suggest the location for the annual residency holiday party. Each learning group took turns playing and received 1 point per correct answer. There was no penalty for wrong answers; however, each group was allowed only one answer. There were 18 diagnoses, and each of the three groups had six turns.

After specifying which group was playing, the faculty showed one of the exemplar diagnoses (Appendix C) on the screen. The group was asked to provide at least one key feature from the history and at least one key feature from the physical examination (System 2). There was no time restriction. After the group provided its answer, faculty showed the case vignette with key features highlighted in red (Appendix C) and engaged the class in a brief interactive discussion to ensure understanding and share clinical examples. This process was repeated for each diagnosis, with three competing groups taking turns.

Posttest: In the next segment, the entire class took an 18-question online posttest (Appendix B) utilizing an institutional learning management system. Posttest questions and answer options were identical to pretest questions and answer options. Ten minutes were allocated for the test, and residents were encouraged to go with their first impression or gut feeling (System 1). Appendix B lists the test questions and answer options in a .docx format that can be easily imported into a local learning management system.

Session evaluation: In the next segment, the entire class completed an online session evaluation consisting of three construct-response questions addressing Kirkpatrick's levels 1 and 2 (Appendix D) and utilizing an institutional learning management system.^{15,16} Fifteen minutes were allocated for the session evaluation.

Outcome Data Analysis

Data analysis was performed using IBM SPSS Statistics 23. Only data from residents who completed the entire test were examined.

Validity of the pre- and posttest was explored in two ways. First, question quality was evaluated by calculating individual items' discrimination index and point-biserial coefficient. The discrimination index indicates how well a question differentiates between high and low performers. It can range from -100% to 100%, with high values indicating a good question and low values indicating a bad question.¹⁵ Similar to the discrimination index, the point-biserial correlation coefficient relates individuals' quiz scores to whether or not they get a question correct. It ranges from -1.00 to 1.00, with high values indicating a good question and low values indicating a bad question.¹⁵

Second, we hypothesized that if our tests accurately measure the construct of musculoskeletal clinical knowledge and if that knowledge indeed increases during training progression, senior residents as a group should perform better on the pretest. We therefore compared means of pretest performance by year of training, visually, and using analysis of variance.

While we expected an improvement in performance between the pre- and the posttest for all years of training, we were also interested in whether there might be a differential effect of the workshop on posttest performance depending on year of training and on the pretest result. Linear multiple regression analysis was performed to explore this question. We used this method because multiple regression allows exploration of both significance and strength of prediction of an independent variable using two or more dependent variables.¹⁷ The theoretical regression model was $\text{posttest} = \text{constant} + \text{beta1}(\text{year of training}) + \text{beta2}(\text{pretest})$.

Results

Twenty-six residents participated in the workshop. Within the allocated time of 10 minutes, 24 and 16 residents were able to complete the neck and shoulder segments of the pretest, respectively. For the posttest, these numbers were 26 and 25, respectively (see the [Table](#)).

Table. Descriptive Statistics, Pre- and Posttest

Test	Percentage		
	<i>M</i>	<i>SD</i>	Range
Neck			
Pretest (<i>N</i> = 24)	86.6	13.9	55.6-100
Posttest (<i>N</i> = 26)	98.7	3.6	88.9-100
Shoulder			
Pretest (<i>N</i> = 16)	76.4	19	44.4-100
Posttest (<i>N</i> = 25)	90.2	12.6	55.6-100

Pretest questions were of reasonable quality for an 18-item test, with a discrimination index average of 39.4% and a point-biserial average of .48.

Pretest performance varied significantly as a function of the year of training (see the [Figure](#)). Analysis of variance demonstrated significant difference between the means by year of training, $F(1, 23) = 11.060$, $p = .003$.

Figure. Pretest score mean by year of training.

A linear regression demonstrated that year of training significantly predicted pretest scores, $b = 0.107$, $t(23) = 3.326$, $p < .001$. Year of training also explained a significant proportion of variance in pretest scores, $R^2 = .335$, $F(1, 23) = 11.060$, $p = .003$.

When controlling for pretest scores, a regression analysis showed year of training did not significantly predict posttest scores independent of pretest scores, $b = -0.023$, $t(23) = 2.596$, $p = .098$.

Twenty-two residents completed an online session evaluation. A majority of residents (86.4%) indicated that they liked the workshop or liked it a lot. Most (90.9%) felt they were better or much better at being

able to recognize common neck and shoulder problems as a result of the workshop. Participants provided a number of narrative comments and recommendations:

- “It’s good.”
- “Fix the submission problem.”
- “Smaller groups, with big groups some people don’t participate.”
- “I’m not entirely sure how to improve the sessions however it was a nice reminder of all of the special tests and histories.”
- “\$5 Starbucks gift cards ;-).”
- “Great interactive MSK workshop! Really helped learn and solidify key topics!”
- “Great.”
- “PGY-4 wouldn’t stop talking so it was difficult to hear what was going on. Perhaps ask them to quiet down.”
- “This was already an improvement over the previous sessions by far. More interactive with a greater learning application and retention. Would be better once technical issues are resolved further.”
- “Different posttest questions.”
- “Keep PGY-2 in front. Forces the seniors to teach.”
- “I think the questions can be harder.”
- “All good.”
- “It was my first time. I was a little lost at the beginning, but then was fine. It was cool.”
- “We can maybe include a short PowerPoint presentation which would include concise clinical presentation, diagnosis (including gold standard) and treatment. Overall, this season was quite helpful.”
- “I preferred today’s format compared to previous CBL sessions. It was much more focused and organized. Allowed for more learning.”
- “Would love to see this format recreated for other aspects of PM&R that we are tested on. The quiz section at the end was my least favorite. It would be nice to wrap up as a large group instead; maybe divide up the questions and have each group present a brief synopsis of each high yield condition. This would really hammer it home, especially after having had a similar session in the smaller groups; it would be nice to hear what the other groups had to say.”
- “When I say that my least favorite part was the quiz section at the end, I mean the part where the PGY-2s are quizzed amongst the three groups, not the actual posttest quiz. Keep the posttest, definitely helps for me to have repetition.”
- “Excellent CBL with pretest and posttest format. Learned a lot and really solidified the topic with this format. Would like it again.”

There were two hallway conversations with the faculty. In one, a year 1 resident spontaneously stated that the workshop format was the most useful the resident had experienced. In another, a different year 1 resident told the faculty that the most valuable part of the workshop was the 1-hour segment where senior residents taught earlier-year residents in a small group.

Discussion

While there is some debate on the relative contributions of clinical reasoning System 1 and System 2 to improving diagnostic accuracy,⁵⁻⁹ it is probably safe to assume that learning to utilize both systems is a helpful skill to add to the residents’ clinical reasoning tool belt. This brief workshop was therefore designed to provide resident physicians with practice in engaging in and toggling between both modes of information processing using exemplar case vignettes.

Pretest performance seemed to suggest that PM&R residents were more capable at recognizing musculoskeletal problems as they progressed through training. Not unexpectedly, year of training explained only 33% of the variance in pretest performance; residency training may be just one of many factors that impact residents’ performance.

Results of multiple regression also suggested that when controlling for pretest results, there was no significant difference in posttest results between residents in different years of training. This may imply that the workshop's effectiveness did not vary based on the year of training.

The design of the educational intervention, as well as of the assessment, is far from perfect. It is philosophically crowded with elements of a number of educational models, such as team-based learning,¹⁸ case-based learning,¹⁹ key features assessment,¹⁴ dual processing theory,¹⁰ and gamification.²⁰ The educational intervention itself is unstructured and is dependent on each group finding its way through the task successfully. The reliability of the assessment is limited by a small number of items in each test. Additionally, the pre- and posttest questions are identical and administered within 2 hours of each other, raising the possibility of participants remembering the answers. We tried to address this by not providing correct-answer feedback immediately after the pretest, and one of the residents mentioned trying to remember the small-group discussion rather than the pretest when taking the posttest. Nevertheless, this is a weakness that future users may want to address by modifying the posttest questions. Another weakness is that both the case vignettes and the questions were written by a single faculty member and may have incorporated biases intrinsic to that person's individual experience. Finally, development and implementation were not resource-free. Design and development required approximately 8 hours of faculty time, most of it spent in writing cases and questions and laboriously uploading them to the learning management system. Implementation required 3 hours of single faculty time, as well as four classrooms to accommodate the group of 26 residents.

At the same time, the workshop has some merits that warrant consideration. It is relevant to residents in several specialties, including PM&R, rheumatology, orthopedic surgery, family medicine, and internal medicine, as well as to senior medical students, physician assistants, and nurse practitioners. It is brief and requires only a single faculty member to implement. It was liked by the residents and resulted in improvement of test scores across the board as well as remarkable and prolonged learning engagement of most of the residents. The workshop was designed based on principles of evidence-based medical education and may enhance development of a community of learners within a residency program.²¹

Several lessons were learned in the process of design and implementation. First, the time limit of 10 minutes was not appropriate, as a number of residents were unable to finish the entire test in time. Not surprisingly, fewer residents had that issue during the posttest compared with the pretest, probably because they both learned the process of taking an online quiz and were more comfortable with the content matter. Second, significant faculty time was necessary to implement the workshop using the institutional learning management system; specifically, entering individual pre- and posttest questions was laborious. This can be delegated to staff, although staff training time is not avoidable, or the test can be done via paper administration. Finally, residents liked this interactive and structured workshop more than other group case-based learning formats previously tried.

Alex Moroz, MD: Associate Professor, Vice Chair for Education, and Residency Program Director, Department of Rehabilitation Medicine, New York University School of Medicine

Disclosures

None to report.

Funding/Support

None to report.

Ethical Approval

Reported as not applicable.

References

1. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med*. 2005;165(13):1493-1499. <https://doi.org/10.1001/archinte.165.13.1493>
2. Norman G. Dual processing and diagnostic errors. *Adv Health Sci Educ Theory Pract*. 2009;14(suppl 1):37-49. <https://doi.org/10.1007/s10459-009-9179-x>
3. Evans JST. Dual-processing accounts of reasoning, judgment, and social cognition. *Annu Rev Psychol*. 2008;59:255-278. <https://doi.org/10.1146/annurev.psych.59.103006.093629>
4. Norman G, Sherbino J, Dore K, et al. The etiology of diagnostic errors: a controlled trial of System 1 versus System 2 reasoning. *Acad Med*. 2014;89(2):277-284. <https://doi.org/10.1097/ACM.0000000000000105>
5. Mamede S, van Gog T, Moura AS, et al. Reflection as a strategy to foster medical students' acquisition of diagnostic competence. *Med Educ*. 2012;46(5):464-472. <https://doi.org/10.1111/j.1365-2923.2012.04217.x>
6. Schmidt HG, Mamede S, van den Berge K, van Gog T, van Saase JLCM, Rikers RMJP. Exposure to media information about a disease can cause doctors to misdiagnose similar-looking clinical cases. *Acad Med*. 2014;89(2):285-291. <https://doi.org/10.1097/ACM.0000000000000107>
7. Sibbald M, de Bruin ABH. Feasibility of self-reflection as a tool to balance clinical reasoning strategies. *Adv Health Sci Educ Theory Pract*. 2012;17(3):419-429. <https://doi.org/10.1007/s10459-011-9320-5>
8. Ilgen JS, Bowen JL, McIntyre LA, et al. Comparing diagnostic performance and the utility of clinical vignette-based assessment under testing conditions designed to encourage either automatic or analytic thought. *Acad Med*. 2013;88(10):1545-1551. <https://doi.org/10.1097/ACM.0b013e3182a31c1e>
9. Weinstein A, Pinto-Powell R. Introductory clinical reasoning curriculum. *MedEdPORTAL Publications*. 2016;12:10370. https://doi.org/10.15766/mep_2374-8265.10370
10. Norman GR, Monteiro SD, Sherbino J, Ilgen JS, Schmidt HG, Mamede S. The causes of errors in clinical reasoning: cognitive biases, knowledge deficits, and dual process thinking. *Acad Med*. 2017;92(1):23-30. <https://doi.org/10.1097/ACM.00000000000001421>
11. Soares S, Wang H, Siddharthan T, Holt S. OSCE-based teaching of the musculoskeletal exam to internal medicine residents and medical students: neck and spine. *MedEdPORTAL Publications*. 2015;11:10120. https://doi.org/10.15766/mep_2374-8265.10120
12. Yardley S, Teunissen PW, Dornan T. Experiential learning: AMEE Guide No. 63. *Med Teach*. 2012;34(2):e102-e115. <https://doi.org/10.3109/0142159X.2012.650741>
13. Bordage G, Lemieux M. Semantic structures and diagnostic thinking of experts and novices. *Acad Med*. 1991;66(9):S70-S72. <https://doi.org/10.1097/00001888-199109000-00045>
14. Hrynchak P, Takahashi SG, Nayer M. Key-feature questions for assessment of clinical reasoning: a literature review. *Med Educ*. 2014;48(9):870-883. <https://doi.org/10.1111/medu.12509>
15. Downing SM, Yudkowsky R. *Assessment in Health Professions Education*. New York, NY: Routledge; 2009.
16. Kirkpatrick DL, Kirkpatrick JD. *Evaluating Training Programs: The Four Levels*. 3rd ed. San Francisco, CA: Berrett-Koehler; 1994.
17. Field A. *Discovering Statistics Using IBM SPSS Statistics*. 4th ed. London, England: SAGE Publications; 2013.
18. Fatmi M, Hartling L, Hillier T, Campbell S, Oswald AE. The effectiveness of team-based learning on learning outcomes in health professions education: BEME Guide No. 30. *Med Teach*. 2013;35(12):e1608-e1624. <https://doi.org/10.3109/0142159X.2013.849802>
19. Thistlethwaite JE, Davies D, Ekeocha S, et al. The effectiveness of case-based learning in health professional education. A BEME systematic review: BEME Guide No. 23. *Med Teach*. 2012;34(6):e421-e444. <https://doi.org/10.3109/0142159X.2012.680939>
20. Hamari J, Koivisto J, Sarsa H. Does gamification work?—a literature review of empirical studies on gamification. In: *2014 47th Hawaii International Conference on System Sciences (HICSS 2014)*. Waikoloa, HI: IEEE; 2014:3025-3034.
21. Rogoff B. Developing understanding of the idea of communities of learners. *Mind Cult Activity*. 1994;1(4):209-229.

Received: December 3, 2016 | Accepted: March 5, 2017 | Published: March 24, 2017