



6. Can scores from English, French and Dutch versions of the Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) be pooled? An assessment of differential item functioning in patients with systemic sclerosis.

Linda Kwakkenbos^{1,2}, Linda M. Willems, MSc³; Murray Baron, MD^{2,4}; Marie Hudson, MD^{2,4}; Cornelia H.M. van den Ende, PhD³; Brett D. Thombs, PhD^{1,2,4,5,6,7}; and the Canadian Scleroderma Research Group.

¹Department of Psychiatry, McGill University, Montréal, Québec, Canada; ²Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada;

³Department of Rheumatology, Sint Maartenskliniek Nijmegen, The Netherlands

⁴Department of Medicine (Division of Rheumatology) McGill University, Montréal, Québec, Canada; Departments of ⁵Epidemiology, Biostatistics, and Occupational Health, ⁶Educational and Counselling Psychology, and ⁷School of Nursing, McGill University, Montréal, Québec, Canada

Background: Increasingly, medical research involves patients who complete outcomes in different languages. This occurs in countries with more than one common language, as well as in international collaborations, which are utilized frequently in rare diseases such as scleroderma (systemic sclerosis, SSc). In order to pool or compare outcomes, instruments should be measurement equivalent across groups, because when measures are not equivalent metrically, it is not possible to determine if any observed differences between groups reflect real differences or are a consequence of measurement artifacts (e.g., linguistic differences). To assess fatigue in scleroderma, several instruments have been used, including the Functional Assessment of Chronic Illness Therapy- Fatigue (FACIT-F).

Objective: To assess the cross-language measurement equivalence of the English, French, and Dutch versions of the FACIT-F in SSc patients.

Methods: The FACIT-F was completed by 871 English-speaking Canadian, 238 French-speaking Canadian and 230 Dutch SSc patients. Confirmatory factor analysis (CFA) was used to assess the factor structure in the three samples. The Multiple-Indicator Multiple-Cause (MIMIC) model was utilized to assess the amount of differential item functioning (DIF), comparing English versus French and English versus Dutch samples separately.

Results: A single-dimensional factor model showed good fit in all three samples. Differential item functioning analysis of the FACIT-F identified statistically significant DIF for 3 of 13 items in French and 4 items in Dutch compared with the original English version. French patients had lower fatigue scores on items 1 and 8, and higher scores on item 4. Dutch patients had higher scores (more fatigue) on items 7, 8, 9, and 13 compared to the English sample. Despite these item differences, overall, there was not evidence that the DIF items influenced fatigue scores substantially.

Conclusions: Minor DIF was found for FACIT-F items for the French and Dutch versions compared to the original English, which had only a small effect on the overall score. Therefore, scores generated with the FACIT-F in English, French, and Dutch SSc patients can be reasonably pooled without adjustment for linguistic differences. If our results are replicated, however, the translations of several items, particularly the Dutch translation of items 7 and 8, should be reconsidered, especially given the influence of the FACIT system in other approaches to measure fatigue in chronic diseases. The importance of assessing cross-language measurement equivalence prior to pooling outcomes obtained in different languages should be emphasized.