The Effect of Curricular Phase on the Clinical Reasoning Skills of Medical Students at the University of Nottingham

Swe Yin Khin-Htun[1], Jake Coles[1]

Corresponding author: Mr. Jake Coles mzyjsc@nottingham.ac.uk
Institution: 1. University of Nottingham
Categories: Comparative Medical Education, Educational Strategies, Teaching and Learning, Research in Health Professions Education, Undergraduate/Graduate

Received: 03/02/2019
Published: 08/03/2019

Abstract

Introduction
There are many factors affecting the development of clinical reasoning (CR) skills. It has been suggested that the stage of undergraduate study of the student can affect their level of CR. The marks of CR questions in the summative exams of students studying medicine at the University of Nottingham (UoN) at different curricular phase were used to ascertain further insight into the theory.

Method
Analysis was conducted to determine whether there are significant differences in outcomes of CR scores within the summative written exam across the three categories of the independent variables of school year (2011-2012, 2012-2013 and 2013-2014) for Clinical Phase (CP) 1; (2013-2014, 2014-2015 and 2015-2016) for CP3.

Results
As a summary, the CR scores of summative written exams of students were significantly higher in the CP3 dataset than in the CP1 dataset for 3 cohorts.

Conclusion
The CR scores from the summative written exams of students from all 3 different periods were significantly higher in the CP3 dataset than in the CP1 dataset. Therefore, the curricular phase and the exposure to clinical practices have an impact on the development of CR.

Keywords: Clinical reasoning; Clinical teaching; Clinical learning; Medical Education; Medical education research
Introduction

There are many factors affecting the development of CR skills. It has been suggested that the stage of Undergraduate (UG) study can affect the medical students' level of CR skills (Nafea, 2015). Neufeld concluded in his study that the CR process remains relatively constant from UG to postgraduate level, with only the measure of content in the students' hypothesis correlating with both outcome and educational level. This implies that medical students' knowledge is developing but the development of their CR skills remains relatively constant during this period (Neufeld et al., 1981). On the other hand, Groves and Da Silva stated that CR skills develop throughout the successive progress of the students (Groves, O'rourke and Alexander, 2003)(Silva, 2013). The studies of (Hmelo, 1998)(Norman and Schmidt, 2000)(Koh et al., 2008) also reported that the clinical phase of the curriculum has a strong impact on students’ CR.

By comparing the results of ‘CR questions in the summative written paper exams' after the first clinical phase (CP1) and after the final (third) clinical phase of students (CP3) at the University of Nottingham studying for a Bachelor of Medicine Bachelor of Surgery degree (M.B.,B.S), it is hoped to ascertain further insight as to whether the level of progression through curricular phase influences the CR skills of the students.

Structure of the Medical Curriculum 2016/2017

There are three routes at Nottingham by which medical students graduate as doctors: the five and six year UG courses and the graduate entry four-year course. On all three courses the educational objectives are the same: to acquire the knowledge, skills and behaviour to allow the graduate to practice as a new doctor on the foundation training programme with the M.B.,B.S. However, all routes of entry congregate and converge at the start of the clinical phase of the course, from that point on they are treated as one single cohort.

Clinical Phases

Clinical Phase 1

Students spend 17 weeks in one of five partner NHS sites. During this time students are taught the basic principles of taking a history and undertaking a clinical examination in both medicine and surgery placements. At the end of the attachment students undertake knowledge based multiple choice question examination and clinical skills examination.

Clinical Phase 2

Currently there are 4 x 10 week placements in this phase covering a range of specialties i.e. Obstetrics and Gynaecology, Child Health, Psychiatry, Care of Later Life, Community Based Medicine, Specials (Ophthalmology, ENT and Dermatology, each 2 weeks) and a special study module (4 weeks).

Clinical Phase 3

CP3 comprises two main components – the Advanced Clinical Experience (ACE) module and the Transition to Practice Module. ACE clinical placements rotate around partner NHS sites and comprise 8 weeks critical illness and general practice, 8 weeks musculoskeletal disorders and disease, 8 weeks medicine, 8 weeks surgery. At the end of the module students take the ACE knowledge and clinical examinations.

In this study, only CP1 and CP3 are compared because the subjects that cover in these two phases are similar whereas CP2 studies more specialities.

Methods

Methodology
The effects of different curricular phases on CR is measured by comparing the students’ marks of those questions in the summative written exam that are considered to be predominantly CR questions in CP1 and CP3. This study looked at three cohort groups in 2 stages: CP1 (2012, 2013, 2014) and CP3 (2014, 2015, 2016). CP1 2012 cohort becomes CP3 in 2014, CP1 2013 cohort becomes CP3 in 2015 and CP1 2014 cohort becomes CP3 in 2016. There are between 318-351 students per cohort. Statistical analysis will then be carried out in order to determine if there is an improvement in student's ability to answer CR questions.

### Table 1: Number of Participants in CP1 and CP3 by Cohort

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

The ethics committee confirmed that ethics approval was not required for this kind of project since it is classed as service evaluation.

**Research Designs and questions**

This leads to a research question and associated hypothesis.

**RQ:** Does the curricular phase of the curricular have an impact on the development of CR?

**H₀:** There is no significant effect on CR from the curricular phase as measured by CR score in the summative written exams.

**H₁:** There is a significant effect on CR from the curricular phase as measured by CR score in the summative written exams.

**Data Collection**

The data for this study is from increasingly challenging clinically orientated summative assessment of CP1 and CP3 written papers in different cohort years.

**Before the exam**

The written papers are reviewed and categorised at standard setting meetings as CR questions or non-CR questions before the exams. The total number of experts who participated in standard setting meetings varies between 15-25. The experts come from a wide range specialities and roles such as, consultants, GPs, clinical teaching fellows, medical educators, Director of clinical skills, module leads and some junior doctors.

When grading, if the experts have different opinions about CR components on the questions, the team maps these questions against the 3 statements from the ‘Outcomes for graduates from Tomorrow’s Doctors from GMC; (8c,8g and 14f), considered in Bloom's taxonomy of learning domains, and the matter is discussed until a mutually acceptable decision was made (GMC, 2015). The Bloom cognitive domain involves knowledge and the development of intellectual skills. This includes the recall or recognition of specific facts, procedural patterns, and concepts that serve in the development of intellectual abilities and skills. There are six major categories of cognitive processes, starting from the simplest to the most complex (Anderson and Krathwohl, 2001).

- Remember
• Understand
• Apply
• Analyse
• Evaluate
• Create

In these summative written exam papers, only the questions that assess the third category (apply) to the sixth category of Bloom’s cognitive processes are accepted as CR questions. In the final data, CR questions can cover: being given a history and being asked to formulate the diagnosis for each case; being given physical findings and being asked to choose the most likely diagnosis; being given investigation results and being asked to find a diagnosis and provide the treatment plan; being given a diagnosis and being asked to choose the matching case vignette or history; being given a history and being asked to match the investigation findings to the interpretation of the findings.

After the exam: Psychometric evaluation
Routine psychometric analysis of the medical exam is carried out for each exam paper in order to provide an assessment tool with high quality. ‘The post examination psychometric analysis of exam data’ is conducted using Item Response Theory (IRT) and Classical Test Theory (CTT). Knowledge papers are analysed using test-score reliability (Cronbach’s alpha), item discrimination index (ID), standard error of the measurement (SEM).

Frequency and discrimination (U-L) analysis and learning objective analysis are done for each paper. Item difficulty (p) and discrimination value (d) is calculated for each item. The reliability of the test was measured using generalizability (G) theory (Student× item). The hope is that students’ marks on the test reflect true or consistent differences between students with respect to their knowledge and skills. G study was used to address error variance among questions.

These papers are also reviewed by the external examiners and internal peers. They gave comment for some questions and these comments are carefully considered and decided whether action is needed to be taken or not. The final scores are provided after these steps are taken. The tables below provide the overall results for the CP1 and CP3. The CP3 has 2 written papers and comparison was made with both papers.

Table 2: CP1 Marks 2012 to 2014

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

CR questions in CP3 written exam are ranging from 85-151 marks worth, it accounts for 50%-79%.

Table 3: CP3 Marks Paper 1 and Paper 2 2014 to 2016

| CP3 | Marks | CR | NCR | Total | CR % of Total | Mark | Paper 2 | CR | NCR | Total | CR % of Total |
|-----|-------|----|-----|-------|---------------|------|---------|----|-----|-------|---------------|
| Paper 1 |       |    |     |       |               |      |         |    |     |       |               |
| Paper 2 |       |    |     |       |               |      |         |    |     |       |               |

Page 1 of 4
Data Analysis Methodology

This study involves the use of parametric statistical tests of independent sample t-test to address the research objective of the study. The parametric test requires the data of the dependent variable to be normally distributed. The dependent variables include the CR scores, the non-clinical reasoning (NCR) scores, and total scores of summative written exams for each of the dataset of CP1 and CP3 datasets. Normality testing was conducted by investigation of the skewness and kurtosis statistics and histogram to check the distribution of data of the different dependent variable.

An independent sample \( t \)-test was conducted to determine whether the outcomes as measured by the CR scores in the summative written exams were significantly different in the each of the three periods between the CP1 and CP3 dataset. For instance, year 2012 CP1 students are the same as the year 2014 CP3 students. This longitudinal analysis was conducted to determine how CP1 developed CR to CP3. A level of significance of 0.05 was used in the independent sample \( t \)-test.

Mean comparisons are conducted if significant differences are observed when comparing cohorts CR results dependent on their clinical phase.

Results/Analysis

Normality

To determine whether the data follows normal distribution, skewness statistics greater than three indicate strong non-normality and kurtosis statistics between 10 and 20 also indicate non-normality (Kline, 2005). As can be seen in Table 4 and 5, the data of all dependent variables in both the CP1 and CP3 datasets exhibited normality. Thus, the parametric statistical analyses can be conducted.

Table 4: Skewness and Kurtosis Statistics of CR Scores, NCR Scores, and Total Scores of Summative Written Exam for CP1 Dataset

|          | N  | Skewness Statistic | Kurtosis Statistic |
|----------|----|--------------------|--------------------|
| CR       | 1022 | 0.63               | -0.15              |
| NCR      | 1020 | 0.04               | -0.79              |
| Summative written exams | 1021 | -0.35              | -0.13              |

Table 5: Skewness and Kurtosis Statistics of Scores of CR Scores, NCR Scores, and Total Scores of Summative Written Exam for CP3 Dataset

|                      | N  | Skewness Statistic | Kurtosis Statistic |
|----------------------|----|--------------------|--------------------|
| CR Score (Paper 1)   | 1003 | 0.26               | -1.02              |
| NCR Paper 1, 41/192  | 1003 | 0.10               | -1.02              |
| Summative written exams (Paper 1) | 1003 | -0.28              | 0.49               |
| CR (Paper 2)         | 1003 | -0.11              | -0.01              |
Results of Analysis of CR Scores in the Summative Exams by Different Years between CP1 and CP3

This section presents the summaries of the data analysis using descriptive statistics, independent sample t-test. IBM® SPSS® Statistics Version 22 was utilized to conduct the data analysis.

Table 6 and 7 summarises the independent sample t-test results of the difference in the CR scores in the summative written exams in the period between the CP1 dataset in the year 2012 and CP3 dataset in the year 2014. The results of the independent sample t-test showed that the CR scores (t(636.93) = -70.31, p < 0.001) were significantly different in the two different periods for the CP1 and CP3 dataset. This was because the p-values were less than the level of significance value of 0.05. Mean comparison showed that the CR scores were significantly higher in the year 2014 of the CP3 dataset (M = 110.17; SD = 10.80) than in the year 2012 of the CP1 dataset (M = 57.66; SD = 8.58) by a mean difference of 52.50.

Table 6: Descriptive Statistics Summaries of CR Scores of the Summative Written Exam in Different Periods of No CR Teaching in Year 2012 for CP1 Dataset and Year 2014 for CP3 Dataset

| Year | N  | Mean  | Std. Deviation | Std. Error Mean |
|------|----|-------|----------------|-----------------|
| 2012 | 351| 57.66 | 8.58           | 0.46            |
| 2014 | 335| 110.17| 10.80          | 0.59            |

Table 7: Independent Sample t-test Results of Difference of CR Scores of the Summative Written Exam between Two Different Periods of No CR Teaching in Year 2012 for CP1 Dataset and Year 2014 for CP3 Dataset

| Levene's Test for Equality of Variances | t-test for Equality of Means |
|----------------------------------------|------------------------------|
| F          | Sig. | t  | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference |
| CR Score   |      |    |    |                |                |                      | Lower | Upper |
| Equal variances not assumed             | 13.69 | 0.00 | -70.31 | 636.93 | 0.00* | -52.50 | 0.75 | -53.97 | -51.04 |

* Significant difference at the 0.05 level of significance

Table 8: Descriptive Statistics Summaries of CR Scores of the Summative Written Exam in Different Periods of Partial CR Teaching in Year 2013 for CP1 Dataset and Year 2015 for CP3 Dataset

| Year | N  | Mean  | Std. Deviation | Std. Error Mean |
|------|----|-------|----------------|-----------------|
| 2013 | 344| 56.32 | 6.73           | 0.36            |
| 2015 | 350| 64.79 | 7.38           | 0.39            |

Table 9: Independent Sample t-test Results of Difference of CR Scores of the Summative Written Exam between Two Different Periods of Partial CR Teaching in Year 2013 for CP1
**Dataset and Year 2015 for CP3 Dataset**

| CR Score | Equal variances assumed | Levene's Test for Equality of Variances | t-test for Equality of Means |
|----------|-----------------------|----------------------------------------|-----------------------------|
|          | F                     | Sig. | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference |
| CR Score | 2.30                  | 0.13 | 692.00 | 0.00* | -7.92 | 0.54 | Lower | Upper |

* Significant difference at the 0.05 level of significance

Table 8 and 9 summarises the independent sample \( t \)-test results of the difference in the CR scores in the summative written exams in the period between the CP1 dataset in the year 2013 and CP3 dataset in the year 2015. The results of the independent sample \( t \)-test showed that the CR scores (\( t(692) = -14.76, p < 0.001 \)) in the summative written exam were significantly different in the two different periods for the CP1 and CP3 dataset. This was because the p-values were less than the level of significance value of 0.05. Mean comparison showed that the CR scores were significantly higher in the year 2015 of the CP3 dataset (\( M = 64.79; SD = 7.38 \)) than in the year 2013 of the CP1 dataset (\( M = 56.32; SD = 6.73 \)) by a mean difference of 7.92.

Table 10 and 11 summarises the independent sample \( t \)-test results of the difference in the CR scores in the summative written exams in the period between the CP1 dataset in the year 2014 and CP3 dataset in the year 2016. The results of the independent sample \( t \)-test showed that the CR scores (\( t(643.13) = -4.17, p < 0.001 \)) of the summative written exam was significantly different in the two different periods of full CR teaching for the CP1 and CP3 dataset. This was because the p-values were less than the level of significance value of 0.05. Mean comparison showed that the CR scores were significantly higher in the year 2016 of the CP3 dataset (\( M = 83.22; SD = 9.89 \)) than in the year 2014 of the CP1 dataset (\( M = 79.71; SD = 11.41 \)) by a mean difference of 3.51.

### Table 10: Descriptive Statistics Summaries of CR Scores of the Summative Written Exam in Different Periods of Full CR Teaching in Year 2014 for CP1 Dataset and Year 2016 for CP3 Dataset

| CR score | Year | N  | Mean | Std. Deviation | Std. Error Mean |
|----------|------|----|------|----------------|-----------------|
|          | 2014 | 327 | 79.71 | 11.41          | 0.63            |
|          | 2016 | 317 | 83.22 | 9.89           | 0.56            |

### Table 11: Independent Sample t-test Results of Difference of CR Scores of the Summative Written Exam between Two Different Periods of Full CR Teaching in Year 2014 for CP1 Dataset and Year 2016 for CP3 Dataset

| CR Score | Equal variances not assumed | Levene's Test for Equality of Variances | t-test for Equality of Means |
|----------|---------------------------|----------------------------------------|-----------------------------|
|          | F                        | Sig. | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference |
| CR Score | 8.85                     | 0.00 | 634.13 | 0.00* | 3.51 | 0.84 | Lower | Upper |

* Significant difference at the 0.05 level of significance

### Discussion

**Does the curricular phase of the curricular have an impact on the development of CR?**

This comparative study (compare: CR marks of CP1 2012 cohort becomes CP3 in 2014, CP1 2013 cohort becomes CP3 in 2015 and CP1 2014 cohort becomes CP3 in 2016) contributes to the body of knowledge regarding the sequential development of CR. This study looked at how students develop from CP1 to CP3 in general, not with curricular model. Analysis was conducted to determine how CR of CP1 developed to CP3: measured by CR scores in the summative written exams. CP3 students score higher than CP1 students in all 3 cohorts and it supported the
statement that CR is improved and developed from CP1 to CP3. Clinical training takes place in the third year and continues until the final year. It is suggested that this comparatively long time during the UG course can help students to develop the experience required for CR.

These results show that clinical practice has an important effect on the development of CR. This supports what has been reported previously in the literature (Hmelo, 1998)(Norman and Schmidt, 2000)(Boshuizen, 2003)(Koh et al., 2008). However, this finding contradicts partially the finding that CR remains relatively constant from medical school entry to practice by Neufeld et al (Neufeld et al., 1981). In order to confirm these findings, a longer longitudinal study would be required, with regular data collection moments, using more in-depth methodologies such as protocol analysis to access students’ CR process.

The previous research and literature shows that learning from the practice is not simple cause and effect phenomenon (Kolb, Boyatzis and Mainemelis, 2000)(Ericsson, 2004); educators should carefully plan students' experience in practice to ensure that these opportunities are created.

**Conclusion**

As a summary, the CR scores of summative written exam of students were significantly higher in the CP3 dataset than in the CP1 dataset for 3 cohorts. With these results, the null hypothesis that there is no significant effect in CR score from the curricular phase on outcomes as measured by CR score in the summative written exams is rejected. CP3 students score higher than CP1 students and it supports the statement that CR is improved and developed from CP1 to CP3.

**Limitation of the study**

In the quantitative study, there are many data sets with different numbers of students, different summative knowledge exam papers and different components of CR marks in each cohort. Due to the nature of the data we have, we are conducting an independent sample t-test wherein we are comparing the group means instead of conducting a pairwise analysis of the difference of scores. The total raw scores reflect the weight of the correct answers.

**Take Home Messages**

1. Clinical Reasoning Skills are shown to improve and develop through progression of the clinical phases of the Bachelor of Medicine Bachelor of Surgery degree at the University of Nottingham.

**Notes On Contributors**

Dr Swe Khin-Htun, M.B.,B.S, MRCS, FICS, MMedSci (Medical Education), MAcdMED, MFST, PhD student, is currently a Medical Education Fellow, Nottingham University Hospitals NHS Trust, Honorary Assistant Professor, School of Medicine, University of Nottingham. Her research interests focus on clinical reasoning.

Jake Coles, BMedSci, is a fourth-year medical student at the University of Nottingham with an interest in medical education.
Acknowledgements

Dr Swe Yin Khin-Htun is very grateful to her supervisor Professor Dennick who is always giving her guidance, advice and encouragement throughout her PhD. Dr Swe Yin Khin-Htun wanted to thank her colleagues and all participants in the standard setting meeting for contributing to the validation of this research. A big thank you to the students and experts from the Medical Education Centre, the School of Medicine UoN who gave time and contributed to this research.

Bibliography/References

Anderson, L. W. and Krathwohl, D. R. (2001) *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*. Longman.

Boshuizen, H. P. A. (2003) *Expertise development: The transition between school and work*, Expertise development: The transition between school and work. ISBN:90 358 2063 0.

Ericsson, K. A. (2004) 'Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains.', *Academic medicine: journal of the Association of American Medical Colleges*, 79(10 Suppl), pp. S70-81. https://doi.org/10.1097/00001888-200410001-00022

GMC (2015) 'Outcomes for Graduates', *General Medical Council*. ISBN:978-0-901458-72-8

Groves, M., O'rourke, P. and Alexander, H. (2003) ‘The association between student characteristics and the development of clinical reasoning in a graduate-entry, PBL medical programme’, *Medical Teacher*, 25(6), pp. 626–631. https://doi.org/10.1080/01421590310001605679

Hmelo, C. E. (1998) ‘Cognitive Consequences of Problem-Based Learning for the Early Development of Medical Expertise’, *Teaching and Learning in Medicine*. Lawrence Erlbaum Associates, Inc., 10(2), pp. 92–100. https://doi.org/10.1207/S15328015TLM1002_7

Kline, T. J. (2005) ‘Classical Test Theory: Assumptions, Equations, and Item Analyses’, in *Psychological Testing: A Practical Approach to Design and Evaluation*. 2455 Teller Road, Thousand Oaks California 91320 United States : SAGE Publications, Inc., pp. 91–107.

Koh, G. C.-H. et al. (2008) ‘The effects of problem-based learning during medical school on physician competency: a systematic review’, *Canadian Medical Association Journal*, 178(1), pp. 34–41. https://doi.org/10.1503/cmaj.070565

Kolb, D., Boyatzis, R. and Mainemelis, C. (2000) ‘Experiential learning theory: Previous research and new directions’, *R. J. Sternberg & L. F. Zhang, eds. Perspectives on thinking, learning, and cognitive styles.*

Nafea, E. (2015) *CR in dental students: A comparative cross-curricular study*. University of Nottingham.

Neufeld, V. R. et al. (1981) ‘Clinical problem-solving by medical students: a cross-sectional and longitudinal analysis.’, *Medical education*, 15(5), pp. 315–22. https://doi.org/10.1111/j.1365-2923.1981.tb02495.x

Norman, G. R. and Schmidt, H. G. (2000) ‘Effectiveness of problem-based learning curricula: theory, practice and
paper darts.’, *Medical education*, 34(9), pp. 721–8. https://doi.org/10.1046/j.1365-2923.2000.00749.x

Silva, A. L. (2013) ‘Clinical reasoning development in medical students: an educational transcultural comparative study’. Available at: http://eprints.nottingham.ac.uk/13623/

### Appendices

None.

### Declarations

*The author has declared that there are no conflicts of interest.*

*This has been published under Creative Commons "CC BY 4.0" ([https://creativecommons.org/licenses/by-sa/4.0/](https://creativecommons.org/licenses/by-sa/4.0/))*

### Ethics Statement

Ethical approval was requested from the UoN to conduct this study. However, the committee confirmed that ethics approval was not required for this kind of project since it is classed as service evaluation. Processing examination data, provided all is kept anonymous, does not directly involve people and is routinely carried out for psychometric purposes and course evaluation. But the application for the student performance dataset, the school of Medicine asked a statement from the ethics committee saying that ‘ethics approval is not required for processing data’. Consultation was made with UoN Faculty of Medicine & Health Sciences Research Ethics Committee and the “ethics” statement is issued.

### External Funding

This paper has not had any External Funding

---

AMEE MedEdPublish: rapid, post-publication, peer-reviewed papers on healthcare professions’ education. For more information please visit [www.mededpublish.org](http://www.mededpublish.org) or contact [mededpublish@dundee.ac.uk](mailto:mededpublish@dundee.ac.uk).