

ORIGINAL ARTICLE

The development and first validation of the Manchester Early Morning Symptoms Index (MEMSI) for patients with COPD

Adam P Garrow,¹ Naimat Khan,¹ Sarah Tyson,² Jørgen Vestbo,³ Dave Singh,⁴ Janelle Yorke⁵

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/thoraxjnl-2014-206410>).

¹The Medicines Evaluation Unit, Manchester, UK

²Stroke & Vascular Research Centre, School of Nursing, Midwifery & Social Work, Jean McFarlane Building, University of Manchester, Manchester, UK

³Respiratory Research Group, 2nd Floor Education and Research Centre, Wythenshawe Hospital, Manchester, UK

⁴Medicines Evaluation Unit, University of Manchester, University Hospital of South Manchester Foundation Trust, Manchester, UK

⁵School of Nursing, Midwifery and Social Work, University of Manchester, Jean McFarlane University, University Place, Manchester, UK

Correspondence to

Professor Janelle Yorke, School of Nursing, Midwifery and Social Work, University of Manchester, Jean McFarlane University, University Place, Room 5.320, Oxford Road, Manchester, UK; janelle.yorke@manchester.ac.uk

Received 14 October 2014

Revised 15 April 2015

Accepted 14 May 2015

Published Online First

3 June 2015

ABSTRACT

Aim Early morning symptoms (EMS) in people with COPD are associated with poor health, impaired activities and increased exacerbation risk. We describe the development and preliminary validation of the Manchester Early Morning Symptom Index (MEMSI) to quantify EMS in COPD.

Methods Focus groups and cognitive debriefing with patients with COPD were used to develop the potential item list, followed by a cross-sectional study to finalise the items for inclusion. In addition to test-retest reliability, comparisons with the St George's Respiratory Questionnaire-C (SGRQ-C), modified Medical Research Council Dyspnoea Scale, Functional Assessment of Chronic Illness Therapy Fatigue Scale (FACIT-F) and Hospital Anxiety and Depression Scale (HADS) evaluated construct validity. Hierarchical methods informed item deletion and Rasch analysis was applied to assess scale unidimensionality.

Results 23 items were identified from the focus groups and debriefings. The cross-sectional study involved 203 patients with COPD (mean age 64.7 SD 7.5 years, male 63%, Global Initiative for Chronic Obstructive Lung Disease (GOLD): 1:14% 2:41% 3:25% 4: 7%). 13 items were removed during item reduction. MEMSI contains 10 items, demonstrates good overall fit to the Rasch model (χ^2 $p=0.26$) and item score distribution; excellent reliability (Person Separation Index: 0.91) and good test-retest repeatability ($r=0.82$). It correlates with the SGRQ-C ($r=0.73$), FACIT-F ($r=-0.65$) and HADS ($r=0.53-0.54$) indicating good construct validity.

Conclusions MEMSI is a reliable and valid unidimensional measure of EMS for patients with COPD. It is simple to use and score supporting its suitability for research and clinical use. Work is underway to determine the minimal clinical important difference and cross-cultural validity.

INTRODUCTION

For many patients with COPD, their symptoms are worse at the start of the day.¹ Common early morning symptoms (EMS) include cough, sputum production, shortness of breath and tiredness,² which are associated with exacerbations, activity limitation and poor health status,³ but there are few rigorously developed and validated patient reported outcome measures (PROMs) to assess EMS in COPD. The aim of this study was to develop and assess the reliability and validity of a new PROM to assess the presence and frequency of EMS in patients with COPD.

Key messages

What is the key question?

- Early morning symptoms in COPD are associated with poor health outcomes but there are currently very few validated instruments that are able to measure this.

What is the bottom line?

- We have developed and validated the Manchester Early Morning Symptoms Index (MEMSI) a simple 10-item unidimensional measure designed to capture symptoms specifically relevant to patients with COPD.

Why read on?

- The article describes the development and validation of MEMSI which, in addition to recording the presence and frequency of early morning symptoms in COPD, can also be used to evaluate treatments.

METHODS

There were four stages in the development process: (1) Item generation. (2) Cognitive debriefing. (3) Hierarchical item reduction and Rasch analysis. (4) Preliminary reliability and validity testing with established measures of disease severity and related constructs. The process was designed to be compatible with US Food and Drug Administration guidance.⁴

Participants were recruited from a research database of primary and secondary care patients with COPD ($n \geq 800$). Potential participants were contacted by telephone to ascertain their interest in taking part and then invited to attend a study appointment.

Inclusion criteria were: male and female patients aged 40 years or older diagnosed with COPD using established criteria:

- $FEV_1 < 80\%$ of predicted normal
- $FEV_1/FVC < 0.70$
- symptoms consistent with COPD.

Airflow limitation was graded according to post-bronchodilator FEV_1 spirometry. Patients with ► $FEV_1 \geq 80\%$ predicted were classified as Global Initiative for Chronic Obstructive Lung Disease (GOLD) Grade 1 (Mild)



To cite: Garrow AP, Khan N, Tyson S, et al. *Thorax* 2015;**70**:757–763.

- ▶ $50\% \leq FEV_1 < 80\%$ predicted GOLD 2 (Moderate)
 - ▶ $30\% \leq FEV_1 < 50\%$ predicted Grade 3 (Severe)
 - ▶ $FEV_1 < 30\%$ predicted Grade 4 (Very Severe).⁵
- Past and current smokers (≥ 10 pack-year smoking history) were eligible to participate.

Patients were excluded if they had a chest infection in the previous 3 months; any other respiratory illness such as asthma, cystic fibrosis and lung cancer; insufficient English skills to give informed consent or regularly engaged in activities that could interrupt sleeping patterns and hence their experience of EMS (eg, night shift workers).

Stage 1: item generation

Using a focus group format, participants were invited to talk about the experience of EMS and their impact including impairments in daily functioning.⁶ The groups were facilitated by two of the authors (JY and APG), audio-recorded, transcribed verbatim and anonymised. The transcripts were analysed using a modified content analysis approach⁷ to produce a provisional item list that highlighted participants' key concerns about EMS.

Stage 2: cognitive debriefing

Additional participants than in stage 1 were invited to participate in cognitive debriefing interviews, based on the PROMIS (Patient-Reported Outcome Measurement Information System) Questionnaire Appraisal System guidelines.⁸ This was to ensure the provisional items were clear and accurately reflected EMS in COPD. Other topics included layout, format acceptability, item scaling and optimal recall period. The interviews were facilitated by one of the authors (APG). Cognitive interview transcriptions were reviewed by one of the authors (APG) for issues related to comprehension, format and response options. Problematic items were considered at a PROM development meeting attended by all authors for revision or deletion. The results of stages 1 and 2 were combined to produce a first draft list of items.

Stage 3: item reduction

A new group of patients with COPD (with the same selection criteria) completed the draft item list to identify the most reliable and valid items to retain.^{9 10} All participants attended the

Medicines Evaluation Unit to complete the study questionnaires. Participants were provided with guidance on how to complete the questionnaire and reminded to focus on their experience of EMS associated with COPD. On completion, all documents were checked for errors and omissions. As we did not have access to recent medical records, information on comorbidities was not available. Items were removed if they demonstrated:

- ▶ significant correlations with age (Pearson's $p < 0.05$)
- ▶ gender differences in individual item responses (Mann-Whitney U test $p < 0.05$)
- ▶ high item-item correlations (> 0.7)
- ▶ item-total correlations < 0.6
- ▶ poor discrimination between different severities of COPD (t test $p < 0.05$).

Items surviving the hierarchical testing were subjected to Rasch analysis to examine how well they contributed to the underlying construct of 'EMS'. Individual item fit was tested using residual and χ^2 fit statistics and items with poorest fit to the Rasch model were removed.^{11 12} The process continued until fit to the Rasch model was achieved, indicated by a non-significant person item trait χ^2 statistic ($p > 0.05$).

Distribution of patient and item thresholds distribution maps were used to show the distribution of frequency of EMS and item frequency (locations) along a linear (logit) scale. Items that span the full range of person estimates with most items located within ± 2 logits indicate a well-targeted scale.¹²

Stage 4: psychometric testing

Stages 3 and 4 were completed simultaneously, so at the same time as completing the draft Manchester Early Morning Symptoms Index (MEMSI), participants also completed questionnaires regarding related constructs to test construct validity. These were the Functional Assessment of Chronic Illness Therapy Fatigue Scale (FACIT-F);¹³ Hospital Anxiety and Depression Scale (HADS);¹⁴ St George's Respiratory Questionnaire-C (SGRQ-C)¹⁵ and the modified Medical Research Council (mMRC) dyspnoea scale.¹⁶ A summary of these measures is provided in online supplementary appendix 1.

To assess the test-retest repeatability, participants were mailed a second copy of the draft item list. A minimum of 50

Table 1 Participant characteristics for stages 1 and 2

	Focus groups (n=36)*	Cognitive debriefing (n=9)
Male n (%)	18 (50.0)	4 (44.4)
Age mean (SD)	68.3 (5.2)	66.3 (6.1)
BMI mean (SD)	27.2 (4.8)	28.5 (7.7)
Current smokers n (%)	11 (30.6)	4 (44.4)
COPD duration in years (SD)	7.7 (4.5)	11.2 (4.1)
FEV1% Predicted (%) (SD)	52.1 (18.7)	53.9 (14.1)
FEV/FVC (%) (SD)	47.7 (12.3)	46.9 (19.2)
GOLD Grade 1 n (%)	2 (5.6)	–
FEV ₁ % predicted (%) mean (SD) Min and Max	92 (9.9) Min 85 Max 99	
GOLD Grade 2 n (%)	14 (38.9)	7 (77.8)
FEV ₁ % predicted (%) mean (SD) Min and Max	61.5 (8.4) Min 50.2 Max 76.0	59.8 (9.0) Min 50.2 Max 74.9
GOLD Grade 3 n (%)	14 (38.9)	2 (22.2)
FEV ₁ % predicted (%) mean (SD) Min and Max	39.8 (4.8) Min 31.5 Max 48.1	33.3 (2.5) Min 31.5 Max 35.1
GOLD Grade 4 n (%)	3 (8.3)	–
FEV ₁ % predicted (%) mean (SD) Min and Max	25.6 (0.7) Min 25.2 Max 26.5	

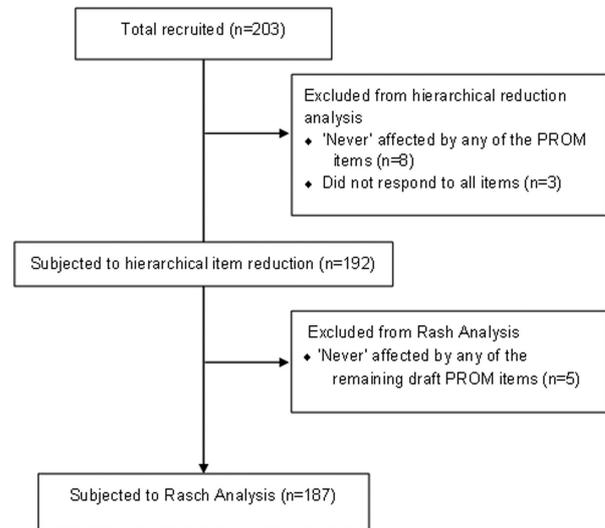
*GOLD grades were available for 33/36 patients. Three patients were diagnosed by a physician with COPD but spirometry data on these patients was not available at the time of the focus groups.
BMI, Body Mass Index; GOLD, Global Initiative for Chronic Obstructive Lung Disease.

Table 2 Participant characteristics—stages 3 and 4—item-reduction and preliminary validation

	COPD (n=203)
Male n (%)	128 (63.1)
Age (years)	64.8 (7.5)
BMI mean (SD)	27.6 (5.2)
Current smokers n (%)	66 (32.5)
Pack years median (IQR)	41 (27 to 57)
COPD duration in years (SD)	7.5 (5.3)
MOS mean (SD)	
Sleep disturbance	39.8 (25.2)
Short of breath/headache	26.2 (27.9)
Sleep adequacy	52.3 (27.4)
Sleep problem index II	41.7 (20.0)
FACIT-F mean (SD)	31.5 (13.1)
HADS mean (SD)	
Anxiety	7.4 (4.5)
% with HADS anxiety score >7	48.6
Depression	5.9 (4.3)
% with HADS depression score >7	35.4
SGRQ-C mean (SD)	
Total	51.9 (20.9)
Symptom	59.6 (19.5)
Activity	66.2 (27.7)
Impact	40.1 (22.8)
mMRC dyspnoea scale n (%)	
0	20 (9.9)
1	64 (31.5)
2	40 (19.7)
3	56 (27.6)
4	23 (11.3)
Lung function	
FEV ₁ % predicted (%)	56.6 (18.6)
FEV ₁ /FVC (%)	48.3 (13.4)
GOLD n(%)	
Grade 1	28 (13.8)
FEV ₁ % predicted (%) mean (SD) Min and Max	87.7 (5.6) Min 80.0-Max 100.1
Grade 2	83 (40.9)
FEV ₁ % predicted (%) mean (SD) Min and Max	64.0 (7.9) Min 51.1-Max 78.9
Grade 3	50 (24.6)
FEV ₁ % predicted (%) mean (SD) Min and Max	43.3 (4.5) Min 34.2-Max 49.9
Grade 4	15 (7.4)
FEV ₁ % predicted (%) mean (SD) Min and Max	23.9 (3.4) Min 21.9-Max 28.6

BMI, Body Mass Index; FACIT-F, Functional Assessment of Chronic Illness Therapy Fatigue Scale; GOLD, Global Initiative for Chronic Obstructive Lung Disease; HADS, Hospital Anxiety and Depression Scale; mMRC, modified Medical Research Council; MOS, Medical Outcomes Study; SGRQ-C, St George's Respiratory Questionnaire-C.

participants is recommended to assess test-retest reliability and 150 to provide robust evidence of validity.^{17 18} Participants also completed a global rating of change questionnaire. This assessed the stability of their health (much better; somewhat better; about the same; somewhat worse; much worse) since their first study visit. Participants were requested to return the completed questionnaires within 1 week using prepaid envelopes.¹⁹ Intraclass correlation coefficients (ICCs) examined the test-retest reliability in participants who indicated that their general health had remained 'about the same'.

**Figure 1** Flow chart showing progress through each stage of item reduction. PROM, patient reported outcome measure.

All participants attended the Medicines Evaluation Unit to complete the study questionnaires. The process was intended to be compatible with the method used in other pharmacological and academic studies carried out at the Unit. Participants were provided with guidance on how to complete the questionnaires and reminded that the focus of the research was their experience of EMS associated with COPD. After completing the questionnaires, all documents were checked for completeness. As we did not have access to recent medical records, information on comorbidities of patients who participated in the psychometric testing stage was not available.

Data analysis

Normally distributed data were analysed using parametrical statistics and presented as means and SDs. Non-parametrical data are presented as medians and IQRs. Categorical data were analysed using χ^2 statistics. p Values <0.05 were considered statistically significant. All statistical analyses were conducted using IBM SPSS Statistics for Windows, V.20.0. Armonk, New York, USA or RUMM2030 (<http://www.eumlab.com>).

Test reliability (ie, the ability of the measure to discriminate between different levels of EMS) was assessed from the Person Separation Index (PSI). PSI coefficients of 0.70 suggest that an outcome measure has 'acceptable', 0.80 'good' and 0.90 'excellent' test reliability.²⁰

ICC examined test-retest repeatability. Values ≥ 0.7 indicate acceptable repeatability.²¹ A Bland-Altman plot was constructed to show the level of agreement between the two scores.²¹ Construct validity was assessed by correlating (Pearson's r) final item-set total score with the other established and construct related instruments.

To assess discriminant validity, our a priori assumption was that there would be a significant incremental increase in total scores for the new measure per mMRC category and GOLD grade.²²

RESULTS

Stage 1 and stage 2 item generation: results of focus group and cognitive debriefing meetings

Thirty-six patients took part in six focus groups involving between two and nine participants (table 1). Twenty-three statements relating to EMS were produced covering six broad areas: (1) Breathlessness; (2) Occurrence and persistence of coughing;

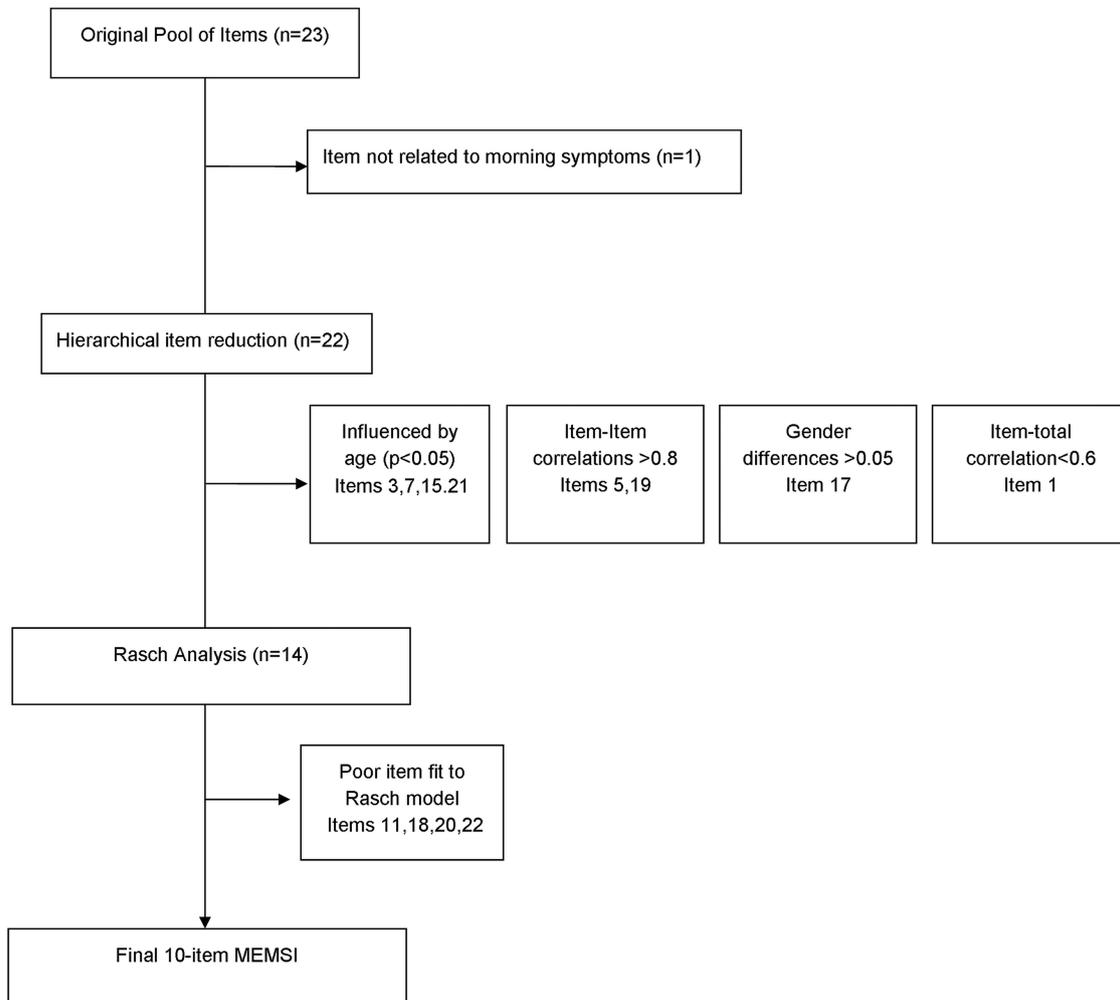


Figure 2 Individual items selected for removal. MEMSI, Manchester Early Morning Symptoms Index.

(3) Presence and removal of sputum; (4) Mouth dryness; (5) Medication use; (6) Fatigue and impaired activity.

A further nine patients with COPD participated in the cognitive debriefing sessions (table 1). The participants considered all the draft items easy to understand and complete, relevant and meaningful to patients with COPD. One statement was removed as participants felt it did not specifically relate to mornings, providing a draft list of 22 items.

Table 3 Individual fit of the 10-item MEMSI to the Rasch unidimensional model

Item	Residual item fit (± 2.5)	χ^2 p Value (0.05)	Logit: item severity (location)
Item 8	-1.47	0.08	1.12
Item 4	-0.95	0.07	0.83
Item 10	0.37	0.59	0.52
Item 13	-0.50	0.70	0.43
Item 9	0.87	0.62	0.07
Item 14	1.14	0.13	-0.25
Item 16	1.02	0.55	-0.31
Item 6	1.46	0.96	-0.57
Item 12	0.86	0.68	-0.88
Item 2	-0.59	0.10	-0.95

MEMSI, Manchester Early Morning Symptoms Index.

Various recall options were considered. An important limitation of longer recall periods is patients' inability to remember changes particularly during periods of disease stability. Shorter recall periods are also problematic due to the episodic nature of COPD. Participants considered a 1-week interval reasonable to detect meaningful changes.

Frequency of symptoms was considered the most appropriate measure of EMS. The participants also agreed that a five-option Likert-type scale with item symptom frequency options 'Never', 'Rarely', 'Sometimes', 'Often' and 'Always' was relevant and easy to understand.

Stage 3 item reduction

Two hundred and three patients with COPD completed the questionnaire pack used for stage 3 and stage 4 (table 2). Eight participants indicated they were 'never' affected by any of 22 EMS items and were removed from the analysis. Three (1.4%) failed to respond to any items and were also deleted leaving 192/203 (94.6%) available for hierarchical item reduction.

Four items influenced by age ($p > 0.05$) and one item which showed a statistically significant gender difference ($p = 0.029$) were deleted. Two items with item-item correlations > 0.8 were also removed as was one with an item-total correlation < 0.6 .

After removal of these eight items, a further five COPD participants were found to have 'never' been affected by any of the remaining 14 items and so their data were deleted leaving 179 complete data sets that were subjected to Rasch analysis.

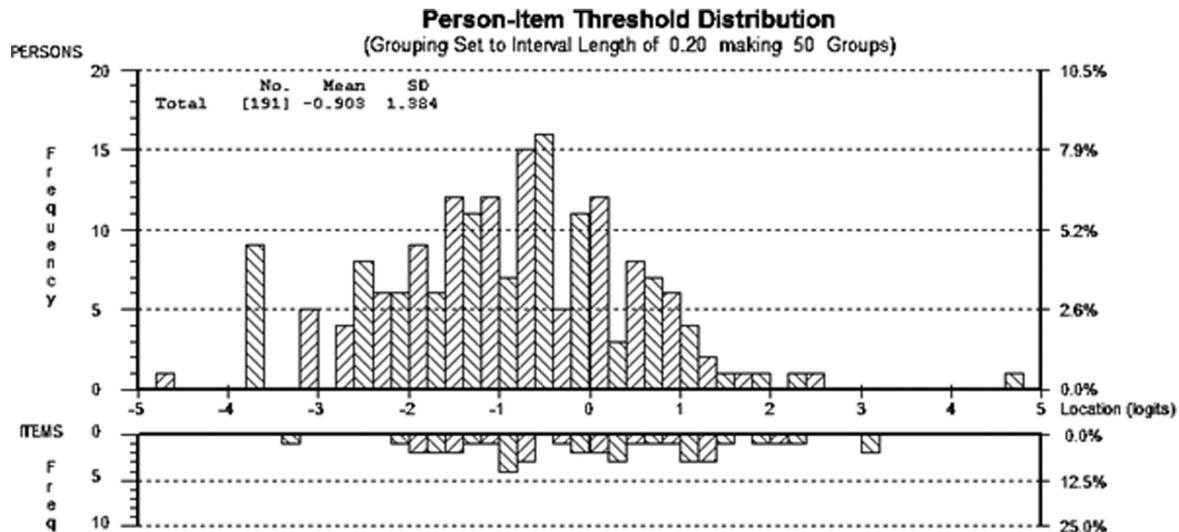


Figure 3 Distribution of item thresholds based on a Rasch logit scale.

Figure 1 shows the progress through each of the item reduction stages and figure 2 summarises the item removal process.

The 14 items did not initially fit the Rasch model. Four items with the poorest fit were deleted resulting in a 10-item scale—MEMSI. The final MEMSI demonstrated good overall fit with the Rasch model (χ^2 26.6; df 20; $p=0.26$) and a good distribution of item scores (table 3 and figure 3). Each item is scored 0 ('never') to 4 ('always'); total score ranges from 0 to 40. The mean total score for MEMSI in this study was 14.2 (SD 8.7).

The last column (logit) shows the mean level of severity for each item. Items are presented in descending order of severity (negative logit indicates less severe and positive indicates more severe). This is also depicted graphically in figure 3.

Stage 4: psychometric assessment—10-item MEMSI

MEMSI demonstrated excellent test reliability (PSI=0.91). Ninety-two (48%) participants completed a second copy of the draft PROM, and 71/92 (77%) reported that their health had not changed since completing the original set of questionnaires. MEMSI showed excellent test-retest repeatability (ICC 0.83 95% CI 0.74 to 0.89) (figure 4).

MEMSI score showed moderate-strong construct validity. Correlations with measures of related constructs were: SGRQ-C total score ($r=0.73$); Symptoms ($r=0.62$); Activity ($r=0.54$)

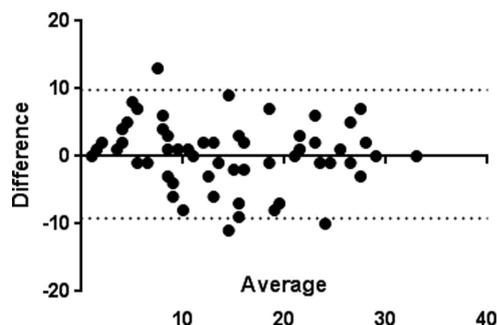


Figure 4 Bland-Altman plot showing differences between the mean scores versus the average. MEMSI, Manchester Early Morning Symptoms Index.

and Impact on daily life ($r=0.72$) (all $p<0.01$); fatigue (FACIT-F $r=-0.65$ $p<0.001$) and mood (HADS Anxiety $r=0.53$; HADS Depression $r=0.54$, both $p<0.001$).

Discriminant validity: There was a progressive increase in MEMSI scores with mMRC scores (ANOVA F (4181)=19.14 $p<0.001$) (figure 5). Although total MEMSI scores increased per GOLD grade the difference was not statistically significant (ANOVA F (3157)=2.21, $p=0.089$) (see online supplementary appendix 2).

DISCUSSION

We have described the development and preliminary validation of MEMSI. The index has excellent test reliability and test-retest repeatability. Total MEMSI scores correlated well with measures of mood (HADS); fatigue (FACIT-F) respiratory problems (SGRQ-C) and dyspnoea (mMRC) confirming construct validity. It is simple to complete and score making it suitable for clinical practice and research.

While there was a significant relationship between MEMSI scores and self-reported disability due to breathlessness the

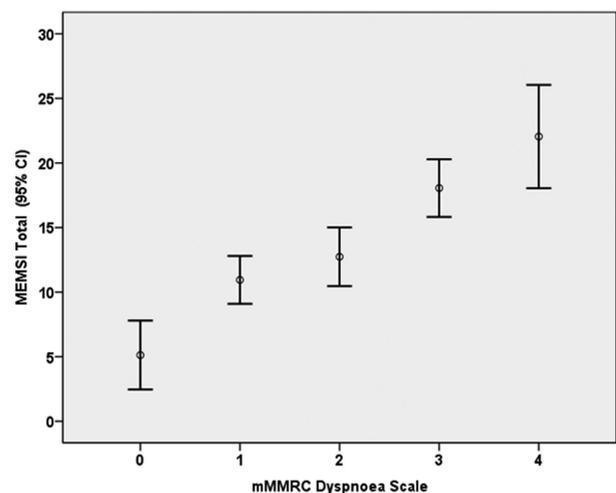


Figure 5 Association between mean MEMSI scores and mMRC scores. MEMSI, Manchester Early Morning Symptoms Index; mMRC, modified Medical Research Council.

association between COPD and disease severity based on GOLD grades, was less clear. This finding is consistent with previous studies which have showed that disability due to breathlessness is common even in patients with mild-moderate disease.^{2,3}

Rasch modelling was used to produce a 10-item unidimensional measure of EMS. The Person-item threshold map and spread of logits indicates that the items represent a higher degree of EMS than that experienced by our participants. This is not unexpected as most of our participants were GOLD grade 2. Patients with more severe disease are more likely to report worse early morning experiences than those with milder disease.² This is also reflected in the low mean total score of the final MEMSI. Future studies should test MEMSI in a population with more severe COPD.

There is a recognised need for validated PROMS to record the occurrence, variability and impact of EMS in patients with COPD.² Patients consider the morning to be the worst time for many COPD symptoms.²⁴ Variability in EMS may also help identify patients at risk of future exacerbations.³ It is also important to capture EMS in a uniform manner to evaluate medicines that target early morning respiratory symptoms.

Other EMS instruments, such as the Capacity of Daily Living during the Morning questionnaire and the Chest Symptom Questionnaire,²⁵ have been developed previously but have several limitations. Few participants (n=15) were involved in the design of these measures which is likely to affect representativeness and may also have hindered the Capacity of Daily Living during the Morning from achieving wider regulatory approval.²⁶

Our development methods were compatible with Food and Drug Administration principles. In particular our use of focus groups to generate EMS items for patients with COPD, the item reduction methods used and the various approaches to assess the reliability and validity follow recommendations.⁴ Our ultimate aim was to develop a PROM with minimal respondent burden without jeopardising the instrument's reliability, validity or sensitivity. This has been achieved.

This study has several limitations. Participants were recruited using a database of patients with COPD living in a predominantly white and socially deprived area and so cultural validity may be limited. However most of the sample characteristics are very similar to a pan-European study of primary care patients¹ which suggests we recruited a representative sample of people with COPD. Such patients may also have other conditions which could produce morning symptoms so the problems they report may not be specific to COPD. However as the final items on MEMSI focus on symptoms such as breathlessness, coughing and sputum production, we are confident that our methods have produced a COPD-specific measure of EMS.

Unfortunately, as we did not have information available on comorbidities that could also affect sleep; future research will be needed to confirm the specificity of MEMSI. This study reports cross-sectional data only and included a sample consisting mainly of GOLD grade 2 patients with COPD. Patients with more severe disease may report worse EMS.² Further assessment of MEMSI's construct validity and sensitivity to change in people with more severe COPD is required.

CONCLUSION

We have outlined the development and preliminary validation of a new patient reported measure to detect the presence and frequency of EMS in patients with COPD that could be used to evaluate medicines and other interventions. MEMSI has a

robust structure with good test reliability, test-retest repeatability and validity. Further work is needed to confirm the usefulness of MEMSI in intervention studies, to confirm the minimal important difference and establish the cross-cultural validity of the new index.

Contributors JY, DS JV conceived the study. APG, NK, ST, JV, DS and JY all contributed to the study design. APG, NK, JY facilitated the focus groups and cognitive debriefing interviews advised by ST, JV and DS. APG and JY carried out the focus group interviews and cognitive debriefing sessions. APG analysed the qualitative data supported by JY. APG and NK recruited the patients for the item reduction and validation stages of the study. APG entered and analysed the data supported by JY and NK. APG produced the first draft of the paper and all authors have critically revised it.

Funding This Partnership received financial support from the Knowledge Transfer Partnerships (KTP) programme. KTP is funded by the Technology Strategy Board along with the other government funding organisations. (Grant number: KTP:8776). The project was also supported by the UK Medical Research Council.

Competing interests None declared.

Ethics approval Ethical approval for the study was provided by the National Research Ethics Committee for Greater Manchester East (ref: 12/NW/0608).

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- Kessler R, Partridge MR, Miravittles M, *et al*. Symptom variability in patients with severe COPD: a pan-European cross-sectional study. *Eur Respir J* 2011;37:264–72.
- Omachi TA, Blanc PD, Claman DM, *et al*. Disturbed sleep among COPD patients is longitudinally associated with mortality and adverse COPD outcomes. *Sleep Med* 2012;13:476–83.
- Roche N, Small M, Broomfield S, *et al*. Real world COPD: association of morning symptoms with clinical and patient reported outcomes. *COPD J Chronic Obstr Pulm Dis* 2013;10:679–86.
- US Department of Health and Human Services Food and Drug Administration. Guidance for Industry Use in Medical Product Development to Support Labeling Claims Guidance for Industry. 2009. <http://www.fda.gov/downloads/Drugs/Guidances/UCM193282.pdf>
- Vestbo J, Hurd SS, Agustí AG, *et al*. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease GOLD executive summary. *Am J Respir Crit Care Med* 2013;187:347–65.
- Katz S. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. *J Am Geriatr Soc* 1983;31:721–7.
- Green J, Thorogood N. *Qualitative Methods for Health Research*. London: SAGE Publications Ltd, 2004.
- Willis GB, Lessler JT. *Question Appraisal System: QAS 99*. National Cancer Institute, 1999.
- Yorke J, Moosavi SH, Shuldham C, *et al*. Quantification of dyspnoea using descriptors: development and initial testing of the Dyspnoea-12. *Thorax* 2010;65:21–6.
- Jones PW, Harding G, Berry P, *et al*. Development and first validation of the COPD Assessment Test. *Eur Respir J* 2009;34:648–54.
- Yorke J, Corris P, Gaine S, *et al*. EmPHasis-10: Development of a health-related quality of life measure in pulmonary hypertension. *Eur Respir J* 2014;43:1106–13.
- Tennant A, McKenna SP, Hagell P. Application of Rasch analysis in the development and application of quality of life instruments. *Value Health* 2004;7(Suppl 1):S22–6.
- Yellen SB, Cella DF, Webster K, *et al*. Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system. *J Pain Symptom Manage* 1997;13:63–74.
- Zigmond AS, Snaith RP. The hospital anxiety and depression scale (HADS). *Acta Psychiatr Scand* 1983;67:361–70.
- Meguro M, Barley EA, Spencer S, *et al*. Development and validation of an improved, COPD-specific version of the St. George respiratory questionnaire. *Chest* 2007;132:456–63.
- Doherty DE, Belfer FMH, Brunton FSA, *et al*. Chronic obstructive pulmonary disease: Consensus Recommendations for Early Diagnosis and Treatment. *J Fam Pract* 2006;55:1–8.
- Altman D, Crc H. *Practical statistics for medical research*. London: Chapman & Hall, 1991.
- Linacre J. Sample size and item calibration stability. *Rasch Meas Trans* 1994;7:28.
- Barber MD, Walters MD, Bump RC. Short forms of two condition-specific quality-of-life questionnaires for women with pelvic floor disorders (PFDI-20 and PFIQ-7). *Am J Obstet Gynecol* 2005;193:103–13.

- 20 Duncan PW, Bode RK, Lai SM, *et al.* Rasch analysis of a new stroke-specific outcome scale: The stroke impact scale. *Arch Phys Med Rehabil* 2003;84:950–63.
- 21 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
- 22 Bestall J, Paul E, Garrod R, *et al.* Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax* 1999;54:581–6.
- 23 Jones PW, Brusselle G, Dal Negro RW, *et al.* Health-related quality of life in patients by COPD severity within primary care in Europe. *Respir Med* 2011;105:57–66.
- 24 Partridge MR, Karlsson N, Small IR. Patient insight into the impact of chronic obstructive pulmonary disease in the morning: an internet survey. *Curr Med Res Opin* 2009;25:2043–8.
- 25 Partridge MR, Miravittles M, Ståhl E, *et al.* Development and validation of the capacity of daily living during the morning questionnaire and the global chest symptoms questionnaire in COPD. *Eur Respir J* 2010;36:96–104.
- 26 Jones P, Miravittles M, van der Molen T, *et al.* Beyond FEV₁ in COPD: a review of patient-reported outcomes and their measurement. *Int J Chron Obstruct Pulmon Dis* 2012;7:697–709.

Supplementary file

Appendix 1: summary of the measures used to assess construct validity of the MEMSI

The FACIT-F assesses the impact of fatigue on daily activities and function[1]. The instrument is scored on a 5-point scale with total scores ranging from 0-52 and has been validated for use in COPD[2].

The HADS is a global measure of psychological distress that assesses Anxiety (7 items) and Depression (7 items). Scores range from 0-21 with values >7 indicating the presence of anxiety or depression [2]. The HADS has been validated for use in patients with COPD[4].

The SGRQ-C measures the impact that COPD has on overall health, daily life and perceived well-being[5]. The scale produces a total score calculated from three component scores: Symptoms; Activity and Impacts on a scale from 0-100.

The mMRC Dyspnea Scale is a widely used instrument containing five statements related to activity limitations associated with breathlessness. Scores range from 0-4 with higher scores indicating greater impairment[6].

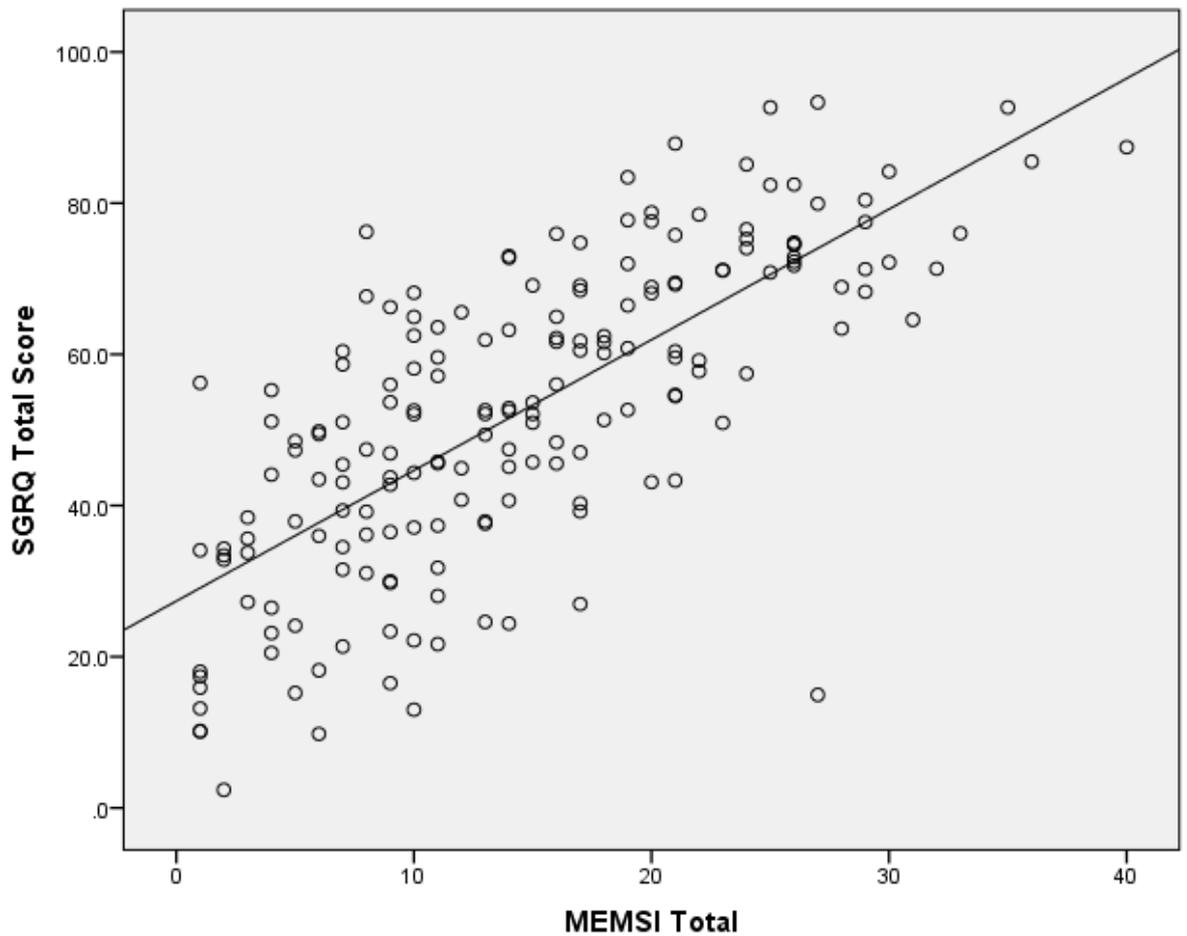
References

- 1 Yellen SB, Cella DF, Webster K. Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system. *J Pain Symptom Manage* 1997;13:63-74.
- 2 Al-shair K, Muellerova H, Yorke J, et al. ECLIPSE investigators. Examining fatigue in COPD:development, validity and reliability of a modified version of FACIT-F scale.*Health Qual Life Outcomes*. 2012;23:100
- 3 Zigmond, AS; Snaith, RP. The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica* 1983;67: 361-70
- 4 Puhan MA, Frey M, Büchi S, et al. The minimal important difference of the hospital anxiety and depression scale in patients with chronic obstructive pulmonary disease *Health and Quality of Life Outcomes* 2008;6:46

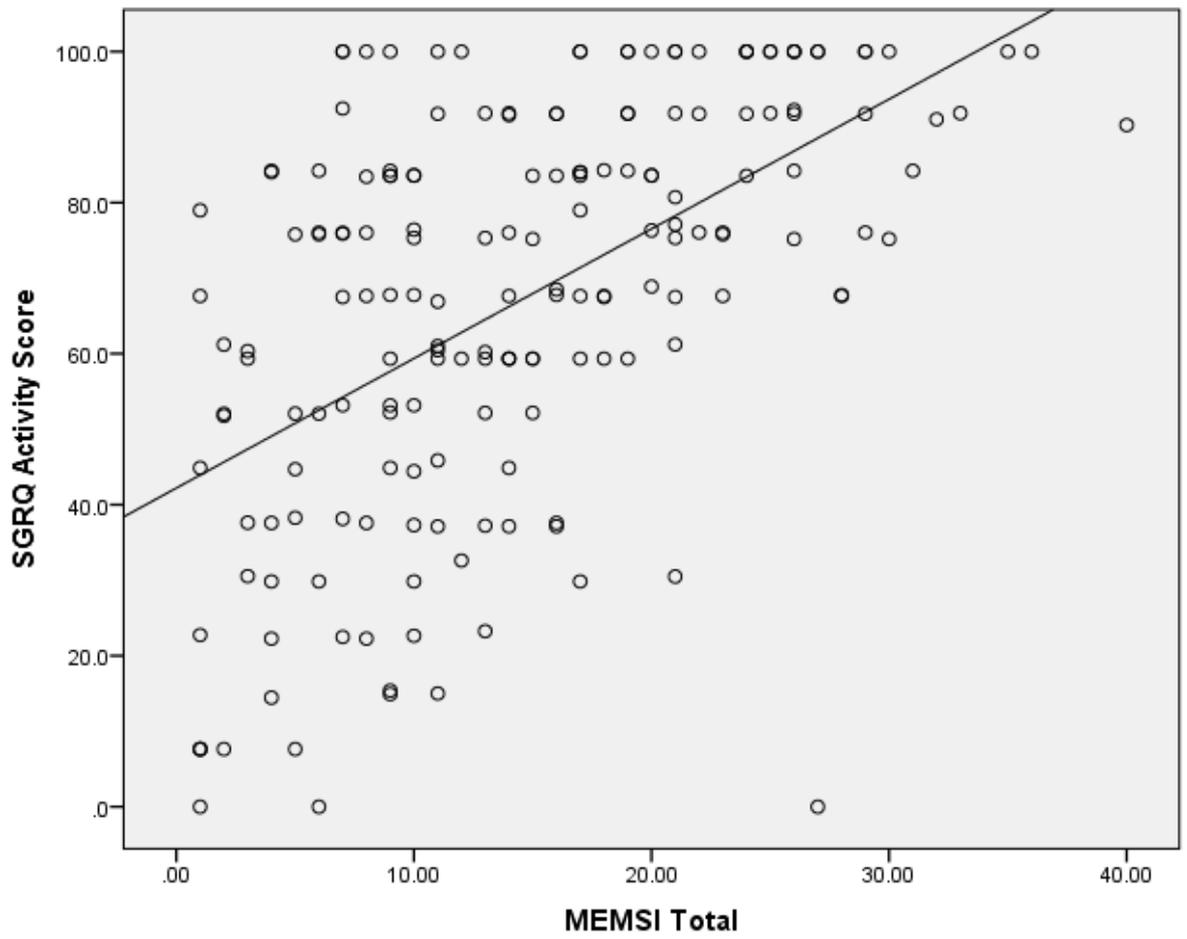
- 5 Meguro M, Barley EA, Spencer S, Jones PW. Development and Validation of an Improved. COPD-Specific Version of the St. George Respiratory Questionnaire. *Chest* 2007;132:456-63.
- 6 Doherty DE, Belfer MH, Brunton SA, et al. Chronic Obstructive Pulmonary Disease: Consensus Recommendations for Early Diagnosis and Treatment. *Journal of Family Practice*, November, 2006

Appendix: 2 Supplementary file showing 1: correlations between the MEMSI and the SGRQ-C Total Score and Subscales; HADS Anxiety and Depression and FACIT-F Total Score. 2: Association between MEMSI score and GOLD Grade.

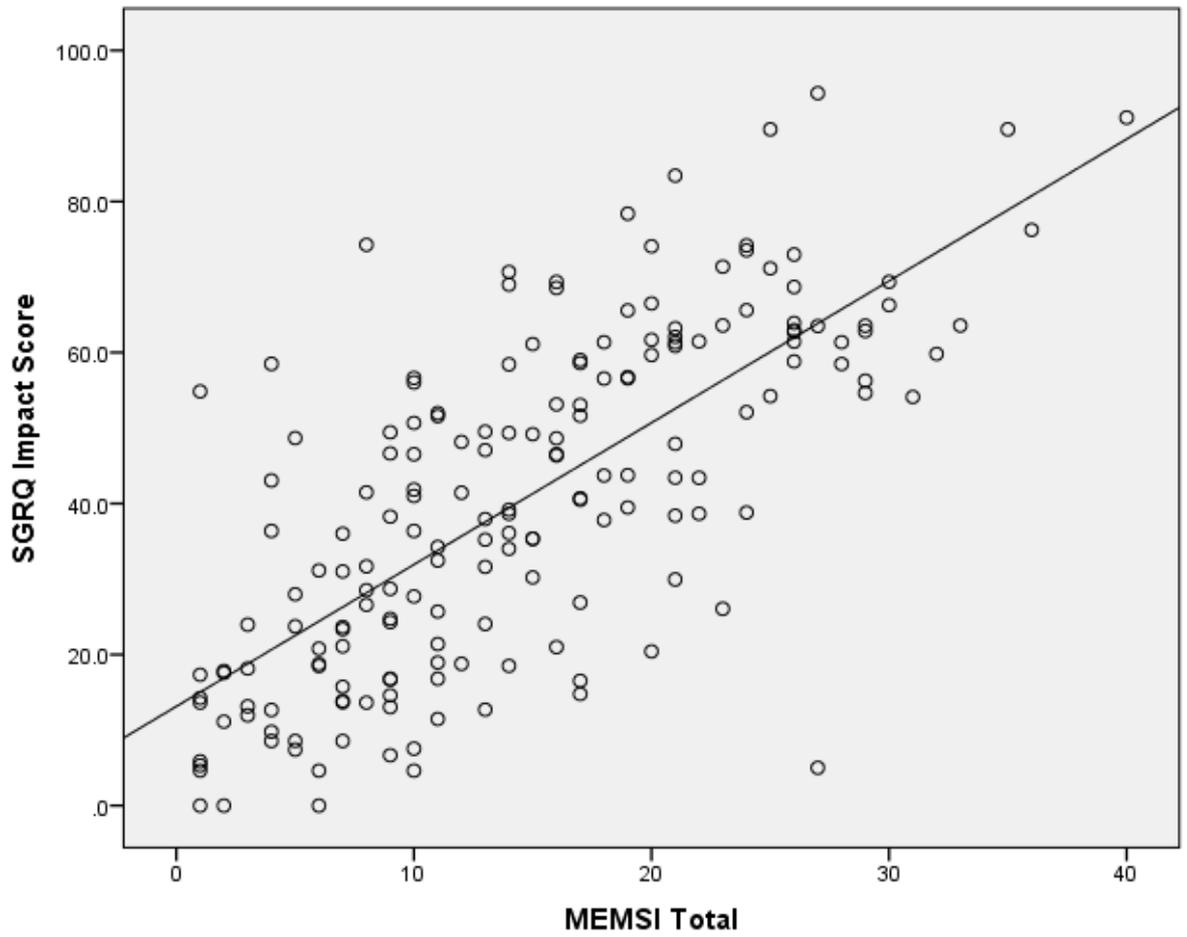
1. Correlation between MEMSI total score and SGRQ-C total score (Pearson's $r=0.73$)



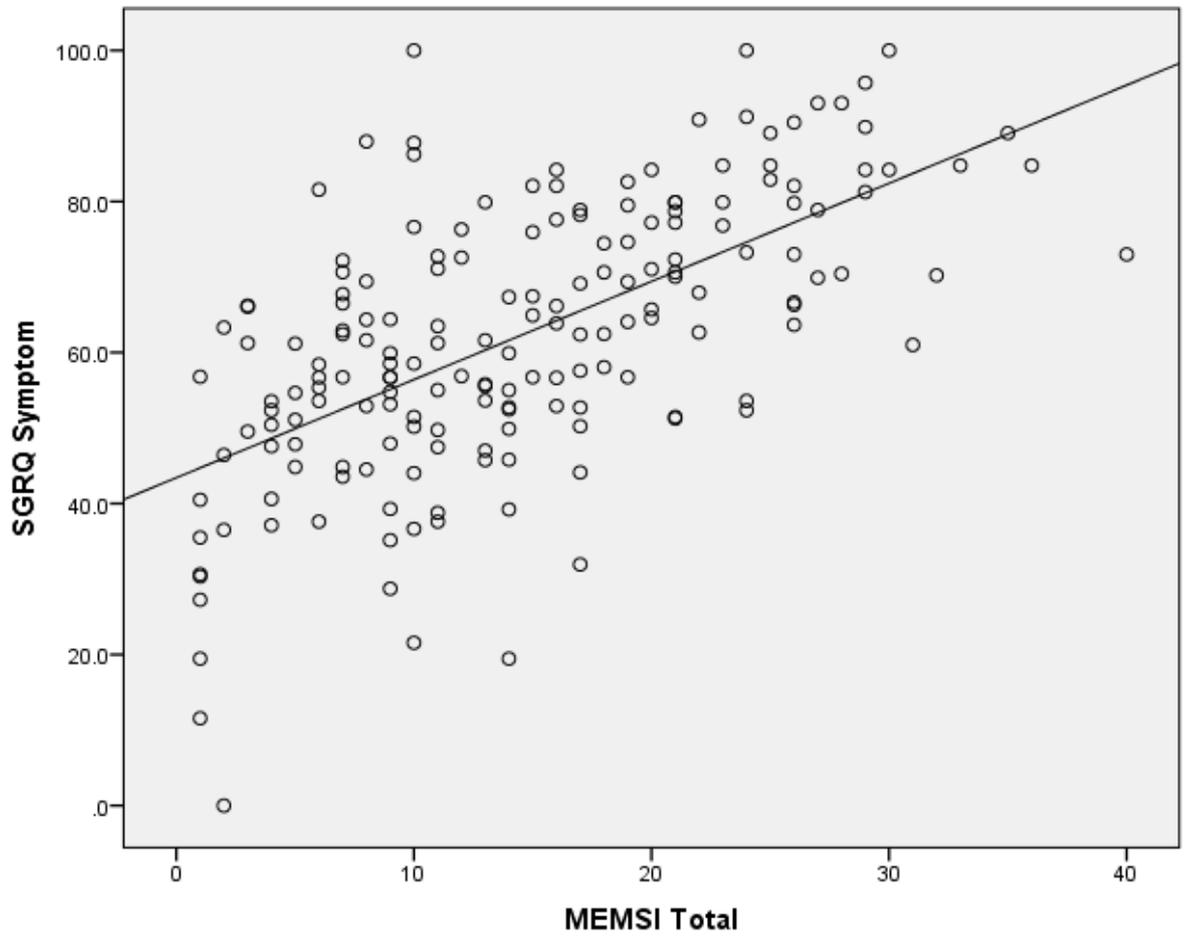
2. Correlation between MEMSI total score and SGRQ-C Activity Score
(Pearson's $r=0.54$)



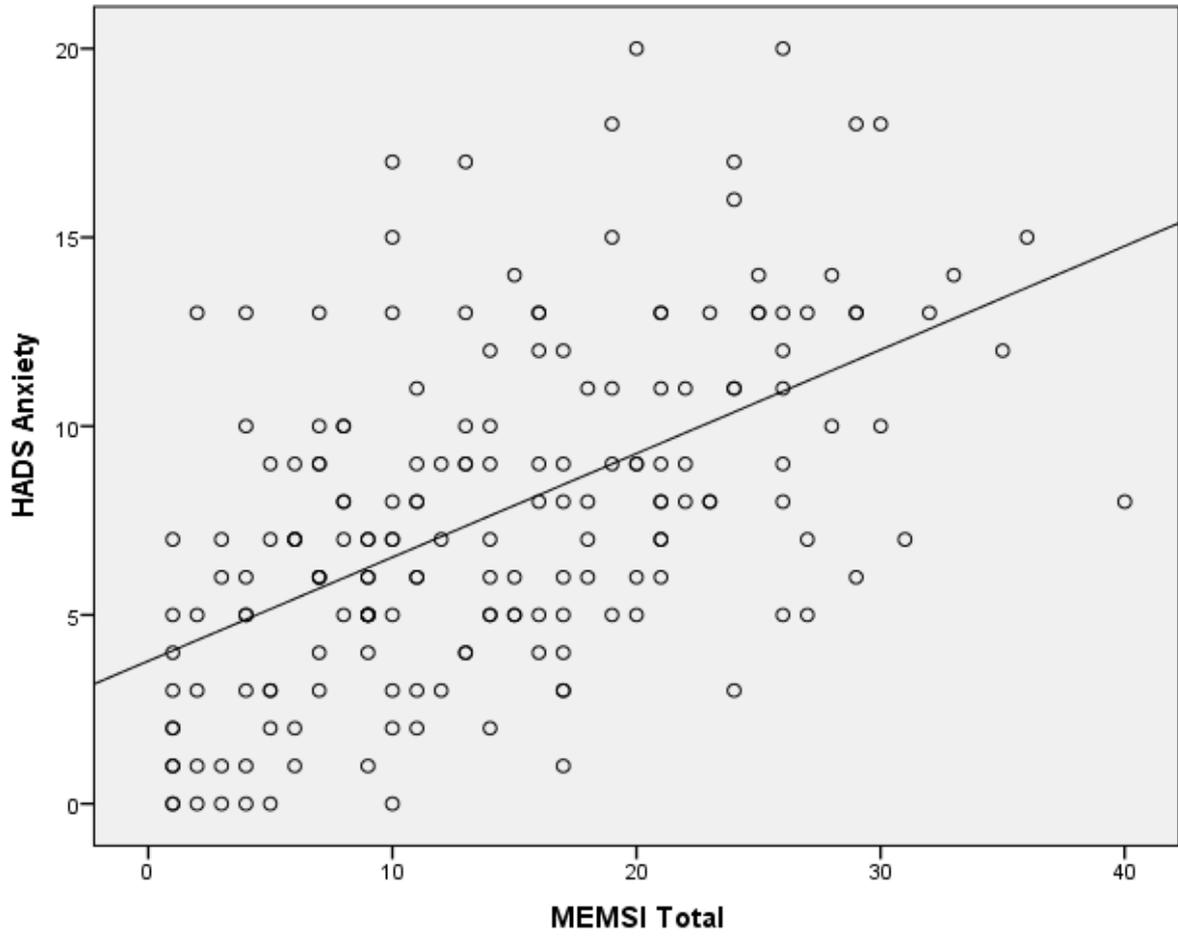
3. Correlation between MEMSI total score and SGRQ-C Impact Score
(Pearson's $r=0.72$)



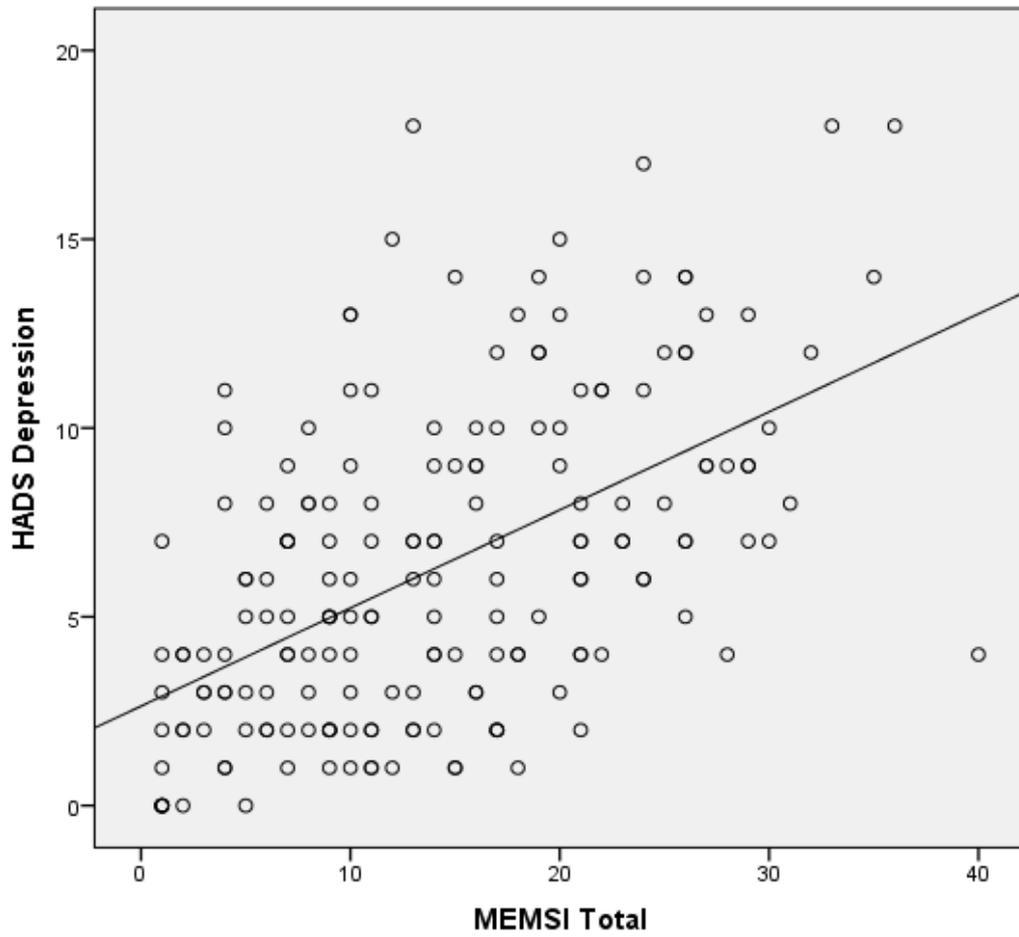
4. Correlation between MEMSI total score and SGRQ-C Symptom Score
(Pearson's $r=0.62$)



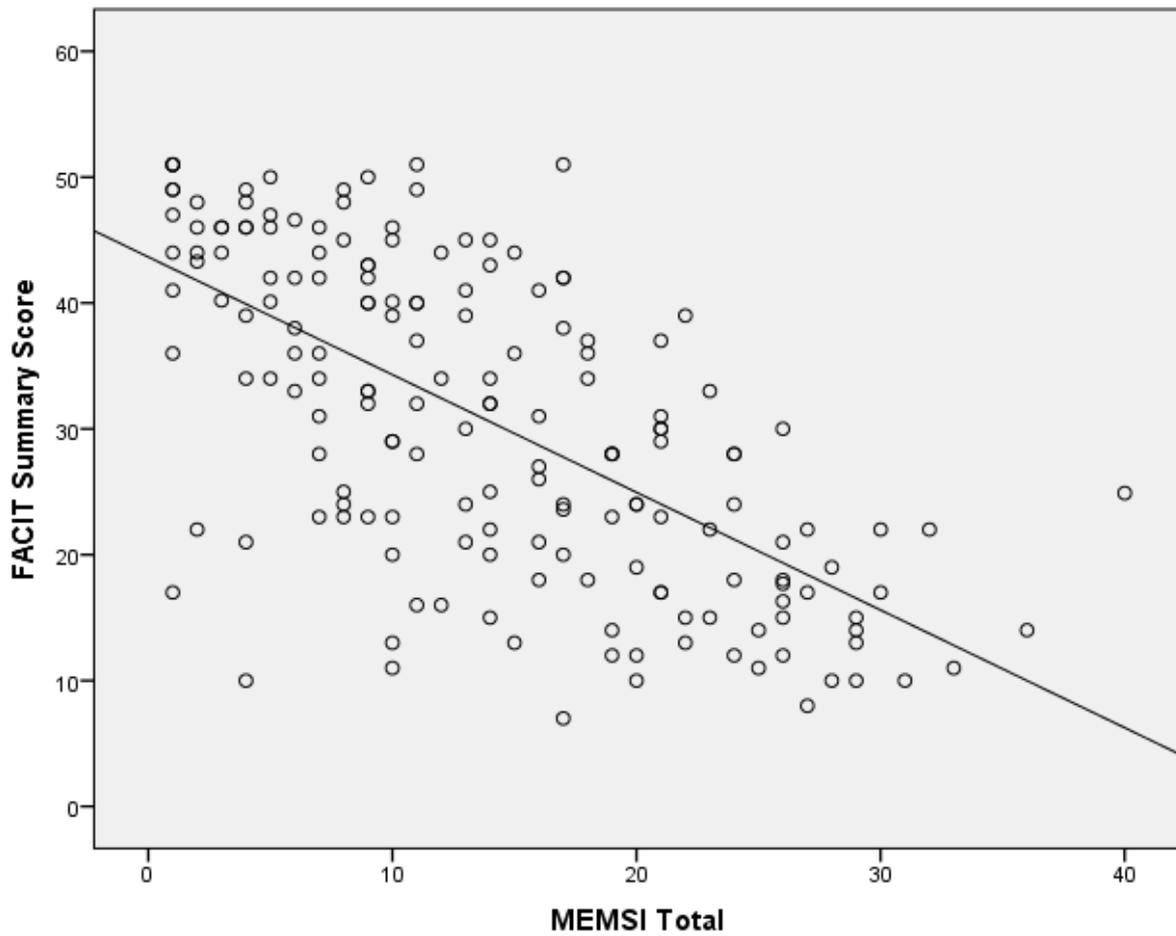
5. Correlation between MEMSI total score and HADS Anxiety score (Pearson's $r=0.53$)



6. Correlation between MEMSI total score and HADS depression (Pearson's $r=0.54$)



7. Correlation between MEMSI total score and FACIT-F Total Score (Pearson's $r=-0.65$)



8. Association between mean MEMSI scores and GOLD grade

