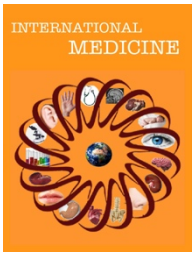
International Medicine 2019; 1(4): 185-192



International Medicine

https://www.theinternationalmedicine.com/

Original Article

The impact of a redesigned clinical course curriculum on the development of clinical reasoning skills among medical students: a quantitative study

Rhys Taylor, Swe Khin Htun

Trent Simulation and Clinical Skills Centre, Nottingham, United Kingdom

Received: 10 May 2019 / Accepted: 12 June 2019

Abstract

Background: Clinical reasoning is an essential skill for healthcare professionals to develop to better manage and treat their patients. It is best described as the ability to analyze, interpret, and investigate a patient appropriately given their constellation of symptoms and signs. Clinical reasoning has come under scrutiny as an area for professional development, with many questionings whether it can be formally taught or if it is simply an organic process which improves with experience. This study aims to investigate whether formalized clinical reasoning teaching during medical school has an impact on the clinical reasoning abilities of its students.

Methods: Data was collected across three cohorts at the University of Nottingham Medical School. Each cohort had had varying degrees of clinical reasoning teaching, with Cohort 1 receiving no formalized clinical reasoning teaching, Cohort 2 receiving some teaching, and Cohort 3 having an embedded clinical reasoning component to their clinical education. The end of year exam results for each cohort were analyzed both in their first and final year of clinical exposure, to determine how well each cohort performed in the clinical reasoning component of the written tests.

Results: On analysis of the results, the clinical reasoning scores were consistently higher in the cohort who received full clinical reasoning teaching during their clinical phases. There was one exception to this, but the data suggests this may itself be due to an anomaly in the overall abilities of that cohort during this time. With these results, the null hypothesis that clinical reasoning teaching has no significant effect on the clinical reasoning of students as measured by the clinical reasoning score in the summative exams was rejected.

Conclusions: The main finding of this research was that the implementation of a clinical reasoning teaching curriculum can have a significant impact on the clinical reasoning of its students. This suggests that clinical reasoning can be developed both before and during clinical practice.

Keywords: clinical reasoning, curriculum, medical education

Address for Correspondence: Rhys Taylor, Trent Simulation and Clinical Skills Centre, C/O Post Grad Centre, QMC Campus, Nottingham, NG7 2UH, United Kingdom. E-mail: theratat25@hotmail.co.uk

DOI: 10.5455/im.46576

Introduction

Clinical reasoning (CR) is the term used to describe the analysis of a patient's clinical presentation by a healthcare professional, so as to better ascertain their likely diagnosis and appropriate management plan. It is an essential skill for healthcare professionals to develop from an early stage in their careers and has thus come under close scrutiny as a possible area of study in medical schools [1].

CR is best broken down into two components that are heavily intertwined. The first being the 'intuitive' component and the second the 'analytical' component [2-4]. The intuitive component draws upon pattern recognition and snaps decision making which cannot often be broken down into a logical thought process either by external sources or indeed the practitioner themselves. The analytical component, however, can be traced relatively easily as it is a methodical, conscious act drawing upon prior knowledge and logical thought processes. Although there is some dissent in the literature, with authors such as Neufeld et al arguing that CR is a relatively constant measure in students, many believe that CR can be taught. Authors such as Groves and Da Silva state that CR skills develop as the student progresses in their learning, whilst others go even further, reporting that the course curriculum itself has a strong impact on the CR of its students [5-7].

Medical education thus focusses on the premise that although CR may develop gradually throughout higher education [8], this is likely to represent the unconscious intuitive element of CR. We can therefore still aim to positively influence the development of a student's CR by targeting the analytical element through focused teaching.

Aims and objectives

In current literature, there is a discrepancy which has previously been alluded to between studies which report the development of CR to be an organic process, which happens naturally over time, and others who believe that it should be formally taught [9-11].

The University of Nottingham (UoN) Medical School redesigned the clinical course to include CR teaching borne on the tenants espoused by authors such as Hmelo et al and Koh et al, that repeated exposure to well-selected cases can nurture and develop a student's CR.

This study aims to look at whether there was a significant difference in the CR skills between the students who received no CR teaching, some CR teaching, and fully integrated CR teaching.

The research question and associated hypothesis are as follows:

Research question: Does the level of CR teaching have an impact on the development of CR?

Hypothesis 1: CR teaching has no significant effect on the CR of students as measured by the CR score in the summative exams.

Hypothesis 2: CR teaching has a significant effect on the CR of students as measured by the CR score in the summative exams.

Methods

Structure of the UoN clinical curriculum

As this study focusses on the UoN and its clinical curriculum, it is important to understand its clinical components.

Clinical Phase 1 (CP1) is a seventeen-week period with rotations on medicine and surgery. The goal during this period is to learn the common conditions associated with these areas and to improve on relevant examination skills and history taking. The attachment culminates in a single knowledge-based multiple choice exam and a single rotation of objective situational clinical examinations.

Clinical Phase 2 (CP2) is a forty-week placement encompassing a greater range of specialties including obstetrics and gynecology, pediatrics, geriatrics, general practice, psychiatry, dermatology, and ophthalmology. The attachment culminates in two knowledge-based multiple choice exams and two rotations of objective situational clinical examinations.

Clinical Phase 3 (CP3) is a thirty-two-week placement covering medicine, surgery, musculoskeletal disorders, general practice, and critical care. The goal during this period is to consolidate the students' knowledge and clinical skills in preparation for their first year as a clinician. The attachment culminates in two knowledge-based multiple choice exams and two rotations of objective situational clinical examinations. After exams students go on to complete a 'Transition to Practice' element which includes seven weeks on elective and seven weeks shadowing on a ward.

CP2 is a specialized year which is difficult to compare to its counterparts. CP1 and CP3, however, are relatively similar in terms of subject material and as such are the focus of this paper.

Since 2012 the UoN medical school has been updating its practices to include CR as part of its curriculum. The role CR teaching played in the clinical curriculum at the UoN was distinctly different between the three following cohorts:

Cohort 1: CP1 2011- 2012 CP3 2013- 2014

Cohort 2: CP1 2012- 2013 CP3 2014- 2015

Cohort 3: CP1 2013- 2014 CP3 2015- 2016

Cohort 1 received no formalized CR teaching during their time at medical school. Cohort 2 received a modicum of CR teaching which started during their clinical phases and was a new addition to the curriculum. Cohort 3 underwent a clinical curriculum into which CR teaching was firmly embedded.

Data collection

Data is collected from the summative knowledge assessments that take place at the end of CP1 and CP3 for each of the three cohorts that passed through the clinical phases of medicine at UoN between 2011 and 2016. This data will then be analyzed to ascertain whether CR teaching has an impact on the CR scores of the different cohorts involved in the study. Table 1 shows the overall number of students per cohort.

Table 1. Number of students in CP1 and CP3 by cohort

CP1	Students in cohort	CP3	Students in cohort
2011- 2012	351	2013- 2014	335
2012- 2013	344	2014- 2015	350
2013- 2014	327	2015- 2016	318

Before the exam: classification of CR and non-CR questions

Written papers were divided into CR and non-CR (NCR) questions by a team of fifteen to twenty-five experts in standard setting meetings. These experts had a range of backgrounds including general practice, consultants in relevant disciplines (e.g. respiratory medicine, cardiology, etc), teaching fellows, junior doctors, and medical educators.

Any discrepancies between the experts as to whether a question is defined as CR or non-CR are resolved through two measures. First is by mapping the question against ‘Outcomes for graduates' from ‘Tomorrow's Doctors' sections 8c, 8g, and 14f from the General Medical Council (GMC) [12]. Secondly, the question must require the student to either apply, analyze, evaluate, or create as per four of Bloom’s six cognitive processes. The questions then continue to be discussed until a mutually agreeable conclusion is reached.

CR questions can cover a range of topics including being given history and formulating an appropriate differential diagnosis; the interpretation of physical signs to diagnose a patient and come up with an appropriate management plan; matching case vignettes to the appropriate pathology; analyze investigations to provide an appropriate follow-up plan.

After the exam: psychometric evaluation

Routine psychometric evaluation is carried out on each medical exam paper using Classical Test Theory (CTT) and Item Response Theory (IRT) to ensure a high-quality assessment tool. Questions identified as too easy or too difficult are identified via student-item maps and then discarded. Each paper is subsequently analyzed using test-score reliability (Cronbach’s alpha), item discrimination index (ID), and standard error of measurement (SEM).

Frequency and discrimination (U-L) analysis and learning objective analysis is also carried out for each paper. Question difficulty (p) and discriminatory value (d) are calculated with questions having a $d < 0.15$ (low discriminatory value) and $p < 0.2$ (very difficult) being excluded. Subsequently, the reliability of the test was measured using generalizability (G) theory and to address error variance among questions. These analyses are carried out so that the summative assessments better reflect the differences between students with regards to their skills and knowledge.

Finally, the papers are reviewed by internal peers and external examiners. Their opinions are carefully considered, and action is decided upon as appropriate. Tables 2 and 3 below describe the overall number of questions in each exam and what proportion were designated as CR questions.

Data analysis

The data was analyzed using analysis of variance (ANOVA) in conjunction with Tukey’s test to evaluate whether there was a statistical difference in the CR scores between each cohort’s CP1 and CP3 score. These parametrical statistical tests require the data to be normally distributed, which was achieved with the use of kurtosis statistics and investigation of the skewness and histogram to confirm the distribution of data.

The longitudinal analysis helped determine whether there was a statistically significant change in CR scores between the cohorts in their CP1 and CP3 exams. $p < 0.05$ was deemed statistically significant.

Table 2. CP1 marks Cohort 1 (2011-12) to Cohort 3 (2013-14)

CP1	Marks			
	CR	NCR	Total	CR % of total
2011- 2012	85	101	186	46
2012- 2013	73	112	185	39
2013- 2014	116	79	195	59

Table 3. CP3 marks in Paper 1 and 2 for Cohort 1 (2013-14) to Cohort 3 (2015-16)

CP3	Marks							
	Paper 1				Paper 2			
	CR	NCR	Total	CR% of total	CR	NCR	Total	CR % of total
2013- 2014	151	41	192	79	87	83	170	51
2014- 2015	85	85	170	50	91	79	170	54
2015- 2016	113	57	170	66	98	77	175	56

Ethics approval

Ethical approval was sought for the conduction of this study. As this research is classified as service evaluation the ethics committee deemed it unnecessary as 'ethical approval is not required for processing data'. This research does not directly involve people and examination data is anonymized throughout.

Results

ANOVA was conducted to determine whether the impact CR teaching had on the final CR scores for each cohort was significant. Total scores in the summative exams were significantly different across the three cohorts which each had different levels of CR teaching.

CP1 results

Table 4 summarizes the ANOVA results from the CP1 dataset. The results showed that the CR scores ($F (2, 1019) = 698.85$, $p < 0.001$) and total score in the summative exam ($F (2, 1018) = 38.62$, $p < 0.001$) were significantly different between the three cohorts.

Table 4. ANOVA results of the difference of CR scores and total scores of summative written exams in different cohorts for CP1 dataset

CP1		Sum of squares	df	Mean square	F	Sig.
CR	Between groups	114866.89	2	57433.44	698.85	0.00
	Within groups	83743.75	10 19	82.18		
	Total	198610.64	10 21			
Summative written exams	Between groups	17200.07	2	8600.04	38.62	0.00
	Within groups	226723.48	10 18	222.72		
	Total	243923.55	10 20			

Cohort 1 (2011- 2012) had no CR teaching, Cohort 2 (2012- 2013) had some CR teaching and Cohort 3 (2013- 2014) had full CR teaching.

Table 5 describes the use of Tukey's test to show that CR scores were significantly higher in the period of full CR teaching (M = 79.71; SD = 11.41) than in the period of no CR teaching (M = 57.66; SD = 8.58) and partial CR teaching (M = 56.35; SD = 6.73) by a mean difference of 22.05 and 23.36 respectively.

Table 5. Post-hoc test results of the difference of CR scores and total scores of the summative written exam in different cohorts for the CP1 dataset

CP1						**95% confidence interval**	
Dependent variable	(I) Year/ Teaching	(J) Year/ Teaching	Mean difference (I-J)	Std. error	Sig.	Lower bound	Upper bound
CR score	0 2011- 2012	1 2012- 2013	1.31	0.69	0.14	-0.30	2.93
	0 2011- 2012	2 2013- 2014	-22.05	0.70	0.00	-23.68	-20.41
	1 2012- 2013	2 2013- 2014	-23.36	0.70	0.00	-25.00	-21.72
Summative written exams	0 2011- 2012	1 2012- 2013	-9.05	1.13	0.00	-11.71	-6.39
	0 2011- 2012	2 2013- 2014	-8.14	1.15	0.00	-10.83	-5.45
	1 2012- 2013	2 2013- 2014	0.91	1.15	0.71	-1.79	3.62

CP3 results

Table 6 summarizes the ANOVA results from the CP3 dataset. The results showed that the CR scores ($F (2, 1000) = 2042.58$, $p < 0.001$) and total score in the summative exam ($F (2, 1000) = 149.07$, $p < 0.001$) for Paper 1 and CR scores ($F (2, 1000) = 229.76$, $p < 0.001$) and total score in the summative exam ($F (2, 1000) = 215.66$, $p < 0.001$) for Paper 2 were significantly different between the three cohorts.

Table 7 describes the use of Tukey's test which showed that the CR score for Paper 1 for Cohort 1 (M = 110.17; SD = 10.80) was significantly higher than the CR score for either Cohort 2 (M = 64.27; SD = 7.39) or Cohort 3 (M = 83.23; SD = 9.88) by a mean difference of 45.90 and 26.94 respectively.

The summative score for Paper 1 in Cohort 1 (M = 141.54; SD = 13.92) was also significantly higher than the summative score for either Cohort 2 (M = 125.39; SD = 11.73) or Cohort 3 (M = 127.96; SD = 13.50) by a mean difference of 16.16 and 13.59 respectively.

Similarly, the CR and summative scores for Paper 1 were significantly higher in Cohort 3 than Cohort 2 by a mean difference of 18.96 and 2.57 respectively.

In Paper 2, Cohort 3 (M = 74.13; SD = 7.94) had a significantly higher CR score than either Cohort 1 (M = 66.82; SD = 6.83) or Cohort 2 (M = 61.63; SD = 7.81) by a mean difference of 5.23 and 12.49 respectively. Cohort 1's CR score also outperformed that of Cohort 2 by a statistically significant mean difference of 5.23.

Null hypothesis

In summary, CR scores were consistently higher in the third cohort who experienced a fully implemented CR teaching curriculum. The exception to this was Paper 1 in CP3 where Cohort 1 had both a significantly higher CR and summative score. However, due to the summative score for Cohort 1 also being so high in both Paper 1 and 2 of CP3 it could be hypothesized that this anomaly is due to the abilities and knowledge of this cohort.

With these results, the null hypothesis that CR teaching has no significant effect on the CR of students as measured by the CR score in the summative exams was rejected. The ANOVA results showed a significant effect which supports the hypothesis that CR teaching has a positive effect on a student's CR abilities.

Table 6. ANOVA results of the difference of CR scores and total scores of summative written exams in different cohorts for CP3 dataset

CP3		Sum of squares	df	Mean square	F	Sig.
CR (Paper 1)	Between groups	363234.94	2	181617.47	2042.58	0.00
	Within groups	88915.64	100 0	88.92		
	Total	452150.58	100 2			
Summative written exams (Paper 1)	Between groups	50847.42	2	25243.71	149.07	0.00
	Within groups	170549.66	100 0	170.55		
	Total	221397.08	100 2			
CR (Paper 2)	Between groups	26125.92	2	13062.96	229.76	0.00
	Within groups	56854.83	100 0	56.86		
	Total	82980.75	100 2			
Summative written exams (Paper 2)	Between groups	69275.41	2	34627.70	215.66	0.00
	Within groups	160610.25	100 0	160.61		
	Total	229885.65	100 2			

Discussion

Even though the literature differs as to the effectiveness of CR teaching in instilling CR in healthcare students, this study supports its effectiveness. CR teaching had a significant effect on the CR of its students and as such course curriculums should have CR embedded into them to better prepare its students for the future.

The research question asked whether CR teaching had a significant effect on the CR of its students. The statistical analysis reported a significant effect from CR teaching and supports its continued implementation at the UoN.

Despite this, it is important that this teaching is intentional and methodical and not opportunistic. It must represent a challenge to students. CR teaching is also most effective if it is supported by timely feedback, coaching and effortful practice as per Ericsson's theory [13]. Thus, the CR teaching will be best received if it is constantly planned, evaluated and redesigned based on the students' needs.

Limitations

This study had three key limitations associated with its methodology, its assessment of CR and the analysis of the data sets.

Due to the limited time and resources with which the study was run, it was not possible to compare the performance of the three cohorts of students using conventional evaluation methodology. This was a key limitation of the overall methodology, alongside the time constraints which meant that the researcher was only able to compare the students' CR during the clinical phase and not throughout the five- year curriculum. A process which would have taken ten years to complete.

The assessment of the students' CR was also limited. The data used throughout the study uses the summative exam performance as a proxy measure of CR, yet that doesn't necessarily reflect all of the aspects of a student's CR abilities. A further area in which the students could have been analyzed was their objective situational clinical examination (OSCE) performance at the end of CP1 and CP3. The researcher was unable to do this, however, as ethical and logistical restrictions applied. To include this data, questions and mark schemes would have had to have been designed and analyzed for decision-making/CR marks and distinguished from the overall score for each student. This information would not only have been difficult to obtain logistically but is also protected and as such should not enter the research domain.

Table 7. Post-hoc test results of the difference of CR scores and total scores of the summative written exam in different cohorts for the CP3 dataset

CP3						95% confidence interval	
Dependent variable	(I) Year/ Teaching	(J) Year/ Teaching	Mean difference (I-J)	Std. error	Sig.	Lower bound	Upper bound
CR score	0 2013- 2014	1 2014- 2015	45.90	0.72	0.00	44.21	47.59
	0 2013- 2014	2 2015- 2016	26.94	0.74	0.00	25.20	28.67
	1 2014- 2015	2 2015- 2016	-18.96	0.73	0.00	-20.68	-17.25
Summative written exams	0 2013- 2014	1 2014- 2015	16.16	1.00	0.00	13.81	18.50
	0 2013- 2014	2 2015- 2016	13.59	1.02	0.00	11.19	15.99
	1 2014- 2015	2 2015- 2016	-2.57	1.01	0.03	-4.94	-0.19
CR score	0 2013- 2014	1 2014- 2015	5.23	0.58	0.00	3.88	6.58
	0 2013- 2014	2 2015- 2016	-7.27	0.59	0.00	-8.65	-5.88
	1 2014- 2015	2 2015- 2016	-12.49	0.58	0.00	-13.87	-11.12
Summative written exams	0 2013- 2014	1 2014- 2015	20.11	0.97	0.00	17.83	22.38
	0 2013- 2014	2 2015- 2016	9.69	0.99	0.00	7.36	12.02
	1 2014- 2015	2 2015- 2016	-10.42	0.98	0.00	-12.73	-8.12

Finally, in the quantitative study, many data sets were collected on a varying number of students completing different summative exams with different CR components in them. As such ANOVA testing was used to compare the group means instead of conducting a pairwise analysis of the difference of scores. This may have had ramifications on the analysis as we cannot assume that there was equal variance among the different groups.

Recommendations for future studies

Future research could focus on a range of areas associated with longevity of CR reasoning improvements, the differences in CR between cohorts with varying clinical experience and an analysis of students' CR using other measurements for analysis.

Future researchers could analyze the CR of participants as students and further in the future to ascertain whether early CR teaching results in a transient or ingrained change in the students' CR decision making and management.

A further recommendation of this study would be for future researchers to use outcomes other than the summative written exams to assess CR. As previously mentioned, OSCE data would be an interesting area to analyze, perhaps with the use of standardized mark sheets with predetermined CR questions and video analysis for the future breakdown. The use of patient outcomes as a marker for CR skills would also make for an interesting study.

Finally, as CR is also thought to improve unconsciously with time, researchers could study the difference in CR between cohorts with different levels of clinical experience to see whether this purported difference is borne out.

Conclusion

The main finding of this research was that the implementation of a CR teaching curriculum can have a significant impact on the CR of its students. This suggests that CR can be developed both before and during clinical practice.

To better facilitate an effective CR curriculum, attention should be focused on: trained and motivated teachers; facilitating changes in the teaching to better respond to the needs of the students; informed and focused feedback alongside effective assessments and strategies for CR teaching to be ingrained into the ongoing curriculum to better embed it into the system.

Acknowledgments

I would like to thank Ms. Nicki Keating, the Moodle team and the Rogo admin team for putting all the tests online. I would also like to thank my colleagues and all participants in the standard-setting meeting for contributing to the validation of this research and Mrs. Louise Sabir and the ethics committee team.

Conflict of interest

The authors declare that they have no conflict of interest.

Funding

The authors declare no funding for this research.

References

1. Drain PK, Mock C, Toole D, Rosenwald A, Jehn M, Csordas T, et al. The emergence of undergraduate majors in global health: systematic review of programs and recommendations for future directions. Am J Trop Med Hyg 2017;96:16-23.
2. Dawson NV. Physician judgment in clinical settings: methodological influences and cognitive performance. Clin Chem 1983;39:1468-80.
3. Stanovich KE. The Robot's Rebellion. Finding meaning in the age of Darwin. Chicago: University of Chicago Press; 2004.
4. Croskerry P. A universal model of diagnostic reasoning. Acad Med 2009;84:1022-8.
5. Neufeld VR, Norman GR, Feightner JW, Barrows HS. Clinical problem-solving by medical students: a cross-sectional and longitudinal analysis. Med Educ 1981;15:315-22.
6. Groves M, O'rourke P, Alexander H. The association between student characteristics and the development of clinical reasoning in a graduate-entry, PBL medical programme. Med Teach 2003;25:626-31.
7. Silva AL. Clinical reasoning development in medical students: an educational transcultural comparative study [dissertation]. University of Nottingham; 2013.
8. Gross M, Moore ZE, Gardner FL, Wolanin AT, Press R, Marks DR. An empirical examination comparing the Mindfulness Acceptance-Commitment approach and Psychological Skills Training for the mental health and sport performance of female student athletes. International Journal of Sport and Exercise Psychology 2018;16 431-51.
9. Hmelo CE. Cognitive consequences of problem-based learning for the early development of medical expertise. Teaching and Learning in Medicine. 1998;10:92–100.
10. Koh GC, Khoo HE, Wong ML, Koh D. The effects of problem-based learning during medical school on physician competency: a systematic review. CMAJ 2008;178:34-41.
11. Norman GR, Schmidt HG. Effectiveness of problem-based learning curricula: theory, practice and paper darts. Med Educ 2000;34:721-8.
12. General Medical Council. Tomorrow's doctors, outcomes and standards for undergraduate medical education. 2009;14-20.
13. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Acad Med 2004;79(10 Suppl):S70-81.