

Received: 2022.07.16

Accepted: 2022.08.16

Available online: 2022.09.07

Published: 2022.09.23

# Reliability and Validity of Patient-Reported Outcome Measures for Ankle Instability in Hebrew

Authors' Contribution:  
Study Design A  
Data Collection B  
Statistical Analysis C  
Data Interpretation D  
Manuscript Preparation E  
Literature Search F  
Funds Collection G

ABCDEF 1 **Uri Gottlieb**  
BDE 2 **Tomer Yona**  
BDE 1 **David Shein Lumbroso**  
E 1 **Jay R. Hoffman**  
ABCDEF 1 **Shmuel Springer**

1 Neuromuscular and Human Performance Laboratory, Department of Physiotherapy, Ariel University, Ariel, Israel  
2 Department of Biomedical Engineering, Technion, Israel Institute of Technology, Haifa, Israel

**Corresponding Author:** Uri Gottlieb, e-mail: [urigott@gmail.com](mailto:urigott@gmail.com)

**Financial support:** None declared

**Conflict of interest:** None declared

**Background:** The International Ankle Consortium has recommended several instruments for assessing and diagnosing chronic ankle instability. These include the Ankle Instability Instrument (AII), Cumberland Ankle Instability Tool (CAIT), Identification of Functional Ankle Instability (IdFAI), and Foot and Ankle Ability Measure (FAAM). This study aimed to translate, culturally adapt, and assess the reliability, validity, discriminative power, and classification agreement of the Hebrew online versions of the AII, CAIT, IdFAI, and FAAM.

**Material/Methods:** After translation and cross-cultural adaptation of the questionnaires, we recruited 87 participants with self-reported ankle disorders to evaluate the psychometric properties of the questionnaires. To evaluate each questionnaire's discriminative power, we recruited 75 healthy participants. Reliability was assessed by calculating internal consistency (Cronbach's alphas) and test-retest intraclass correlation coefficients. Convergent and divergent validity were assessed by Spearman's correlation between each instrument and the Short-Form-12 (SF-12) score for physical and mental components, respectively.

**Results:** All instruments had acceptable internal consistency ( $\alpha > 0.7$ ) and good test-retest reliability ( $ICC_{2,1} > 0.8$ ), except for the IdFAI ( $ICC_{2,1} = 0.73$ ). All instruments had moderate convergent validity ( $\rho > 0.4$  with SF-12 physical component) except for AII ( $\rho = 0.36$ ). No instrument was correlated with the SF-12 mental component score (good divergent validity,  $\rho < 0.3$ ). All instruments had excellent discriminative power (area under the receiver operator characteristic curve  $> 0.9$ ). Strong correlations were found between all instruments.

**Conclusions:** The Hebrew online versions of the AII, CAIT, IdFAI, and FAAM are valid, consistent, and reliable and may be used in research or clinical settings.

**Keywords:** **Ankle Injuries • Joint Instability • Patient Reported Outcome Measures • Reproducibility of Results**

Full-text PDF: <https://www.medscimonit.com/abstract/index/idArt/937831>



## Background

Chronic ankle instability (CAI) is a function-restricting condition characterized by repeated episodes or perception of the ankle giving way and may be accompanied by persistent symptoms such as pain, weakness, or limited ankle range of motion [1]. The prevalence of CAI ranges from 9% to 76% after an acute ankle sprain [2]. This wide range can be explained by different measurement tools, differences in the studied populations, and different diagnostic criteria for CAI.

The diagnostic criteria for CAI were proposed by the International Ankle Consortium [3] and primarily consist of personal history and subjective reports of the ankle giving way or a feeling of ankle instability lasting more than 1 year after the initial injury. The subjective report should be accompanied by a validated patient-reported outcome measure (PROM), such as the Ankle Instability Instrument (AII) [4], the Cumberland Ankle Instability Tool (CAIT) [5], or the Identification of Functional Ankle Instability (IdFAI) [6] which was developed to consolidate the various elements of the 2 previous CAI instruments (AII and CAIT) into a simple and concise means of identifying individuals with CAI [6]. However, it is unknown to what extent these tools are consistent with each other and whether they are interchangeable. Furthermore, in addition to these discriminative PROMs, which are primarily used to identify individuals with CAI [7], it is recommended that a self-reported general foot and ankle function PROM, such as the Foot and Ankle Ability Measure (FAAM) [8] be used to capture individuals' perceptions of their limitations. A recent study evaluating the content, criterion, construct validity, reliability, and responsiveness of 17 PROMs relevant as instruments for assessing ankle instability concluded that the FAAM is the best outcome measure for patients with CAI [9].

Although PROMs are tested for reliability and validity in their original language, they need to be translated and validated to allow comparisons between national and international study results and improve clinical practice by assessing patients using internationally recognized instruments [10,11]. To date, the AII, CAIT, IdFAI, and FAAM have not been translated and validated into Hebrew.

Therefore, the current study aims to (i) translate and culturally adapt the AII, CAIT, IdFAI, and FAAM into Hebrew, (ii) assess their internal consistency and test-retest reliability, (iii) assess their convergent and divergent validity, and (iv) evaluate their discriminative power.

## Material and Methods

### Ethical Considerations

The study was approved by the Ariel University ethical committee (approval number: AU-HEA-SS-20211102, dated November 2, 2021). The study followed the COSMIN study design checklist for patient-reported outcome measurement instruments [12].

### Instruments

The AII consists of 9 *yes/no* items. According to the International Ankle Consortium [3], an affirmative answer to item #1 (*Have you ever sprained an ankle?*) plus at least 4 additional *yes* answers is considered positive for CAI. The CAIT consists of 9 multiple response questions and is completed separately for each ankle. Its score ranges from 0 to 30, with a score of less than 24 considered positive for CAI [3]. The IdFAI, which combines questions from the 2 previous instruments, consists of 10 multiple response questions and is completed separately for each ankle. Its score ranges from 0 to 37, and scores greater than 11 are considered positive for CAI [3]. The FAAM is a self-reported functional questionnaire used to assess subjective changes in perceived pain and ankle function. It consists of 29 items in 2 subscales: Activities of Daily Living (FAAM<sub>ADL</sub>) and Sports (FAAM<sub>Sport</sub>). Each subscale is scored between 0 to 100, with higher scores indicating better ankle condition. These 4 PROMs (AII, CAIT, IdFAI, and FAAM) were previously translated and culturally adapted into several languages.

The Global Rating of Change (GRC) scale was used to identify subjective deterioration or improvement in the ankle condition [13]. The GRC consisted of a 7-point Likert scale with the following attributes: *significantly worse*; *worse*; *mildly worse*; *very small change or no change at all*; *mildly better*; *better*; *significantly better*.

### Translation

The translation of the AII, CAIT, IdFAI, and FAAM followed the 6-step guidelines for the process of cross-cultural adaptation [11]. These guidelines require 1) initial translation into Hebrew by 2 independent translators fluent in English and Hebrew; 2) combining the 2 translations into a single agreed version; 3) back-translation into English by a native English speaker who is not a medical professional; 4) discussing the translation by an expert committee consisting of 3 physical therapists until reaching a final consensus version of each instrument; 5) completion of each instrument by a group of 30 participants with CAI, who were interviewed regarding ambiguous or unclear words or syntax; and 6) the documentation of the final version of each instrument.

## Psychometric Evaluation

To assess the translated psychometric properties of the questionnaires, we conducted an online survey completed by the participants at 2 different time points, approximately 14 days apart [12]. The survey included all 4 questionnaires (All, CAIT, IdFAI, and FAAM) and the Hebrew version of the 12-item Short Form Survey (SF-12), which was used to assess the construct validity of the ankle-specific PROMs. The SF-12 is a 12-item questionnaire designed to evaluate the respondent's general, physical, and mental health. The summary score of the SF-12 consists of a physical aspect (PCS) and a mental aspect (MCS). Higher scores indicate better health. The Hebrew version of the SF-12 is reliable and valid for the Hebrew-speaking population [14].

The survey was launched online (<https://www.qualtrics.com/>), and participants were recruited through social media groups of Israeli amateur athletes. Inclusion criteria included Hebrew language fluency, age between 18 and 65 years, and a self-reported chronic ankle condition, such as chronic ankle instability. Additionally, a healthy control group with similar inclusion criteria except for self-reported chronic ankle conditions was recruited via social media to analyze the discriminative power of these questionnaires, as well as to find the optimal cut-off score to differentiate between individuals with and without ankle disorders. The healthy control participants completed the same survey in a single session.

Data collection took place between November 2021 and January 2022. Potential respondents were presented with a landing page describing the study, the inclusion criteria, the researchers' contact details, and informed consent information. Responders who were willing to participate were asked to provide contact details (phone number and email address) and preferred method of communication (WhatsApp message, text message, email, or phone call; multiple selections were allowed). Based on this information, a reminder to complete the survey for the second time was sent to the participants 13 days after the first time point. Respondents' personal information was kept separately and could not be matched with the survey responses. To match participants' responses between the first and second time points, respondents were required to provide a personal code at each time point. The respondents were asked to complete the survey regarding their injured ankle. If both ankles were affected, the participants were asked to choose the more severely injured ankle.

## Statistical Analysis

Statistical analyses were performed using Python v.3.8. Internal consistency, test-retest reliability, standard error of the measurement (SEM), minimal detectable change (MDC), and construct validity were calculated for each PROM. Next, we evaluated the classification agreement between the PROMs and

assessed their discriminative power. Internal consistency, which reflects the extent to which the questionnaire items are inter-correlated or whether they measure the same construct consistently, was calculated using Cronbach's alpha ( $\alpha$ ). Acceptable consistency is indicated when  $\alpha > 0.7$  [15]. Test-retest reliability was calculated using 2-way random effects, absolute agreement, and interclass correlation coefficient ( $ICC_{2,1}$ ). The following scale was used to determine the ICC score: moderate ( $>0.5$ ), good ( $>0.75$ ), and excellent ( $>0.9$ ) [16].

SEM was calculated as  $SEM = SD \cdot \sqrt{1 - ICC}$ , where SD is the pooled standard deviation. MDC was calculated as  $SEM = SD \cdot \sqrt{2}$  [16]. Lastly, we considered a floor and ceiling effect if more than 15% of the participants had the highest or lowest score for each instrument [15].

Construct validity was assessed using Spearman's rank correlation coefficient ( $\rho$ ) between each pair of questionnaires. To evaluate the agreement among the 3 CAI classification questionnaires (All, CAIT, and IdFAI), each response was classified as a binary result (CAI/non-CAI) according to the International Ankle Consortium cut-off recommendations [3]. Matthews' correlation coefficients were calculated to assess the agreement between each pair of classifiers [17,18]. Matthews' correlation, also known as the Phi coefficient, is a measure of association between 2 binary variables. Agreement was interpreted as weak ( $<0.3$ ), moderate (0.4-0.7) or strong ( $>0.7$ ) [19].

We hypothesized that all 4 PROMs would be moderately correlated with the SF-12 PCS (construct validity,  $\rho > 0.5$ ) but only weakly correlated with the SF-12 MCS (divergent validity,  $\rho < 0.3$ ). Furthermore, we assumed strong correlations between the different subscales of the FAAM ( $\rho > 0.6$ ) and between the CAIT and IdFAI scores ( $\rho > 0.6$ ), and moderate correlation ( $\rho > 0.4$ ) between the All and all other questionnaires. Regarding classification agreement, we hypothesized a strong agreement between the CAIT and IdFAI, and moderate agreement between the All and the other 2 classifiers.

Finally, a receiver operator characteristic (ROC) curve was calculated for each instrument, and Youden's index was used to find the optimal cut-off score for reporting ankle-related symptoms [20]. Youden's index is a single statistic that captures the performance of a dichotomous diagnostic test; namely, it finds the cut-off score which maximizes the test sensitivity and specificity.

## Results

There were no difficulties in the translation process. Hebrew is a grammatically gendered language, and care was taken to ensure that the translated questionnaires were gender neutral. The Hebrew word for squatting (item #10 in the FAAM<sub>ADL</sub>)

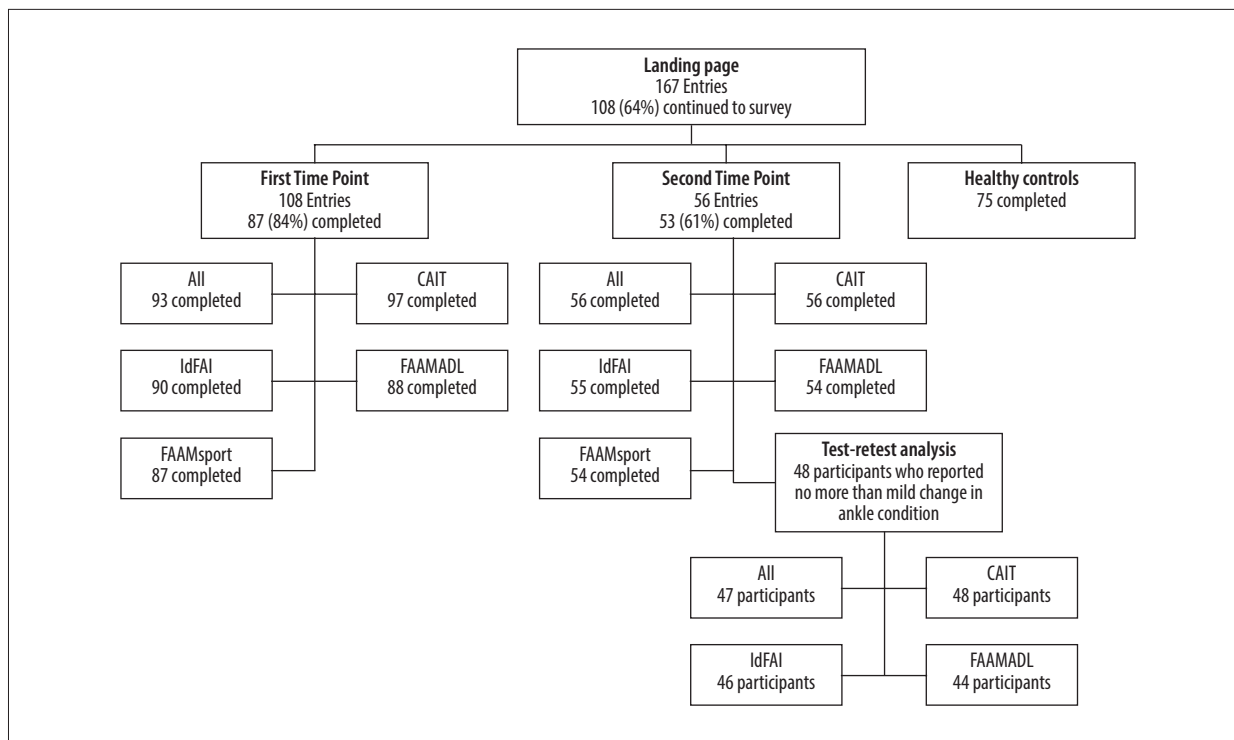


Figure 1. Participants' flow diagram. Figure created using MS-Word, version 2207, Microsoft.

Table 1. Participant's demographic characteristics.

	Participants with ankle conditions (n=87)	Healthy participants (n=76)
Age (median, IQR)	32 (13)	36 (14.0)
Sex (men/women, % women)	34/53 (60.9%)	29/46 (61.3%)
Reported diagnosis (n, %)		
CAI	70 (80.5%)	
Severe ankle injury in the last months	10 (11.5%)	
Chronic pain	2 (2.3%)	
Other	5 (5.7%)	

CAI – chronic ankle instability; IQR – inter-quartile range.

is often misunderstood. Therefore, a short description of the action (“bending your knees while the feet are supported on the ground”) followed by the non-Hebrew term (squat) was used in the translated version.

A total of 167 people with ankle conditions visited the landing page, of whom 108 (64%) provided contact information

Table 2. Outcomes of the instruments. All values are median (IQR).

	Participants with ankle condition	Healthy participants
All	6.0 (3.0)	1.0 (2.0)
CAIT	15.0 (8.0)	27.0 (3.5)
IdFAI	22.0 (8.5)	5.0 (6.0)
FAAM <sub>ADL</sub>	90.6 (15.3)	100.0 (1.3)
FAAM <sub>Sport</sub>	67.9 (33.9)	100.0 (5.4)
SF-12 PCS	54.8 (11.9)	57.2 (2.3)
SF-12 MCS	46.1 (15.1)	50.0 (12.2)

All – Ankle Instability Instrument; CAIT – Cumberland Ankle Instability Tool; IdFAI – Identification of Functional Ankle Instability; FAAM – Foot and Ankle Ability Measure; ADL – Activities of Daily Living; SF – short-form; PCS – Physical Component Score; MCS – Mental Component Score.

and continued to the questionnaires. Eighty-seven participants (84%) completed all questionnaires at the first time-point with a median completion time of 10.8 min (IQR: 5.5). Fifty-three participants (61%) completed all questionnaires at the second time-point. The median between-completion duration was 14 days (IQR: 2). Seventy-five healthy participants

**Table 3.** Psychometric characteristics of all instruments.

Instrument	Sample size (n)	Internal consistency ( $\alpha$ )	Test-retest reliability (ICC)	Construct validity	Floor and ceiling effects (CAI group)	Cut-off score	SEM	MDC
All	168	0.86 (0.82-0.89)	0.82 (0.71-0.90)	CAIT: 0.85 IdFAI: 0.90 FAAM <sub>ADL</sub> : 0.75 FAAM <sub>Sport</sub> : 0.83 SF-12 PCS: 0.36 SF-12 MCS: 0.12	Floor: None Ceiling: 7.5%	$\geq 5$	0.71	1.96
CAIT	172	0.89 (0.86-0.91)	0.88 (0.79-0.93)	All: 0.85 IdFAI: 0.87 FAAM <sub>ADL</sub> : 0.84 FAAM <sub>Sport</sub> : 0.87 SF-12 PCS: 0.44 SF-12 MCS: 0.17	None	$\leq 23$	1.85	5.14
IdFAI	162	0.91 (0.89-0.93)	0.73 (0.56-0.84)	All: 0.90 CAIT: 0.87 FAAM <sub>ADL</sub> : 0.75 FAAM <sub>Sport</sub> : 0.84 SF-12 PCS: 0.49 SF-12 MCS: 0.11	None	$\geq 13$	3.1	8.6
FAAM <sub>ADL</sub>	161	0.95 (0.94-0.96)	0.84 (0.73-0.91)	All: 0.75 CAIT: 0.84 IdFAI: 0.76 FAAM <sub>Sport</sub> : 0.84 SF-12 PCS: 0.49 SF-12 MCS: 0.20	Floor: None Ceiling: 5.7%	$\leq 97.5$	4.5	12.5
FAAM <sub>Sport</sub>	161	0.96 (0.95-0.97)	0.88 (0.79-0.93)	All: 0.83 CAIT: 0.88 IdFAI: 0.84 FAAM <sub>ADL</sub> : 0.84 SF-12 PCS: 0.42 SF-12 MCS: 0.19	Floor: None Ceiling: 1.1%	$\leq 87.5$	7.0	19.5

ICC – Intraclass Correlation Coefficient; CAI – Chronic Ankle Instability; SEM – Standard Error of the Measurement; MDC – Minimal Detectable Change; All – Ankle Instability Instrument; CAIT – Cumberland Ankle Instability Tool; IdFAI – Identification of Functional Ankle Instability; FAAM – Foot and Ankle Ability Measure; ADL – Activities of Daily Living; SF – short-form; PCS – Physical Component Score; MCS – Mental Component Score.

completed the survey in a median time of 7.7 min (IQR: 4.8). **Figure 1** describes the participant flow diagram, and **Table 1** describes the participants' demographic characteristics. **Table 2** describes the outcomes of each instrument.

According to the GRC scale, most participants reported very small or no change at all between the 2 time points (n=41, 73.2%). Mild improvement or deterioration was reported by 5 participants (8.9%) each, and an improvement or worsening in ankle condition was reported by 1 (1.8%) and 2 (3.6%) participants, respectively. No participants reported significant changes in ankle condition. The 8 participants (14%) who reported more than a mild change in their ankle condition at the second time point were excluded from the test-retest analysis.

The psychometric characteristics of the Hebrew translations are displayed in **Table 3**. All instruments had acceptable internal consistency ( $\alpha > 0.7$ ). Further information regarding Cronbach's  $\alpha$  analysis can be found in the **Appendix**. Good test-retest reliability was demonstrated for all instruments except the IdFAI (ICC<sub>2,1</sub> = 0.73). All instruments had moderate convergent validity ( $\rho > 0.4$  with SF-12 PCS), except for the All ( $\rho = 0.36$ ), and good divergent validity ( $\rho < 0.3$  with SF-12 MCS). No significant floor or ceiling effects were found for any of the instruments. The CAI classifications agreed moderately between the All and CAIT classifications ( $\phi = 0.63$ ) and strongly between the All and the IdFAI ( $\phi = 0.71$ ) and between the CAIT and the IdFAI ( $\phi = 0.76$ ).

## Discussion

In this study, we described the psychometric properties of the Hebrew version of 4 PROMs used to diagnose and assess CAI and provided important information about the relationship between these instruments.

The Hebrew All's internal consistency ( $\alpha=0.86$ ) was similar to the original English version ( $\alpha=0.89$ , [4]), as well as to the French and Persian translations ( $\alpha=0.86-0.87$ , [21,22]). The ROC analysis yielded an optimal cut-off score of 5, which is identical to the cut-off score reported by Locquet et al in the French translation of the All [21] and the International Ankle Consortium recommendations for patient selection [3]. While the test-retest reliability of the Hebrew All was considered good ( $ICC_{2,1}=0.82$ ), it was lower than that of previous versions ( $ICC=0.93-0.98$ ) [4,21,22].

The Hebrew CAIT internal consistency ( $\alpha=0.89$ ) was similar to the values previously reported in the French, Dutch, and Chinese translations ( $\alpha=0.86-0.89$ , [23-25]). The ROC analysis yielded an optimal cut-off score of 23, which is comparable with the cut-off score reported by Geerinck et al in the French translation of the CAIT [23], as well as in the International Ankle Consortium recommendations for patient selection [3]. Once again, the test-retest reliability of the Hebrew version ( $ICC_{2,1}=0.88$ ) was slightly lower than the previously reported reliability scores of  $ICC=0.93-0.96$  [5,23-25].

The Hebrew IdFAI internal consistency ( $\alpha=0.91$ ) was similar to the values reported in the Greek and Spanish translations ( $\alpha=0.88-0.94$ , [26,27]). The ROC analysis yielded an optimal cut-off score of 13, slightly higher than the recommended score of 11 points [3,26]. Only moderate test-retest reliability was achieved for the Hebrew IdFAI ( $ICC_{2,1}=0.73$ ), which is considerably lower than previously reported scores of  $0.85-0.95$  [26-28]. It is unclear why this instrument resulted in the lowest test-retest reliability scores. One possible explanation may be that the questions of the IdFAI are a combination of the 2 previous instruments, and responders were less focused when they realized they were being asked the same questions repeatedly. A second explanation could be related to its items and responses. While all instruments ask about functional tasks, the IdFAI's responses include detailed frequencies that may be less intuitive to responders. For example, the seventh item of the CAIT (*My ankle feels unstable when...*) requires respondents to select a specific, unique task (eg, running on uneven surfaces, walking on uneven surfaces, walking on a flat surface). In contrast, responses to most functional items on the IdFAI include frequencies such as *never, once a year, or once a month*.

Finally, the 2 FAAM subscales had comparable internal consistency and test-retest reliability scores with the original English

version ( $\alpha=0.98$ ,  $ICC=0.87$ ) [8], as well as with the French and Brazilian-Portuguese versions ( $\alpha=0.90-0.97$ ,  $ICC=0.82-0.94$ ) [29,30]. Interestingly, while other versions reported strong correlation between the SF-36 PCS and  $FAAM_{ADL}$  ( $\rho=0.78-0.84$ ) and  $FAAM_{Sport}$  ( $0.65-0.80$ ), our results showed moderate correlation only ( $\rho=0.49$  and  $\rho=0.42$ , respectively). The studied populations may explain these differences: the French and Brazilian versions recruited participants with "various foot and ankle disorders", while our sample consisted mostly of individuals with CAI, a condition which may be less function-limiting in daily life. Although the FAAM is not intended to be used as a classifier between individuals with and without ankle disorders, Youden's indices were computed for both subscales. These results can be used for this purpose in clinical scenarios.

In general, the Hebrew versions of all instruments had acceptable internal consistency, and construct validity was verified both in terms of convergence and divergence. However, except for the 2 FAAM subscales, all other instruments demonstrated lower test-retest reliability scores than previous translations. This may be attributed to the study design as an online survey, which may have resulted in higher variability. The convergent and divergent validity results were in line with our hypothesis that moderate and weak correlations would be found between all instruments and the SF-12 PCS and SF-12 MCS, respectively.

To the best of our knowledge, this is the first study to compare all 3 ankle instability classification instruments. As expected, the All and CAIT had only moderate classification agreement, whereas the IdFAI had a strong classification agreement with both the All and the CAIT. This is not surprising since the IdFAI was constructed as a combination of the All and the CAIT. On the other hand, due to the IdFAI's slightly lower test-retest reliability ( $ICC_{2,1}=0.73$ ), no clear recommendation can be made as to which instrument should be used by researchers and clinicians.

This online survey was open to the general population and was based on a self-defined ankle condition. Therefore, the sample of this study may represent a different population than samples drawn from outpatient clinic visitors, which is a common approach for recruiting participants in validation studies. However, our approach may better represent the utilization of PROMs in a research setting. Recruitment of study participants often involves selecting relevant individuals from the general population. Online questionnaires are more time- and cost-efficient when inviting potential participants to research institutions for inclusion testing. Thus, this study supports the feasibility of conducting online surveys for questionnaire validation.

Another aspect to consider is the relatively long completion time for the entire survey, which may have led to a phenomenon known as respondent fatigue [31], in which participants' tiredness may affect the study results. This may explain the

relatively large dropout rate (more than 10% of participants who completed the first questionnaire did not finish the entire survey). Nonetheless, there did not seem to be a decrease in the questionnaires' psychometric characteristics, which mitigates the presence of respondent fatigue. Nevertheless, responding to that many questionnaires at once is rarely utilized in research or clinical settings, where participants will usually complete only 2 of the 4 instruments.

The study design had some limitations. No long-term follow-up was performed, and it is unknown whether participants attended any form of rehabilitation. Therefore, it was impossible to measure the responsiveness, the minimal clinically important difference, or the predictive validity of the instruments. Additionally, since the participants were not formally diagnosed, it may be difficult to distinguish between participants who had CAI and those who had other ankle-related disorders. Pain was not assessed (for example, using a numeric pain rating scale), and therefore further construct validity was not performed.

## Appendix – Description of Cronbach's $\alpha$ Analysis

### Ankle Instability Instrument

Mean (variance)	3.7 (7.4)
Median (IQR)	4.0 (5.0)
Cronbach's $\alpha$ (95% CI)	0.86 (0.82-0.89)

	Scale mean if removed	Scale variance if removed	Cronbach alpha if removed	Item-scale correlation
Item 1	2.98	6.05	0.84	0.64
Item 2	3.16	5.58	0.83	0.75
Item 3	3.48	6.16	0.85	0.59
Item 4	3.12	5.58	0.83	0.76
Item 5	3.61	6.45	0.85	0.56
Item 6	3.23	5.41	0.82	0.81
Item 7	3.23	5.51	0.83	0.77
Item 8	3.65	6.70	0.86	0.48
Item 9	3.44	5.82	0.84	0.70

## Conclusions

The translated AII, CAIT, IdFAI, and FAAM are valid, consistent, and reliable for use in Hebrew speaking patients with ankle instability. Therefore, clinicians and researchers can use these questionnaires to better assess functional deficits and severity of ankle instability. Furthermore, the calculated cut-off scores and MDC values provide important psychometric information that may improve patient diagnosis and clinical progress assessment.

### Declaration of Figures' Authenticity

All figures submitted have been created by the authors, who confirm that the images are original with no duplication and have not been previously published in whole or in part.

**Cumberland Ankle Instability Tool**

Mean (variance)	19.7	(61.9)
Median (IQR)	21.0	(13.0)
Cronbach's $\alpha$ (95% CI)	0.89	(0.86-0.91)

	Scale mean if removed	Scale variance if removed	Cronbach alpha if removed	Item-scale correlation
Item 1	15.99	43.58	0.88	0.78
Item 2	17.02	47.75	0.87	0.77
Item 3	17.73	48.93	0.87	0.81
Item 4	17.44	50.66	0.87	0.77
Item 5	18.47	53.70	0.88	0.67
Item 6	17.61	49.27	0.87	0.81
Item 7	16.77	47.79	0.87	0.79
Item 8	18.02	52.81	0.88	0.63
Item 9	18.24	51.95	0.89	0.58

**Identification of Functional Ankle Instability**

Mean (variance)	14.1	(99.3)6
Median (IQR)	13.0	(17.8)
Cronbach's $\alpha$ (95% CI)	0.91	(0.89-0.93)

	Scale mean if removed	Scale variance if removed	Cronbach alpha if removed	Item-scale correlation
Item 1	12.32	77.50	0.90	0.77
Item 2	12.90	84.02	0.91	0.69
Item 3	13.31	83.65	0.92	0.59
Item 4	11.59	67.54	0.90	0.88
Item 5	12.73	77.38	0.89	0.90
Item 6	12.56	87.22	0.91	0.75
Item 7	11.88	79.54	0.90	0.76
Item 8	13.00	78.52	0.90	0.84
Item 9	12.70	78.62	0.90	0.83
Item 10	12.32	77.50	0.90	0.77



### Foot and Ankle Ability Measure – Activities of Daily Living

Mean (variance)	77.4	(93.4)
Median (IQR)	81.0	(8.3)
Cronbach's $\alpha$ (95% CI)	0.96	(0.95-0.97)

	Scale mean if removed	Scale variance if removed	Cronbach alpha if removed	Item-scale correlation
Item 1	73.59	85.88	0.95	0.78
Item 2	73.58	86.31	0.95	0.81
Item 3	73.64	85.60	0.95	0.74
Item 4	73.82	82.20	0.95	0.82
Item 5	74.14	79.33	0.95	0.84
Item 6	73.65	85.69	0.95	0.74
Item 7	73.88	82.42	0.95	0.95
Item 8	74.14	80.19	0.95	0.82
Item 9	73.66	85.56	0.95	0.74
Item 10	73.91	83.52	0.96	0.67
Item 11	73.98	81.52	0.95	0.78
Item 12	73.50	90.88	0.96	0.47
Item 13	73.47	90.90	0.96	0.52
Item 14	73.54	87.37	0.95	0.71
Item 15	73.75	81.45	0.95	0.83
Item 16	73.54	88.41	0.96	0.66
Item 17	73.61	86.61	0.95	0.79
Item 18	73.50	90.00	0.96	0.51
Item 19	73.65	85.86	0.95	0.79
Item 20	74.03	79.08	0.95	0.87
Item 21	73.78	85.01	0.95	0.73

### Foot and Ankle Ability Measure – Sports

Mean (variance)	21.8	(46.1)
Median (IQR)	24.0	(10.0)
Cronbach's $\alpha$ (95% CI)	0.96	(0.95-0.97)

	Scale mean if removed	Scale variance if removed	Cronbach alpha if removed	Item-scale correlation
Item 1	18.69	33.65	0.95	0.91
Item 2	18.66	33.55	0.95	0.91
Item 3	18.87	32.73	0.94	0.94
Item 4	18.52	35.44	0.95	0.87
Item 5	18.85	33.35	0.95	0.91
Item 6	18.41	36.87	0.96	0.82
Item 7	18.80	33.32	0.95	0.89

## References:

- Hertel J, Corbett RO. An updated model of chronic ankle instability. *J Athl Train*. 2019;54(6):572-88
- Lin CI, Houtenbos S, Lu YH, Mayer F, Wippert PM. The epidemiology of chronic ankle instability with perceived ankle instability – a systematic review. *J Foot Ankle Res*. 2021;14(1):41
- Gribble PA, Delahunt E, Bleakley C, et al. Selection criteria for patients with chronic ankle instability in controlled research: A position statement of the International Ankle Consortium. *Br J Sports Med*. 2014;48(13):1014-18
- Docherty CL, Gansnedder BM, Arnold BL, Hurwitz SR. Development and reliability of the ankle instability instrument. *J Athl Train*. 2006;41(2):154-58
- Hiller CE, Refshauge KM, Bundy AC, et al. The Cumberland ankle instability tool: A report of validity and reliability testing. *Arch Phys Med Rehabil*. 2006;87(9):1235-41
- Simon J, Donahue M, Docherty C. Development of the Identification of Functional Ankle Instability (IdFAI). *Foot Ankle Int*. 2012;33(9):755-63
- Houston MN, Van Lunen BL, Hoch MC. Health-related quality of life in individuals with chronic ankle instability. *J Athl Train*. 2014;49(6):758-63
- Martin RL, Irrgang JJ, Burdett RG, et al. Evidence of validity for the Foot and Ankle Ability Measure (FAAM). *Foot Ankle Int*. 2005;26(11):968-83
- Hansen CF, Obionu KC, Comins JD, Krogsgaard MR. Patient reported outcome measures for ankle instability. An analysis of 17 existing questionnaires. *Foot Ankle Surg*. 2022;28(3):288-93
- Wild D, Grove A, Martin M, et al. Principles of Good practice for the translation and cultural adaptation process for Patient-Reported Outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health J Int Soc Pharmacoeconomics Outcomes Res*. 2005;8(2):94-104
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25(24):3186-91
- Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Qual Life Res Int J Qual Life Asp Treat Care Rehabil*. 2010;19(4):539-49
- Kamper SJ, Maher CG, Mackay G. Global Rating of Change Scales: A review of strengths and weaknesses and considerations for design. *J Man Manip Ther*. 2009;17(3):163-70
- Amir M, Lewin-Epstein N, Becker G, Buskila D. Psychometric properties of the SF-12 (Hebrew Version) in a primary care population in Israel. *Med Care*. 2002;40(10):918-28
- Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42
- Foundations of Clinical Research: Applications to evidence-based practice – F.A. Davis Company. Accessed February 6, 2022. <https://www.fadavis.com/product/physical-therapy-foundations-clinical-research-portney-4>
- Delgado R, Tibau XA. Why Cohen's Kappa should be avoided as performance measure in classification. *PLoS One*. 2019;14(9):e0222916
- Chicco D, Warrens MJ, Jurman G. The Matthews Correlation Coefficient (MCC) is more informative than Cohen's Kappa and Brier Score in Binary Classification Assessment. *IEEE Access*. 2021;9:78368-81
- Dancey CP, Reidy J. *Statistics Without maths for psychology*. 7th ed. Pearson; 2017. Accessed February 6, 2022. <https://www.pearson.com/uk/educators/higher-education-educators/program/Dancey-Statistics-Without-Maths-for-Psychology-7th-Edition/PGM1768952.html>
- Ruopp MD, Perkins NJ, Whitcomb BW, Schisterman EF. Youden Index and optimal cut-point estimated from observations affected by a lower limit of detection. *Biom J Biom Z*. 2008;50(3):419-30
- Locquet M, Benhotman B, Bornheim S, et al. The "Ankle Instability Instrument": Cross-cultural adaptation and validation in French. *Foot Ankle Surg*. 2021;27(1):70-76
- Mohamadi S, Ebrahimi I, Dadgoo M, et al. Translation, cross-cultural adaptation and factor analysis of the Persian version of ankle instability instrument. *Med J Islam Repub Iran*. 2018;32:79
- Geerinck A, Beaudart C, Salvan Q, et al. French translation and validation of the Cumberland Ankle Instability Tool, an instrument for measuring functional ankle instability. *Foot Ankle Surg*. 2020;26(4):391-97
- Vuurberg G, Kluit L, van Dijk CN. The Cumberland Ankle Instability Tool (CAIT) in the Dutch population with and without complaints of ankle instability. *Knee Surg Sports Traumatol Arthrosc*. 2018;26(3):882-91
- Wang W, Liao D, Kang X, et al. Development of a valid Chinese version of the Cumberland Ankle Instability Tool in Chinese-speaking patients with chronic ankle instability disorders. *Sci Rep*. 2021;11(1):9747
- Cervera-Garvi P, Ortega-Avila AB, Marchena-Rodriguez A, Gijon-Nogueron G. Transcultural adaptation and validation of the Spanish version of the Identification of Functional Ankle Instability questionnaire (IdFAI-Sp). *Disabil Rehabil*. 2022;44(13):3221-27
- Tsekoura M, Billis E, Samada EK, et al. Cross cultural adaptation, reliability and validity of the Greek version of Identification of Functional Ankle Instability (IdFAI) questionnaire. *Foot Ankle Surg*. 2021;27(8): 906-910.
- Beaudart C, Demoulin C, Mehmeti K, et al. Validity and reliability of the French translation of the Identification of Functional Ankle Instability (IdFAI). *Foot Ankle Surg*. 2022;28(6):756-62
- Borloz S, Crevoisier X, Deriaz O, et al. Evidence for validity and reliability of a french version of the FAAM. *BMC Musculoskelet Disord*. 2011;12(1):40
- Moreira TS, de Castro Magalhães L, Silva RD, et al. Translation, cross-cultural adaptation and validity of the Brazilian version of the Foot and Ankle Ability Measure questionnaire. *Disabil Rehabil*. 2016;38(25):2479-90
- O'Reilly-Shah VN. Factors influencing healthcare provider respondent fatigue answering a globally administered in-app survey. *Peer J*. 2017;5:e3785