Original Article

Reliability and Validation of the Persian Version of JUICE Study Questionnaire

Maryam Rayatpisheh, MD[1]; Saba Nasseri-Moghaddam[2]; Afsoon Kherad[3]; Takeshi Kanno, MD, PhD[4]; Ali Reza Sima, MD[5]; Paul Moayyedi, MD, MPH, PhD[6]; Siavosh Nasseri-Moghaddam, MD, MPH[1*]

[1]Digestive Disease Research Institute, Tehran University of Medical Sciences, Tehran, Iran
[2]Psychology Division of York University, Toronto, Canada
[3]Islamic Azad University, Central Tehran Branch, Tehran, Iran
[4]Department of Medicine, Division of Gastroenterology, McMaster University, Hamilton, Canada Division of Gastroenterology, Tohoku University Graduate school of Medicine, Sendai, Japan
[5]Medicine Digestive Disease Research Institute, Tehran University of Medical Sciences, Tehran, Iran
[6]Department of Medicine, Division of Gastroenterology, McMaster University, Hamilton, Canada

Abstract

Background: This study is part of a prospective international cohort study on the composition of microbiota living in the upper gastrointestinal (GI) tract and its correlations to the patients' symptoms, and their psychological and physical health status in three different populations (JUICE Study). Our study evaluates the reliability and validity of the Persian version of a three-part questionnaire which will be used in this study.

Methods: The original English version of the questionnaire was translated to Farsi and then back translated to English by an expert Iranian English teacher. The back-translation was edited by a native English speaker and then retranslated to Farsi. The questionnaire consists of three parts; the first part includes demographic data, the second part is the EQ-5D questionnaire which is an instrument developed by the EuroQol group to measure five dimensions of quality of life and health status, and the third part is the Hospital Anxiety and Depression Scale questionnaire (HADS). By convenience sampling, 22 participants were enrolled. To check reliability, they were asked to complete the questionnaire and repeat this two weeks later. For validity, one of the researchers completed the questionnaire for each of the participants after interviewing them and this was compared with the questionnaire completed by the participant. Each participant was also asked to comment on the content and structure of the questionnaire and these points were considered for improvement of the questionnaire.

Results: The first six questions were demographic, and completely the same in the test and retest phases. All of the other questions in the first part of the questionnaire had Kappa values above 0.6 for both reliability and validity. Four questions in this section were assessed with percentage agreement. Percentage agreements were 0.86, 0.54, 0.66 and 0.9 for reliability and 1, 0.81, 0.81 and 1 for validity for the 8th, 12th, 13th and 14th questions, respectively. In the second part, i.e. the EQ5D questionnaire, percentage agreements of the first and second items for reliability equaled 1. Percentage agreements of the three last items were 0.90, 0.63 and 0.72, respectively. Validity of the EQ5D questionnaire was 100% for the first three items and 95% for the last one. For the HADS questionnaire, percentage agreement averaged 0.63 in the reliability phase and 0.78 in the validation phase.

Conclusion: According to our data, this three-part questionnaire has acceptable reliability and validity to be used as an instrument in Farsi-speaking populations for the JUICE study.

Keywords: JUICE study, Kappa coefficient, Persian, Reliability, Validity

Cite this article as: Rayatpisheh M, Nasseri-Moghaddam S, Kherad A, Kanno T, Sima AR, Moayyedi P, et al. Reliability and validation of the persian version of JUICE study questionnaire. Arch Iran Med. 2022;25(1):1–5. doi: 10.34172/aim.2022.01

Introduction

Our understanding of microbes living in our body has changed in the past decade and has gone from only a few species to a large and diverse community that develops in humans along with their chronologic age.[1] The largest number of microbes exist in the lower gastrointestinal (GI) tract (large intestine); however, there is a variety of microbiota throughout the human body including the upper GI tract. The microbiota in the GI tract has been associated with various aspects of health and disease, covering a broad spectrum from local GI conditions to psychiatric and behavioral ones. Therefore, better

understanding of the GI microbiota, and carefully manipulating it, may enable us to find new treatments for many conditions.[2] Most of our knowledge about the GI microbiota comes from studies performed on the stool which will not let us know where exactly the identified microbes live and come from. We know little about the composition and correlations of the upper GI tract microbiota.

Therefore, we decided to perform a prospective multicenter cohort study to compare upper GI symptoms and endoscopy findings in Iran (Tehran University of Medical Sciences, Masoud Clinic) with Canada (McMaster

*Corresponding Author: Siavosh Nasseri-Moghaddam, MD, MPH, AGAF; Digestive Disease Research Institute, Tehran University of Medical Sciences, Shariati Hospital, Tehran, Iran. Tel:+98-21-82415104; Email: sianasseri@yahoo.com

University) and Japan (Tohoku University) and correlate these findings with the upper GI microbiota (The JUICE Study, Japanese upper GI symptoms compared with Iranian and Canadian patients presenting for Endoscopy). We will recruit adult patients who undergo esophago-gastro-duodenoscopy (EGD) for any reason and consent to be enrolled in the study during the study period. If the patient is eligible for this study, we will record demographics and the main reason for endoscopy. Patients will be asked to complete the following questionnaires: upper GI symptoms measured by Short form Leeds Dyspepsia Questionnaire (SFLDQ), quality of life measured by EQ-5D, and anxiety and depression measured by the Hospital Anxiety and Depression Scale (HADS).

Reliability and validity of a measuring instrument are the key indicators of the quality of that instrument. The current study reports the reliability and validity of the questionnaires which will be used in the JUICE study (listed above) in Iran.

Reliability is defined as the extent to which results are consistent over time; if the results of a study can be reproduced using a similar methodology, then the research instrument is considered to be reliable. An instrument can be reliable without being valid. Validity or trustworthiness is defined as the degree to which the results of a questionnaire agree with the real world.[3]

Materials and Methods

The original English questionnaire was translated to Farsi by one of the researchers (SaNM). It was then back translated to English by an expert Iranian English teacher. The back-translated questionnaire was then reviewed by one of the researchers (PM, a native English speaker) and corrections and refinements were made for the final Farsi translation. This latter version was used for the validity and reliability study performed at the Digestive Disease Research Institute (DDRI) of Tehran University of Medical Sciences. This questionnaire contains three sections: The first part has 17 questions: of those, the first six are demographics (name, sex, birth date, education level, birth place and race), questions number 7 and 8 ask about any current drug consumption, question 9 asks if the client has had any abdominal surgery before and the 10th question is about history of *Helicobacter pylori* infection. The questions number 11 to 16 ask about using PPIs (proton pump inhibitors), aspirin products, NSAIDS (non-steroidal anti-inflammatory drugs), and any other painkiller or antibiotic use during the previous month, respectively. The 17th question asks about the present and past smoking history.

The second section is the EQ5D questionnaire which asks about five items related to the same- day's quality of life including "movement", "self care", "daily activities", "pain or discomfort", "stress and depression" and a chart is designed to indicate the interviewee's health status on the day of study.

The third and last part is the HADS questionnaire which consists of 14 questions (asking about one's psychological health in the last week).

At the end of our questionnaire, a list of all the available brand names of PPIs and NSAIDs in Iran is attached which is extracted from the official Iranian pharmaceutical statistics (11th version) (https://www.fda.gov.ir/fa/).

Reliability of this questionnaire was assessed, using the standard (test-retest) method.[4] For the first (test) phase, 22 volunteers from DDRI were asked to complete a coded questionnaire anonymously. During this process, any ambiguity was clarified by the oriented distributor; most of the participants returned the completed form on the same day, but a few needed reminding. After gathering all the forms, data codes were entered on an Excel worksheet. Two weeks later, the same questionnaire was given to the same people (in order to reduce the bias of memory recall, a time interval of two weeks was selected between the test and retest phases and the participants were not informed about the retest phase before its time). Data codes of these forms were entered on the same Excel sheet, the same questions were arranged beside each other with different numbers (1st and 2nd).

To assess validity, 22 participants completed the questionnaire and one of the researchers (MR) interviewed them and completed the questionnaire for them accordingly. The paired questionnaires were then transferred to the Excel sheet.

Cohen's kappa coefficient and 95% confidence interval (CI) were used to estimate inter-rater reliability for categorical variables. Kappa coefficient shows agreement beyond chance. Kappa values range between 0 and 1 and a value above 0.6 is considered acceptable (Table 1).[5,6]

Percentage agreement was used for variables without a normal distribution, where kappa statistics cannot be used. For the purpose of this study, a percentage agreement of 0.6 and above was considered adequate and percentage agreement of 0.5–0.6 was acceptable.

The Bland-Altman plot was used for assessment of agreement of the chart scores on current health status. We analyzed the data using IBM SPSS Statistic software, version 22 for Windows (released 2013, Armonk, NY, USA).

Results

Table 2 shows the reliability and validity results for the categorical variables of the first part of the questionnaire. As shown, in reliability assessment, only question number

Table 1. Kappa Coefficient Classification

Kappa Statistics	Strength of Agreement
<0.00	Poor
0.00–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost perfect

Table 2. Results of Reliability and Validity of the First Part of the Questionnaire

No.	Questions of First Part	Reliability Assessment			Validity Assessment		
		Kappa Statistics	95% Confidence Interval	Percentage Agreement	Kappa Statistics	95% Confidence Interval	Percentage Agreement
7	Are you currently on any medication?	0.80	0.55–1.00			0.89	0.70–1.00
8	Have you ever had abdominal surgery?			0.86	1.00	1.00–1.00	
9	Have you ever been infected with *Helicobacter pylori* (a bacterium that lives in the stomach)?	0.91	0.76–1.00			0.76	0.51–1.00
10	Have you ever used proton pump inhibitors for gastric acid?	1.00	1.00–1.00			1.00	1.00–1.00
11	Have you ever used aspirin or medications containing it?	0.71	0.41–1.00			0.90	0.75–1.00
12	Are you currently using anti inflammatory drugs?			0.54			0.81
13	Have you ever used any other painkillers?			0.66			0.81
14	Have you used any antibiotics during the last month?			0.90	1.00	1.00–1.00	
15	Do you smoke?	1.00	1.00–1.00			0.83	0.52–1.00

12 had a percentage agreement of less than 0.60 (i.e. 0.54) and the rest of the questions had Kappa values or percentage agreements in the acceptable range. Kappa value of the 10th question in both reliability and validity evaluations was 1.

The reliability and validity test results for the EQ-5D questionnaire are depicted in Table 3. All values were in the acceptable range.

In the last part (the HADS questionnaire), percentage agreement was calculated for all of the 14 items which is summarized in Table 4.

Discussion

This study was performed to establish the reliability and validity of the Persian version of the three-part questionnaire which will be used in the JUICE Study in Iran. The questions in the 1st and 2nd parts of the questionnaire were found to be adequately valid and reliable. The third part, which consists of the HADS questionnaire, was moderately reliable and substantially valid.

In order to assess the sixth item of the EQ5D questionnaire in our study, we drew a Bland–Altman plot because this item is a chart that categorizes the same-day's health status as 0 (worst) to 100 (best). This plot can show the limits of agreements but gives no information on whether these limits are clinically acceptable or not.[5,7] As we did not claim to prove the acceptability of each individual's health status and the main goal is to compare two scores of each phase, it seems that this plot can fulfill our purpose (Figure 1 and 2).

Most of the studies performed in this field have assessed the reliability and validity of a questionnaire in a special and selected group of patients, but we enrolled the participants in DDRI by convenience sampling and this

Table 3. Results of Reliability and Validity of the EQ-5D Questionnaire

No.	EQ-5D	Reliability Assessment Percentage Agreement	Validity Assessment Percentage Agreement
1	Movement	1.00	1.00
2	Self care	1.00	1.00
2	Daily activities	0.90	1.00
4	Pain or discomfort	0.63	0.86
5	Stress or depression	0.72	0.95

can be the strength of our study. However, our sample size is smaller than other studies which explains the wide range of 95% confidence intervals for kappa values.

In a study performed to assess the reliability and validity of the Persian version of the HADS questionnaire in Iran among 261 depressed and/or anxious patients and 261 healthy controls, acceptable validity and reliability were found for this questionnaire.[5]

A cross-sectional study, which was performed in an emergency department in Saudi Arabia to assess the reliability and validity of the Arabic version of the HADS Questionnaire among 257 participants, showed acceptable results in 95% of subjects. They used Cronbach's alpha coefficient to evaluate reliability, and it indicated a significant correlation with both the anxiety (0.73) and depression (0.77) subscales of HADS, thereby supporting the validity of the instrument. By means of factor analysis, they obtained a two-factor solution according to the two HADS subscales (anxiety and depression), and they observed a statistically significant correlation ($r=0.57$; $P<0.0001$) between the two subscales.[7]

In a research on reliability, validity and responsiveness of EQ-5D to evaluate health status in patients with social phobia in Germany, a sample of 445 patients with social phobia were studied with five measurement points over

Table 4. Reliability and Validity Results of the HADS Questionnaire

No.	HADS Questionnaire	Percentage Agreement Reliability Assessment	Percentage Agreement Validity Assessment
1	I feel like I am very angry and worried	0.72	0.54
2	I still find the stuffs that used to be attractive, fascinating almost the same amount	0.59	0.63
3	I feel like something bad is about to happen	0.63	0.77
4	I can laugh and see the funny sides of things	0.54	0.86
5	I have worrying thoughts	0.68	0.59
6	I feel happy	0.72	0.86
7	I can sit still and feel relaxed	0.54	0.81
8	I feel very passive	0.68	0.54
9	I feel anxious	0.77	0.81
10	I have lost the desire to take care of my appearance	0.68	0.77
11	I feel tired whenever I need to do something	0.59	0.86
12	I look at things with enthusiasm	0.54	0.68
13	I have a sudden feeling of anxiousness and fear	0.63	0.86
14	I am attracted to a good book or radio or TV program	0.54	0.68
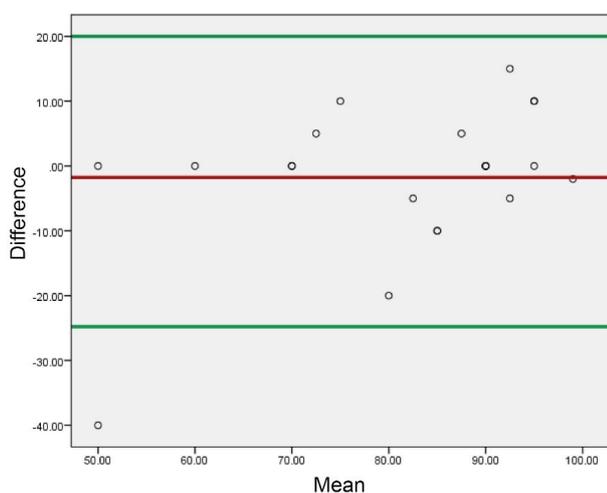
Figure 1. Bland-Altman Plot (Showing Reliability of Scores in the Chart of "Today's Health Status").
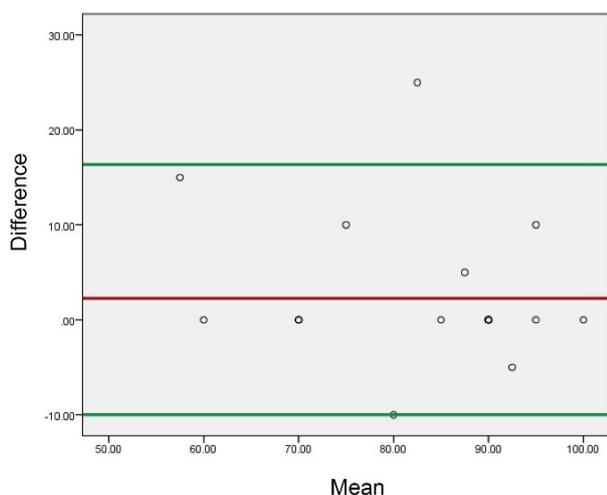
Figure 2. Bland-Altman plot (for validity of scores in the chart of "Today's health status").

a 30-month period. The discriminative ability of EQ-5D was analyzed by comparing the patients' responses with the general population and between different disease severity levels. For test-retest reliability, they assessed the level of agreement in patients' responses over time, which was moderate (intra class correlation coefficient > 0.6). Construct validity was analyzed by identifying correlations of EQ-5D with more specific instruments which was limited. [8]

Hereby, we report the validation procedure of the questionnaire to be used in the Iranian section of the JUICE study. According to our data, the Persian version of this questionnaire is reliable and valid for use in this study.

Authors' Contribution
SNM, and TK had an effective role in the implementation of this project. SaNM translated the questionnaire to Farsi and AK back translated the Farsi questionnaire to English. PM reviewed and corrected the back translated one. SNM and MR reviewed different methods of questionnaire validation and reliability. MR performed the test retest reliability method and interviewed as a part of validation assessment. Analysis was done by MR; SNM and MR participated in writing the manuscript and editing.

Conflict of Interest Disclosures
The authors have no conflicts of interest.

Ethical Statement
Not applicable.

References
1. Durack J, Lynch SV. The gut microbiome: Relationships with disease and opportunities for therapy. J Exp Med. 2019;216(1):20-40. doi: 10.1084/jem.20180448.
2. Pinto-Sanchez MI, Hall GB, Ghajar K, Nardelli A, Bolino C, Lau JT, et al. Probiotic *Bifidobacterium longum* NCC3001 reduces depression scores and alters brain activity: a pilot study in patients with irritable bowel syndrome. Gastroenterology. 2017;153(2):448-59.e8. doi: 10.1053/j.gastro.2017.05.003.
3. Golafshani N. Understanding reliability and validity in qualitative research. Qual Rep. 2003;8(4):597-607. doi:

10.46743/2160-3715/2003.1870.

4. De Moraes ACF, Vilanova-Campelo RC, Torres-Leal FL, Carvalho HB. Is self-reported physical fitness useful for estimating fitness levels in children and adolescents? a reliability and validity study. Medicina (Kaunas). 2019;55(6):286. doi: 10.3390/medicina55060286.

5. Vetter TR, Schober P. Agreement analysis: what he said, she said versus you said. Anesth Analg. 2018;126(6):2123-8. doi: 10.1213/ane.0000000000002924.

6. Cyr L, Francis K. Measures of clinical agreement for nominal and categorical data: the kappa coefficient. Comput Biol Med. 1992;22(4):239-46. doi: 10.1016/0010-4825(92)90063-s.

7. Al Aseri ZA, Suriya MO, Hassan HA, Hasan M, Sheikh SA, Al Tamimi A, et al. Reliability and validity of the Hospital Anxiety and Depression Scale in an emergency department in Saudi Arabia: a cross-sectional observational study. BMC Emerg Med. 2015;15:28. doi: 10.1186/s12873-015-0051-4.

8. Sonntag M, Konnopka A, Leichsenring F, Salzer S, Beutel ME, Herpertz S, et al. Reliability, validity and responsiveness of the EQ-5D in assessing and valuing health status in patients with social phobia. Health Qual Life Outcomes. 2013;11:215. doi: 10.1186/1477-7525-11-215.