RESEARCH

Psychometric properties of the 12-item WHODAS applied through phone survey: an experience in PERSIAN Traffic Cohort

Nasrin Shahedifar2,3, Homayoun Sadeghi-Bazargani2*, Mohammad Asghari-Jafarabadi1,2,6, Mostafa Farahbakhsh4 and Shahrzad Bazargan-Hejazi5

Abstract

Background: Due to limited capability to function in post-injury daily life injury, survivors need to be reliably assessed without need to commute more than necessary. The key action is to determine the level of functioning difficulties. Having the opportunity of conducting a national post-crash traffic safety and health cohort study, we aimed to translate into Persian and assess the psychometric properties of the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) through phone surveys six month post injury.

Methods: First, having World Health Organization permission, we tested the translation validity by forward translation, expert panel evaluation, back-translation, pre-testing and cognitive interviewing, and finalizing the Persian WHODAS. Then, through a psychometric study within a national cohort platform, the validity, reliability and applicability of the 12-item WHODAS was assessed through phone surveys. We included data of 255 road traffic injury patients enrolled from the cohort at six-month follow-up. The psychometric assessment (internal consistency reliability and stability reliability) was conducted on test–retest data of 50 patients with an average 7-day time span. An exploratory factor analysis tested the construct validity using extraction method of principal component factor and oblique rotation on data from 255 patients. Regarding the multiple criteria including an eigenvalue > 0.9, Cattell’s scree test, cumulative variance, and the theoretical basis, the minimum number of factors were retained. Data were analyzed using STATA statistical software package.

Results: The respondents were mostly male (81%), employed (71%), educated (87%), and with a mean age of 37.7(14.9). The Persian version had high internal consistency reliability (Cronbach’s α = 0.93) and excellent stability reliability (ICC = 0.97, 95% CI: 0.92–0.98). An exploratory factor analysis retained four factors defining 86% of all the variance. Factors of Self-care, Mobility, and Cognition were completely retained.

Conclusions: The brief Persian WHODAS 2.0 was highly reliable and valid to be applied through phone interviews post injury.

Keywords: WHODAS, Disability, Post-crash, Psychometrics, Road traffic injury, Post injury, Cohort study

Introduction

Road traffic crashes (RTC) result in 1.35 million deaths annually and leave 20 to 50 million people with non-fatal injuries, with many suffering a disability [1]. The injuries account for 41.2 million years of healthy life lost [2–4], and 90% of disability-adjusted life years in
Low and Middle Income Countries [5]. In Iran, road traffic injuries (RTIs) are the second leading cause of death [6, 7], and the third cause of disability adjusted life years [8]. Over 2600 person years of life lost are attributed to RTIs [9]. The people hospitalized due to RTIs mainly suffer severe injuries resulting in short- or long-term difficulties in functioning [10]. Only with the functioning information would we recognise their right health care needs [11, 12].

The concepts “functioning” and “disability” in public health and clinical settings represent information on how the population accomplishes its activities and participation [13]. Functioning is “generic term which includes body functions and structures, activities and participation. It indicates the positive aspects of the interaction between the individual (with a health condition) and its contextual factors (personal and environmental factors)” [14]. Disability is an umbrella term for deficiencies of body functions, activity limitations or participation restrictions [13]. The controversial issue is how to measure the functioning level, and many tools have been developed. However, the tool by the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) was developed on an inclusive categories of International Classification of Functioning, Disability and Health (ICF) [14]. Although there are many tools developed to measure functioning, WHODAS 2.0 is reputable for several chief characteristics. For instance, the 11th International Classification of Disease—ICD-11 added the section “V Supplementary section for functioning assessment” to form functioning profiles and overall functioning scores [15, 16].

Tool’s psychometric properties may be affected through the method of data collection such as phone surveys comparing with in-person interviews. So it is critical to assess the tool’s psychometric properties for various data collection. Then, very little similar work has been done so far. One of the major field is assessing disability in injuries post-crash and post discharge. For such assessing, there is rare opportunity to have in-person visits, in some settings particularly in low and middle income countries (LMICs) and then the resources do not support the researchers to do in-person interviews. It is critical to have robust tools to use for phone surveys. To assess this, we had a great opportunity of a cohort study. This setting helps to apply the tool and assess its psychometric properties through national cohort platform of Post-crash Traffic Safety and Health Cohort Study, Prospective Epidemiological Research Study in Iran (PERSIAN) [17, 18]. Then the key action was to apply reliable and valid tools in measuring the outcomes of RTIs such as functioning difficulties, a common consequence. So, the expert team chose the inclusive tool of WHODAS 2.0 to be administered to specifically RTI survivors through phone, with great care, sensitivity, and searching.

Although the available translations are neither specified for the target group with road traffic injuries, nor applied for phone-based surveys [19–22], the academic expert panel meticulously reviewed and piloted it among road traffic injury patients in the setting of the prospective PERSIAN Traffic Safety and Health Cohort Study. As its validity was not approved in the setting, we communicated with WHO permissions team and obtained its permission to retranslate the original version to apply in the PERSIAN Traffic Cohort Study (ID: 380480). As the short form of this globally approved tool has a good explanation of the full-version tool and takes five minutes on average, as well as considering the physical health conditions of RTI interviewees, we utilized and psychometrically assessed the 12-item WHODAS 2.0, through phone interviews six months post injury. Also, to our knowledge, there are not published psychometric works in which the short-version tool was used in phone surveys and in RTI patients in the world, except for one phone survey among critically-ill patients admitted to intensive care units, and one interview-based study among patients with road traffic injuries. The paucity of such all-inclusive tool on creating the functioning profile of people involved in RTCs is sensible reason to conduct the translation, and to examine internal consistency, reproducibility, time consistency, and construct validity of brief WHODAS 2.0 among study population through phone surveys six months post injury. Then, it would pave the way for study of the functioning problems caused by road traffic injuries. We expected that the short version would demonstrate excellent internal consistency and test–retest reliability, defined as values of 0.80 or greater and 0.90 or greater, respectively.

Materials and methods

Phase one: translation of the 12-item WHODAS 2.0

The World Health Organization (WHO) permitted the Road Traffic Injury Research Center (RTIRC) to translate the 12-item WHODAS 2.0 into Persian language for use specially in the context of the Traffic Safety and Health Cohort Study (permission ID: 380480). Our expert team composed of members from RTIRC, and physical medicine and rehabilitation Research Center, Tabriz University of Medical Sciences. It conducted an iterative, rigid, and meticulous process of translation, according to the translation guideline recommended by the WHO in five steps: (1) forward translation; (2) Expert panel; (3) Back-translation; (4) Pre-testing and cognitive interviewing, and (5) Final version [23, 24].
Phase two: assessment of reliability and factor structure

The psychometric properties of the brief WHODAS 2.0 was evaluated at six months after crash through phone-based interviews. This study was conducted in a clinical population. The source population of the study were RTIs survivors enrolled from post-crash traffic safety and health cohort study [18]. All participants hospitalized in one of two referral trauma centers, Shohada and Imam Reza, in Tabriz, were followed up through the cohort study at six months after crash. We conducted on a representive sample of RTIs survivors 18 years and above, between 21 May 2020 and 20 March 2021. The whole hospitalized patients have been recorded in the database of Integrated Road Traffic Injury Registry System (IRIIRS) since 2019 [25]. To examine the reliability of the instrument, the minimum number of sample required was 50 for detecting the value of 0.4 for ICC, with alpha and power defined at 0.05 and lower than 90%, respectively [26]. In order to conduct factor analysis, a maximum number of 21 participants per item was taken to attain the best model fit [27]. So, 255 patients were included in our sample. The selected participants were replaced by the same sex and age as the refusal to participate. As inclusion criteria, participants must have been hospitalized due to RTIs six months prior to their interview date for at least 24 h, registered, aged 18 or over, and informed consent was required. People unable to complete the questionnaire due to any disorders were excluded from the study. The study was carried out by an appropriately trained interviewer with broad experience (≥ 4 years) routinely collects data from registered patients on phone calls [17]. The completeness of tools was mostly guaranteed since the interviewer filled them out.

Measures

The 12-item WHODAS 2.0

As a general measure, WHODAS 2.0 was developed based on an inclusive set of categories considered in the framework of ICF to measure disability during the preceding 30 days [14]. It can be used for epidemiological studies and other purposes. This questionnaire was researched by the WHO in many countries, where its reliability, convergent validity with other assessment instruments, constant factor structure, and other psychometric properties were analysed [28].

The short version describes 81% of the variance of the full version. Researchers can compute overall functioning scores, administer through an interview during the mean interview time of about 5 min [29]. The response choices for each item ranged on an ordinal scale, from zero denoting “none” to four denoting “extreme/cannot do”. The individual scores in each subscale were estimated by simply adding up the results of the two relevant items [30]. Each subscale contains two questions from the corresponding subscale in the full version: Subscale 1: Cognition (items 3 (learning) and 6 (concentrating) of the questionnaire); Subscale 2: Mobility (items 1 (standing) and 7 (walking)); Subscale 3: Self-care (items 8 (washing) and 9 (getting dressed)); Subscale 4: Getting along with people (items 10 (dealing with people) and 11 (maintaining a friendship)); Subscale 5: Life activities (items 12 (day-to-day work) and 2 (household responsibilities)); Subscale 6: Participation (items 4 (community activities) and 5 (emotions)) [24, 31]. A simple way of calculating the results was used for the psychometric evaluation of the tool. The total score is a percentage as follows:

\[
\text{Total score} = \left( \sum \text{item scores}/48 \right) \times 100
\]

It ranged from 0 (no disability) to 48 (complete disability). A higher result indicates a higher level of disability [29]. To draw a comparison, raw total scores and subscale scores were then converted into a 0–100 scale as normalized scores, using the complex method of scoring [32]. We categorized the total disability scores according to the ICF severity ranges (no problem, 0–4; mild disability, 5–24; moderate disability, 25–49; severe/extreme disability, 50–100). We considered a disability score of greater than or equal to 25 to indicate ‘disability’ based on the WHODAS ICF [33]. When only one item from the 12 items was missing, the average of the remaining 11 items was assigned to the missing item. If more than one item was missing, the survey was rejected [34].

Socio-demographic data

Basic socio-demographic data (sex, age, marital status, education, and job) and selected crash-related information [35] (used vehicle, counterpart vehicle, crash mechanism, and injured person’s role) were collected.

Statistical analysis

Initial analyses of the WHODAS and baseline characteristics data were performed using descriptive statistics. First an inter-item correlation matrix was visually observed. Correlations were defined as excellent (> 0.75), moderate (0.50–0.74), fair (0.25–0.49), and no meaningful correlation (< 0.25) [36]. With a sample size of 255 patients, this study has 80% power (two-sided \( P = 0.05 \)) to detect a correlation coefficient of 0.2 or greater. The data were analyzed using STATA statistical package version 15 (StataCorp LLC, Texas) and adopted a significance level of 0.05.
Scale score reliability

Internal consistency reliability
Internal consistency indicates the extent to which the items in a subscale are correlated. The homogeneity was evaluated using Cronbach’s Alpha statistic [37]. A value between 0.5 and 0.75 was considered moderate while values between 0.75 and 0.9 indicate good and values greater than 0.90 specify excellent reliability [38].

Test–retest reliability
Another measurement used to directly assess the scale score reliability of the tool was repeatability assessment using the test–retest method. Comparing the results of the first and second measurements examined the consistency over time. A group of 50 people was interviewed on telephone, between 21 May 2020 and 15 June 2020. All of them participated in both test and retest measurements. They did not have any treatment during that week period or changes in their medication. The average time between two measurements was 7 days (range from 6 to 8 days). Generally, a one-week time frame is applied on functional measurements [39]. In this study, inconsistencies might be produced by change in functional conditions of patients (physical and mental conditions) [40]. Moreover, the minimum detectable consistency is reachable by 7-day time interval with regard to lessen the potential for learning effect [41]. Other sources of inconsistencies such as dissimilarities in testing procedure and testers were not true as we took consistency, fixed protocol for testing, and one tester. Along with the criteria of the study, specific exclusion criteria for this step included having psychiatric problems affect the validity of the responses or an event experienced during the test–retest interval significantly increases the level of disability (e.g. disease, another car crash, and fall).

The normality of total scale and subscales’ distributions were tested based on the value of skewness and kurtosis. The test–retest correlation was measured by tau Kendall b. The reliability of the method was assessed by intra-class correlation coefficient (ICC). Scores of test–retest were analysed using the “single-rater type (k = 1), absolute-agreement, 2-way mixed-effects model to produce the intra-class correlation with 95% confidence intervals [38]. An ICC coefficient of ≥0.75 was considered as evidence of measurement stability. ICC between 0.4 and 0.75 indicates fair to good reliability, and ICC <0.4 indicates poor reliability [42].

Factor structure

Exploratory factor analysis (EFA)
The exploratory factor analysis is a fundamental tool to define the optimal count of latent variables in the validation of 12-item WHODAS 2.0 completed by 255 patients six months after crash on phone calls. As the correlation matrix was not compound symmetric, we did not run the method of Maximum Likelihood. Since 5-point Likert scales, are subject to lack of normality, we extracted factors using method of principal component factor and oblique rotation on data of 255 patients, from May 2020 to March 2021. To test the assumptions of EFA, the overall KMO (Kaiser–Meyer–Olkin) to measure sampling adequacy was 0.81, and the Bartlett’s test of sphericity proved an appropriate model (X²=4278.422 and P <0.0001) [43].

In the EFA, it is aimed to explain the maximum variance using a small number of factors. It is essential to apply multiple criteria and wisely judge each acceptable solution to retain the most appropriate factor numbers. In fact, no single method has been created to be correct in all conditions [43]. Moreover, it is not recommended to use Kaiser-Guttman method of “eigenvalue 1” anymore, and the researchers could act with more flexibility of preferred cut-off when applying multiple criteria [43–45]. Then, we used multiple criteria of eigenvalue, Cattell’s scree plot, cumulative variance, and theory, to determine the number of factors (Fig. 1).

Floor and ceiling effects
The floor and ceiling effects were derived from the percentage of respondents with the lowest and highest possible scale scores for individual items of the tool. Ceiling and floor effects were considered present when it is higher than 15% [46].

Results
The research data were data for the brief-version WHO-DAS 2.0, and crash-related and socio-demographic characteristics. A total of 255 subjects were approached by a research assistant. The majority of participants were male (81%). The mean age (SD) was 37.7(14.9; 18–86 yrs). The respondents were mostly employed (71%) and educated (87%) (Table 1). The mechanism of half of the crashes was vehicle-vehicle collision. The majority of subjects were motorcycle and car drivers. A fourth of patients were passenger or pillon passenger when they had crash (Table 1).

The average disability score for the study population (N=255) was 26.87 ±22.1 on a scale of 0 to 100 points, with 50 percentile at 25.

Phase one: translation of the 12-item WHODAS 2.0
The forward translation was done by three qualified translators as native-like English speakers and one of them is a professor living in the USA. They were familiar with health and disability terms. The tool with 15 items and 402 words was translated into
Persian language with 15 items and 380 words. Then the panel discussed on the term “emotionally affected” in item 5 (emotions) which takes several Persian meanings, and finalized. An independent linguist blind to the original English terms performed back translation with 15 items and 365 words. Three experts examined compatibility of the forward and back translations per each item, separately on a 4-level Likert scale. Measuring the compatibility, modified Kappa rate (> 0.76) revealed that two items were completely compatible and the others’ compatibilities were acceptable, except for the item 5 (emotions) lower than 0.76, due to the semantic complexity of the used word in forward translation. Following the expert panel and translators’ discussion and minor amendments, the translated questionnaire was completed by several people as a pre-test to examine if the questions were translated and written in a clear and comprehensible manner. The final translation is presented in Table 2.
### Table 2  Item reliability of the short form of WHODAS 2.0

| Items | ICC (95% CI) | Items in Persian Language |
|-------|--------------|---------------------------|
| S1    | 0.98 (0.97–0.99) | مدت طولانی مربی ایمن‌اند (مثلاً 30 دقیقه؟) |
| S2    | 0.96 (0.94–0.98) | آیا اینکه دیده که لمبم در منزل؟ |
| S3    | 0.86 (0.77–0.93) | چگونه کار جدید (مثل یا آموزش در حال نوآوری) |
| S4    | 0.93 (0.89–0.96) | چگونه کمک کنید چگونه مشکلات را حل کنید؟ |
| S5    | 0.94 (0.91–0.97) | هر چاهی و احساس شما چگونه تاثیر وضیفه؟ |
| S6    | 0.93 (0.87–0.96) | چه درک حکمی از بر انجام کاری؟ |
| S7    | 0.97 (0.96–0.99) | چیزی از شیوه‌های طولانی‌رسیده (مثل بازی از کیسه‌هایی) |
| S8    | 0.90 (0.83–0.94) | برگزاری جلسات تیمی (مثل 1 ساعت کنک) |
| S9    | 0.88 (0.79–0.94) | چگونه دیده‌ای؟ |
| S10   | 0.85 (0.71–0.92) | رابطه برقرار کردن با چه‌گونه؟ |
| S11   | 0.88 (0.79–0.94) | حفظ رابطه دوستی؟ |
| S12   | 0.85 (0.73–0.92) | انجام وظایف روزانه شغلی/تماسی؟ |

### Table 3  Factor loading matrix for WHODAS 2.0 items (N = 255)

| Items (number of item) | Factor1 | Factor2 | Factor3 | Factor4 |
|------------------------|---------|---------|---------|---------|
| Maintaining a friendship? (11) | 0.9438 |         |         |         |
| Dealing with people you do not know? (10) | 0.9393 |         |         |         |
| How much of a problem did you have in joining in community activities (for example, festivities, religious or other activities) in the same way as anyone else can? (4) | 0.9266 |         |         |         |
| How much have you been emotionally affected by your health problems? (5) | 0.8671 |         |         |         |
| Your day-to-day work/school? (12) | 0.8224 |         |         |         |
| Taking care of your household responsibilities? (2) | 0.6231 |         | -0.5900 |         |
| Concentrating on doing something for ten minutes? (6) | 0.4884 | 0.3985 |         |         |
| Walking a long distance such as a kilometre [or equivalent]? (7) | 0.5777 | 0.7656 |         |         |
| Standing for long periods such as 30 min? (1) | 0.6050 | 0.7171 |         |         |
| Washing your whole body? (8) | 0.6102 | 0.7529 |         |         |
| Getting dressed? (9) | 0.6140 | 0.7500 |         |         |
| Learning a new task, for example, learning how to get to a new place? (3) | 0.3865 | 1.25 | 1.07 | 0.8528 |
| Eigen values | 7.12 | 6.09 | 0.78 | 0.86 |
| Cumulative variance (%) | 0.59 | 0.69 | 0.78 | 0.86 |
**Phase two**

**Factor structure**

An exploratory factor analysis was conducted for 12 item WHODAS 2.0. We retained four factors based on multiple criteria. Based on the Cattell's scree plot, we recognized the point of break at which the shallow scree initiates, then maintained four factors that do not join the scree [47] (Fig. 1). Based on the cumulative percent of variance, the first four factors account for a proportion 0.86 of the total variance. More than 80% of variance was explained by four factors, 1/3 as many factors as variables [48]. Kaiser-Guttman rule allowed us to keep only three factors. Although the fourth factor added 0.14 variance which might seem little, it was theoretically sensible and meaningfully interpretable, too [48]. Then, we retain four factors as stated by criteria of eigenvalue, Cattell's scree plot, cumulative variance, and theory, which led us to extract factors with eigenvalues above 0.9.

All items showed uniqueness less than 0.2, except for the item 6 for “concentrating doing something for ten minutes” with 0.44 [49]. Accordingly, four factors with factor loadings equal or > 0.3 on 255 people were retained. Items grouping on the same factor presented four factors as follows: factor 1, “social/self-activities” included items involving “Life activities”, “Participation”, and “Getting along” as in the original instrument; factors 2, 3 and 4 indicated activities of “Mobility”, “Self-care”, and “Cognition”, respectively as in the original tool. The loading matrix indicated the item 2 cross loaded onto the factors one and four, though the factor loading was negative for factor four. Each and every item of factors 2 to 4 cross-loaded onto the factor one, too. The highest loadings were considered the most representative, except for the item 6 (Table 3). Since the difference of its loadings was less than 0.1, we categorized it in factor 4 in accordance with the theory of original WHODAS. The total Cronbach’s α was 0.92 ranged from 0.91 for item 5 to 0.93 for both items of Cognition.

The PCA-based factor loading plot depicted clusters containing items with similar correlations of more than 0.30 with a factor (Fig. 2) [50]. There were three items separated from other clusters (learning, concentrating and household responsibilities), cross loaded onto two factors of one and four.

**Scale score reliability**

**Floor and ceiling effects**

For the total data (N=255), the floor effect values (an answer of “no problem”) ranged from 24% (item 12 (day-to-day work)) to 91% (item 6 (concentrating)). All values were above 25% except for items 5 (emotions) and 12 (day-to-day work). The ceiling effect values (an answer of
“extremely large” or “cannot do”) ranged from 0.4% (item 6 (concentrating)) to 6% (item 1 (standing)) except for item 7 (walking) with 32% ceiling effect.

Internal consistency
For the total data (N = 255), the Cronbach’s α coefficients resulted in 0.93 (total scale) ranged from 0.92 (items 1 (standing), 4 (community activities), 5 (emotions), 10 (dealing with), 11 (maintaining a friendship), and 12 (day-to-day work)) to 0.93 (items 2 (household responsibilities), 3 (learning), 6 (concentration), 7 (walking), 8 (washing) and 9 (getting dressed)). The Cronbach’s α coefficients for subscales fluctuated between 0.82 (for Participation) and 0.88 (Cognition).

Internal consistency was assessed for the test–retest data (N = 50) by estimating Cronbach’s α coefficient for the whole scale (0.89), subscales, and items ranged from 0.88 for items 8 (washing) and 9 (getting dressed) to 0.90 for item 7 (walking). The Cronbach’s α coefficients for individual subscales were high (Table 4).

Test–retest reliability
Consistency over time was examined by comparing the results of the first and second measurements in 50 people. The skewness-kurtosis test showed the normal distribution for some subscales 4, 5, and 6, on test data. Kendall’s tau-b indicated no significant differences between test and retest responses (Table 4).

The reliability of the test–retest method was confirmed using the ICC. Time consistency became very high as ranged from 0.89 (subscale 4) to 0.99 (subscale 2) indicating good to excellent reliability and stability. For the overall result, the ICC was 0.97, demonstrating excellent test–retest reliability (Table 4). The ICC for items ranged from 0.85 to 0.97, good to excellent stability (Table 2).

The test data showed the lowest and the highest mean disability score in Cognition (3.1), and Mobility (7.0), respectively (Table 4).

Table 4  Reliability indices of the Persian version of 12-item WHODAS 2.0 in phone-based survey (N = 50)

| Subscales (items)       | Mean (SD) | Cronbach’s alpha | ICC (95% CI) | Agreement | Kendall’s tau-b* |
|-------------------------|-----------|------------------|--------------|-----------|------------------|
|                         | test      | Retest           |              |           |                  |
| Cognition (3 & 6)       | 3.10 (1.90) | 2.86 (1.80)     | 0.79         | 0.93 (0.87–0.96) | Excellent         | 0.71 |
| Mobility (1 & 7)        | 7.02 (2.69) | 7.06 (2.62)     | 0.84         | 0.99 (0.98–0.99) | Excellent         | 0.91 |
| Self-care (8 & 9)       | 3.64 (2.26) | 3.36 (2.07)     | 0.77         | 0.92 (0.85–0.95) | Excellent         | 0.75 |
| Getting along (10 & 11) | 5.56 (2.34) | 4.94 (2.19)     | 0.78         | 0.89 (0.78–0.94) | Good              | 0.75 |
| Life activities (12 & 2)| 6.68 (1.47) | 6.50 (1.43)     | 0.75         | 0.94 (0.89–0.97) | Excellent         | 0.87 |
| Participation (4 & 5)   | 5.90 (1.84) | 5.68 (1.74)     | 0.76         | 0.96 (0.93–0.98) | Excellent         | 0.84 |
| Total score             | 31.90 (9.52) | 30.40 (8.68)   | 0.81         | 0.97 (0.92–0.98) | Excellent         | 0.82 |

Each subscale has two items. *At significance level of 0.0001

Discussion
This paper developed the Persian WHODAS 2.0 by translation, then described psychometric properties of the Persian version of phone-based interviewer-administered tool for road traffic injury survivors in Tabriz, Iran, six months after crash, between 21 May 2020 and 20 March 2021. In this study, we translated the tool into Persian, and indicated that Persian WHODAS 2.0 is psychometrically reliable to the great extent in phone-based assessment of disability. Satisfactory internal consistency reveals the Persian WHODAS to measure the construct of disability and deliver high reliable score at item level and total score in our representative sample. The schedule is robust to changes in the same subjects over time and less prone to measurement error caused by time elapse, as ICC showed excellent reliability. Exploratory factor analysis discloses the high validity of the schedule as items hang consistently together to create the single construct of disability. The final model represented four factors of social/self-activities, cognition, mobility, and self-care, accounted for 86% of the total variance.

This study worked on the translation of the brief WHO-DAS, although, it was translated and validated in different languages even in Persian [20, 22]. However, the Persian version was applied neither specifically through phone survey nor in RTIs patients. The psychometric properties of a questionnaire could be affected through the method of data collection such as phone surveys comparing with in-person/self-administered interviews. The available Persian translation was used in interview surveys [20, 22], not phone surveys. Moreover, in data collection by a questionnaire, there may be misunderstanding of questions and recall bias because of language/cultural barriers. Also, questionnaires are not exactly developed for people with certain conditions [51]. The results may be influenced by subjective measures, too [51]. Because the team planned to apply the tool later in a national prospective post-crash cohort study, it was so critical to have
a robust Persian version in informal Persian language and easy to understand. Thus, the team decided to translate the available version. To do so, we obtained the permission from the WHO (permission ID: 380480).

According to the exploratory factor analysis, the Persian WHODAS 2.0 measured disability with four components accounted for 86% of total variance and demonstrated better results compared with a study in trauma patients [22], and some other studies [52–54]. The factor loadings of the individual items were strong, as eleven items’ factor loadings were above 0.60 and one item showed 0.39, consistent with factor loadings in a study [52]. One factor, titled social/self-activities, contains items pertaining to the subscales of Getting along, Life activities, and Participation, similar to a study in terms of the items of the extracted factor, not the number of extracted factors in the study of Abedzadeh–kalahroudi [22]. Despite similar number of sample in both studies, her study gathered data by interview with patients categorized in five groups (Traffic, Home, Assault, Work, and Other). The data collection method and uniformity of sample may cause in contradiction and fewer number of extracted factors. Three separate factors certainly distinguished cognition, mobility, and self-care subscases. The structure is quite similar to the structure found in the original English version [29]. A study indicated different groupings, attained a two-factor solution in older people from different dementia sites, with subscales of mobility, life activities, cognition, and participation in its first dominant factor and the self-care and getting along subscales in its second factor [55]. It seems higher number of sample size would help to extract the remaining three subscales separately as the same three distinct factors as in the original version. However, it is better to consider the “most appropriate” rather than the “correct” number of factors.

No measurement difference was indicated between the test and retest results. In spite of importance of using the spectrum of non-extreme in face-to-face interviews, it is not practical in phone-based surveys. However, we attained high consistency between total and subscales’ scores in phone-based surveys.

The Cronbach’s α coefficient for the Persian version indicated the ideal internal consistency and better results compared with a study among Syrian refugees (α=0.74)[52]. The Cronbach’s α coefficient was similar to several studies with comparable target groups (Ethiopian road traffic patients α=0.88 [56], in trauma patients α=0.91[22]). Although both studies used the tool through in-person interviews, not phone interviews, we reached the high internal consistencies as Ustun et al. reported [29]. The result was the same as the result of studies with different target groups (a global study [34], Brazilian study in patients with chikungunya (α=0.93) [57], in only phone-based psychometric study of WHO-DAS among critically ill patients (α=0.91) [36], Spanish version in depressed patients (α=0.89)[53], and Portuguese version in patients with musculoskeletal pain (α=0.84) [58]). It is necessary to note that the internal consistency between WHODAS scores of patients with different injuries such as traumatic brain injury and spinal cord injury was high [59, 60]. Then, comparing the results of such different studies could be reasonable. The highest Cronbach’s α coefficient was related to Cognition in consistent with a study in Ethiopia [56], as both worked on RTI patients.

The study confirmed that the scale was consistent over time representing excellent test–retest reliability as ICC fell in a range from 0.89 (Getting along) to 0.99 (Mobility) which is in line with a study in people with and without hand injuries (ICC value of 0.88) [61]. A study in patients with Kashin-Beck disease affirmed a good repeatability of the 12-item WHODAS 2.0 [62]. We reached higher time consistency at total and subscale levels compared with a study in Polish society (ICC value of 0.91, ranged from 0.50 (Self-care) to 0.97 (Mobility)) [24]. Both studies showed the highest ICC for the subscale of Mobility with the items of “Standing for long periods such as 30 min?” and “Walking a long distance such as a kilometre [or equivalent]?” It could be explained by the fact that these activities are of physical activities rarely affected by any actions through the time between test and retest. Generally, the results might differ regarding the mode of administration (in-person interviews versus our phone-based interview) and characteristics of the study population such as age, health status, country of origin) [63]. Moreover, the quality of translation definitely influence the understanding of questions by interviewees.

Data showed floor effects but no ceiling effects, except for the item 7 (walking). It was almost in line with baseline result of a study in critically ill patients [36] as well as a study of polish society [24]. Overall, study setting and way of using a tool might bring about floor and ceiling effects in data. In this study, the tool was used through phone interviews. However, the content of the tool was examined by asking the opinion of interviewees about the clarity and comprehensibility [64].

This tool is short, and have less complex questions, so it is a proper choice to use in such prospective cohort study with various questionnaires and examinations. The tool was completed via phone calls. This technique is essential for a couple of reasons. It takes around five minutes to complete it on phone calls [65]. It enabled data collection without unnecessary travel of patients to the study site/hospital or travel of interviewers to the place of patients’ residence, due to patients’ movement restrictions. Consecutively it helped to avoid their exposure to pathogens.
in the current situation of the Pandemic COVID-19. Also, in-person interview is not an easy-to-access opportunity for every researcher/participant. Our data were attained by one assistant interviewer with more than four years of experience in running phone interviews to fill out questionnaires determined in the IRTIRS. So, lack of such skills would also lead to different results in phone-based interviews. Since conducting telephone interviews requires special communication skills, it is necessary to repeat the research on the basis of phone-interview conducted by different interviewers. On the other hand, some patients such as elderly patients and those with hearing problems may have problem with phone survey. So it is recommended to reassess the tool in specific target population such as elderly people.

In this study, subjects were recruited at six months after crash. It is important to ensure that their function was partially stable and patients have got recovered from their minor injuries causing slight physical restrictions [17].

As different works on gender difference in road traffic injuries reported, male users have higher risk of injuries [66–71] due to various reasons such as their higher exposure to driving, patterns of risky traffic behaviours, and so on. The current study also indicated the same pattern, especially the common proportion in Iran.

Strengths and limitations
To the best of our knowledge, this is the first study in which the psychometric properties of the Persian version of the 12-item WHODAS 2.0 have been evaluated through phone interviews in road traffic injury patients above 18 years in the region and even the region. This study enables assessing functioning, and disability in injured population unable to attend the study site and conducting studies to compare results between different settings.

The age limitation in the tool assessment is defined owing to some feasibility issues. So, the psychometric assessment of brief WHODAS in teenagers is recommended. We plan to extend the research by developing and administrating the tool to the population with Azeri language, the most common language in the northwest Iran. Since it is necessary to use the valid and reliable tool for people whose mother language is Azeri, and there are difficulties in running the Persian version for aged and illiterate people. Regarding the effect of various communication skills, we recommend running phone-based survey by different interviewers.

The potential sources of bias might be considered when making inferences based on our findings. This study as an exploratory cross-sectional one, six months after crash, never provides evidence on the longitudinal performance of the tool for short or long term timing after crash. Moreover, this was a cohort study involving patients with road traffic injuries hospitalized in the only two referral hospitals in Northwest Iran and registered. Thus, it mainly represents the patients needed to be hospitalized post-crash. Conducting the research on the basis of phone interviews in the same and different contexts but in larger sample size is needed. We suggest using a longitudinal design to evaluate responsiveness to change and test–retest reliability.

Conclusion
The current study translated into Persian and examined the psychometric properties of the 12-item WHODAS 2.0 in a study population experienced road traffic crashes and injuries approximately six months prior to study time and hospitalized in one of two referral trauma centres located in the city of Tabriz, Iran. We realized that the disability assessment tool is short and easy-to-use, and it is suitable to apply in phone-based interviews during follow-ups of injured patients above 18 years.

Abbreviations
WHODAS 2.0: World Health Organization Disability Assessment Schedule 2.0; RTIs: Road traffic injuries; ICF: International Classification of Functioning, Disability and Health; ICD-11: 11Th International Classification of Disease; LMICs: Low and middle income countries; PERSIAN: Prospective Epidemiological Research Study in Iran; RTIRC: Road Traffic Injury Research Center; IRTIRS: Integrated Road Traffic Injury Registry System; ICC: Intra-class correlation coefficient; EFA: Exploratory factor analysis.

Acknowledgements
The authors would like to thank the participants in the study.

Author contributions
HSB conceived and designed the study, HSB and NS conducted the analyses and, HSB, NS interpreted the results. HSB and NS drafted the manuscript. MAJ, MF, and SBH contributed in the critical drafting and revising of the manuscript for important intellectual content. All authors read and approved the final manuscript.

Funding
Not applicable.

Availability of data and materials
Data are available upon reasonable request. The interested researchers may contact Dr. Homayoun Sadeghi-Bazargani (PI). The applications will be reviewed upon approval by the research council and the regional ethics committee of Tabriz University of Medical Sciences. The study website is https://cohortsafety.tbzmed.ac.ir/.

Declarations
Ethics approval and consent to participate
This study valued ethical precepts in research and was performed in agreement with the principles of the Declaration of Helsinki. The Ethics Committee of the Road Traffic Injury Research Center, Tabriz University of Medical Sciences (cohort ethics code: IR.TBZMED.REC.1398.543) evaluated and approved the study. An informed consent was required from all participants.
Consent for publication
All authors have read and announced their consent for publication of this work.

Competing interests
The authors declare that they have no competing interests.

Author details
1 Cabrini Research, Cabrini Health, VIC 3144, Australia. 2 Road Traffic Injury Research Center, Tabriz University of Medical Sciences, Tabriz 5167846311, East Azerbaijan, Islamic Republic of Iran. 3 Injury Epidemiology and Prevention Research Group, Turku Brain Injury Center, Turku University Hospital and University of Turku, Turku, Finland. 4 Research Center of Psychiatry and Behavioural Sciences, Tabriz University of Medical Sciences, Tabriz, Islamic Republic of Iran. 5 College of Medicine, Charles Drew University of Medicine and Science, Los Angeles, CA, USA. 6 School of Public Health and Preventative Medicine, Faculty of Medicine, Nursing and Health Sciences, Monash University, VIC 3800, Australia.

Received: 7 April 2022   Accepted: 29 June 2022
Published online: 09 July 2022

References
1. Sadeghi-Bazargani H, Samadiad B, Shahedifar N, Golestani M. Epidemiology of road traffic injury fatalities among car users; a study based on forensic medicine data in east azerbaijan of iran. Bull Emerg Trauma. 2018;6(2):146–54. http://doi.org/10.29522/betar‑60209.
2. Palmira‑Suárez R et al. Disability related to road traffic crashes among adults in Spain. Gac Sanit. 2015;29:493–8.
3. Peden M, McGee K, Krug E. Injury: A leading cause of the global burden of disease, 2000. Geneva: World Health Organization. 2002.
4. World Health Organization. Global Status Report on Road Safety: time for action. Geneva: World Health Organization, 2009.
5. Abegaz T, Gebremedhin S. Magnitude of road traffic accident related injuries and fatalities in Ethiopia. PLoS ONE. 2019;14(1); e0202240.
6. Bazaran‑Hejazi S et al. The Burden of Road Traffic Injuries in Iran and 15 Surrounding Countries: 1990–2016. Arch Iran Med. 2018;21(12):556–65.
7. Homayoun SB, Bahram S, Mina G, Nasrin S, Milad J. Analysis of provincial mortalities among bus/minibus users over twelve years, East Azerbaijan, Iran. J Med Life. 2018;11(4):312–9. https://doi.org/10.25122/jml‑2018‑0051.
8. Institute for Health Metrics and Evaluation. 2010. http://ghdx.healthdata.org/record/ihme‑data‑gbd‑2010‑results‑cause‑1990‑2010.
9. Saadat S et al. The most important causes of death in iranian population; a retrospective cohort study. Emerg (Tehran). 2015;3(1):16–21.
10. Nakjavani‑Shahrahi B et al. Worthing physiological score vs revised trauma score in outcome prediction of trauma patients; a comparative study. Emerg (Tehran). 2017;5(1):e31.
11. Stuck G, Bickenbach J. Functioning: the third health indicator in the health system and the key indicator for rehabilitation. Eur J Phys Rehabil Med. 2017;53(1):134–8.
12. Sadeghi‑Bazargani H, Hasanzadeh K, Salarilak S, Amir S, Golestani M, Shahedifar N. Evaluating the relationship between adult attention‑deficit/hyperactivity disorder and riding behavior of motorcyclists. J Injury Viol Res. 2019; https://doi.org/10.5249/jivr.v11i1.1098.
13. Castro S et al. The World Health Organization Disability Assessment Schedule 2 (WHODAS 2.0): remarks on the need to revise the WHODAS. Cadernos Saúde Pública. 2019;35:37.
14. World Health Organization. International Classification of Functioning, Disability and Health (ICF). 2021 [cited 2021 24/05/2021]. https://www.who.int/standards/classifications/international‑classification‑of‑functioning‑disability‑and‑health.
15. World Health Organization. WHO Disability Assessment Schedule 2.0 (WHODAS 2.0). 2018. https://www.who.int/classifications/icf/whodasi/ en/, retrieved at 29.03.2020.
16. Papadaki M et al. Psychological distress and physical disability in patients sustaining severe injuries in road traffic crashes. Results from a one‑year cohort study from three European countries. Injury. 2017;48(2):297–306.
17. Marin S et al. The protocol for validating phone interview tools on post‑discharge consequences of road traffic injuries. J Injury Violence Res. 2020;12(3):1368.
18. Sadeghi‑Bazargani H et al. PERSIAN Traffic Safety and Health Cohort: A Study Protocol on Post‑Crash Mental and Physical Health Consequences. J Injury Prev. 2022;28:269–79.
19. Norouzi H et al. Psychometric properties of the world health organization disability assessment scale 2.0 among iranian cancer patients. Middle East J Cancer. 2020;11(3):333–42.
20. Rajaeiezafar S et al. Validity of the 36‑item Persian (Farsi) version of the world health organization disability assessment schedule (WHODAS) 2.0. Int J Mental Health. 2019;48(1):14–39.
21. Salehi R et al. Validity and reliability of the world health organization disability assessment schedule 2.0 36‑item persian version for persons with multiple sclerosis. Korean J Fam Med. 2020;41(3):195–201.
22. Abedzadeh‑Kalâhroodi M et al. Psychometric properties of the world health organization disability assessment schedule II‑12 Item (WHODAS II) in trauma patients. Injury. 2016;47(5):1104–8.
23. World Health Organization. Process of translation and adaptation of instruments. [cited 2021 17 May 2021]. http://www.who.int/substance‑Abuse/Research‑Tools/Translation/En/.
24. Civrile‑Sozariška A et al. Psychometric properties and validation of the polish version of the 12‑item WHODAS 2.0. BMC Public Health. 2020;20(1):1203.
25. Sadeghi‑Bazargani H et al. Developing a national integrated road traffic injury registry system: a conceptual model for a multidisciplinary setting. J Multidiscip Healthc. 2020;13:983–96.
26. Bujaing MA, Baharan N. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. 2017.
27. Lynn MR. Determination and quantification of content validity. Nurs Res. 1986;35(6):382–5.
28. T. BedihamNST, et al. Chapter 58: The World Health Surveys, in Health Systems Performance Assessment. World Health Organization, Geneva; 2003. p. 797–808.
29. Ustun, T.B., et al., Measuring health and disability: manual for WHO Disability Assessment Schedule (WHODAS 2.0) / edited by T.B. Ustun, N. Kostanjsek, S. Chatterji, J.Rehm. 2010, World Health Organization: Geneva. 30. Silva AG et al. Self‑reported disability: association with lower extremity performance and other determinants in older adults attending primary care. Phys Ther. 2015;95(12):1626–37.
31. Mousazadeh Y, Sadeghi‑Bazargani H, Janati A, Poustchi H, Zakeri R, Shahedifar N. Evaluating the relationship between adult attention‑deficit/hyperactivity disorder and riding behavior of motorcyclists. J Injury Viol Res. 2019; https://doi.org/10.5249/jivr.v11i1.1098.
32. Lee HH et al. Is WHODAS 2.0 Useful for Colorectal Cancer Survivors? Ann Rehab Med. 2017;41(4):667–76.
33. Ustün TB, et al. Developing the World Health Organization disability assessment scale, WHODAS II: reliability and validity in the measurement of activity and participation in a spinal cord injury population. J Rehabil Med. 2012;44(9):747–55.
34. Higgins A et al. The psychometric properties and minimal clinically important difference for disability assessment using WHODAS 2.0 in critically ill patients. Critical Care Resuscit. 2021;23:103–12.
35. Ando‑sal F, et al. The World Health Organization disability assessment scale. J Phys Ther Sci. 2010;22(4):881–8.
36. Ćwirlej‑Sozańska A et al. Psychometric properties and validation of the polish version of the 12‑item WHODAS 2.0. BMC Public Health. 2020;20(1):1203.
37. Sadeghi‑Bazargani H et al. Developing a national integrated road traffic injury registry system: a conceptual model for a multidisciplinary setting. J Multidiscip Healthc. 2020;13:983–96.
38. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients: a review. 2017.
39. Lynn MR. Determination and quantification of content validity. Nurs Res. 1986;35(6):382–5.
40. World Health Organization. International Classification of Functioning, Disability and Health (ICF). 2021 [cited 2021 24/05/2021]. https://www.who.int/standards/classifications/international‑classification‑of‑functioning‑disability‑and‑health.
41. World Health Organization. WHO Disability Assessment Schedule 2.0 (WHODAS 2.0). 2018. https://www.who.int/classifications/icf/whodasi/ en/, retrieved at 29.03.2020.
42. Papadaki M et al. Psychological distress and physical disability in patients sustaining severe injuries in road traffic crashes. Results from a one‑year cohort study from three European countries. Injury. 2017;48(2):297–306.
41. Amick R, Charro P, Patterson J. Test–retest reliability of the sway balance mobile application. J Mobile Technol Med. 2015;4:40–7.
42. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess. 1994;6:40–284.
43. Watkins MW. Exploratory factor analysis: a guide to best practice. J Black Psychol. 2018;44(3):219–46.
44. Izquierdo I, Olea J, Abad FJ. Exploratory factor analysis in validation studies: uses and recommendations. Psicothema. 2014;26(3):395–400.
45. Fabrigar LR, Wegener DT. Exploratory factor analysis: Oxford; Oxford University Press, 2011.
46. Terwee CB, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42.
47. Hayton JC, Allen DG, Scarpello V. Factor retention decisions in exploratory factor analysis: a tutorial on parallel analysis. Organ Res Methods. 2004;7(2):191–205.
48. Neill J. Exploratory Factor Analysis. Lecture S: Survey Research & Design in Psychology, 2018.
49. MacCallum RC, et al. Sample size in factor analysis: the role of model error. Multivar Behav Res. 2001;36(4):361–7.
50. Tavakol M, Wetzel A. Factor Analysis: a means for theory and instrument development in support of construct validity. Int J Med Educ. 2010;1:245–7.
51. Doma K, et al. Comparison of psychometric properties between recall methods of interview-based physical activity questionnaires: a systematic review. BMC Med Res Methodol. 2019;19(1):43.
52. Akhtar A, et al. Exploratory and confirmatory factor analysis of the 12-item Arabic World Health Organization Disability Assessment Schedule (WHODAS 2.0) as a screening tool for Syrian refugees. BJPsych Open. 2021;7(6):e190.
53. Luciano JV, et al. Psychometric properties of the twelve item World Health Organization Disability Assessment Schedule II (WHODAS II) in Spanish primary care patients with a first major depressive episode. J Affect Disord. 2010;121(1–2):52–6.
54. Saltchev M, et al. A study of the psychometric properties of 12-item World Health Organization Disability Assessment Schedule 2.0 in a large population of people with chronic musculoskeletal pain. Clin Rehabil. 2017;31(2):262–72.
55. Sousa RM, et al. Measuring disability across cultures—the psychometric properties of the WHODAS II in older people from seven low- and middle-income countries. The 10/66 Dementia Research Group population-based survey. Int J Methods Psychiatr Res. 2010;19(1):1–17.
56. Denu ZA, et al. The 12 items Amharic version WHODAS-2 showed cultural adaptation and used to measure disability among road traffic trauma victims in Ethiopia. BMC Psychol. 2021;9(1):1.
57. Jéssica, A., et al., Psychometric properties of WHODAS for use in patients with chikungunya in Brazil. Fisioterapia e Pesquisa, 2019.
58. Silva C, et al. Adaptation and validation of WHODAS 2.0 in patients with musculoskeletal pain. Rev Saude Publica. 2013;47(4):752–8.
59. Tarvonen-Schroeder S, Kaljonen A, Laimi K. Utility of the World Health Organization Disability Assessment Schedule and the World Health Organization minimal generic set of domains of functioning and health in spinal cord injury. J Rehabil Med. 2019;51(1):40–6.
60. Tarvonen-Schroeder S, et al. Comparing disability between traumatic brain injury and spinal cord injury using the 12-item WHODAS 2.0 and the WHO minimal generic data set covering functioning and health. Clin Rehabil. 2018;32(2):1676–83.
61. Marom BS, et al. Cross-cultural adaptation of the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) for Hebrew-speaking subjects with and without hand injury. Disabil Rehabil. 2017;39(12):1155–61.
62. Younus MI, et al. Reliability and validity of the 12-item WHODAS 2.0 in patients with Kashin-Beck disease. Rheumatol Int. 2017;37(9):1567–73.
63. Kirberger I, et al. Feasibility and psychometric properties of the German 12-item WHO Disability Assessment Schedule (WHODAS 2.0) in a population-based sample of patients with myocardial infarction from the MONICA/KORA myocardial infarction registry. Popul Health Metrics. 2014;12(1):27.
64. Anjos DBMD, et al. Reliability and construct validity of the Instrument to Measure the Impact of Valve Heart Disease on the Patient’s Daily Life. Rev Lat Am Enfermagem. 2016;24:e2730–e2730.
65. White MC, et al. Measurement of patient reported disability using WHO-DAS 20 before and after surgical intervention in Madagascar. BMC Health Serv Res. 2018;18(1):305–305.
66. González-Sánchez G, et al. Traffic injury risk based on mobility patterns by gender, age, mode of transport and type of road. Sustainability. 2021;13(18):10112.
67. World Health Organization. Gender and road traffic injuries. Geneva: World Health Organization, 2002.
68. Santamaria-Rubio E, et al. Gender differences in road traffic injury rate using time travelled as a measure of exposure. Accid Anal Prev. 2014;65:1–7.
69. Shams M, et al. The level and trend of road traffic injuries attributable mortality rate in Iran, 1990–2015: a story of successful regulations and a roadmap to design future policies. BMC Public Health. 2021;21(1):1722.
70. Yousefifard M, et al. Risk factors for road traffic injury-related mortality in Iran, a systematic review and meta-analysis. Arch Acad Emerg Med. 2021;9(1):e61.
71. Sadeghi-Bazargani H, et al. Epidemiology of Road Traffic Mortalities among bus/minibus users in East Azerbaijan. Iran. J Injury Viol Res. 2019;11(4):1484.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.