6-2020

Reliability and validity of WATCH: Warwick assessment instrument for clinical teaching

Sonia Ijaz Haider
Aga Khan University, sonia.haider@aku.edu

Khalid Masood Gondol
King Edward Medical University, Lahore, Pakistan

Muhammed Tariq
Aga Khan University, muhammed.tariq@aku.edu

Muhammad Furqan Bari
Dow University of Health Sciences, Karachi, Pakistan

Iqbal Azam Syed
Aga Khan University, iqbal.azam@aku.edu

Recommended Citation

Haider, S. I., Gondol, K. M., Tariq, M., Bari, M. F., Syed, I. A. (2020). Reliability and validity of WATCH: Warwick assessment instrument for clinical teaching. *Journal of the College of Physicians and Surgeons--Pakistan : JCPSP, 30*(6), 633-637.
Available at: https://ecommons.aku.edu/pakistan_fhs_mc_ded/40

Reliability and Validity of WATCH: Warwick Assessment InsTrument for Clinical TeacHing

Sonia Ijaz Haider[1], Khalid Masood Gondol[2], Muhammad Tariq[3], Muhammad Furqan Bari[4] and Iqbal Azam[5]

[1]Department of Educational Development, Faculty of Health Sciences, The Aga Khan University, Karachi, Pakistan
[2]Department of Surgery, King Edward Medical University, Lahore, Pakistan
[3]Department of Medicine, Faculty of Health Sciences, The Aga Khan University, Karachi, Pakistan
[4]Department of Pathology, Dow University of Health Sciences, Karachi, Pakistan
[5]Department of Community Medicine, Faculty of Health Sciences, The Aga Khan University, Karachi, Pakistan

ABSTRACT

Objective: To determine the reliability, validity, feasibility, acceptability and perceived educational impact of WATCH: Warwick Assessment insTrument for Clinical teacHing among doctors in Pakistan.

Study Design: Cross-sectional research study.

Place and Duration of Study: The College of Physicians and Surgeons, Pakistan, from September 2018 to August 2019.

Methodology: Postgraduate trainees were asked to rate the clinical teaching sessions, using WATCH, which consists of 15 items. Percentage was used to calculate gender and participation from different specialties. Inter-item correlations of 15 items with individual mean scores, standard deviations and Cronbach's Alpha were reported, including Friedman test, in order to observe the scores across multiple conditions. The Hotelling's T^2 test was used to test whether the answers provided by the study participants to the questionnaire were equal. Construct validity was determined using factor analysis while feasibility, acceptability, and educational impact was evaluated by seeking participants' feedback on five semi-structured questions.

Results: More than 80% ranked WATCH from good to excellent. Oveall 8 items were perceived as excellent, while 7 items received rating of good. Inter-item correlation ranged from 0.61 to 0.81. Cronbach Alpha was reported to be 0.975, with significant difference in mean scores of different items (Friedman's Chi-Square=4285.54; p<0.001). The Hotelling's T^2 test (21598.871 with F=185.249, df=14,2654; p<0.001) indicated that the mean values of the responses of different questions in the instrument were statiscally different. Factor analysis indicated one factor accounting for 73.97 of variance. The majority (93%) of the participants found the instrument easy to complete, most participants (91.5%) indicated it as an acceptable method of assessment, and majority (90.8%) perceived that it can improve clinical teaching.

Conclusion: WATCH demonstrated valid, reliable, feasible, and acceptable results for assessment of teaching of medical doctors and it can be used for providing feedback and rewarding teachers who excel in teaching.

Key Words: Clinical teaching, Validity, Reliability, Feasibility, Medical students, Residents, Doctors.

INTRODUCTION

Clinical teaching, also known as bedside teaching, is a significant component of the learning and teaching in medicine.[1] If a clinical teaching session is well planned and delivered effectively, it can enhance learning. However if not, it can raise a number of problems related to teaching and learning, including lack of critical thinking, analytical reasoning, passive learning, limited feedback and reflection.[2,3]

In the medical profession, doctors are responsible to teach junior trainees, and as they are increasingly involved in teaching, it is important to assess their teaching skills.[4] Evidence indicates that in addition to doctors, postgraduate trainees / residents are involved in teaching activities.[5] During their training, doctors are required to undertake a detailed assessment to determine their clinical competence;[6] however, in terms of teaching, no comprehensive assessment is conducted to determine their teaching effectiveness.[7] Effective clinical teaching is a challenging task and is influenced by a number of common problems, such as lack of clear objectives, focus on factual recall, inadequate feedback, and lack of opportunity for reflection and discussion.

In Pakistan, doctors and residents are facing similar problem as many regions of the world in formal teaching activities; and presently, no formal structured method is in place to

assess their teaching abilities.[8] Therefore, the need arises to assess clinical teaching so that efforts can be undertaken to facilitate teachers by providing feedback, and rewarding teachers who excel in teaching.

Thus, the aim of the current study was to determine the reliability and validity of WATCH: Warwick Assessment insTrument for Clinical teacHing among teaching doctors in Pakistan. The development and testing of WATCH has been previously published;[9] however, its reliability, validity, feasibility, acceptability and educational impact have not been explored among doctors working within Pakistan.

METHODOLOGY

The current research study was cross-sectional and completed from September 2018 to August 2019. The instrument, WATCH, consisted of 15 items and each item was supplemented with a rating scale. The ethical institutional committee approved the study. Data collection was facilitated by the CPSP. Informed consent was obtained from all participants and participation was voluntary. All postgraduate trainees enrolled in FCPS and MCPS programmes were asked to rate the teaching ability of their teachers using WATCH. Participants were also required to complete a semi-structured 5-item questionnaire to determine feasibility, acceptability and educational impact of the administered instrument. It was an online assessment, which was administered by the College of Physicians and Surgeons (CPSP), Pakistan from September 2018 to August 2019. All the trainees completed the assessment with constant reminders and follow-ups. Only those trainees, who left their training or took a break from their training for a specified period of time, did not complete the assessment.

Data was analysed using SPSS software. For data analysis, frequencies and percentages of countries, states/provinces and cities of institutes by number of enrolled students was generated. Similarly, frequencies and percentages of gender, specialty, residency year and on different items were also generated. Inter-item correlations of 15 items were computed and Cronbach's Alpha was also reported, including Friedman test, in order to observe the scores across multiple conditions. The Hotelling's T^2 test was used to test whether the answers provided by the study participants to the questionnaire were equal. Factor analysis of these 15 items with factor coefficients using PCA were also generated. Extraction of factors were done using explained variations and Eigen values. The criteria of factor extraction was Eigen value of greater than and equal to one. A p-value of less than 5% was taken as significant. Feasibility was calculated by time taken to complete the instrument; whereas, acceptability and educational impact were determined by eliciting participants' responses to the five semi structured questions.

RESULTS

A total of 6,268 doctors from different specialties and sub-specialties participated in the study. Among them, 3,427 (54.7%) were males and 2841 (45.3%) were females. The majority 6,218 (99.2%) of the participants were from Pakistan, followed by Nepal 27 (0.4%), Kingdom of Saudi Arabia 10 (0.2%), and lastly Ireland 8 (0.1%). Postgraduate trainees were in different years of training and specialties. The majority of the residents were from year III, 2,294 (36.6%), followed by year II, 1,234 (19.7%), year IV, 989 (15.8%) and finally, year V, 917 (14.6%). The greater number of the residents were from medicine 889 (14.2%), followed by obstetrics and gynecology 808 (12.9%), surgery 720 (11.5%) and pediatrics 624 (10%).

All the participants answered the 15 items of the instrument. More than 80% of the students ranked the instrument from good to excellent on different items. Eight items received excellent rating in regard to teaching, and included the following items:

A) Q02: Communicates effectively with trainees; B) Q03: Maintains polite and considerate attitude with trainees; C) Q04: Expresses enthusiasm towards teaching and learning; D) Q05: Teaches concepts and skills in an organised manner; E) Q12: Demonstrates professional and ethical conduct; F) Q13: Avoids favouritism, criticism and discrimination; G) Q14: Remains up to date with knowledge of developments in the field; H) Q15: Is a good role model for trainees.

Following eight excellent rating items, seven received good rating and included the following items:

A) Q01: Promotes active engagement of trainees during learning; B) Q06: Demonstrates clinical competence (sound analytical, diagnostic, therapeutic and reasoning skills) appropriate for the stage of training; C) Q07: Adjusts teaching to learning needs of trainees; D) Q08: Demonstrates appropriate use of teaching aids and resources (powerpoint, flipcharts, paper handouts etc); E) Q09: Provides regular feedback to trainees about their performance; F) Q10: Stimulates reflective skills among students; G) Q11: Is able to teach in diverse settings (bedside, operating theatre, wards) and involves patients in teaching, if relevant).

Inter-item correlation ranged from 0.61 to 0.81, indicating that there is good correlation among all items (Table I). A very strong value of Cronbach's Alpha (0.975) was observed with significant difference in the mean scores of different items (Friedman's Chi-Square=4285.54; p<0.001). Similarly the results from the Hotelling's T^2 test (21598.871 with F=185.249, df=14,2654; p<0.001) indicated that the mean values of the responses of different questions in the instrument was statiscally different implying that the participants showed different approaches to answer the items and that the responses were reliable (Table II). Factor analysis indicated only one factor, which accounted for 73.97 of the total variance of the 15 items (Table III).

In terms of feasibility, majority 5,831(93%) of the participants indicated that it was easy to complete the instrument and most of the participants 3,975 (63.4%) took less than 5 minutes to complete the instrument. In addition, majority 5,562 (88.7%) of the participants indicated that completing the instrument was not time consuming.

Table I: Inter-item correlation matrix.

	Q01	Q02	Q03	Q04	Q05	Q06	Q07	Q08	Q09	Q10	Q11	Q12	Q13	Q14
Q02	.795	1.000												
Q03	.677	.736	1.000											
Q04	.790	.745	.679	1.000										
Q05	.772	.747	.667	.782	1.000									
Q06	.765	.739	.668	.765	.816	1.000								
Q07	.783	.765	.685	.775	.803	.811	1.000							
Q08	.701	.674	.604	.702	.720	.735	.777	1.000						
Q09	.721	.709	.606	.712	.720	.725	.761	.751	1.000					
Q10	.766	.736	.659	.761	.773	.783	.800	.757	.798	1.000				
Q11	.744	.723	.650	.752	.770	.769	.776	.701	.713	.779	1.000			
Q12	.711	.725	.736	.721	.731	.736	.738	.669	.686	.741	.759	1.000		
Q13	.629	.663	.718	.630	.646	.647	.671	.593	.621	.654	.643	.719	1.000	
Q14	.675	.668	.633	.685	.708	.713	.689	.630	.629	.687	.709	.721	.651	1.000
Q15	.747	.759	.731	.747	.767	.756	.765	.682	.692	.751	.759	.778	.723	.768

Table II: ANOVA with Friedman's test.

		Sum of squares	df	Mean square	Friedman's Chi-square	Sig
Between people		93371.777	6267	14.899		
Within people	Between items	1705.277[a]	14	121.805	4285.542	<0.001
	Residual	33212.457	87738	.379		
	Total	34917.733	87752	.398		
Total		128289.510	94019	1.365		

Grand Mean = 3.64. [a]Kendall's coefficient of concordance W = .013.

Table III: Factor analysis: Total variance explained.

Component	Initial Eigen values			Extraction sums of squared loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	11.096	73.973	73.973	11.096	73.973	73.973
2	.646	4.307	78.280			
3	.419	2.792	81.071			
4	.402	2.677	83.749			
5	.287	1.912	85.661			
6	.284	1.890	87.551			
7	.272	1.814	89.365			
8	.240	1.600	90.965			
9	.229	1.526	92.491			
10	.209	1.392	93.883			
11	.201	1.338	95.220			
12	.192	1.280	96.500			
13	.179	1.194	97.694			
14	.174	1.159	98.852			
15	.172	1.148	100.000			

Extraction Method: Principal component analysis.

Sonia Ijaz Haider, Khalid Masood Gondol, Muhammad Tariq, Muhammad Furqan Bari and Iqbal Azam

In terms of acceptability, 5,733 (91.5%) of the participants indicated that it is an acceptable method of assessment. In terms of educational impact, most 5,694 (90.8%) of the participants agreed that the use of the instrument will improve their clinical teaching.

DISCUSSION

The aim of the present study was to validate a clinical teaching instrument (questionnaire) among doctors in Pakistan. This instrument, WATCH, was validated in England, United Kingdom, to measure clinical teaching of postgraduate trainees;[9] however, in order to be used for assessment of clinical teaching, it is required to be validated in the context to be used to determine its relevance.[10]

The findings of the present study indicate that all the items of the instruments can be used to assess clinical teaching among doctors in Pakistan. Overall participants rated 8 items as excellent in terms of assessment of clinical teaching of their teachers. These items are related to overall conduct, communication skills, and professional attitude of their clinical teachers. The remaining 7 items were rated good pertaining to effective delivery of clinical teaching. It would have been ideal if participants had also rated them as excellent, but the rating of 'good' indicates that teachers need to work on their teaching skills in order to promote effective teaching. This is not surprising because doctors are not trained to be teachers, they are trained to be doctors.[11] They are not required to undertake any formal teaching certification or course to be able to comprehend the effective teaching strategies. In their professional career of becoming a doctor, they learn teaching through ad-hoc methods.[12-14] Although professionalism, communication skills and ethics are taught formally in the medical school,[15,16] no teaching strategies are taught and the need for it is well supported by the existing literature.[17,18]

The results of inter-item correlation indicated internal consistency of all the items — that is the scores of the individual items of the survey are measuring a single concept only — which is clinical teaching. The reliability coefficient Cronbach Alpha was 0.975, further confirming the reliability of the instrument used. This is similar to the findings of the WATCH, which was used in England in which the reliability coefficient was 0.92.[9]

Factor analysis was used to determine the construct validity, which showed that there is only one factor which the instrument is measuring and that factor is clinical teaching only. This corroborates the findings of the earlier study, in which clinical teaching was among the three factors found on the WATCH.[9]

Majority of the participants found the instrument as a feasible method to assess clinical teaching because it takes less than 5 minutes to complete it. In a busy clinical setting in which patient care is priority to clinical teaching, anything which takes more than 5 minutes is less likely to be used. This is confirmed by existing evidence in which clinical teaching instru-

ments which have demonstrated valid and reliable findings, but those which took more than 5 to 10 minutes were not used by the stakeholders.[19]

Moreover, a greater number of the participants found WATCH an acceptable method of assessment. Clinical teaching assessment can be challenging to measure when residents and physicians are working in busy environment. The fact that 90.8% and above indicated its acceptability, it shows that it can be used easily in a busy clinical setting.

In addition, most of the participants also agreed in terms of perceived educational impact of the instrument. The implication is that the long term aim is to use this instrument as a formative method of assessment of postgraduate trainees;[20] and therefore, the results of present study and existing literature[9,21] indicate that in absence of formal training course, this instrument can be used to facilitate clinical teachers in acquiring effective clinical teaching skills, thorough assessment, and feedback.

CONCLUSION

Warwick Assessment insTrument for Clinical teacHing (WATCH) can be used for assessment of clinical teaching in Pakistan. The instrument has demonstrated valid, reliable, feasible and acceptable results and has potential for positive educational impact on doctors in Pakistan. It can be used for providing feedback and rewarding teachers who excel in teaching.

ETHICAL APPROVAL:
The ethical institutional committee approved the study.

PATIENTS' CONSENT:
Informed consents were obtained from all participants and the participation was voluntary.

AUTHORS' CONTRIBUTION:
SIH, KMG: Substantial contribution to conception, drafting and final approval.
MT: Substantial contribution to acquisition, drafting and final approval.
MFB: Substantial contribution to interpretation, drafting and final approval.
IA: Substantial contribution to data analysis, drafting and final approval.

CONFLICT OF INTEREST:
Authors declared no conflict of interest.

REFERENCES

1. Peters M, Ten Cate O. Bedside teaching in medical education: A literature review. *Perspect Med Educ.* 2014; **3(2)**: 76-88.

2. Hoffman KG, Donaldson JF. Contextual tensions of the clin-

ical environment and their influence on teaching and learning. *Med Educ* 2004; **38(4)**:448-54.

3. Dolmans DH, Wolfhagen IH, Heineman E, Scherpbier AJ. Factors adversely affecting student learning in the clinical learning environment: A student perspective. *Educ Health (Abingdon)* 2008; **21(3)**:32.

4. Council GM. The Doctor as Teacher. *London* 1999.

5. Edwards JC, Friedland JA, Bing-You R. Residents teaching skills. New York: Springer Publishing; 2002.

6. GMC. Tomorrow's doctors: General Medical Council; 2003. Available from: http://www.gmc-uk.org/education/ under graduate/tomorrows_doctors.asp.

7. Zabar S, Hanley K, Stevens DL, Kalet A, Schwartz MD, Pearlman E, *et al.* Measuring the competence of residents as teachers. *J Gen Int Medicine* 2004; **19(5)**:530-3.

8. Siddiqui FG, Shaikh NA. Challenges and Issues in medical education in pakistan. *J Liaquat Uni Med Health Sci* 2014; **13(3)**:91-2.

9. Haider SI, Johnson N, Thistlethwaite JE, Fagan G, Bari MF. WATCH: Warwick assessment instrument for clinical teaching: Development and testing. *Med Teach* 2015; **37(3)**:289-95.

10. Al Ansari A, Strachan K, Hashim S, Otoom S. Analysis of psychometric properties of the modified SETQ tool in under-graduate medical education. *BMC Med Educ* 2017; **17(1)**:56.

11. British, Medical, Association. Doctors as Teachers. BMA Marketing & Publications 2006.

12. Seabrook MA. Medical teachers' concerns about the clinical teaching context. *Med Educ* 2003; **37(3)**:213-22.

13. Khan IA. Bedside teaching-making it an effective instruc-tional tool. *J Ayub Med Coll Abbottabad* 2014; **26(3)**:286-9.

14. Ah-Kee EY, Scott RA, Shafi A, Khan AA. How can junior doctors become more effective teachers? *Adv Med Educ Pract* 2015; **6**:487-8.

15. Jagzape TB, Jagzape AT, Vagha JD, Chalak A, Meshram RJ. Perception of medical students about communication skills laboratory (csl) in a rural medical college of central india. *J Clin Diagn Res* 2015; **9(12)**:JC01-4.

16. Mahajan R, Aruldhas BW, Sharma M, Badyal DK, Singh T. Professionalism and ethics: A proposed curriculum for under-graduates. *Int J Appl Basic Med Res* 2016; **6 (3):**157-63.

17. Busari JO, Scherpbier, AJJA. Why residents should teach: A literature review. *J Postgrad Med* 2004; **50(3)**:205-10.

18. Boursicot K, Etheridge L, Setna Z, Sturrock A, Ker J, Smee S, *et al*. Performance in assessment: Consensus statement and recommendations from the Ottawa conference. *Med Teach* 2011; **33(5)**:370-83.

19. Beckman TJ, Lee MC, Mandrekar JN. A comparison of clinical teaching evaluations by resident and peer physicians. *Med Teach* 2004; **26(4)**:321-5.

20. AlHaqwi AI, Taha WS. Promoting excellence in teaching and learning in clinical education. *J Taibah University Med Sci* 2015; **10(1)**:97-101.

21. Nejatdarabi H, Amini M, Yarahmadi J, Bikineh P, Kaveh M, Gholampoo H. Validity and reliability of the Persian version of WATCH questionnaire in assessing the clinical learning environment. *Biomed Res* 2018; **29(17)**:3356-61.

••••☆••••