ORIGINAL ARTICLE

RELIABILITY OF A DIRECT OBSERVATIONAL TOOL TO ASSESS CLINICAL REASONING (DOT-CR)

[1]TABASSUM NAVEED
[2]NAVEED MAZHAR BHATTI

ABSTRACT

Medical students are usually left to learn clinical reasoning in the clinical years with direct patient interaction. Research evidence shows that the skill can be learned in preclinical years by employing paper case based methods of teaching without actual patient interaction. A direct observational tool to assess clinical reasoning was developed after literature search in consultation with and by receiving feedback from educationalists. Cronbach alpha reliability of the tool was excellent at 0.92. Interrater reliability by t-test showed lack of agreement between two raters.

The tool is reliable, useful and feasible. Further studies can be carried out on its validity and utility. Regular training of raters with timely feedback is required to improve interrater reliability.

Key Words: *Assessment tool, Case based learning, Clinical reasoning, Script theory and assessment tool.*

INTRODUCTION

Case based learning is a form of small group collaborative learning that commonly involves the use of patient cases associated with real life situations.[1] A case is used to stimulate learning and acquisition of knowledge, skills and attitudes. It is structured so that students develop the skills of clinical reasoning and critical thinking. In CBL, the learning objectives for the case are explicitly stated at the beginning of the case, and the learners can focus their learning on attaining the specific outcomes while developing the skills of clinical reasoning and critical thinking.[2] Cases are designed and organized to allow students to approach their learning as "Discovery Learning"[3] bringing together related content from the different domains of medical knowledge. Assessment of clinical reasoning abilities should be done throughout the training of students in different settings. Scenario based multiple choice questions, key feature test and script concordance test are some ways of theoretically assessing clinical reasoning ability. In the clinical years of students, these skills can be tested in most forms of workplace based assessment by instruments including the mini-CEX, Direct observation of procedural skills (DOPS), Chart stimulated recall (Case based discussion) and mini peer assessment (Mini-PAT) for multisource feedback.[4] Case based learning (CBL) is one instructional strategy used for integrating learning horizontally and vertically at Institute of Medical Sciences, Heavy Industries Education City Taxila (HITEC-IMS). CBL's are preceded by interactive lectures and directed self-study sessions. Cases for the sessions were developed and approved by National University of Medical Sciences (NUMS) academic council and were adopted by HITEC-IMS with addition of learning objectives in some cases. Literature showed paucity of evidence to assess clinical reasoning in preclinical MBBS students. We developed a Direct Observational Tool (DOT-CR) to assess clinical reasoning. This current study was planned to determine the reliability of Direct Observational Tool (DOT-CR) to assess clinical reasoning. We hypothesized that DOT-CR is a reliable tool to asses clinical reasoning in CBL sessions. We also hypothesized that there will be no difference between measurements of the two independent raters.

METHODOLOGY

A cross-sectional quantitative study was conducted at HITEC-IMS from January to March 2017. Approval for the study was received from institutional head. Participants in this study gave written informed consent according to the principles of the Declaration of Helsinki. It was ensured that all participants understood that their participation is voluntary and they can leave the study at any time without causing them any disadvantage.

Participants of the study were fresh graduates appointed as demonstrators. It was assumed that training

[1] **For Correspondence:** Tabassum Naveed, Assistant Professor and Head of Department of Medical Education, HITEC-Institute of Medical Sciences, Taxila Email: tabassumnaveed@hotmail.com
[2] Naveed Bhatti, Assistant Professor and Head of Orthodontics Department, HITEC - Institute of Medical Sciences, Taxila Email: naveedbhatti@hotmail.com

Received for Publication: May 4, 2017
Revised: May 13, 2017
Approved: May 24, 2017

of staff would be translated into students learning. Twenty-four items for the construct of clinical reasoning were identified from literature search. The items were reviewed by a group of six educationists and after discussions only ten items were kept. The items were piloted for readability, language and comprehension. Modifications recommended were incorporated and the tool was named as Direct Observational Tool to assess Clinical Reasoning (DOT-CR). Series of five workshops were conducted to train the young graduates working as teachers to conduct the CBL sessions. There were thirty participants. All participants were independently rated by two assessors on the DOT-CR assessment form in five sessions from January to March 2017. DOT-CR is a Likert scale with ten items. The ratings were on a scale of 1-6. Where 1=poor, 2=fair, 3= fairly good, 4=good, 5=very good, 6=excellent. Minimum marks for each participant were 10, Maximum were 60. Data was analyzed on SPSS version 22.0 for reliability of the scale by Cronbach's alpha and one sample t- test for inter rater reliability.

RESULTS

All sixty-two cases were included in analysis no case was excluded due to missing values. Minimum score on the scale was 17, maximum was 41. Mean score of the scale was 27.35±1 Standard error (Table 1). Cronbach's alpha reliability (α) of the DOT-CR scale was 0.92, which indicates a high level of internal consistency for our ten-item scale. Deleting item labelled as identification of key and new words may slightly increase the reliability to 0.93. This item also had a slightly low item total correlation of 0.21. The remaining items have high corrected item total correlation (>0.3) and high squared multiple correlation and should be kept in the scale. Deleting item "Manages time" may increase reli-

TABLE 1: DESCRIPTIVE STATISTICS OF DOT-CR SCALE

	N	Minimum	Maximum	Mean		Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic
Total marks	62	17.00	41.00	27.5	1	7.9
Valid N (list wise)	62					

TABLE 2: TOTAL STATISTICS

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Identification of key and new words	25.21	60.92	.21	.85	.93
Explained key and new words	24.97	56.09	.76	.92	.92
Generates hypothesis	24.34	46.98	.95	.91	.90
Explains mechanism	24.19	47.07	.85	.86	.91
Asks questions from the group	24.90	52.48	.73	.86	.92
Manages time	25.32	56.62	.51	.95	.93
interpersonal skills	23.67	47.40	.70	.97	.92
Maintains group dynamics	24.41	55.10	.69	.91	.92
Comes prepared with objectives and details of the course	24.5968	48.310	.850	.943	.907
Provides feedback	24.5645	44.348	.915	.929	.903

TABLE 3: ONE-SAMPLE STATISTICS

	N	Mean	Std deviation	Std. Error Mean
Difference	31	10.58	8.94	1.61

TABLE 4: ONE-SAMPLE t-TEST

	Test Value = 0					
					95% Confidence Interval of the Difference	
	T	Df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Difference	6.59	30	.000	10.58	7.30	13.86

ability of the scale to 0.93. Deleting the items "explains mechanism", "comes prepared with details and learning objectives of the course" and provides feedback" may decrease the reliability of the scale to 0.90, 0.91 and 0.90 respectively and hence should be kept. Items labelled as "generates hypothesis", "Explained key and new words", "asks questions from the group", "interpersonal skills" and "group dynamics" if deleted will not affect the reliability of the scale, it would remain the same at 0.92 (Table 2). *For the second objective,* we applied a one sample T-test to identify if there is a significant difference between measurements of the two raters. Results are shown in Table 3 and Table 4.

A one sample t-test was run to determine whether DOT-CR score in subjects was different between two raters. Mean difference between the two raters=10.58 and Standard deviation of the difference=8.94 (Table 3) This score is lower than the mean 27.35±1 Standard error of DOT-CR scale for the sample studied (Table 1). There is a statistically significant difference of 10.58 (95% confidence interval, 7.30,13.86), t (30) =6.59, p=.00, with effect size d=1.2 between the two measurements (Table 4). They do not show a useful level of agreement. The two raters don't agree. Further rater training is required.

DISCUSSION

Reliability can be estimated in several ways.[5,6] The method will depend upon the type of assessment instrument. Sometimes reliability is referred to as internal validity or internal structure of the assessment tool. For internal consistency questions or items are created that measure the same concept, and the difference among the answers is calculated. That is, the correlation among the answers is measured. The reliability of DOT-CR tool in this study was excellent at 0.92 which is comparable to a study in which a tool was developed to measure clinic social case studies to measure clinical reasoning. It demonstrated excellent reliability (Cronbach's alpha of 0.92) and interrater reliability between 0.2 and 0.3.[7]

The study showed significant differences between the two raters. The difference in inter rater reliability may be because the raters have different notions for assessing clinical reasoning. Without calibration of rating, changes might be due to differences in scoring principles[8] and not the intervention (DOT-CR) itself. These issues are usually addressed through on-going training and regular interrater reliability checks.

The construct of clinical reasoning can be further defined through further studies for its validity in different contexts and those involved in the data collection should be trained to ensure precision of opinion when measuring the construct of clinical reasoning and critical thinking. This will help to ensure satisfactory validity and interrater reliability.[9]

Clinical reasoning is the ability to integrate and apply different types of knowledge, weigh evidence critically and reflect upon the process used to arrive at a diagnosis.[10] Problems with clinical reasoning often occur because of inadequate knowledge, flaws in data gathering and improper approach to information processing. Some of the educational strategies which can be used to encourage acquisition of clinical reasoning skills are: exposure to a wide variety of clinical cases, activation of previous knowledge, development of illness scripts,[11] sharing expert strategies to arrive at a diagnosis, forcing students to prioritize differential diagnoses; and encouraging reflection, metacognition, deliberate practice[12] and availability of formative feedback.[13]

CONCLUSION

DOT-CR tool is a reliable, feasible and useful tool to measure clinical reasoning in case based learning in preclinical years. It is recommended that DOT-CR tool can be used for low-stakes summative or formative assessment. Faculty training should be an ongoing process to remove individual biases in scoring.

REFERENCES

1 ABC of learning and teaching in medicine. 2nd edition. Edited by Peter Cantillon, Diana Wood. Wiley-Blackwell; 2010. 10-14.

2 Wu B, Wang M, Grotzer TA, Liu J, Johnson JM. Visualizing complex processes using a cognitive-mapping tool to support the learning of clinical reasoning. BMC Med Educ. 2016; 22; 16(1): 216.

3 Ardi F, Lesleyanne H, Geoff McColl, Neville C. How clinical reasoning is taught and learned: Cultural perspectives from the University of Melbourne and Universitas Indonesia. BMC Med Educ. 2016; 16: 185.

4 Norcini J. Work place based assessment in clinical training. In Swanwick T, ed. Understanding Medical education. Edinburgh: AMEE, 2007.

5 Rafdzah Z, Awang B, Noorhaire N, Noor AI. A Systematic Review of Statistical Methods Used to Test for Reliability of Medical Instruments Measuring Continuous Variables. Iran J Basic Med Sci. 2013 Jun; 16(6): 803-07.

6 Bolarinwa OA. Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. Niger Postgrad Med J [serial online] 2015 [cited 2017 Apr 20]; 22: 195-201. Available from: http://www.npmj.org/text.asp?2015/22/4/195/173959.

7 Gohel M, Singh US, Bhanderi D, Phatak A. Developing and pilot testing of a tool for "clinicosocial case study" assessment of community medicine residents. Educ Health. 2016; 29: 68-74.

8 Wu SM, Whiteside U, Neighbors C. Differences in inter-rater reliability and accuracy for a treatment adherence scale. Cogn Behav Ther. 2007; 36(4): 230-39.

9 Im S, Kim DK, Kong HH, Roh HR, OH YR, Seo JH. Assessing clinical reasoning abilities of medical students using clinical performance examination. Korean J Med Educ. 2016 Mar; 28(1): 35-47.

10 Eva KW. What every teacher needs to know about clinical reasoning. Med Educ. 2005 Jan; 39(1): 98-106.

11 Charlin B, Boshuizen HP, Custers EJ, Feltovich PJ. Scripts and clinical reasoning. Med Educ. 2007; Dec; 41(12): 1178-84.

12 Norman G. Research in clinical reasoning: past history and current trends. Med Educ. 2005 Apr; 39(4): 418-27.

13 Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. Med Educ. Oct 2015; 49(10): 961-73.

CONTRIBUTIONS BY AUTHORS

1 Tabassum Naveed: Topic selection, Literature search, write up, results.

2 Naveed Mazhar Bhatti: Topic selection, Literature review, write up, references.