RESEARCH ARTICLE

Validation of Patient Perception Instruments for Junior Doctor Performance: a Factor Analysis

Mia Kusmiati,[1,2] Rafidah Bahari,[3] Noor Aini Abdul Hamid,[4] Suhaila Sanip,[5] Ova Emilia[6]
[1]Cyberjaya University College of Medical Sciences, Cyberjaya, Malaysia,
[2]Department of Biochemistry, Faculty of Medicine, Universitas Islam Bandung, Bandung, Indonesia,
[3]Department of Psychiatry, Cyberjaya University College of Medical Sciences, Cyberjaya, Malaysia,
[4]Department of Biochemistry, International Medical School, Management and Science University,
Shah Alam, Malaysia, [5]Department of Medical Education, Universiti Sains Islam Malaysia,
Kuala Lumpur, Malaysia, [6]Department of Obstetrics and Gynecology, Faculty of Medicine,
Public Health and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia

Abstract

The patient is an essential stakeholder within the medical healthcare system and an important stakeholder of the medical education program. The patients should be able to assess the performance of junior doctors in general practitioner residency to ensure their competency. Some instruments of patient assessment are available, but they do not adapt to local needs and context. This study aims to validate newly developed evaluation instruments from the patient's perspective against the performance of a junior doctor in a teaching hospital. Fifty patients from outpatient clinics of internal medicine of two teaching hospital Faculty of Medicine Universitas Islam Bandung were selected to fill out the questionnaire in September–October 2018. The tool consists of 20 items and used a 4-point Likert scale of strongly disagree, disagree, agree, and strongly agree. The SPSS version 21 have used to extract the data as the principal axis factoring of analysis. Oblimin rotation method was applied with Kaiser normalization to simplify and describe the data structure. The detailed analysis identified five factors based on the initial eigenvalue >1. Patient perception instruments of junior doctor performance (PIJDP) showed that five constructs extracted explained 81.27% of the variance of them. Constructs were namely: humanism, responsibility-accountability, communication-empathy, altruism, and pleasant manner. Construct validity achieved after the PIJDP run fifteen times, and consistency internal with Cronbach's alpha was 0.95. In conclusions, the PIJDP could be used to assess the performance of junior doctors and could make a novel contribution to the development of medical education.

Key words: Construct, evaluation, patient, performance, psychometric

Validasi Instrumen Persepsi Pasien terhadap Kinerja Dokter Muda: Sebuah Analisis Faktor

Abstrak

Pasien merupakan *stakeholder* kunci dalam sistem pelayanan kesehatan dan *stakeholder* penting dalam program pendidikan kedokteran. Pasien dapat menilai kinerja dokter muda dalam pemagangan umum untuk memastikan kompetensi mereka. Beberapa instrumen penilaian pasien sudah dibuat, namun mereka tidak diadaptasi terhadap kebutuhan dan konteks lokal. Penelitian ini bertujuan memvalidasi instrumen evaluasi yang baru dikembangkan menurut perspektif pasien terhadap kinerja dokter muda di rumah sakit pendidikan. Lima puluh pasien dari klinik rawat jalan penyakit dalam dua rumah sakit pendidikan Fakultas Kedokteran Universitas Islam Bandung dipilih untuk mengisi kuesioner pada September–Oktober 2018. Kuesioner berisi 20 item yang menggunakan Skala Likert empat poin dari sangat tidak setuju, tidak setuju, setuju, dan sangat setuju. SPSS versi 21 digunakan untuk menganalisis data melalui *principal axis factoring*. Metode rotasi oblimin dengan normalisasi Kaiser diaplikasikan untuk menyederhanakan dan menjelaskan struktur data. Hasil analisis mengidentifikasi lima faktor berdasar atas *eigenvalue* awal >1. Instrument persepsi pasien terhadap kinerja dokter muda (PIJDP) menggambarkan 5 *construct* yang diekstraksi sebesar 81,27% dari varian indikator dapat dijelaskan oleh faktor yang terbentuk. Faktor tersebut adalah humanisme, tanggung jawab-akuntabilitas, komunikasi-empati, altruisme, dan sifat menyenangkan. Kesahihan *construct* dicapai setelah PIJDP diulang lima belas kali dan konsistensi internal dengan *Cronbach's alpha* sebesar 0,95. Simpulan, PIJDP dapat digunakan untuk menilai kinerja dokter muda dan dapat memberi kontribusi baru dalam pengembangan pendidikan kedokteran.

Kata kunci: *Construct*, evaluasi, kinerja, pasien, psikometrik

Correspondence: Mia Kusmiati. Department of Biochemistry, Faculty of Medicine, Universitas Islam Bandung. Jln. Tamansari No. 22, Bandung 40116, West Java, Indonesia. Mobile: +6281221524771. E-mail: mia.kusmiati@unisba.ac.id

Introduction

An instrument of the junior doctor's performance is necessary to assess the outcome of education. The result of education is one of the curriculum components in addition to the teaching method, assessment method, and the course content.[1,2] Patient satisfaction's questionnaire has been used to assess health services in a hospital. Assessment of patient satisfaction usually covers things relates to the hospital. They include clinical services, hour or day of operation, waiting time, friendliness of staff, amount of time with students officer and supervising doctor, lab services, and medication provided.[3] However, the assessment of the performance of a young physician taking a clinical clerkship program by the patients has not commonly performed. The patient is one of the key stakeholders within the medical curriculum, besides medical student, alumni, academic staff and employers.[4]

Evaluation of medical graduates' performance on the education needs assessment from the patient's perspective because they are the main stakeholder in a health care system.[5] Moreover, patient satisfaction is an essential aspect of the health care system to improve patient outcome. The success of health care services quality has determined by the satisfaction level of the patient as a user of the health provider.

The questionnaire is subjective measurement scales providing one or several scores based on sum (or mean) of responses to items (binary or ordinal variables). The construct validity is one of validation type that used to assess whether an instrument valid or not as a measurement tool. This validity obtained by conducting the exploratory factor analysis, and also can be used to evaluate the psychometric properties of an instrument.[6,7] Similarly, concurrent validity often was undertaken to validate the subjective measurement scale of the same matters from data of the other design. Validity refers to the degree to which a questionnaire measures the concepts of interest accurately.[8]

This study aims to validate newly developed evaluation instruments from the patient's perspective against the performance of a junior doctor in a teaching hospital.

Methods

The study used cross-sectional design for two months in September–October 2018 in two teaching hospital of the Faculty of Medicine Universitas Islam Bandung; Al Ihsan Regional General Hospital and Al Islam Hospital. Al Ihsan Hospital is the primary referral hospital in the West Java province, Indonesia, while Al Islam Hospital is the type B hospital in Bandung city that serves both public patient and patient of insurance health coverage.

This study is part of the exploratory mixed method design, where the qualitative phase has already finished. From the qualitative finding, we achieved four themes in terms of humanism, responsibility, communication, and altruism. There is 2-factor analysis conducted, exploratory factor analysis (EFA) and confirmatory factor analysis (CFA).

Fifty patients at the outpatient clinic of internal medicine at two teaching hospital; Al Ihsan Regional General Hospital and Al Islam Hospital were selected to conduct an EFA. The sample size is determined based on a factor to participant ratio of 10:1, yielding 50 samples for five elements.

The 153 patients of Al Ihsan Regional General Hospital and Al Islam Hospital filled out the questionnaire of EFA result to conduct CFA. Determination of sample size based on the estimation proportion population as follows.

$$n = \frac{z_{1-\alpha/2}^2 P(1 - P)N}{d^2(N - 1) + z_{1-\alpha/2}^2 P(1 - P)}$$

Notes: N: number of population, n: number of sample needed (sample size), $z_{1-\alpha/2}$: confidence interval 95% (1.96), P: anticipated population proportion of patient (0.54),[9] d: absolute precision required (0.07)

Based on the result of the qualitative finding, five variables have observed. The five variables are humanism, responsibility-accountability, communication, altruism, and pleasant attribute. The four patients of surgery inpatient clinic on the recovery phase was selected to conduct a focus group discussion (FGD) for 60 minutes in the primary teaching hospital of Faculty of Medicine Universitas Islam Bandung, Al Ihsan Regional General Hospital.

Figure 1 described the flow of the research. In determining whether the item on the instrument adequately represents the factor assessed, an exploratory structured analysis conducted. Fifty patients at the outpatient clinic of internal medicine at two teaching hospital were randomly selected to participate in this study. They were
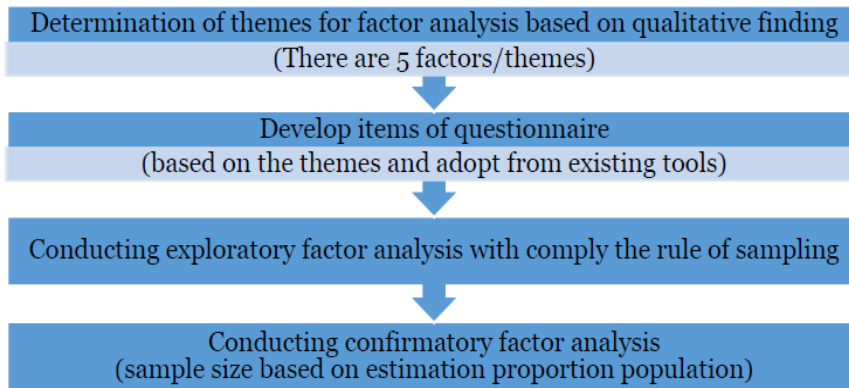
Figure 1 Flow of the Research

asked to fill out the questionnaire that consists of 20 items using four-point of Likert scale, namely strongly disagree, disagree, agree, and strongly agree. The Kaiser-Myer-Olkin (KMO) statistic and Bartlett's test of sphericity were implemented to obtain sampling adequacy (factorability of the data).[7] The KMO value, which was above 0.6, showed an adequate sample and suggested suitable for factor analysis.[10] Using SPSS version 21, a principal axis factoring was conducted to simplify and describe the structure of data. An oblique rotation method, oblimin rotation with Kaizer normalization, performed to confirm the data extraction that grouped in the same factor. The early four factors or domains were humanism, accountability, excellence, and
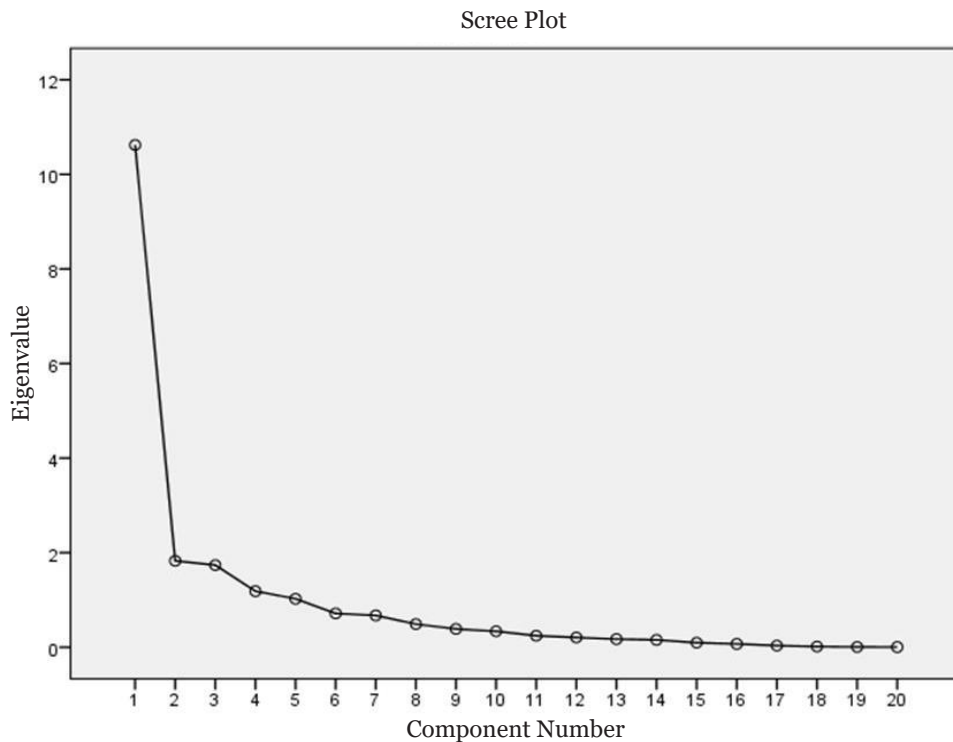

Figure 2 Factors in the 5 Components that Explained Variable based on the Initial Eigenvalue >1

altruism—the items with a higher value than 0.6-factor loading grouped in the same factor or domain. Factor loading has been used to predict the correlation between factors and to determine the highest correlation factor.[7,11] The composition of the domains was five items of humanism, five issues of accountability, five elements of altruism, and five items of excellence.

Stages of scale development of the factor analysis included a) determining the construct of an instrument based on the qualitative finding on three themes, following competencies of skills as a doctor, communication ability, and professional behavior when serving the patient. The primary scales of the instrument has extracted from the result of FGD (4 patients of the surgery ward on the recovery phase). In this study, the development of questionnaires of patient perspective based on previous qualitative research. According to Creswell and Guetterman,[12] and Creswell,[13] qualitative findings could be used to develop an instrument of evaluation.

There are 4 subscales, altruism, humanism, accountability, and excellence based on article,[6] b) developing the 20 items of initial instrument onto 4 subscales which based on 4-point Likert scale, namely strongly disagree, disagree, agree, and strongly agree, c) completing the questionnaire by 50 patients of outpatient clinic of internal medicine at two teaching hospital of the Faculty of Medicine Universitas Islam Bandung. They are 18 patients of Al Ihsan Regional General Hospital and 32 patients of Al Islam Hospital. They were asked to rate items of questionnaire based on a 4-point Likert scale (from 1-strongly disagree to 4-strongly agree).

Confirmatory factor analysis was conducted to 153 patients at the outpatient clinic of internal medicine and surgery in two teaching hospitals. The 20 items of a patient perception instruments of junior doctor performance (PIJDP) was distributed to them to rate item based on a 4-point Likert scale and then analyzed by using the Linear Structural Relations (LISREL) program.

Collecting data used purposive convenience sampling until the quota of 50 samples reached. Data has collected in 1 September–30 October 2018 at two teaching hospital; Al Ihsan Regional General Hospital and Al Islam Hospital. The instrument used is a PIJDP as a result of scale development of EFA. For a sampling of CFA, it used purposive convenience sampling in two months from October to November 2018.

An EFA conducted using SPSS version 21. The data extracted by principal axis factoring method and oblimin rotation method. Before extracting the factor, the KMO and Bartlett's test of sphericity were tested respectively to get sampling adequacy and to determine the probability that correlations in a matrix are zero.[10] Some article gave that the sample size was determined based on factor and not variable. The sample size is ideally 10-20 participant per element,[14] but the minimum sample 30 to 100 can be accepted.[12] This study used 10 participants for five factors, yielding 50 participants. The total variance explained, and scree plots (Figure 2) were used to extract the factor. The total variance explained is in Table 1.

Confirmatory factor analysis used the LISREL program. The data confirmed for each factor and indicator variable of their construct. The item that had value more than 1.96 (t value>1.96) was

Table 1 Total Variance Explained based on the Initial Eigenvalue >1

Factors	Initial Eigenvalues			Extraction Sums of Squared Loadings			Total of Rotation Sums of Squared Loadings
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	
1	10.621	53.104	53.104	10.621	53.104	53.104	10.621
2	1.828	9.141	62.245	1.828	9.141	62.245	1.828
3	1.736	8.680	70.925	1.736	8.680	70.925	1.736
4	1.186	5.928	76.853	1.186	5.928	76.853	1.186
5	1.025	5.124	81.977	1.025	5.124	81.977	1.025

Factor 1: altruism; Factor 2: communication and empathy; Factor 3: pleasant manner; Factor 4: humanism and medical treatment; and Factor 5: responsibility-accountability

Table 2 Five Factors Solution based on Extraction Method of Oblimin Rotation of 50 Participants

Items		Factors				
		1	**2**	**3**	**4**	**5**
Q1	They are capable deal with my disease under supervision.	0.616				0.927
Q2	In serving patients, they do wholeheartedly until completely appropriate given the hospital authority.					0.800
Q3	The Ministry is done very nice.	0.517	−0.547			0.761
Q4	They are highly skilled and trained.	0.595				0.636
Q5	I do not hesitate against their ability in performing medical action standard.	0.673		0.564	−0.721	0.654
Q6	They are very friendly and enthusiastic when listening our complaints.		−0.842			0.652
Q7	They are very painstaking served us.	0.525	−0.895			0.501
Q8	They always give good advice related to our health.	0.600	−0.762	0.530		
Q9	They were polite when serving us.	0.553	−0.552	0.644		0.563
Q10	They know well the health problems we face.	0.865		0.524		0.595
Q11	During treatment, communication is easy to understand.	0.559	−0.709			0.508
Q12	Instructions given are clear and simple.	0.911	−0.574			0.639
Q13	They always involve our family in decision making during treatment if needed considerations of family.	0.781				
Q14	They look professionally serving us.	0.693	−0.500		−0.794	0.661
Q15	They are patient and friendly in serving us.	0.634	−0.691		−0.699	0.563
Q16	Young doctors always give precedence to the interests and our safety.	0.759		0.664		0.726
Q17	They do not hesitate to refer us when felt not his authority.	0.529		0.837		
Q18	I am pleased involved in research that they do for the development of medical science.			0.573		
Q19	I am not afraid of my complaints will be known by the other doctor.	0.564		0.928		
Q20	I served by them in ways that are fun.	0.870		0.666		0.639

Factor 1: altruism; Factor 2: communication and empathy; Factor 3: pleasant manner; Factor 4: humanism and medical treatment; and Factor 5: responsibility-accountability

considered significant and valid to describe its indicator constructed.

This study already had ethical approval from the Health Research Ethics Committee of Faculty of Medicine Universitas Islam Bandung with letter-number: 005/Komite Etik.FK/VI/2017.

Results

Demographic information of EFA yielded 62% were males (n=31) and 38% were females (n=19). The average age was 51.96 years old in male respondent and 42.89 years old in the female
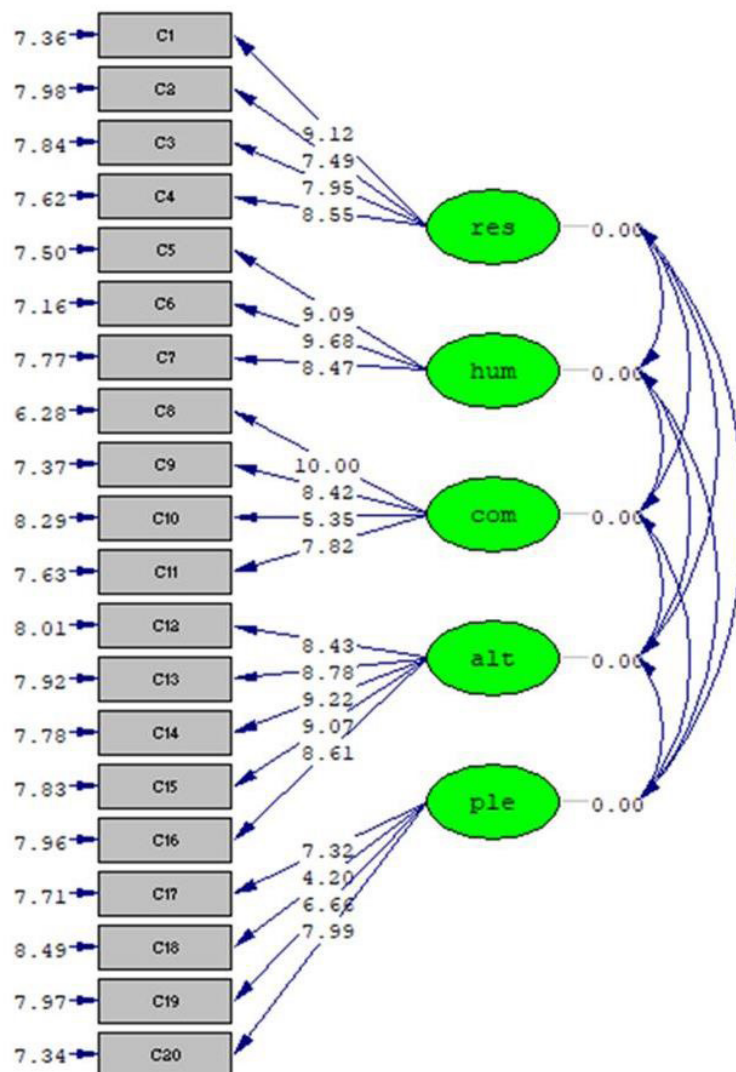
Figure 3 Result of Confirmatory Factor Analysis to 5 Latent Variables of 153 Patients

respondent. Demographic information of CFA resulted 44.44% were females (n=68), and 55.56% were males (n=85) with the mean of age was 53.97 years old.

As saw in Table 1, five factors had eigenvalue more than one. The factors that had eigenvalue more than one retained in this study. Then, the rotated component matrix was implemented to extract item in each factor.[14] Items inserted into each factor according to their values and nature. Items that had values more than 0.6 in every factor loaded at the same.[11]

Review of the factor structures suggested that 5 factors composition was the most appropriate structure. The 5-factor solution initially showed a minimal cross loading and higher commonalities than the 4-factor solution. The KMO value was 0.810, and Bartlett's test of sphericity was 0.000 ($p<0.001$, $X^2=1{,}018.134$) indicating sampling adequacy and significance of this construct validity. The minimum sample size of the performed analysis was 10:1 for the participant to factor ratio. The five factors at a 10:1 participant to factor ratio have been extracted, creating a sample size of 50.[14–16] Extraction of five elements composing 20 items of evaluation instruments showed 81.97% of the variance between them after 15 iterations (Table 1). Table 2 demonstrated

that items below were grouped based on factor loading value more than 0.6, except item q18 that has factor loading value 0.573. This result is following the study of Velicer and Fava[17] that item has a factor loading between 0.4 to 0.7 is low to moderate, and the item has a factor loading ≥0.8 is considered high. The item has the factor loading <0.40 showed that that factor either not related to the other items. It might be an additional factor that should explore more.[7]

Based on the result of component number from the initial eigenvalue more than one, thus it was found five factors, namely 1) altruism, 2) communication and medical treatment, 3) a pleasant manner, 4) humanism, and 5) responsibility-accountability. The composition of factors was five items for altruism, four items for responsibility, four items for communication and medical treatment, four items for a pleasant manner, and three items for humanism.

Construct validity achieved after the PIJDP run fifteen times, and consistency internal with Cronbach's alpha was 0.95. The Cronbach's alpha value of each factor was: 0.948 of humanism, 0.947 of accountability, 0.947 of excellence, and 0.948 of altruism, which showed all factors have high reliability and can replicate for another study.

As shown in Figure 3, all items showed a significant relationship to each latent variable (responsibility-accountability, humanism and medical treatment, communication-empathy, altruism, and pleasant manner) with t value was almost similar (around 7–9). Every group of latent variables represented by item or indicator in that group (t value>1.96).

Discussion

The analysis resulted in 5 scales or factor, namely altruism, responsibility-accountability, communication skill, humanism, and pleasant manner. Central to understanding the essential factor in assessing performance seems to be represented by indicators within the latent variable (potential variable). The indicator representing for each scale as following: latent variable of responsibility-accountability was most represented by indicator item c1 (t value=9.12), "they are capable deal with my disease under supervision." It understood that the clinical clerkship student was still under supervision

because their authority limited by regulation.

For the latent variable of the humanism and medical treatment, the most indicator represented with item c6 and c5 respectively (t value=9.68 and 9.09), which are an item of "they look professionally serving us" and item of "I do not hesitate regarding their ability in performing medical action standard." While latent variable of communication and empathy represented by item c8 (t value=10.0), "they are very friendly and enthusiastic when listening to our complaints." For the latent variable of altruism, the most indicator or observable variable represented by item c14 and c15. Item c14, if needed consideration of family in decision making during treatment, they always involve our family. Item 15, young doctor always giving the precedence to the interest and our safety. The last, latent variable was pleasant manner, this variable represented by item c20, I am not afraid the other doctor will know my complaints; and item c17, they were polite when serving us.

The evaluation of the performance of a junior doctor was a crucial aspect of the patient perspective. This study has developed a prototype instruments to evaluate the junior doctor's performance. This study used the participant to factor ratio of 10:1, thus create 50 of sample size. Amount of sample for an exploratory factor of analysis is ideally ten subjects for one item, but it depends on the purpose of research.[18] The instrument of this study was created to assess the performance of junior doctor from patients' perspective in terms of professional behavior, the particular aspects of the responsibility-accountability, the humanism, the communication, the altruism, and a pleasant manner. For principal component analysis, a factor with less than three items is generally weak and unstable, while factor with more than or the same as five items is strong.[7] Almost all factors in this study categorized either as strong and strong enough, except the factor of altruism that consists of three pieces. It still accepted, since the factor that has three items were stable.

There were four items for responsibility, three items for humanism, four items for communication and medical treatment, five items for altruism, and four items for pleasant manner. Humanism refers to an entity consisting of respect for others, compassion, and integrity. In this questionnaire, humanism consists

of trusted in performing a corrective action, professional, patient, and friendly. Responsibility at multiple levels refers to fulfilling the contract governing the doctor-patient relationship, the profession, and society. While altruism entails the consequence that the best interest of patient, not self-interest guide the doctors to do their job. Excellence involves a commitment to exceeding the ordinary expectations and the commitment to lifelong the learning.[19] The altruism of the questionnaire refers to some aspect, such as clear and straightforward instruction, involving family into decision making, give precedence to patient interest, know well problem of the patient, and fun services.

Communication skills have high clinical relevance. Studies have found that students' performances in communication increased the through time.[20] In this context, communication with the patients have some components, namely communication easy to understand, enthusiastic and friendly, painstaking in health service, and giving good advice of health. Doctor-patient communication is one aspect that increased adherence to medical treatment and enhanced clinical outcomes.[9] The element of the pleasant manner in the questionnaire refers to politeness in serving patients, known limitation and authority, involving in research, and trustworthiness.

Accordingly, research done by Donini-Lenhoff and Hedrick[21] shows that the ways of communication are an integrated pattern of learned beliefs and behaviors shared among groups. These include thoughts, styles of communicating, ways of interacting, views of roles and relationships, values, practices, and customs.[22,23] Therefore, communication skill is one of the essential factors that must be reserved by a physician in addition to clinical skill and professional behavior. This study is in line with Betancourt's[22] research showing that a doctor ought to has communication ability and clinical reasoning ability because evidence also suggests that doctor-patient communication highly linked to patient satisfaction and adherence and subsequently to health outcomes.

In this study, responsibility-accountability aspects assessed by knowing self-capability, able to deal with patient's disease, wholeheartedly whole serving, skillful, and trained. Studies assessing pertain to professionalism divided into three categories: as a part of the clinical

program, as a comprehensive entity, and as separate elements of professionalism, such as humanism and ethical decision making.[24] We conducted a study of assessing professionalism as a part of the clinical performance because it was done by determining aspects of being evaluated. Medical professionalism, in the performance context, is defined as the ability to meet the relationship-centered expectation required to practice medicine competently from the patient's perspective. Similarly, assessment of integrated inevitably have to be based on the patient evaluation in the workplace. The human assessment has three main fields of research in terms of bias and heuristic, original decision making, and social cognitive theory. Original decision making refers to how human can reach a satisfactory decision. The social cognitive theory stated decision making of individual highly influenced by the interpersonal and social environment. In other words, decision making depends on both the actual problem, motivation, and personal goal of the individual and local situation.[25]

Validity and reliability are an essential consideration to apply an evaluation tool. Based on the result of exploratory factor analysis, this instrument is acceptable as a tool of evaluation because this instrument was valid and reliable to assess the performance of junior doctor in a teaching hospital. This tool has similarities with the existing instrument developed by the American Board of Internal Medicine (ABIM) that consists of a component of altruism, humanism, accountability, and excellence. In other words, this result has construct validity based on exploratory factor analysis and concurrent validity based on the similarity with the existing instrument.[24] The results of Cronbach's alpha 0.95 indicates that the questionnaire is replicable for other research because of the value of Cronbach's alpha more than 0.7. Reliability determines reproducibility and consistency the result of measurement yields, and it is defined by a coefficient ranging from 0; no reliability to 1; absolute reliability.[8]

Conclusion

The patient perception instruments of junior doctor performance (PIJDP) could be used to assess the performance of junior doctors and could make a novel contribution to the

development of medical education.

Conflict of Interest

The authors have no conflict of interest to declare.

References

1. Prideaux D. Curriculum development in medical education: from acronyms to dynamism. Teach Teach Educ. 2007;23(3):294–302.
2. Sanchez-Reilly S, Ross JS. Hospice and palliative medicine: curriculum evaluation and learner assessment in medical education. J Palliat Med. 2012;15(1):116–22.
3. Ellett JD, Campbell JA, Gonsalves WC. Patient satisfaction in a student-run free medical clinic. Fam Med. 2010;42(1):16–8.
4. Ruhe V, Boudreau JD. The 2011 Program Evaluation Standards: a framework for quality in medical education programme evaluations. J Eval Clin Pract. 2013;19(5):925–32.
5. Spiel C, Schober B, Reimann R. Evaluation of curricula in higher education: challenges for evaluators. Eval Rev. 2006;30(4):430–50.
6. Dilmore TC, Rubio DM, Cohen E, Seltzer D, Switzer GE, Bryce C, et al. Communications psychometric properties of the mentor role instrument when used in an academic medicine setting. Clin Transl Sci. 2010;3(3):104–8.
7. Costello AB, Osborne JW. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis [Internet]. Pract Assess Res Eval. 2005 [cited 2018 September 20];10(7). Available from: https://pareonline.net/getvn.asp?v=10&n=7.
8. van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: from methods to programmes. Med Educ. 2005;39(3):309–17.
9. Francis MD, Warm E, Julian KA, Rosenblum M, Thomas K, Drake S, et al. Determinants of patient satisfaction in internal medicine resident continuity clinics: findings of the Educational Innovations Project Ambulatory Collaborative. J Grad Med Educ. 2014;6(3):470–7.
10. Worthington RL, Whittaker TA. Scale development research: a content analysis and recommendations for best practices. Couns Psychol. 2006;34(6):806–38.
11. DeVellis RF. Scale development: theory and applications. 4th Edition. Thousand Oaks, California: SAGE Publication; 2016.
12. Creswell JW, Guetterman TC. Educational research: planning, conducting, and evaluating quantitative and qualitative research. 6th Edition. Boston, Massachusetts: Pearson Publication; 2018.
13. Creswell JW. A concise introduction to mixed methods research. Thousand Oaks, California: SAGE Publication; 2015.
14. Shirali G, Shekari M, Angali KA. Assessing reliability and validity of an instrument for measuring resilience safety culture in sociotechnical systems. Saf Health Work. 2018;9(3):296–307.
15. Boerebach BCM, Lombarts KMJMH, Arah OA. Confirmatory factor analysis of the System for Evaluation of Teaching Qualities (SETQ) in graduate medical training. Eval Health Prof. 2016;39(2):21–32.
16. Anthoine E, Moret L, Regnault A, Sébille V, Hardouin JB. Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures. Health Qual Life Outcomes. 2014;12:176.
17. Velicer WF, Fava JL. Effects of variable and subject sampling on factor pattern recovery. Psychol Methods. 1998;3(2):231–51.
18. Cabrera-Nguyen P. Author guidelines for reporting scale development and validation results in the Journal of the Society for Social Work and Research. J Soc Social Work Res. 2010;1(2):99–103.
19. Arnold L. Assessing professional behaviour: yesterday, today, and tomorrow. Acad Med. 2002;77(6):502–15.
20. Klamen DL, Williams RG. The effect of medical education on students' patient-satisfaction ratings. Acad Med. 1997;72(1):57–61.
21. Donini-Lenhoff FG, Hedrick HL. Increasing awareness and implementation of cultural competence principles in health professions education. J Allied Health. 2000;29(4):241–5.
22. Betancourt JR. Cross-cultural medical education: conceptual approaches and frameworks for evaluation. Acad Med. 2003;78(6):560–9.

23. Aragaw A, Yigzaw T, Tetemke D, G/Amlak W. Cultural competence among maternal healthcare providers in Bahir Dar City Administration, Northwest Ethiopia: cross sectional study. BMC Pregnancy Childbirth. 2015;15:227.
24. Lynch DC, Surdyk PM, Eiser AR. Assessing professionalism: a review of the literature. Med Teach. 2004;26(4):366–73.
25. Berendonk C, Stalmeijer RE, Schuwirth LW. Expertise in performance assessment: assessors' perspectives. Adv Health Sci Educ. 2013;18(4):559–71.