**Validation of a patient-reported outcome measure for fatigue in patients receiving hemodialysis: the SONG-HD Fatigue instrument**

**Authors’ full names and highest degree:**

Angela Ju1,2,3, Armando Teixeira-Pinto1,2, Allison Tong1,2, Alice C Smith4, Mark Unruh5, Sara N Davison6, Juan Dapueto7, Mary Amanda Dew8, Richard Fluck9, Michael J Germain10, Sarbjit V Jassal11, Gregorio T. Obrador12, Donal O’Donoghue13, Andrea K Viecelli14,15, Giovanni Strippoli1,2,16, Marinello Ruospo16, Delia Timofte17, Ankit Sharma1,2, Eric Au1,2, Martin Howell1,2, Daniel S.J. Costa18, Samaya Anumudu19, Jonathan C Craig20, Claudia Rutherford3,21

**Author affiliations:**

1Sydney School of Public Health, University of Sydney, Sydney, NSW, Australia

2Centre for Kidney Research, The Children’s Hospital at Westmead, Sydney, NSW, Australia

3Quality of Life Office, School of Psychology, University of Sydney, Sydney, Australia

4Department of Health Sciences, University of Leicester, Leicester, United Kingdom

5Division of Nephrology, University of New Mexico School of Medicine, Albuquerque, NM, USA

6Division of Nephrology and Immunology, Department of Medicine, University of Alberta, Edmonton, Alberta, Canada

**7**Departamento de Psicología Médica. Hospital de Clínicas. Facultad de Medicina. Universidad de la República. Montevideo, Uruguay

8Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA

**9**Department of Renal Medicine, Royal Derby Hospital, Derby, United Kingdom

10Renal and Transplant Associates of New England, Division of Nephrology, Baystate Medical Center, University of Massachusetts School of Medicine, Springfield , MA, USA.

11Division of Nephrology, Department of Medicine, University Health Network, University of Toronto, Toronto, Ontario, Canada.

12Universidad Panamericana School of Medicine, Mexico City, Mexico

13Department of Renal Medicine, Salford Royal Hospital, Salford, United Kingdom

14Department of Nephrology, Princess Alexandra Hospital, Brisbane, QLD, Australia.

15Australasian Kidney Trials Network, Centre for Health Services Research, University of Queensland, Brisbane, QLD, Australia.

16Department of Emergency and Organ Transplantation, University of Bari, Bari, Italy.

17Emergency University Hospital, Bucharest, Romania

18Pain Management Research Institute, Royal North Shore Hospital

19Selzman Institute for Kidney Health, Section of Nephrology, Baylor College of Medicine, Houston, Texas

20College of Medicine and Public Health, Flinders University, Adelaide, Australia

21The University of Sydney, Susan Wakil School of Nursing and Midwifery, Cancer Nursing Research Unit (CNRU), Faculty of Medicine and Health, Sydney, Australia

**Corresponding author:**

Angela Ju

Quality of Life Office

Level 6 North, Chris O’Brien Lifehouse C39Z

University of Sydney NSW 2006 Australia

E-mail: [angela.ju@sydney.edu.au](mailto:angela.ju@sydney.edu.au)

P: +612 8627 9177

F: +612 90365292

**Word count (abstract):**

**Word count (body):**

**Short title:** Core outcome instrument for fatigue in hemodialysis

**Financial conflict of interest:**

Angela Ju none

Armando Teixeira-Pinto none

Allison Tong none

Alice C Smith none

Mark Unruh none

Sara Davison none

Juan Dapueto none

Mary Amanda Dew none

Richard Fluck none

Michael J Germain none

Vanita Jassal none

Gregorio T Obrador none

Donal O’Donoghue none

Andrea K Viecelli none

Giovanni Strippoli none

Marinello Ruospo none

Delia Timofte none

Ankit Sharma none

Eric Au none

Martin Howell none

Daniel S.J. Costa none

Jonathan C Craig none

Claudia Rutherford none

**Abstract**

**Background:** Fatigue is a very common and debilitating symptom and identified by patients as a critically important core outcome to be included in all trials involving patients receiving hemodialysis. A valid, standardized measure for fatigue is needed to yield meaningful and relevant evidence about this outcome.

**Objectives:** To validate a core patient-reported outcome measure (PROM) for fatigue in hemodialysis.

**Design**: A longitudinal cohort study was conducted to assess the validity and reliability of a new fatigue measure (SONG-HD Fatigue). Eligible and consenting patients completed the measure at three time points: baseline, a week later and twelve days following the second time point. Cronbach’s α and Intraclass correlation coefficient were calculated to assess internal consistency and Spearman’s rho was used to assess convergent validity. Confirmatory factor analysis was also conducted.

**Setting**: Hemodialysis units in the United Kingdom, Australia and Romania participated in this study.

**Participants**: Adult patients aged 18 years and over, English-Speaking, and receiving maintenance hemodialysis were eligible to participate.

**Measurements**: SONG-HD Fatigue, visual analogue scale for fatigue, 12-Item Short Form survey, Functional Assessment of Chronic Illness Therapy-Fatigue were used.

**Results:** In total, 485 participants completed the study across the United Kingdom, Australia, and Romania. Psychometric assessment demonstrated that the SONG-HD Fatigue is internally consistent (Cronbach’s α =0.81- 0.86) and stable over a one-week period (Intraclass correlation coefficient =0.68-74). The measure demonstrated convergence with Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) and had moderate correlations with other measures that assessed related but not the same concept (12-Item Short Form Survey and Visual Analogue Scale). Confirmatory factor analysis supported the one-factor model.

**Conclusions:** The SONG-HD Fatigue appears to be a reliable and valid measure to be used in trials involving patients receiving hemodialysis.

**INTRODUCTION**

Fatigue is a highly prevalent and debilitating symptom that affects 60-97% of patients receiving hemodialysis[1](#_ENREF_1), and is associated with an increased risk of cardiovascular disease, mortality, depression and impaired quality of life[2](#_ENREF_2) [3](#_ENREF_3). The experience of fatigue may be unique in hemodialysisbecause the causes are complex and multifactorial, exacerbated by the burden of hemodialysis, anemia, uremia, depression, anxiety and other comorbidities[1](#_ENREF_1),[4](#_ENREF_4),[5](#_ENREF_5). Fatigue can severely limit patients’ physical and social functioning, life satisfaction, and ability to fulfil their roles and goals.

Through the Standardized Outcomes in Nephrology-Hemodialysis (SONG-HD) initiative, which involved over 1300 patients and health professionals from more than 70 countries, fatigue was identified as a critically important core outcome to be assessed in all trials involving patients receiving hemodialysis[6-8](#_ENREF_6). In order to facilitate the uptake of this core outcome in trials, SONG-HD initiative sought to endorse a short, yet meaningful measure of fatigue. Despite 18 different measures of fatigue that had previously been used in hemodialysis[9](#_ENREF_9), none matched the characteristics of an ideal core outcome measure: short (<5 items) and assesses content deemed as important and meaningful by patients on hemodialysis, caregivers and health professionals.

To establish a core outcome measure for fatigue, the SONG-HD Fatigue measure was developed using the Core Outcome Measures Effectiveness Trials methodological framework and other relevant frameworks for developing patient-reported outcome measures (PROMs)[10-12](#_ENREF_10). This process involved patient input throughout all stages, both as research partners and participants: a systematic review of PROMs used to report fatigue in hemodialysis to identify all measures that had been used to evaluate fatigue 9; an international survey with patients receiving hemodialysis, caregivers and health professionals to identify the critically important content and measurement dimensions of fatigue to include in a core outcome measure13; a multi-stakeholder consensus workshop involving patients, caregivers, health professionals, researchers and policy makers to discuss potential items14; selection of items from existing measures for the initial measure; and pilot cognitive interviews with patients using a response model[13](#_ENREF_13).

As the final phase in establishing the SONG-HD Fatigue as a core outcome measure, this study aimed to assess the validity of its use in the hemodialysis population. We report the first psychometric evaluation of the core PROM for fatigue in hemodialysis and present the SONG-HD Fatigue Measure.

**METHODS**

*Participant selection and recruitment*

Participants were eligible if they were adult patients aged 18 years or over receiving maintenance hemodialysis for more than 6 months who could read and speak English and able to provide informed consent. Patients were excluded if they were cognitively impaired such that they would not be able to complete a patient-reported outcome measure on their own. Through the authors’ professional network, we used convenience sampling to recruit patients across seven sites in the United Kingdom (UK), three in Australia and one in Romania. All measures were administered in English across all three countries. A sample size of 450 participants was calculated to allow the estimation of the intraclass correlation coefficient with a precision of 0.05, assuming an expected intraclass correlation coefficient of 0.7 for the agreement between the SONG-HD Fatigue measured at times time point 1 and time point 2 (one week apart). This sample size also allows us to estimate the Pearson’s correlation coefficient between the SONG-HD Fatigue measure and FACIT-F**,** with a precision superior to 0.07 (i.e, <0.07), for correlations stronger than 0.7. Ethics approval was obtained from East Midlands Nottingham 1, Princess Alexandra Hospital, Western Sydney Local Health District and University of Sydney Human Research Ethics Committee.

*Measures*

The SONG-HD Fatigue**:** The SONG-HD Fatigue measure consists of three items, which assess: 1) the impact of fatigue on life participation, 2) tiredness, and 3) level of energy. These dimensions are assessed on a 4-point Likert scale indicating increasing severity, ranging from 0 (not at all) to 3 (severely). Patients respond based on their experience of fatigue in the past week. An overall score for fatigue is obtained by summing the responses across the three questions resulting in a scale ranging from 0 (no fatigue) to 9 (maximum fatigue). We also assessed the possibility of weighting each question differently in the calculation of the overall score. The final instrument is provided in Figure 1.

The Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F)[14](#_ENREF_14): This measure was selected to assess the convergent validity of the SONG-HD Fatigue total scores. Although FACIT-F has not been validated in the HD population, it assesses the construct of fatigue with multiple items and has been tested for psychometric robustness in the general population[15](#_ENREF_15) as well as other chronic conditions such as cancer and rheumatoid arthritis[15](#_ENREF_15),[16](#_ENREF_16). We determined that FACIT-F had good face validity for HD because the items and wording most closely operationalized dimensions of fatigue as described by patients receiving HD[17](#_ENREF_17). Owing to the similarity in construct, we hypothesized that the FACIT-F scores would be highly correlated to the SONG-HD Fatigue scores.

The visual analogue scale (VAS) for fatigue: The VAS was used to assess the patients’ level of fatigue at the time of assessment. We anticipated a moderate to high correlation between the SONG-HD Fatigue and VAS, as both measures assess the same construct (fatigue) but with different recall periods and dimensions.

The 12-item Short Form Survey version 2 (SF-12 v2)[18](#_ENREF_18): This measure was chosen as it is a relatively brief, simple measure, which is regarded to be of minimal burden to respondents. Although this measure assesses constructs related to fatigue such as vitality and physical functioning, the physical and mental composite scores (PCS and MCS) produced are not indicative of the same construct and thus we hypothesized that the scores will be moderately correlated.

*Data collection*

Each participant was given the completed the SONG-HD Fatigue measure and a VAS at three time points (at time point1: baseline, at time point2: a week later and at time point3: twelve days after T2). At baseline, they were also asked to complete either the SF-12 or the FACIT-F. All measures were self-completed before the start of hemodialysis on each occasion to minimize the confounding effects of the hemodialysis process itself on experiences of fatigue. While limited assistance was provided to facilitate the reading of the questions, no assistance was provided to aid the interpretation of the questions, nor the retrieval of responses.

*Data analysis*

A psychometric evaluation was conducted in SPSS Version 25.0 Armonk, NY: IBM Corp. and MPlus Version 8.3 to assess data quality (completeness of item- and scale-level data), reliability (internal consistency and test re-test between time point1 and time point2) and validity (convergent validity, and exploratory known-groups for dialysis vintage).

Mean scores of the SONG-HD Fatigue were compared between countries to identify any systematic differences. Due to statistically significant mean differences, a post-hoc decision was made to test and report psychometric properties separately by country with the following criteria:

Content validity: Previous work[9](#_ENREF_9),[17](#_ENREF_17),[19](#_ENREF_19) was conducted to ensure content validity.

Internal consistency: Cronbach’s alpha ≥0.70 indicates adequate scale consistency.

Test-retest reliability: Intra-class coefficient between time point1 and time point2 >0.70 indicates reliable stability.

Construct validity: For construct-convergent validity, spearman’s rho >0.7 indicates high, 0.3-0.7 moderate and <0.3 low correlation

A confirmatory factor analysis was undertaken to examine the hypothesized one-factor structure of SONG-HD Fatigue for each country. A multi-group confirmatory factor analysis was conducted to examine whether the same structure of SONG-HD Fatigue held for each country, by analyzing the complete dataset, with each country included in the model as a group but loadings and thresholds constrained to be equal across groups. The variance of the latent variable was fixed to one to identify the model. Adequacy of the model fit was determined based on the following criteria: ratio of χ 2 to degrees of freedom (χ 2/df) < 3, the root mean square error of approximation (RMSEA) < 0.08, comparative fit index > 0.9, and standardized root mean square residual < 0.08.

Known groups: Hemodialysis vintage was selected as a grouping variable for an exploratory known groups analysis as there is mixed evidence concerning the relationship between hemodialysis vintage and fatigue. Relatively more recent studies have found a significant association between these two outcomes[20](#_ENREF_20),[21](#_ENREF_21) while earlier studies did not[22](#_ENREF_22),[23](#_ENREF_23).

Scoring: We assessed the possibility of using weights when summing the answers of the 3 questions as our previous work indicated greater importance of one dimension (impact of fatigue) above the other two (tiredness and level of energy)[17](#_ENREF_17),[19](#_ENREF_19). We used a linear regression with the FACIT-F score as the dependent variable and the three questions of the SONG-HD Fatigue as covariates. The regression coefficients rounded to the nearest integer were used as weights in the sum of the answers and compared to the score based on the raw (non-weighted) sum of the answers.

**RESULTS**

**Participant characteristics**

In total, 485 patients from Australia (n=106), UK (n=303) and Romania (n=76) participated, with 289 (60%) male. Most of the participants were aged from 51 to 70 years (322, 66%), and were white (l84, 79%). Also, the majority had been on hemodialysis for up to five years (405, 84%). The participant characteristics are presented in Table 1. The number of eligible, recruited patients and those lost to follow up can be seen in Figure 2.

**Data quality**

Data quality was high, with a complete response rate of 93% and mean scores were near scale mid-points for all three items. Scores spanned the scale range with no notable skew, and floor and ceiling effects were negligible (Supplementary Table 1). Main reason for missing data is losing patients to follow up due to change in their hemodialysis sessions.

**Mean scores across countries**

The SONG-HD Fatigue mean scores for the UK were different to the mean scores in Australia (mean difference = 0.93, p=0.004) and Romania (mean difference = 1.38, p<0.001). There were no differences between Australia and Romania. Mean scores for all measures by countries are presented in Table 2.

**Psychometric evaluation**

Internal consistency: Cronbach’s alpha for the total SONG-HD-Fatigue score demonstrated good consistency of the items for UK (0.86), Australia (0.81) and Romania (0.83) (Table 3). None of the items had a significant impact on the Cronbach’s alpha when removed.

Test re-test reliability: The intraclass correlation coefficients were 0.68, 0.74 and 0.72 for UK, Australia and Romania, respectively, indicating stability between time point 1 and time point 2 (Table 3).

Convergent validity: Convergent validity was assessed with FACIT-F in 244 patients, and with SF-12 in 241 patients. There was a high correlation between the FACIT-F and SONG-HD Fatigue total scores for all three countries ranging from -0.72 to -0.83. The SONG-HD Fatigue total scores had a moderate correlation with the MCS and PCS scores for all three countries, with correlations ranging from -0.30 to -0.53. The VAS scores for fatigue had a slightly higher and moderate correlation with the SONG-HD Fatigue total scores for all three countries (0.61-0.67). Table 3 presents the correlation values by country. We have explored the possibility of weighting differently each item of the SONG-HD Fatigue when summing the answers. The linear regression predicting the FACIT scores, gave more weight to item 3 followed by item 1 and item 2 (data not shown). However, the overall correlation of the weighted score was only slightly higher that the unweighted one (0.80 vs 0.79). Thus, we chose the more parsimonious approach of the unweighted sum.

Exploratory known-groups validity: There were five categories for hemodialysis vintage (less than 1 year, 1 to 5 years, 6 to 10 years, 11 to 15 years and 16 or more years). Due to limited number of patients who had been on hemodialysis for more than 6 years, the categories were recoded and two analyses were performed to assess exploratory known-groups validity: one with two categories (less than 1 year and 1 or more years) and another with three categories (less than 1 year, 1 to 5 years and 6 to more years). Patients who had been on hemodialysis for less than a year had higher SONG-HD Fatigue scores than those who had been on hemodialysis for more than a year (mean difference = 0.66, p=0.004, 95% CI 0.21, 1.10). Similarly, scores were significantly higher for patients who had been on hemodialysis for less than a year compared to those with dialysis vintage of 1-5 years (mean difference = 0.67, p=0.03, 95% CI 0.06, 1.29), but not when compared to those with dialysis vintage greater than 6 years. There were no significant differences between groups by country.

Confirmatory factor analysis (CFA): Results of the multi-group CFA analysis demonstrate that all of the goodness of fit indices supported the one-factor model for the SONG-HD Fatigue with variance of fatigue and thresholds constrained to be equal across countries: χ 2/df = 1.89, p=0.02, RMSEA = 0.07, standardized root mean square residual =0.04, and comparative fit index = 0.99. Thus, we conclude that measurement of SONG-HD fatigue is invariant across countries. Factor loadings were high for each item, ranging from 0.82-0.91 (Supplementary Table 2).

**Discussion**

Fatigue is a critically important outcome for patients receiving hemodialysis. The SONG-HD Fatigue instrument is intended to facilitate consistent, meaningful and feasible assessment of fatigue in trials involving patients receiving hemodialysis. The instrument includes three items that addresses the severity of the impact of fatigue on life participation, tiredness and energy – which are the three most important content dimensions of fatigue as indicated by patients receiving hemodialysis. The measure was developed through a multi-phase, patient-centered approach that allowed for extensive input from patients, caregivers and health professionals to ensure content validity. The psychometric evaluation found the SONG-HD Fatigue measure satisfies criteria for reliability and validity in accordance with the recommended US Food and Drug Administration guidelines for measurement[24](#_ENREF_24).

The psychometric properties were assessed by country because of the differences in means between the UK, Australia and Romania. However, similar patterns of correlationintraclass correlation coefficient, Spearman’s rho), equal factor loadings across countries and the good model fit indicate similar measurement properties across countries. There are several possible explanations for these differences in mean scores. Cultural differences in the way patients respond to questions about fatigue are possible, as has been identified for other patient-reported outcomes such as quality of life, symptoms and functional outcomes in oncology and cardiology[25-27](#_ENREF_25). It may also be related to differences in healthcare systems and practices, or the population that completed the survey. We note that the majority of the participants in the UK had been on hemodialysis for less than one year, which may contribute to the higher mean score observed. Furthermore, the generalizability of our results beyond the English-speaking population is unknown. The SONG-HD Fatigue was developed in English and thus we were able to only recruit patients who were able to read, speak and understand English.

Fatigue in the hemodialysis population is complex, and the many possible causes are difficult to delineate. For this reason, it was difficult to identify a definitive known groups a priori for comparison. However, we identified a grouping variable through the exploratory known group analysis conducted with dialysis vintage to assess whether patients reported less fatigue the longer they were on hemodialysis[4](#_ENREF_4),[5](#_ENREF_5),[20](#_ENREF_20). Similar to previous work, we found that patients who had been on hemodialysis for a shorter period of time had higher mean scores for fatigue. Also, our findings suggest that perhaps there may be a non-linear relationship between dialysis vintage and fatigue. Although the limited data collection did not permit further investigation of this relationship, we will aim to include it in future work building on the findings of this study.

We acknowledge that we were unable to demonstrate the responsiveness of the SONG-HD Fatigue due to the uncertainty around time points at which we could be certain that a clinically meaningful change in fatigue had occurred and the lack of an intervention expected to change fatigue. The SONG-HD Fatigue scores were similar between time point 1 and 3(results presented in Supplementary Table 3) ), and this was not unexpected. The measure asks the patients to report on the overall fatigue they experienced in the past week. As a result, the similar scores between two time points may be attributed to patients responding to the questions as instructed (to report overall fatigue in the past week), rather than focussing on fluctuations in fatigue. For the same reason, we were unable to use pre- and post-dialysis as time points to assess responsiveness in this study, as responses to a measure with a recall period of one week are not expected to change from before to after one hemodialysis session. We will conduct further studies to assess responsiveness in a clinical trial for an intervention designed to improve fatigue. Further work will also be done to determine a minimally clinically significant difference, and to assess other psychometric properties including cross-cultural validity.

The SONG-HD Fatigue measure was developed to assess the average fatigue over a period of one week. Patients give higher priority to overall fatigue that can limit their day to day functioning and social roles[17](#_ENREF_17),[19](#_ENREF_19). However, we recognize that post-dialysis fatigue has been identified as an important symptom among patients receiving hemodialysis. Trials investigating the effects of an intervention on post-dialysis fatigue may need to consider other measures. Post-Dialysis Fatigue Index[28](#_ENREF_28) was developed for this purpose, but a comprehensive psychometric evaluation is yet to be conducted. Ecological momentary assessment of fatigue can provide a way to assess post-dialysis fatigue with minimal risk of recall bias[1](#_ENREF_1),[29](#_ENREF_29). Our future work may also examine the suitability of SONG-HD Fatigue for assessing post-dialysis fatigue by testing an adjusted recall period.

SONG-HD Fatigue is a patient-reported outcome measure that demonstrates evidence for validity for use in patients receiving hemodialysis. The SONG-HD Fatigue is a three-item measure intended for patients to self-report fatigue in the past week and is suitable for use as a core outcome measure in hemodialysis trials. Items are summed to produce a total fatigue score. Implementing this short, validated and relevant core outcome measure in all trials involving patients on hemodialysis will enable the standardised and consistent assessment of fatigue in a way that is meaningful for patients. This will contribute to the generation of evidence that will better support patients and caregivers in decision-making by providing information about the effect of interventions on this critically important outcome of fatigue.

**Acknowledgements**

We acknowledge the Princess Alexandra Research Foundation and the Estate of the late Don and Lorraine Jacquot for funding data collection at the Princess Alexandra Hospital. In the UK, this research was funded by National Institute of Health Research (NIHR) Applied Research Collaboration East Midland (ARC-EM) and the Stoneygate Trust, and supported by the NIHR Leicester Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health. We would also like to thank Jared Palmer for coordinating the UK data collection.

**References**

1. Jhamb M, Weisbord SD, Steel JL, Unruh M. Fatigue in patients receiving maintenance dialysis: a review of definitions, measures, and contributing factors. *Am J Kidney Dis.* 2008;52(2):353-365.

2. Koyama H, Fukuda S, Shoji T, et al. Fatigue is a predictor for cardiovascular outcomes in patients undergoing hemodialysis. *Clin J Am Soc Nephrol.* 2010;5(4):659-666.

3. van Sandwijk MS, Al Arashi D, van de Hare FM, et al. Fatigue, anxiety, depression and quality of life in kidney transplant recipients, haemodialysis patients, patients with a haematological malignancy and healthy controls. *Nephrology Dialysis Transplantation.* 2018;34(5):833-838.

4. Artom M, Moss-Morris R, Caskey F, Chilcot J. Fatigue in advanced kidney disease. *Kidney International.* 2014;86(3):497-505.

5. Horigan AE. Fatigue in hemodialysis patients: a review of current knowledge. *J Pain Symptom Manage.* 2012;44(5):715-724.

6. Evangelidis N, Tong A, Manns B, et al. Developing a Set of Core Outcomes for Trials in Hemodialysis: An International Delphi Survey. *American Journal of Kidney Diseases.* 2017;70(4):464-475.

7. Tong A, Manns B, Hemmelgarn B, et al. Establishing Core Outcome Domains in Hemodialysis: Report of the Standardized Outcomes in Nephrology-Hemodialysis (SONG-HD) Consensus Workshop. *Am J Kidney Dis.* 2017;69(1):97-107.

8. Urquhart-Secord R, Craig JC, Hemmelgarn B, et al. Patient and Caregiver Priorities for Outcomes in Hemodialysis: An International Nominal Group Technique Study. *Am J Kidney Dis.* 2016;68(3):444-454.

9. Ju A, Unruh ML, Davison SN, et al. Patient-Reported Outcome Measures for Fatigue in Patients on Hemodialysis: A Systematic Review. *Am J Kidney Dis.* 2018;71(3):327-343.

10. Boers M, Kirwan JR, Wells G, et al. Developing Core Outcome Measurement Sets for Clinical Trials: OMERACT Filter 2.0. *Journal of Clinical Epidemiology.* 2014;67(7):745-753.

11. Rothrock NE, Kaiser KA, Cella D. Developing a valid patient-reported outcome measure. *Clin Pharmacol Ther.* 2011;90(5):737-742.

12. Wilson IB, Cleary PD. Linking Clinical Variables With Health-Related Quality of Life: A Conceptual Model of Patient Outcomes. *JAMA.* 1995;273(1):59-65.

13. Collins D. Pretesting survey instruments: an overview of cognitive methods. *Qual Life Res.* 2003;12(3):229-238.

14. Yellen SB, Cella DF, Webster K, Blendowski C, Kaplan E. Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system. *J Pain Symptom Manage.* 1997;13(2):63-74.

15. Cella D, Lai JS, Chang CH, Peterman A, Slavin M. Fatigue in cancer patients compared with fatigue in the general United States population. *Cancer.* 2002;94(2):528-538.

16. Cella D, Yount S, Sorensen M, Chartash E, Sengupta N, Grober J. Validation of the Functional Assessment of Chronic Illness Therapy Fatigue Scale relative to other instrumentation in patients with rheumatoid arthritis. *J Rheumatol.* 2005;32(5):811-819.

17. Ju A, Unruh M, Davison S, et al. Establishing a Core Outcome Measure for Fatigue in Patients on Hemodialysis: A Standardized Outcomes in Nephrology-Hemodialysis (SONG-HD) Consensus Workshop Report. *Am J Kidney Dis.* 2018;72(1):104-112.

18. OPTUM. SF-12v2 Health Survey <https://www.optum.com/solutions/life-sciences/answer-research/patient-insights/sf-health-surveys/sf-12v2-health-survey.html>. Accessed 28 January, 2018.

19. Ju A, Unruh M, Davison SN, et al. Identifying dimensions of fatigue in haemodialysis important to patients, caregivers and health professionals: An international survey. *Nephrology (Carlton).* 2019.

20. Jhamb M, Pike F, Ramer S, et al. Impact of fatigue on outcomes in the hemodialysis (HEMO) study. *Am J Nephrol.* 2011;33(6):515-523.

21. Karadag E, Kilic SP, Metin O. Relationship between fatigue and social support in hemodialysis patients. *Nursing & Health Sciences.* 2013;15(2):164-171.

22. Brunier GM, Graydon J. The influence of physical activity on fatigue in patients with ESRD on hemodialysis. *ANNA J.* 1993;20(4):457-461; discussion 462, 521.

23. McCann K, Boore JRP. Fatigue in persons with renal failure who require maintenance haemodialysis. *Journal of Advanced Nursing.* 2000;32(5):1132-1142.

24. *US Department of Health & Human Services FDA: Patient reported outcome measures: Use in medical product development to support labelling claims.* MD2009.

25. Fischer MJ, Inoue K, Matsuda A, et al. Cross-cultural comparison of breast cancer patients’ Quality of Life in the Netherlands and Japan. *Breast Cancer Research and Treatment.* 2017;166(2):459-471.

26. King MT, Stockler MR, O'Connell RL, et al. Measuring what matters MOST: validation of the Measure of Ovarian Symptoms and Treatment, a patient-reported outcome measure of symptom burden and impact of chemotherapy in recurrent ovarian cancer. *Qual Life Res.* 2018;27(1):59-74.

27. Park J, Johantgen ME. A Cross-Cultural Comparison of Symptom Reporting and Symptom Clusters in Heart Failure. *Journal of Transcultural Nursing.* 2017;28(4):372-380.

28. Sklar AH, Riesenberg LA, Silber AK, Ahmed W, Ali A. Postdialysis fatigue. *Am J Kidney Dis.* 1996;28(5):732-736.

29. Abdel-Kader K, Jhamb M, Mandich LA, et al. Ecological momentary assessment of fatigue, sleepiness, and exhaustion in ESKD. *BMC Nephrol.* 2014;15:29.

**Table 1. Patient characteristics at baseline**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Demographics** | **UK  (n=303)** | **Australia (n=106)** | **Romania (n=76)** | **Total  (n=485)** |
| Sex *n* (%) | | | |  | |
| Male | 181(60) | 67(63) | 41(54) | 289(60) |
| Female | 113(37) | 36(34) | 35(46) | 184(38) |
| Missing | 9(3) | 3(3) | 0 (0) | 12(2) |
| Age, years *n* (%) | | | |  | |
| 18-40 | 26(8.6) | 10(9.4) | 25(33) | 66(14) |
| 41-50 | 33(11) | 8(7) | 21(20) | 62(13) |
| 51-60 | 69(23) | 26(25) | 17(22) | 112(23) |
| 61-70 | 68(22) | 30(28) | 11(14) | 109(22) |
| 71+ | 106(36) | 32(30) | 2(2) | 140(29) |
| Missing | 1(0) | 0(0) | 0(0) | 1(0) |
| Ethnicity *n* (%) | | | |  | |
| Caucasian | 253(83) | 55(52) | 76(100) | 384(80) |
| Asian | 30(10) | 15(14) | 0(0) | 45(10) |
| Black British/Afro-Caribbean | 14(5) | 0(0) | 0(0) | 14(3) |
| Aboriginal/Torres Strait/First Nation | 1(0) | 10(10) | 0(0) | 11(10) |
| Other | 3(1) | 25(24) | 0(0) | 28(6) |
| Missing | 2(0) | 1(1) | 0(0) | 3(1) |
| Marital Status *n* (%) | | | |  | |
| Single | 87(29) | 25(24) | 20(26) | 132(3) |
| Married | 133(44) | 49((46) | 39(51) | 221(46) |
| Partner (living with/not living with) | 19(18) | 5(5) | 7(10) | 31(6) |
| Widowed | 39(13) | 13(12) | 2(3) | 54(11) |
| Divorced/separated | 25(8) | 12(11) | 8(11) | 45(10) |
| Missing | 0(0) | 2(2) | 0(0) | 2(0) |
| Employment *n* (%) | | | |  | |
| Full time/part time/casual | 36(12) | 18(17) | 22(29) | 76(16) |
| Student | 2(1) | 0(0) | 1(1) | 3(1) |
| Unemployed | 259(85) | 87(82) | 53(70) | 399(82) |
| Missing | 6(2) | 1(1) | 0(0) | 7(1) |
| Education level *n* (%) | | | |  | |
| Did not complete primary school | 23(8) | 8(8) | 1(1) | 32(7) |
| Completed high school before or up to 10th grade | 181(60) | 33(31) | 12(11) | 226(47) |
| Completed high school 12th grade/professional certificate | 75(25) | 36(34) | 29(27) | 140(29) |
| Undergraduate/postgraduate | 23(8) | 26(25) | 34(32) | 83(17) |
| Missing | 1(0) | 3(3) | 0(0) | 4(1) |
| Years on dialysis n (%) | | | |  | |
| <1 | 226(75) | 34(32) | 10(13) | 270(56) |
| 1-5 | 44(15) | 43(41) | 48(63) | 135(28) |
| 6-10 | 18(6) | 18(17) | 9(12) | 45(10) |
| 11+ | 14(5) | 11(10) | 9(12) | 34(7) |
| Missing | 1(0) | 0(0) | 0(0) | 1(0) |

**Table 2. Mean scores for SONG-HD Fatigue, SF-12, FACIT-F and VAS by country at baseline**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Country** | **Mean scores** | | | |  |
| SF-12 (n=240)a | | FACIT-F (n=245) b (SD) | VAS (n=485) c (SD) | SONG-HD Fatigue (n=485)d (SD) |
| MCS (SD) | PCS (SD) |
| UK (n=303) | 46.41(11.71) | 36.65 (9.19) | 26.34 (12.88) | 4.83 (2.68) | 5.22 (2.54) |
| Australia (n=106) | 48.33 (12.84) | 41.53 (9.88) | 30.59 (12.23) | 4.90 (2.70) | 4.29 (2.23) |
| Romania (n=76) | 51.69 (10.21) | 41.90 (8.20) | 35.44 (11.38) | 3.86 (2.67) | 3.84 (2.37) |

aSF-12 (MCS, PCS) scores range from 0-100 where higher scores indicate better functioning

bFACIT-F scores range from 0 -52 where higher scores indicate better functioning/less fatigue

cVAS scores range from 0-10 where higher scores indicate higher fatigue

dSONG-HD Fatigue scores range from 0-9 where higher scores indicate higher worse functioning/greater fatigue

SF-12, 12-Item Short Form Survey; MCS, Mental Composite Score; PCS, Physical Composite Score; FACIT-F, Functional Assessment Chronic Illness Therapy-Fatigue; VAS, Visual Analogue Scale

**Table 3. Reliability for the SONG-HD Fatigue, and Spearman’s rho correlations between SONG-HD Fatigue and SF-12, FACIT-F and VAS**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Country** | **Cronbach’s α for SONG-HD Fatigue** | **ICC a**  **(95% CI)** | **Spearman Correlationb** | | | |
| **SF-12** | | **FACIT-F** | **VAS** |
| **MCS** | **PCS** |  |  |
| UK (n=303) | 0.86 | 0.68 (0.61, 0.73) | -0.53 | -0.45  (-0.57, -0.31) | -0.83  (-0.87, -0.77) | 0.61  (0.54, 0.68) |
| (-0.64, -0.40) |
| Australia (n=106) | 0.81 | 0.74 (0.63, 0.82) | -0.30  (-0.64, 0.14) | -0.46  (-0.74, -0.05) | -0.75  (-0.85, -0.60) | 0.67  (0.55, 0.76) |
| Romania (n=76) | 0.83 | 0.72 (0.57, 0.82) | -0.52  (-0.72, -0.25) | -0.53  (-0.72, -0.26) | -0.72  (-0.88, -0.43) | 0.65  (0.50, 0.76) |

a ICC – intraclass correlation for time point 1 and time point 2. The SONG-HD Fatigue was recorded in two different time points (time point 1 and time point 2), one week apart, prior to commencing the second dialysis session of the week.  
b95% CI in parentheses

SF-12, 12-Item Short Form Survey; MCS, Mental Composite Score; PCS, Physical Composite Score; FACIT-F, Functional Assessment Chronic Illness Therapy-Fatigue; VAS, Visual Analogue Scale

**Supplementary Table 1. Data completeness and targeting**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Scale** | **Data completeness -** Computable scale score (%) | **Targeting** | | | | | |
| **Possible score range** | **Observed score range** | **Mean score** | **SD** | **Floor/Ceiling effect\*** | **Skewness** |
| SONG-HD Fatigue time point 1 | 485 (100) | 0 – 9 | 0 – 9 | 4.80 | 2.5 | 5.4/11.3 | 0.012 |
| SONG-HD Fatigue time point 2 | 467 (96.3) | 0 – 9 | 0 – 9 | 4.41 | 2.6 | 7.7/9.0 | 0.050 |
| SONG-HD Fatigue time point 3 | 446 (92.0) | 0 – 9 | 0 – 9 | 4.41 | 2.5 | 6.7/9.8 | 0.100 |

\*Floor effect =% scoring 0 (least fatigue/burden); ceiling effect = % scoring 9 (greatest fatigue/burden)

**Supplementary Table 2. Standardized estimates of factor loadings and thresholds for SONG-HD Fatigue items**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Estimate** | **SE** | **p-value** | **Threshold for response option 1** | **Threshold for response option 2** | **Threshold for response option 3** |
| UK item 1 | 0.88 | 0.02 | <0.001 | -1.48 | -0.38 | 0.71 |
| UK item 2 | 0.91 | 0.02 | <0.001 | -1.36 | -0.33 | 0.71 |
| UK item 3 | 0.82 | 0.02 | <0.001 | -0.99 | -0.16 | 0.80 |
| Australia item 1 | 0.88 | 0.02 | <0.001 | -1.48 | -0.38 | 0.71 |
| Australia item 2 | 0.91 | 0.02 | <0.001 | -1.36 | -0.33 | 0.71 |
| Australia item 3 | 0.82 | 0.02 | <0.001 | -0.99 | -0.16 | 0.80 |
| Romania item 1 | 0.88 | 0.02 | <0.001 | -1.48 | -0.38 | 0.71 |
| Romania item 2 | 0.91 | 0.02 | <0.001 | -1.36 | -0.33 | 0.71 |
| Romania item 3 | 0.82 | 0.02 | <0.001 | -0.99 | -0.16 | 0.80 |

Supplementary Table 3. Correlation between SONG-HD Fatigue and VAS scores at time point 1 and 3

|  |  |  |  |
| --- | --- | --- | --- |
|  | Intraclass correlation coefficient | 95% CI | Significance |
| Australia | 0.79 | 0.67, 0.86 | P<0.001 |
| UK | 0.71 | 0.63, 0.77 | P<0.001 |
| Romania | 0.79 | 0.66, 0.86 | P<0.001 |