Is the Arabic version of the WeeFIM instrument reliable and feasible enough to measure functional independence in Egyptian children with burns?

Eman othman (dr.emanothman@yahoo.com)
Cairo university

Research Article

Keywords: WeeFIM instrument, functional independence, cross-cultural adaptation, pediatric burns, reliability, feasibility

Posted Date: March 17th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1460850/v1

Abstract

Purpose

The adaptation of the Arabic WeeFIM instrument to an interview format suitable for burned children with communicative and/or cognitive problems with evaluation of its feasibility and reliability.

Methods

An observational study with two measurements within a 2-week time period in a sample of 53 patients with healed burns who were aged 3 to 16 years. Feasibility was evaluated by the assessment of the frequency of missing answers per item and administration time. The reliability was assessed by a test-retest procedure. For content validity, experts were asked to complete all questions in the Arabic Index of Content Validity (ICV).

Results

The Arabic version of WeeFIM has borderline reliability (Cronbach's Alpha = 0.619 and Pearson correlation coefficient: r = 0.986). There was an acceptable percent to have a feasible test as 83.963% of the filled questionnaires had no missing answers and 16.037% had missing answers. The ICV showed that all the questions were relevant except for two questions: Question 15, for expression, and question 18, for memory.

Conclusion

The Arabic version of WeeFIM has high test-retest reliability, borderline internal consistency, and excellent feasibility in measuring the functional outcome of burned children.

Main Text

Normative WeeFIM statistics for American, Chinese, Dutch, Turkish, and Indian children have already been validated. Previous research has found that ethnic, cultural, and environmental differences between countries affect the pattern of independence [9, 10, 12–14]. Due to cultural and environmental differences among countries and the dearth of WeeFIM applications in Egypt, normative data for Egyptian kids is needed due to cultural and environmental differences among countries. Therefore, the aim of this study was to investigate the reliability and feasibility of the Arabic version of the WeeFIM instrument in Egyptian kids from 3 years to 16 years of age with healed burns. The final results of this study showed that the Arabic version of WeeFIM has high test-retest reliability, borderline internal consistency, and excellent feasibility in measuring the functional outcome of burned children.

Introduction

Burn wounds in children under the age of seven account for 1/4 to 1/2 of all burn wounds worldwide. [1, 2]. Males aged 2.5 to 18 years are more likely to be admitted to the hospital with a burn injury as they're more subjected to harmful activities and play more violently [3]. Because children are the most vulnerable victims of burns and may suffer long-term problems as a result of their injuries, special attention should be paid to their evaluation and treatment procedures throughout research [4, 5]. Parents must be aware of specific cognitive and physical limitations in children of various ages in order to develop safety guidelines that can withstand children's requests for participation in situations or activities with a high risk of injury [6]. Until now, the extent to which burn injury impacts a child's functional independence has been unknown. To find the risk factors for reduced functioning due to burns and to encourage early intervention, it is critical to focus on the actual degree of functional independence in pediatric burned patients [7]. A valid evaluation and a competent rehabilitation program are required to improve ADL performance [8]. The WeeFIM test is useful for assessing basic ADL in kids. The outcome of using this instrument is a measurement of the child's ability to complete the task safely, precisely, and without assistance [9]. The WeeFIM instrument is based on and modified from the Functional Independence Measure (FIM) for adults, but with habitability, developmental, and pediatric interdisciplinary perspectives [10]. Reliability, agreement, and responsiveness are all relative terms, as are circumstance-specific terms that are highly dependent on the study population and measurement conditions [9, 11]. Normative WeeFIM statistics for American, Chinese, Dutch, Turkish, and Indian children have already been validated. Previous research has found that ethnic, cultural, and environmental differences between countries affect the pattern of independence [9, 10, 12–14]. Because children develop at various rates in different cultures, transferring a set of cultural norms to another society may misrepresent a child's true developmental position [15]. Due to cultural and environmental differences among countries and the dearth of WeeFIM applications in Egypt, normative data for Egyptian kids is needed due to cultural and environmental differences among countries. With a normative database, the progression of independence at home and in the community can be assessed. Therefore, the aim of this study was to investigate the reliability and feasibility of the Arabic version of the WeeFIM instrument in Egyptian kids from 3 years to 16 years of age with healed burns.

Materials And Methods

Study design

An observational, double-center, cross-sectional study design was used to investigate the reliability and feasibility of the Arabic version of the WeeFIM instrument in Egyptian kids from 3 years to 16 years of age with healed burns. Ethical approval was obtained from the institutional review board at the Faculty of Physical Therapy, Cairo University before the beginning of the study, commencement with the number P.T. REC/012/003578. This study was registered at Clinical trials.gov with the reference number NCT05276869. The procedures used in this study adhere to the tenets of the Declaration of Helsinki. All the children and their parents give their informed consent before the beginning of the study, after

explaining the nature and purpose of the measurement tools, informing them of their right to refuse or withdraw at anytime, as well as about the confidentiality of any obtained information.

Participants

Enrollment of participants was done for all children who meet the inclusion criteria in the outpatient burn clinic in the Faculty of Physical Therapy-Cairo University, and Om Elmasryeen Hospital, Cairo, Egypt. The demographic details about each parent and child were noted, and screening was done for adherence to inclusion and exclusion criteria. The tool was given to all children aged 3 to 16 years who were hospitalized in one of the two places for more than two weeks after the burn. Inclusion began in September 2021 and ended in February 2022. It is possible that some children were not assessed based on the extent of their injuries. As a result, children with more severe burns were less likely to be included. A total of fifty-three children (26 males and 27 females) who had a healed burn with an age range of 3 years to 16 years were found to be eligible for the study. All participants or their parents had to be able to read, write, and speak Arabic. They had passed two weeks post-burn. Children with a history of neurological impairment or orthopaedic surgery to the limbs or spine, or those with severe burns (4th-degree burns), extensive burns greater than 60%, infected burn wounds, or other medical conditions that might impair physical performance, as well as children with any medical illness present in the previous month, were excluded from the study. The exclusion of medical conditions was based only on history and examination.

Materials

The WeeFIM tool can be used in a wide range of fields. Self-care (6 items), sphincter control (2 items), transfers (3 items), mobility (2 items), communication (2 items), and social cognition (2 items) were all included in the WeeFIM instrument (3 items). There were 13 items on the motor sub-scale, which covered self-care, sphincter control, transfer, and mobility. The cognitive domain is made up of the remaining two components (communication and social cognition). The WeeFIM is a standard ordinal scale with seven levels. Level 7 involves no aid from the child, and the work is completed by the child without the use of technology. There is no concern about safety or taking an excessive amount of time throughout the work. Level 6 indicates a degree of limited independence, which may include the use of assistive technology or the failure to complete a task in a timely or safe manner. The performance was rated at level 1 (complete help) [16].

Procedures

According to Niemeijer et al., [9], the WeeFIM instrument was translated into Arabic and culturally altered before being used with Egyptian-burned children. The WeeFIM evaluation was administered for 15–20 minutes by directly interviewing parents about the child’s performance on tasks in Arabic. The researchers had asked two translators who were fluent in English and Arabic, to provide the first translator with a forward translation. The first translator had been asked to give us the overall meaning of each item, which should be easy to understand at the reading level of an approximately 7-year-old candidate. The other translator was asked to do backward translation (i.e., translate the new Arabic version into English).

The researchers reviewed the translated instrument and checked for any modifications. Finally, the researchers had made a simple form for the questionnaire to be suitable for practical use, and then the instrument was applied to the study group and the researcher assessed its feasibility, reliability, and content validity.

A-Feasibility was evaluated by the assessment of the frequency of missing answers per item and administration time. Where the missing answers count was taken in the 1st week and compared to that of the 2nd week, with missing-1 referring to the number of missing answers in the 1st week and missing-2 referring to the number of missing answers in the 2nd week. And the administration time was taken using a clock or a stopwatch, with time-1 referring to the time taken to fill out the questionnaire in the 1st week, and time-2 referring to the time taken to fill out the questionnaire in the 2nd week.

B-Test–retest reliability: First, the researcher gathered a randomly selected sample of participants. Second, the researchers explained the procedure to the participants. Third, the researchers explained the items of the questionnaire to the participants. Fourth, participants were asked to fill out the scale with a score of 1. Fifth, one week apart, they were asked to re-fill the scale with a score of 2. Sixth, test-retest reliability was analyzed by using Pearson's correlation coefficient (PCC) and internal consistency was measured by using Cronbach's alpha-coefficient.

C-Content validity

Content validity deals with the extent to which an instrument reflects the meaningful elements of the content without extraneous elements [17]. In the process of content validity, experts were asked to complete all questions in the Arabic Index of Content Validity (ICV) by circling the number that represented their opinion of each one. Each expert rated each item as either 1 (agreed), 0 (undetermined), or 1 (disagreed). The ICV for each item was calculated using the summation of scores from each expert divided by the number of experts. There is no quantitative index available for this type of validity. Content validity is often judged simply by comparing the content of an instrument with the domains that are intended as the areas to be measured, and it is judged by seeking expert opinion from health professionals involved in the study and subjects who participated in the study [17]. In the current study, the researcher used the Arabic index of content validity (ACV) to determine content validity.

Final adaptation of the questionnaire

The researcher had made a final adaptation of the questionnaire after the study to take clinical practice into account, creating an adapted form that combines experts' opinions (which were taken using the index of content validity) with practical experience.

Independent Variables

Bias can occur when assumed independent variables such as age and gender affect responses [18]. In the current study, the researchers performed a correlation study for age, gender and both measured

scores and measured times.

Data analysis

Statistical analyses were performed using SPSS Statistics version 22.0 (IBM Corp., Armonk, NY, USA). The level of statistical significance was set at 0.05% with a 95% confidence interval. Feasibility was measured by using the mean of administration time and the count of missing answers. Test-retest reliability was measured by using the Pearson Correlation Coefficient (PCC) to correlate the 1st score with the 2nd score. Internal consistency was measured by using Cronbach's alpha coefficient. Correlation analysis between both measured scores and both measured times was performed by using PCC and P-value. A Wilcoxon signed-rank test was performed to detect a p-value for both the measured scores and both measured times. A Mann-Whitney test (two independent samples) was performed to detect the effect of the independent variables (gender and age) on both measured scores. Content validity was assessed by performing simple statistics of ICV values. P-value of 0.05 was considered statistically significant.

Results

Descriptive Statistics

1-Qualitative variables

Fifty-three patients of both sexes with healed burns who were from 3 to 16 years of age were divided into two groups; the age group (A) from 3 to 7 years included 36 patients (67.92%), and the age group (B) from 8 to 16 years included 17 patients (32.08%). The patients were also classified into two groups according to gender; the gender group (A) for boys included 26 patients (49.06%), and the gender group (B) for girls included 27 patients (50.94%), making the total number of participants (N) = 53 patients.

2- Quantitative variables

Table 1 showed that the mean score_1 for the first week was of high value (92.96) and the mean score_2 was also high (94.04), with a slight increase in mean score, which may be attributed to some improvement in patients' functions. The mean values of time_1 and time_2 were low, indicating the feasibility of filling, and the mean of time_2 (5.57) was less than that of time_1, which may be attributed to a better understanding of the questionnaire by the parents.

Wilcoxon signed ranks test for comparison between both scores, both times and both frequencies of missing answers of WeeFIM

- Between score_1 and score_2:

The P-value was 0.092 ($P \leq 0.1$), which has statistical significance at level alpha = 0.1, indicating improvement in the second assessment than the first one, which can be attributed to some improvement in the patients' function (Table 2).

- Between time_1 and time_2:

The P-value was 0.000 ($P \leq 0.05$), indicating significant improvement in the second assessment compared to the first, which may be attributed to the parents' greater understanding of the questionnaire items (Table 3).

-Between missing_1 and missing_2:

The P-value was 0.014 ($P \leq 0.05$), which indicates fewer missing answers in the second assessment than in the first, which could be attributed to the parents' better.

Correlations (using Pearson correlation coefficient and P-value)

- Between score_1 and score_2:

There was a very strong direct relationship between score_1 and score_2 as r = 0.986 (r is Pearson correlation coefficient), which is close to the perfect score, and there was a high statistical significance of score_1 in relation to score_2 with a P-value of 0.000, indicating improvement of the patients' scores in the second test compared to their scores in the first test (Table 5).

-Between score_1 and time_1:

There was a weak inverse relationship between score_1 and time_1 as r = -0.324 and there was a high statistical significance of score_1 in relation to time_1 with a P-value of 0.018, indicating decreasing the time of filling of the questionnaire in relation to high patients' scores in the first test (Table 5).

- Between score_1 and time_2:

There was a weak inverse relationship between score_1 and time_2 as r = -0.233 and there was no statistical significance of score_1 in relation to time_2 with a P-value of 0.092 (Table 5).

-Between score_2 and time_1:

There was a weak inverse relationship between score_2 and time_1 as r = -0.340 and there was a high statistical significance of score_2 in relation to time_1 with a P-value of 0.013, indicating that the increase in the score in the second test is relevant to less time of filling in the first test (Table 5).

-Between score_2 and time_2:

There was a weak inverse relationship between score_2 and time_2 as r = -0.239 and there was no statistical significance for score_2 in relation to time_2 with a P-value of 0.085 (Table 5).

-Between time 1 and time 2:

There was a weak direct relationship between time_1 and time_2 as r = 0.516 and there was a high statistical significance of time_1 in relation to time_2 with a P-value of 0.000, indicating improvement in the time of filling of the questionnaire in the second test compared to that in the first test (Table 5).

The Mann-Whitney test (two independent samples) for testing the effect of age and gender on both scores and both times of WeeFIM

Gender and score_1

The P-value was 0.341 (P > 0.05), indicating that there was no significant difference between the scores of boys and girls, as the mean rank of the scores of boys was 29.06 and that of girls was 25.02 (Table 6).

-Gender and score_2

The P-value was 0.403 (P > 0.05), indicating that there was no significant difference between the scores of boys and girls, as the mean rank of the scores of boys was 28.81 and that of girls was 25.26 (Table 7).

-Gender and time_1

The P-value was 0.763 (P > 0.05), indicating that there was no significant difference between the time of filling out the questionnaire for boys and girls, as the mean rank of the time of filling out the questionnaire for boys was 27.62 and that for girls was 26.41 (Table 8).

-Gender and time_2

The P-value was 0.451 (P > 0.05), indicating that there was no significant difference between the time of filling out the questionnaire for boys and girls, as the mean rank of the time of filling out the questionnaire for boys was 25.46 and that for girls was 28.48 (Table 9).

-Age and score_1

The P-value was 0.182 (P > 0.05), indicating that there was no significant difference between the scores of both groups as the mean rank of the score of group (A) (children from 3 to 7 years) was 25.06 and that of group (B) (children from 8 to 16 years) was 31.12 (Table 10). This result may be attributed to the greater number of children in group (A) (36 children) than in group (B) (17 children), and that 25% of group (A) has a score < 70 while 75% of group (A) has a score > 70, and that 11.76% of group (B) has a score < 70 while 88.24% of group (B) has a score > 70. So, the researchers recommended more research for better results.

-Age and score_2

The P-value was 0.195 (P > 0.05), indicating that there was no significant difference between the scores of both groups as the mean rank of the scores of group (A) was 25.11 and that of group (B) was 31.00 (Table 11).

-Age and time_1

The P-value was 0.864 (P > 0.05), indicating that there was no significant difference between the time of filling of groups as the mean rank of the time of filling the questionnaire for group (A) was 27.24 and that of group (B) was 26.50 (Table 12).

-Age and time_2

The P-value was 0.755 (P > 0.05), indicating that there was no significant difference between the time of filling of groups as the mean rank of the time of filling the questionnaire for group (A) was 26.57 and that of group (B) was 27.91 (Table 13).

Reliability

*Internal consistency:

The value of > 0.6 Cronbach's Alpha refers to acceptable reliability, but it is still a borderline value, so the researcher prefers more specific studies on this issue (Table 14). According to Fan and Lê [19], "all Cronbach's Alpha coefficient values above 0.6 are considered to be acceptable". And according to George [20], a commonly accepted rule of thumb for the value of Cronbach's alpha for describing internal consistency is as follows: "_ > .9 – Excellent, _ > .8 – Good, _ > .7 – Acceptable, _ > .6 – Questionable, _ > .5 – Poor, and _ < .5 – Unacceptable".

*Test-retest reliability:

The Pearson correlation coefficient showed a very strong direct relationship (r = 0.986) between score _1 and score_ 2 which showed very good intra-rater reliability.

Feasibility

Count of missing answers

83.963% of the filled questionnaires had no missing answers and 16.037% had missing answers, which was an acceptable percent to have a feasible test.

Table (15) showed that the count of questionnaires with missing answers were 16 with one missing item and only one with two missing items, for a total of 17 questionnaires with missing answers (16.03%). The questionnaires with no missing answers were 42 in the 1st week, with 79.25%, and in the 2nd week, they were 47, with 88.68%, an increase in the number of questionnaires with no missing answers in the 2nd week compared to the 1st week.

Administration time

The mean of T-1 "time_1" was 7.4528 (S.D. = 2.52347) and that of T-2 "time_2" was 5.5755 (S.D. = 2.07178), a good average of time, indicating the feasibility of filling. Time elapsed since the beginning of the study showed that more experience administering the WeeFIM instrument resulted in less time needed to complete the test. Administration time was not influenced by the age of the child (P-value = 0.864 <

table 14>, P-value = 0.755 < table 15>, not significant) and not influenced by the gender of the child (P-value = 0.763 < table 10>, P-value = 0.451 < table 11>, not significant).

Content Validity

The researchers sought the opinion of 20 experts as follows: One to sixteen physical therapy burn specialists, 17 of whom are staff members at the Faculty of Physical Therapy, Cairo University. A medical researcher who had experience in filling out questionnaires, a physical therapy pediatric specialist, a professor of statistics then, the researcher made simple statistics for ICV as shown in Fig. 1. The Index of Content Validity (ICV) of adapted WeeFIM items showed that all the questions were relevant [Their ICVs ranged from 0.55 to 1, with a mean of 0.7277, except 2 questions: 1. Question number 15, which referred to an expression, was irrelevant because (ICV = 0.4). 2-Question number 18, which referred to memory, was irrelevant because (ICV = 0.4). Experts who disagreed on those two items have attributed their opinion to the irrelevance of the two items to burn cases. The relevant questions have four questions with low ICV, which are: 1-Question number 8, which referred to bowel control, has low relevance because (ICV = 0.6), 2-Question number 11, which referred to tub/shower transfers, has low relevance because (ICV = 0.55), 3-Question number 14, which referred to comprehension, has low relevance because (ICV = 0.55), and 4-Question number 17, which referred to problem-solving, has low relevance because (ICV = 0.6). All other questions had an ICV > 0.75.

Discussion

In this study, the Arabic adaptation of the WeeFIM questionnaire was performed following a systematic standardized approach to provide population-based information from a parental perspective about the functional skills of children with healed burns. The study was conducted in two main steps. The first was the translation process from the original English version of the questionnaire into an Arabic one according to published guidelines, and the second was the determination of its psychometric characteristics. In the current study, the researchers chose to assess the feasibility and reliability of the WeeFIM instrument as the time-consuming factor was too important in terms of clinical basis, especially for patients with burn injuries, aiming to have a simple, reliable instrument for measuring the functions of children with burns. The time of filling out the questionnaire and the number of missing answers were the measurements of the assessment of feasibility, and the 1st and 2nd scores were the indicators of intra-rater reliability. Interviewing a child and the parents is considered feasible and takes approximately 5.5 to 7.5 minutes; this range of time is considered less than the average of most researchers (about 10 to 15 min.), which can be attributed to the feasibility of the Arabic version of the WeeFIM instrument, in addition to some limitations that have obligated the researcher to ask the questions in a too simple way, which leads consequently to less time of filling. The mean administration time of the second assessment (5.57 min.) was less than that of the first assessment (7.45 min.), which can be attributed to the time elapsed since the beginning of the study, which showed that more experience administering the WeeFIM instrument resulted in less time needed to complete the test and more understanding of the items of the questionnaire by the parents. The administration time was not influenced by the age of the child nor the

gender, which can be attributed to the method of filling, as it is a parent-filled questionnaire, and to the domains assessed as the WeeFIM is used to assess general function. Another aspect of feasibility is the number of missing answers. In this study, a high answer rate was expected as the WeeFIM is a 7-rank scale that nearly includes most of the varieties of functional independence and also assesses general functional activities that are used daily by every child. The results have shown a low percent of questionnaires that had missing answers (16.03%), indicating that WeeFIM is a highly feasible instrument in measuring functional independence in Egyptian children with burns (Table 15). The number of missing answers per each measurement decreased from the first to the second tests (Table 4), which can be attributed to improvement in the patients' function or a better understanding of the items of the questionnaire. Moreover, the reliability of the WeeFIM scores was good. In terms of reliability, the current study has established that the WeeFIM was reliable with excellent test-retest reliability as there was a strong direct correlation between the first score and the second one. It was evaluated by calculating the Pearson Correlation Coefficient (r = 0.986).

The average interval between both measures was one week. Compared with the original English WeeFIM, the Arabic version showed a stronger correlation coefficient than the English one. Ottenbacher et al., [21] used PCC to assess test-retest reliability and they found that (r = 0.67), in the Chinese one (r = 0.8 approximately) [13], and also in the Dutch one (r = 0.8 approximately) [9]. When the WeeFIM instrument was administered over a one-week period, the researchers noticed some minor test effects. In other circumstances, for example, the interview caused parents to have a different opinion of their child's talents, which resulted in different responses within a few days (memory of the previous interview leading to additions or a more realistic view of the child's performances). The Arabic WeeFIM had worse internal consistency than the original English one. It was assessed using the Cronbach alpha coefficient (alpha = 0.619), indicating acceptable internal consistency, whereas in the original English version, alpha = 0.9 [22], in the adapted version for Turkish children with C.P., alpha = 0.93 approximately [23], and in the adapted version for Dutch children with burns, alpha = 0.984 [9], all suggesting high internal consistency in these studies. Low Cronbach alpha values can be linked to 1-a small number of variables (only 4 items were included in the test). Two-internal consistency is lower with basic questions than with more specific ones, and the questions in this questionnaire are designed to assess general function [24].

Other studies have used a 0.6 Cronbach's alpha for other instruments, including the Pediatric Evaluation of Disability Inventory (PEDI) [25], the Instrument to Assess Knowledge Sharing Quality [26], the General Health Questionnaire (GHQ-12) [27], and the Adult Learning Inventory (AL-i) among first-year medical students in a Malaysian medical school [28]. This study examined the effects of two independent variables on the measured scores and times: Gender: There were no significant differences in the mean scores (and mean times) of the eighteen items, so there was no gender effect on both the 1st and 2nd scores (and times). Age: There were no significant differences in the mean scores (and mean times) of the eighteen items, so there was no age effect on both the 1st and 2nd scores (times).

The results of the study showed that the majority of the sample had achieved high scores, which can be attributed to the fact that the study was limited to patients with healed burns. Furthermore, the mean of

the 2nd score (94.04) was slightly greater than the mean of the 1st score (92.96), which can be attributed to some improvement in some patients' functions (Table 1). The low mean of time for filling can be attributed to the feasibility of filling, and the mean of the second time (5.57) was less than that of the first time (7.45), which can be attributed to a better understanding of questions by parents (Table 1). A high SD indicated non-homogeneity, which could be attributed to a wide age range (Table 1). The minimum scores were for 3-year-old boys (a low score may be attributed to their small age) and 15-year-old girls with burns to the lower jaw, chest, abdomen, and both upper limbs (a low score may be attributed to a high extent). The maximum score was achieved by a 15-year-old girl with a chronic burn in her thigh that had no effect on her functions (Table 1). The minimum time was for the older children, the smart and agile parents, and the cases with excellent functional independence levels. The maximum time was for the younger children, the highly educated parents, the parents with slow comprehension or less concentration, and the cases with poor functional independence levels.

The Wilcoxon-signed-rank test showed that score _2 was statistically significant with respect to score _1 with a p-value of 0.92 at level alpha = 0.1, which can be attributed to the improvement of the patients' functions (Table 2). Time_2 was statistically significant with respect to time _1 with a p-value of 0.000, which can be attributed to more understanding of the items in the questionnaire by the parents (Table 3).

The Mann-Whitney test showed that all of score_1, score_2, time_1 and time_2 were not statistically significant with respect to both gender and age.

It was predicted that age would be statistically significant to scores, but the results showed no statistical significance, which can be attributed to the unequal frequency distribution of children in both age groups and the severity of the cases, as the score of a 15-year-old child with a high extent may be less than the score of a 3-year-old child with a small burn area or away from joints, which has no effect on function. Correlations showed a very strong direct relationship between scores _1 and score_2 as r (Pearson correlation coefficient) = 0.986, which can be attributed to the improvement of the patients' functions (Table 5). There was also a weak inverse relationship with r = -0.233 between score_1 and time_1, which can be attributed to the feasibility of completing the questionnaire by patients with good functions who scored high (Table 5).score_2 and time_1 as r = -0.340, which can be explained by the fact that the less time in the 1st test refers to cases with good function, which in turn would have good scores in the 2nd time (Table 5). Score_2 and time_2 had a weak inverse relationship with r = -0.239, which had no statistical significance (Table 5). There was a weak inverse relationship between score_1 and time_2 with r = -0.233 and no statistical significance (Table 5). Time_1 and time_2 as r = 0.516, which can be attributed to more understanding of the items in the questionnaire by the parents (Table 5).

This study was limited by the high percentage of illiteracy of parents and children, the small sample size, the bad psychological status of some parents and children, some distracted parents who had filled out the questionnaire randomly, and some patients who had not continued in the study. From the properties of the WeeFIM instrument that the researchers experienced throughout the study, it nearly had no ceiling or floor effects, as it starts with complete independence and complete dependence, and any function of

an ordinal individual will not have more or less of a score. So, it can perfectly assess improvement and deterioration throughout the rehabilitation program. The researchers made these final adaptations to overcome the limitations that they experienced in the practical application of the study, hoping to do better research in the future and to provide the researchers with good data for further research. The findings of the present study have indicated the need to consider the following recommendations: Studying the psychometric properties of the WeeFIM in relation to certain body parts. For more extensive studies in the validation process of the Arabic version of the original WeeFIM questionnaires in Egypt, different countries, and other Arabic countries are needed for results to be generalized. Other psychometric properties of the WeeFIM questionnaire (construct validity, criterion validity, responsiveness to change, agreement, and inter-rater reliability) should also be evaluated in future research. The design of future studies should include the division of the children into groups according to age for more accurate results of the effect of age on scores. Further studies should be undertaken with a larger sample size to provide a better statistical analysis of the data. Future studies should validate other questionnaires that measure function in children with burn injuries based on the results of the current study, and the researcher suggested replication of the current study for verification of results.

Conclusion

The Arabic version of WeeFIM has high test-retest reliability, borderline internal consistency, and excellent feasibility in measuring and reporting the functional outcome of children with healed burns. The results also proved that patients' ages and genders were not related to both measured scores and both measured times.

Declarations

Acknowledgements

The authors would like to thank all the patients and their parents who took part in this research.

Declaration of interest

The authors report that there are no competing interests to declare.

Availability of data and materials

Original data not included in the manuscript can be obtained from the authors. Please contact the corresponding author via email.

References

1. Ansari-Lari, M., & Askarian, M. (2003). Epidemiology of burns presenting to an emergency department in Shiraz, South Iran. Burns, *29*(6), 579–581. https://doi.org/10.1016/S0305-4179(03)00066-4
2. Komolafe, O. O., James, J., Kalongolera, L., & Makoka, M. (2003). Bacteriology of burns at the queen elizabeth central hospital, Blantyre, Malawi. Burns, *29*(3), 235–238. https://doi.org/10.1016/S0305-4179(02)00273-5

3. Piazza-Waggoner, C., Dotson, C., Adams, C. D., Joseph, K., Goldfarb, I. W., & Slater, H. (2005). Preinjury behavioral and emotional problems among pediatric burn patients. Journal of Burn Care & Rehabilitation, *26*(4), 371–378. https://doi.org/10.1097/BCR.0b013e31827217a9
4. Lollar, D. J., Simeonsson, R. J., & Nanda, U. (2000). Measures of outcomes for children and youth. Archives of physical medicine and rehabilitation, *81*, S46-S52. https://doi.org/10.1053/apmr.2000.20624
5. Steinvall, I., Fredrikson, M., Bak, Z., & Sjoberg, F. (2011). Mortality after thermal injury: no sex-related difference. Journal of Trauma and Acute Care Surgery, *70*(4), 959–964. Doi: 10.1097/TA.0b013e3181e59dbe
6. Zuckerman, B. S., & Duby, J. C. (1985). Developmental approach to injury prevention. Pediatric Clinics of North America, *32*(1), 17–29. DOI: 10.1016/s0031-3955(16)34753-8
7. Disseldorp, L. M., Niemeijer, A. S., Van Baar, M. E., Reinders-Messelink, H. A., Mouton, L. J., & Nieuwenhuis, M. K. (2013). How disabling are pediatric burns? Functional independence in Dutch pediatric patients with burns. Research in developmental disabilities, *34*(1), 29–39. https://doi.org/10.1016/j.ridd.2012.07.012
8. Sien, N. Y., & Jung, H. (2014). Assessment of the six activities of daily living in adults. The Singapore Family Physician, 26–36.
9. Niemeijer, A. S., Reinders-Messelink, H. A., Disseldorp, L. M., & Nieuwenhuis, M. K. (2012). Feasibility, reliability, and agreement of the WeeFIM instrument in Dutch children with burns. Physical therapy, *92*(7), 958–966. https://doi.org/10.2522/ptj.20110419
10. Msall, M. E., DiGaudio, K., Duffy, L. C., LaForest, S., Braun, S., & Granger, C. V. (1994). WeeFIM: normative sample of an instrument for tracking functional independence in children. Clinical pediatrics, *33*(7), 431–438. https://doi.org/10.1177/000992289403300709
11. Chokshi, K. P., Tedla, J. S., Narayan, A., Ganesan, S., & Reddy, R. S. (2021). Functional independence measure (WeeFIM) reference values in Indian children aged 3–7 years: A cross-sectional study. National Medical Journal of India, *34*(2). DOI: 10.4103/0970-258X.3267488
12. Kottner, J., Audigé, L., Brorson, S., Donner, A., Gajewski, B. J., Hróbjartsson, A., … Streiner, D. L. (2011). Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. International journal of nursing studies, *48*(6), 661–671. https://doi.org/10.1016/j.ijnurstu.2011.01.016
13. Wong, V., Wong, S., Chan, K., & Wong, W. (2002). Functional independence measure (WeeFIM) for Chinese children: Hong Kong cohort. Pediatrics, *109*(2), e36-e36. https://doi.org/10.1542/peds.109.2.e36
14. Aybay, C., Erkin, G., Elhan, A. H., Sirzai, H., & Ozel, S. (2007). ADL assessment of nondisabled Turkish children with the WeeFIM instrument. American journal of physical medicine & rehabilitation, *86*(3), 176–182. Doi: 10.1097/PHM.0b013e31802b8f8d
15. Schneider, E., Parush, S., Katz, N., & Miller, L. J. (1995). Performance of Israeli versus US preschool children on the Miller Assessment for Preschoolers. The American Journal of Occupational Therapy, *49*(1), 19–23. https://doi.org/10.5014/ajot.49.1.19

16. Jongjit, J., Komsopapong, L., Saikaew, T., Wanich, U., Chewapanich, S., Udomsubpayakul, U., & Ruangdaraganon, N. (2006). Reliability of the Functional Independence Measure for children in normal Thai children. Pediatrics international, *48*(2), 132–137. https://doi.org/10.1111/j.1442-200X.2006.02183.x
17. Goldstein MS, Elliott SD, Guccione AA. The development of an instrument to measure satisfaction with physical therapy. Physical therapy. 2000 Sep 1;80(9):853 – 63.
18. Khan, W. S., Dillon, B., Agarwal, M., Fehily, M., & Ravenscroft, M. (2009). The validity, reliability, responsiveness, and bias of the Manchester-Modified Disability of the Arm, Shoulder, and Hand score in hand injuries. Hand, *4*(4), 362–367. DOI 10.1007/s11552-009-9191-x
19. Fan, S., & Lê, Q. (2011). Developing a valid and reliable instrument to evaluate users' perception of web-based learning in an Australian university context. *Journal of Online Learning and teaching*, *7*(3), 366–379. ISSN 1558–9528 (2011)
20. George, D. (2011). *SPSS for windows step by step: A simple study guide and reference, 17.0 update, 10/e*. Pearson Education India.
21. *Archives of physical medicine and rehabilitation*, *81*(10), 1317–1326. https://doi.org/10.1053/apmr.2000.9387
22. *Pediatric Research*, *39*(4), 378–378.
23. Tur, B. S., Küçükdeveci, A. A., KUTLAY, Ş., Yavuzer, G., Elhan, A. H., & Tennant, A. (2009). Psychometric properties of the WeeFIM in children with cerebral palsy in Turkey. Developmental Medicine & Child Neurology, *51*(9), 732–738.-8. https://doi.org/10.1111/j.1469-8749.2008.03255.x
24. Juul, L., Van Rensburg, J. A., & Steyn, P. S. (2012). Validation of the king's health questionnaire for South Africa in English, Afrikaans and isiXhosa. South African Journal of Obstetrics and Gynaecology, *18*(3). DOI: 10.7196/sajog.498
25. Harvey, A., Robin, J., Morris, M. E., Graham, H. K., & Baker, R. (2008). A systematic review of measures of activity limitation for children with cerebral palsy. Developmental Medicine & Child Neurology, *50*(3), 190–198. DOI: 10.1111/j.1469-8749.2008.02027.x
26. Ismail, M. B., & Yusof, Z. M. (2010). The impact of individual factors on knowledge sharing quality. Journal of Organizational Knowledge Management, *13*(1), 1–12. DOI: 10.5171/2010.327569
27. Nordin, N. M., Talib, M. A., Yaacob, S. N., & Sabran, M. S. (2010). A study on selected demographic characteristics and mental health of young adults in public higher learning institutions in Malaysia. Global Journal of Health Science, *2*(2), 104. DOI: 10.5539/gjhs.v2n2104
28. Yusoff, M. S. B. (2011). Reliability and validity of the Adult Learning Inventory among medical students. Education in Medicine Journal, *3*(1).;3(1). Doi:10.5959/eimj.3.1.2011.or3

Tables

Table (1): Statistical analysis of relating the WeeFIM 1st and 2nd scores and the time elapsed to fill the questionnaire in the 1st and 2nd weeks

	Score_1	Score_2	Time_1	Time_2
Mean	92.9623	94.0377	7.4528 min,	5.5755 min.
S.D.	±23.05092	±22.57207	±2.52347	±2.07178
Min.	47.00	46.00	3.00 min.	2.50 min.
Max.	126.00	126.00	13.00 min.	12.00 min.

* The scores basic range is from 18 to 126, SD= standard deviation.

Table (2): Statistical analysis of score_1 and score_2

Variables	Negative ranks	Positive ranks	Ties	Total
N	10[a]	19[b]	24[c]	53
Mean	14.00	15.53		
Sum of ranks	140.00	295.00		
Z-value	-1.684-[a]			
P-value	0.092			
Level of significance	Significant at level alpha= 0.1			

*N=Number of patients.

Table (3): Statistical analysis of time_1 and time_2

Variables	Negative ranks	Positive ranks	Ties	Total
N	31[a]	2[b]	20[c]	53
Mean	17.32	12.00		
Sum of ranks	537.00	24.00		
Z-value	-4.606-[a]			
P-value	0.000			
Level of significance	Significant			

Table (4): Statistical analysis of missing_1 and missing_ 2

Variables	Negative ranks	Positive ranks	Ties	Total
N	6[a]	0[b]	47[c]	53
Mean	3.50	0.00		
Sum of ranks	21.00	0.00		
Z-value	-2.449-[a]			
P-value	0.014			
Level of significance	Significant			

Table (5): Correlations between scores and times

	Score_1		Score_2	Time_1
Score_2	0.98 0.000			
Time_1	-0.324 0.018		-0.340 0.013	
Time_2	-0.233	0.092	-0.239 0.085	0.516 0.000

Cell Contents: *Pearson correlation

*P-Value

Table (6): Statistical analysis of the association of gender to score_1

Variables	Group A (boys)	Group B (girls)
N	26	27
Mean rank	29.06	25.02
Sum of ranks	755.50	675.50
Z-value	-0.952-	
P-value	0.341	
Level of significance	Not significant	

Table (7): Statistical analysis of the association of gender to score_2

Variables	Group A (boys)	Group B (girls)
N	26	27
Mean rank	28.81	25.26
Sum of ranks	749.00	682.00
Z-value	-0.837-	
P-value	0.403	
Level of significance	Not significant	

Table (8): Statistical analysis of the association of gender to time_1

Variables	Group A (boys)	Group B (girls)
N	26	27
Mean rank	27.62	26.41
Sum of ranks	718.00	713.00
Z-value	-0.301-	
P-value	0.763	
Level of significance	Not significant	

Table (9): Statistical analysis of the association of gender to time_2

Variables	Group A (boys)	Group B (girls)
N	26	27
Mean rank	25.46	28.48
Sum of ranks	662.00	769.00
Z-value	-0.753-	
P-value	0.451	
Level of significance	Not significant	

Table (10): Statistical analysis of the association of age to score_1

Variables	Group A (3-7 yrs)	Group B (8-16 yrs)
N	36	17
Mean rank	25.06	31.12
Sum of ranks	902.00	529.00
Z-value	-1.334-	
P-value	0.182	
Level of significance	Not significant	

Table (11): Statistical analysis of the association of age to score _2

Variables	Group A (3-7 yrs)	Group B (8-16 yrs)
N	36	17
Mean rank	25.11	31.00
Sum of ranks	904.00	527.00
Z-value	-1.296-	
P-value	0.195	
Level of significance	Not significant	

Table (12): Statistical analysis of the association of age to time_1

Variables	Group A (3-7 yrs)	Group B (8-16 yrs)
N	36	17
Mean rank	27.24	26.50
Sum of ranks	980.50	450.50
Z-value	-0.171-	
P-value	0.864	
Level of significance	Not significant	

Table (13): Statistical analysis of the association of age to time_2

variables	Group A (3-7 yrs)	Group B (8-16 yrs)
N	36	17
Mean rank	26.57	27.91
Sum of ranks	956.50	474.50
Z-value	-0.313-	
P-value	0.755	
Level of significance	Not significant	

Table (14): Cronbach's Alpha

Cronbach's Alpha	N of Items
0.619	4

*The items used in the test were score_1, score_2, time_1 and time_2.

Table (15): Frequency distribution of missing 1 and missing 2

Missing-1	Frequency	%	Missing-2	Frequency	%	Total frequency	Total %
0	42	79.25%	0	47	88.68%	89	83.96%
1	10	18.87%	1	6	11.32%	16	15.09%
2	1	1.89%	2	0	0%	1	0.94%
N	53	100%	N	53	100%	106	100%
Total missing	11	20.76%	Total missing	6	11.32%	17	16.03%

Missing-1 = (1st week), Missing-2 = (2nd week).*

Figures

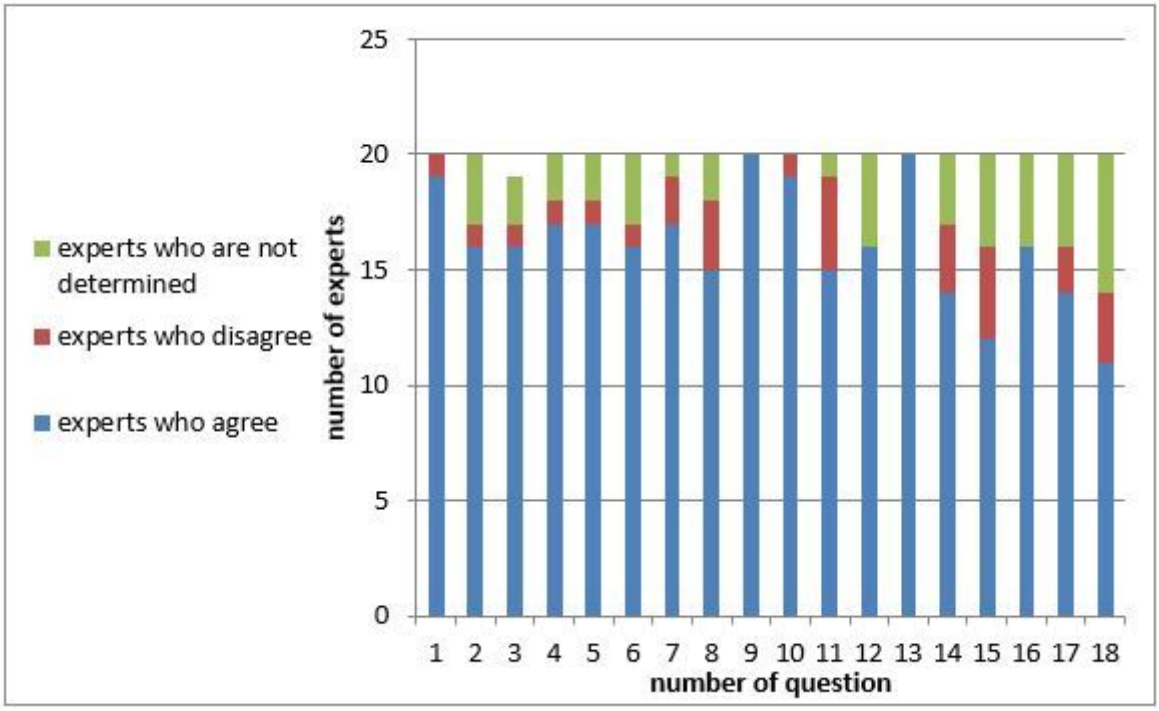

Figure 1

Experts' opinions.