Reproducibility of the Brazilian version of the Edmonton Frail Scale for elderly living in the community

Suzele Cristina Coelho Fabrício-Wehbe
Idiane Rosset Cruz
Vanderlei José Haas
Marina Aleixo Diniz
Rosana Aparecida Spadoti Dantas
Rosalina Aparecida Partezani Rodrigues

Objective: to verify the inter and intra-rater reproducibility of the Brazilian adapted version of the Edmonton Frail Scale (EFS) in an elderly group of residents. Method: in order to test the inter-rater reproducibility, two assessments were independently conducted by two researchers on the same day but at different times, in a sample of 103 elderly. Concerning the intra-rater reproducibility, the instrument was administered to 83 elderly (80.6% of the initial sample) by the same researcher in a time gap of 15 days between the two assessments. Results and Discussion: in relation to the inter-rater test, the Kappa was 0.81 (CI 0.61-1.00) and the Intraclass Correlation Coefficient (ICC) corresponded to 0.87 (CI 0.82-0.91, p<0.001). In relation to the intra-rater test, the Kappa was 0.83 (CI 0.72-0.94) and the ICC 0.87 (CI 0.81-1.00, p<0.001). Conclusion: the results show that the EFS is reliable and can be used as a tool to improve geriatric nursing care in Brazil.

Descriptors: Frail Elderly; Nursing Methodology Validation; Validation Studies; Reproducibility of Results; Geriatric Nursing.

1 Paper extracted from doctoral dissertation "Cross-Cultural Adapation and Validation of the "Edmonton Frail Scale (EFS)" – elderly frailty scale", presented to Escola de Enfermagem de Ribeirão Preto, Universidade de São Paulo, WHO Collaborating Centre for Nursing Research Development, Ribeirão Preto, SP, Brazil.

2 Pós doutoranda, Escola de Enfermagem de Ribeirão Preto, Universidade de São Paulo, WHO Collaborating Centre for Nursing Research Development, Ribeirão Preto, SP, Brazil. Scholarship holder from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

3 PhD, Adjunct Professor, Escola de Enfermagem, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil.

4 PhD, Visiting Professor, Universidade Federal do Triângulo Mineiro, Uberaba, MG, Brazil.

5 Doctoral student, Escola de Enfermagem de Ribeirão Preto, Universidade de São Paulo, WHO Collaborating Centre for Nursing Research Development, Ribeirão Preto, SP, Brazil.

6 PhD, Associate Professor, Escola de Enfermagem de Ribeirão Preto, Universidade de São Paulo, WHO Collaborating Centre for Nursing Research Development, Ribeirão Preto, SP, Brazil.

7 PhD, Full Professor, Escola de Enfermagem de Ribeirão Preto, Universidade de São Paulo, WHO Collaborating Centre for Nursing Research Development, Ribeirão Preto, SP, Brazil.
Introduction

In the opinion of researchers and healthcare professionals, frailty can undoubtedly cause a negative impact in the life of the elderly, their families, caregivers and the society. In the national and international literature, there is a consensus that frailty represents a non-optimal and multifactorial clinical condition that is vulnerable to adverse effects upon lesser impact stressors(1-3).

Currently, this syndrome emerges as an important event to public health because it is associated to adverse health outcomes, such as functional decline, dependency, recurrent falls, fractures, institutionalization, hospitalization and mortality of the elderly from both genders. It is still believed that there is a strong relationship between frailty and comorbidities, which can cause people with this syndrome to be more susceptible to diseases(4-6).

Due to these factors, researchers defend the theory that early detection of frailty is very important, that the disabilities resulting from it are a lot better treated and have better prognosis when detected in the first months of occurrence. Interventions are more effective when performed in the initial stages of frailty(7-9).

For this early detection, however, it is necessary to know an instrument that is easy for healthcare professionals to understand and administer, so that it is possible to accurately and safely detect the frailty indicators in elderly people.

To assist with this search, a group of Canadian researchers(10) studied a clinical proposal to detect frailty in elderly people and proposed a scale to assess it: the Edmonton Frail Scale (EFS). It comprises nine domains and 11 items and its scores are grouped according to the frailty degree, which ranges from no frailty to severe frailty.

This scale has been administered to various populations and cultures with different purposes. It was adapted and validated in Australia with elderly patients over 70 years of age, admitted to a teaching hospital. It was considered valid and reliable within the elderly group subject of the study and a valuable frailty research tool(11-13).

In Taiwan, researchers investigated the prevalence of frailty in people aged between 65 and 80 living in the community using the EFS, and compared it with the prevalence showed by another instrument, the Fried Frailty Index (FFI)(14). In England, researchers from the Oxford Project to Investigate Memory and Aging (OPTIMA) used, amongst other instruments, the EFS to verify the relationship between neurocognitive speeds and elderly frailty, during three years of study(15).

In Brazil, researchers performed the cultural adaptation of the scale with people aged 65 or over living in the community. The construct and criteria validity of the scale, as well as the internal consistency of the items, were verified and considered valid in relation to the sample studied(16).

The reproducibility of an instrument, however, should also be analyzed. Reliability, reproducibility and accuracy are terms used to assess an important psychometric property of assessment instruments of subjective constructs, which is the reliability of the measure(17). The analysis of the inter-rater reliability is performed to estimate possible errors during the administration that may be caused by the differences between evaluators (inter-rater test), while in the intra-rater reliability the same evaluator administers the instrument more than once (test-retest). In the first case, if the two evaluators properly follow the instructions for the use of the instrument, the results should be consistent between them. In the second case, if the construct to be measured does not change, the measures obtained should be similar(17-18).

Given these considerations and the increasing elderly population, it can be noted that studies suggesting action strategies for this population in relation to frailty are of utmost importance. The identification of the frail elderly through the administration of instruments with confirmed psychometric properties should assist with the development of interventions for frailty, thereby preventing impairments, disabilities and drawbacks in many elderly people.

Therefore, this study was aimed at assessing the intra and inter-rater reproducibility of the adapted Brazilian version of the Edmonton Frail Scale (EFS) in relation to elderly Brazilians living in the community.

Methods

Prior to this research stage, a process of cultural adaptation of the EFS to Portuguese was carried out and its validity verified(19).

This study, which is part of the project “Life and health conditions of the elderly living in Ribeirao Preto, Sao Paulo”, was undertaken in the urban area of Ribeirao Preto, in the state of Sao Paulo, involving people aged 65 or more who lived in the community. For the larger project, a probability, cluster and double stage sampling
process was used. In the first stage, the census tract was considered as the Primary Sampling Unit (PSU), and 30 census tracts were randomly selected, with probability proportional to the number of homes, among the 600 tracts in the city. The second stage involved visiting a fixed number of homes (110) from each tract in order to ensure the sample self-weighting, and the streets and blocks where this search process was initiated were randomly selected.

Data collection took place in the period from September 2007 to June 2008. Five hundred and fifteen elderly people were interviewed in a first stage, but a subsample of 137 elderly people was selected for the psychometric analysis of the EPS, using the simple random sampling (SRS) and the data was collected in June 2008. To calculate the sample size, the reliability criterion was considered using the overall score of the EFS, based on the Intraclass Correlation Coefficient (ICC) and inter-rater \(\alpha=0.01\) and \(\beta=0.1\), with statistical power of 0.90, \(H_0: ICC=0.90\) and \(H_1: ICC=0.8\), with two evaluators. This resulted in a sample of 109 elderly people but, considering the 20% loss, a sample of 137 elderly people was needed to determine the reproducibility of the adapted instrument. The Statistical Package for Social Science (SPSS) version 15.0 was used for the random selection of this subsample from a random number generator.

To check the reproducibility of the scale, the inter-rater test was performed. From the sample of 137 elderly people used during validation, one died, five refused to participate in the test and 28 were not located due to change of address. Therefore, the scale was administered to 103 elderly people, simultaneously by two different evaluators on the same day, in the same period but at different times.

The intra-rater test was performed with the re-administration of the adapted instrument by the same evaluator with a time gap of 15 days. Concerning the administration of the scale for the second time, of the 103 elderly people who participated in the inter-rater test, only 83 agreed to participate again in the research.

The interviews were conducted in the homes of the elderly with the use of an instrument that contained the adapted version of the Edmonton Frail Scale\(^{16}\). This scale assesses nine domains, represented by 11 items, which are Cognition (Clock drawing test (CDT), two points), General health status (number of hospital admissions in the last year, two points and Health description, two points), Functional Independence (Need for assistance with 8 daily activities, two points), Social support (can rely on the help of someone to attend to their needs, two points), Use of medication (Use of five or more prescribed medications, one point and forgetting to take the medication, one point), Nutrition (recent weight loss, one point), Humor (often feeling depressed, one point), Continence (loss of urine control, one point) and functional performance (timed “get up and walk” test). The maximum score in this scale is 17, which represents the highest degree of frailty. The scores for frailty analysis are: 0-4 no frailty, 5-6 visibly vulnerable, 7-8 mild frailty, 9-10 moderate frailty, and 11 or over, severe frailty\(^{16}\).

The data were entered into EXCEL through the double entry validation technique. When data entry and data consistency were concluded, the data were imported into the software SPSS version 15.0.

The reproducibility of the scale was assessed through three interviews. For this analysis, the concordance correlation coefficients were calculated for simple kappa (for nominal answers) and weighted kappa (for ordinal answers) in relation to the 11 items of the EFS, as well as the calculated frailty diagnosis. To assess the agreement between the interviewers during the tests (intra and inter-rater), the intraclass correlation coefficient (ICC) of the raw score of frailty was calculated. The following values were used: \(\kappa<0.20\) poor agreement; \(\kappa\) between 0.21-0.40 low agreement; \(\kappa\) between 0.41-0.60 moderate agreement; \(\kappa\) between 0.61-0.80 good agreement and \(\kappa\) between 0.81-1.00 very good agreement (18). Still for these authors: \(\kappa=1\) when there is 100% agreement; \(\kappa=0\) when agreement is not better than if the items were randomly answered and negative \(\kappa\) indicates agreement below expected for the items randomly answered.

The project was approved by the Ethics Committee of the Ribeirão Preto College of Nursing – USP, procedure number 0825/2007. Before the start of all the interviews, an Informed Consent Form was read and signed by the elderly and/or caregiver/family member of the elderly in two copies, one of which was given to the participant.

**Results**

Most participants selected for the validation of the scale were female (102; 74.5%), widowed (58; 42.3%), with an average age of 75.3 (ranging from 65 to 100, sd 8.01) and with an average time of formal education from one to four years (75; 54.8%). The results about the reproducibility of the instrument, taking into consideration the results of the administrations
performed by two evaluators (inter-raters herein called A1 and T1), are presented in Table 1, where the agreement between the results obtained by the two evaluators is shown ($\kappa=0.81;\ CI\ 0.61-1.00$).

### Table 1 – Distribution of the inter-rater frailty diagnosis. Ribeirao Preto, SP, Brazil, 2008

<table>
<thead>
<tr>
<th>Frailty Diagnosis T1</th>
<th>No frailty</th>
<th>Visibly vulnerable</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td></td>
</tr>
<tr>
<td>No frailty</td>
<td>40 (90.9%)</td>
<td>4 (9.1%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>44 (100%)</td>
</tr>
<tr>
<td>Visibly vulnerable</td>
<td>6 (22.2%)</td>
<td>15 (55.6%)</td>
<td>5 (18.5%)</td>
<td>1 (3.7%)</td>
<td>0 (0.0%)</td>
<td>27 (100%)</td>
</tr>
<tr>
<td>Light</td>
<td>1 (6.3%)</td>
<td>5 (31.3%)</td>
<td>9 (56.3%)</td>
<td>1 (6.3%)</td>
<td>0 (0.0%)</td>
<td>16 (100%)</td>
</tr>
<tr>
<td>Moderate</td>
<td>0 (0.0%)</td>
<td>1 (7.7%)</td>
<td>2 (15.4%)</td>
<td>9 (69.2%)</td>
<td>1 (7.7%)</td>
<td>13 (100%)</td>
</tr>
<tr>
<td>High</td>
<td>1 (33.3%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>2 (66.7%)</td>
<td>3 (100%)</td>
</tr>
<tr>
<td>Total</td>
<td>48 (46.6%)</td>
<td>25 (24.3%)</td>
<td>16 (15.5%)</td>
<td>11 (10.7%)</td>
<td>3 (2.9%)</td>
<td>103 (100%)</td>
</tr>
</tbody>
</table>

* A1 – evaluator 2
† T1 – evaluator 1

The agreement of the 11 items of the EFS in the inter-rater test suggested that the highest scores were obtained in items 8 – Nutrition ($\kappa=0.95;\ CI\ 0.88-1.00$), 2 – General health status ($\kappa=0.87;\ CI\ 0.52-1.00$), 5 – Social support ($\kappa=0.87;\ CI\ 0.65-1.00$), 6 – Use of medication ($\kappa=0.87;\ CI\ 0.77-0.98$). In contrast, the lowest scores were obtained in items 7 – Use of medication ($\kappa=0.53;\ CI\ 0.35-0.72$) and 9 – Mood ($\kappa=0.59;\ reliability\ interval\ 0.42-0.75$), (Table 2).

### Table 2 - Agreement coefficient (Kappa index) for inter-rater reliability measure of the 11 items of the Edmonton Frail Scale administered to 103 elderly people living in the community. Ribeirao Preto, SP, Brazil, 2008

<table>
<thead>
<tr>
<th>Edmonton Frail Scale Total</th>
<th>Edmonton Frail Scale Items</th>
<th>Kappa</th>
<th>Reliability interval 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.Cognition</td>
<td>Clock drawing test</td>
<td>0.77</td>
<td>0.62-0.91</td>
</tr>
<tr>
<td>2.General health status</td>
<td>How many times have you been hospitalized in the last 12 months?</td>
<td>0.87</td>
<td>0.52-1.00</td>
</tr>
<tr>
<td>3.General health status</td>
<td>Generally, how would you describe your health?</td>
<td>0.74</td>
<td>0.47-1.00</td>
</tr>
<tr>
<td>4.Functional Independence</td>
<td>In how many of the following activities do you require assistance with?</td>
<td>0.80</td>
<td>0.74-0.86</td>
</tr>
<tr>
<td>5.Social Support</td>
<td>When you need assistance, is there anyone you can rely on to attend to your needs?</td>
<td>0.87</td>
<td>0.65-1.00</td>
</tr>
<tr>
<td>6.Use of medication</td>
<td>Usually, do you use five or more different and prescribed medications (by a doctor)?</td>
<td>0.87</td>
<td>0.77-0.98</td>
</tr>
<tr>
<td>7.Use of medication</td>
<td>Do you sometimes forget to take your medication?</td>
<td>0.53</td>
<td>0.35-0.72</td>
</tr>
<tr>
<td>8.Nutrition</td>
<td>Have you been losing weight recently, in a way that your clothes have become looser?</td>
<td>0.95</td>
<td>0.88-1.00</td>
</tr>
<tr>
<td>9.Mood</td>
<td>Do you often feel sad or depressed?</td>
<td>0.59</td>
<td>0.42-0.75</td>
</tr>
<tr>
<td>10.Continence</td>
<td>Do you have trouble losing control your urine? (hold urine?)</td>
<td>0.82</td>
<td>0.70-0.93</td>
</tr>
<tr>
<td>11.Functional Performance</td>
<td>Timed ‘get up and walk’ test</td>
<td>0.76</td>
<td>0.58-0.94</td>
</tr>
</tbody>
</table>

In relation to the intra-rater test (evaluators herein called T1 and T2), the agreement coefficient for the 11 items of the EFS, with reliability interval of 95%, could be observed, with the highest scores obtained in items 4 - Functional independence ($\kappa=0.91;\ CI\ 0.87-0.95$) and 5 – Social support ($\kappa=0.84;\ CI\ 0.68-1.00$). The lowest scores were obtained in items 3 – General health status ($\kappa=0.58;\ CI\ 0.25-0.91$) and 8 – Nutrition ($\kappa=0.59;\ CI\ 0.39-0.79$), (Table 3).

To calculate the ICC, a variance model of analysis with two factors (two-way mixed model) was used: a fixed factor (evaluators) and a random factor (elderly people). Based on this, the result was ICC T1/A1=0.87; CI 0.82-0.91 p<0.001 and ICC T1/T2=0.87; CI 0.81-0.91 p<0.001.
Table 3 - Agreement coefficient (Kappa index) for intra-rater reliability measure of the 11 items of the Edmonton Frail Scale administered to 83 elderly people living in the community. Ribeirão Preto, SP, Brazil, 2008

<table>
<thead>
<tr>
<th>Edmonton Frail Scale Items</th>
<th>Kappa</th>
<th>Reliability interval 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cognition</td>
<td>Clock drawing test</td>
<td>0.74</td>
</tr>
<tr>
<td>2. General health status</td>
<td>How many times have you been hospitalized in the last 12 months?</td>
<td>0.76</td>
</tr>
<tr>
<td>3. General health status</td>
<td>Generally, how would you describe your health?</td>
<td>0.58</td>
</tr>
<tr>
<td>4. Functional independence</td>
<td>In how many of the following activities do you need assistance with?</td>
<td>0.91</td>
</tr>
<tr>
<td>5. Social Support</td>
<td>When you need assistance, is there anyone you can rely on to attend to your needs?</td>
<td>0.84</td>
</tr>
<tr>
<td>6. Use of medication</td>
<td>Usually, do you use five or more different and prescribed medications (by a doctor)?</td>
<td>0.68</td>
</tr>
<tr>
<td>7. Use of medication</td>
<td>Do you sometimes forget to take your medication?</td>
<td>0.64</td>
</tr>
<tr>
<td>8. Nutrition</td>
<td>Have you been losing weight recently, in a way that your clothes have become looser?</td>
<td>0.59</td>
</tr>
<tr>
<td>9. Mood</td>
<td>Do you often feel sad or depressed?</td>
<td>0.72</td>
</tr>
<tr>
<td>10. Continence</td>
<td>Do you have trouble losing control your urine? (hold urine?)</td>
<td>0.77</td>
</tr>
<tr>
<td>11. Functional Performance</td>
<td>Timed “get up and walk” test</td>
<td>0.69</td>
</tr>
</tbody>
</table>

Discussion

In a recent systematic review of the international literature, it could be noted that there still does not exist a single frailty assessment model for the elderly that is accepted by the researchers. But there are instruments using different parameters and concepts(19).

In Brazil, there is extensive use of the frailty phenotype developed by the Cardiovascular Health Study (CHS)(18). The EFS started to be used after recent publications about its cultural adaptation to Brazil(16).

An instrument can, however, be valid with low reliability, and there could be instruments with high reliability without validity(20), thus the need to analyze the maximum possible number of psychometric properties in a scale. Therefore, the reproducibility of the EFS was assessed in this study. In this study, the reproducibility of the scale was verified through the test-retest and intra-rater.

In relation to the intra and inter-rater reliability, it could be noted that the lowest agreement in the answers was obtained for the items that depended on the answer of the elderly, especially in more subjective items, and not on the evaluators’ assessment, showing that the commitment of the participants interviewed may alter the answers. In addition, it is possible that the demographic, socioeconomic and cultural characteristics of the participants affect, even if partially, some of the answers.

The authors of the original scale only determined the inter-rater reliability, which was re-administered within 24 hours(12). To test the reliability between the evaluators of the EFS, these authors used the Kappa coefficient (κ) with excellent agreement indicated by a score ≥ 80. According to the authors, the EFS showed good reliability between the evaluators (κ=0.77, p=0.0001, n=18). Taking into consideration the larger period and sample used in the present study, when compared to that on the original scale(12), good reliability of the scale was also observed, showing its applicability among elderly Brazilians living in the community.

During the analysis of the item cognition of the scale, showed by the clock drawing test, the researchers were hesitant because, due to it being a test that depends on the assessment and interpretation of the evaluator, it could have a questionable applicability. However, based on the analysis of the agreement coefficient of this item, it was verified that this did not occur, both in relation to the inter-rater and the intra-rater, presenting good agreement with statistically significant results in the two situations.

Although some differences were detected in the Kappa test, it was noted that there was not a statistically significant change in the results obtained in the inter-rater and intra-rater tests when the influence of each item on the frailty diagnosis was checked. It is important to highlight that providing appropriate training to interviewers contributes to obtaining reliable results in a study.

The EFS addresses multidimensional aspects that can be related to frailty. This approach may assist healthcare professionals with assessing health related or care needs situations, addressing not only physical aspects, but also those related to mood, cognitive status and social support, amongst others(21). This practice should be also extended to patients’ home care and support proposals of care network adjustments to
understand the care practices directed to the elderly by the family caregiver. The care practices vary according to the needs of the elderly, the environment, the family structure, and the knowledge of each family\(^{(22)}\). The use of the EFS is a strategy used to assess the elderly and can be used by various healthcare professionals to better direct the care practices to be implemented according to the specific needs of the elderly.

Furthermore, according to the scores proposed by the scale, there is the possibility of identifying the elderly with conditions that precede frailty, that is, those visibly vulnerable to this syndrome. These are groups at risk of developing frailty who, when detected early, can be subject to interventions focused on the promotion of health and prevention of the syndrome. The identification of these elderly people should be mainly considered by primary healthcare professionals who directly work with this population, since there is the possibility of an immediate intervention in order to keep this status unchanged\(^{(23)}\).

One limitation of this study is related to the lack of assessment of the sensitivity and responsiveness of the EFS. The authors are already elaborating these analyses, as well as its predictive validity.

It is necessary to further investigate the administration of this scale in the elderly severely compromised in our environment and those living in institutions, since it has only been administered to the elderly living in the community. Also, during the study, the elderly clinically and severely compromised were not interviewed.

**Conclusions**

Through the statistical analysis of the results of this research, the EFS was considered reliable and with good reproducibility. Given that this scale is easy and practical to administer, it can be used by a multidisciplinary team and even by people who are not specialized in the area, but properly trained to assess frailty in elderly people.

The use of a scale to identify frailty can permit a less conceptual and more operational understanding of this syndrome in clinical practice. Its use by healthcare professionals will support the prevention of diseases and promotion of health, since it can be used to detect pre-frailty in elderly people.

The scale was shown to be useful when administered to healthy elderly people or those with Non-transmissible Chronic Diseases (DCNTs) living in the community. However, its administration could include scenarios of chronic health situations, such as outpatient care centers or long stay institutions.

The translation of the EFS to Portuguese and its cultural adaption to a sample of elderly Brazilians, as well as the demonstration of its validity and reproducibility, make the use of this scale feasible in the frailty diagnosis of elderly people.

**References**