Comparison of seven performance scales for differentiating individuals with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) from other patient groups

Authors

Ruth R. Miller, PhD[1] (Corresponding)

British Columbia Centre for Disease Control, 655 West 12th Avenue, Vancouver, BC, V5Z 4R4, Canada. Phone: +1604 707 2491. Fax: +1604 707 2516. Email: ruth.miller@bccdc.ca

Jennifer L. Gardy, PhD[1,2]

British Columbia Centre for Disease Control, 655 West 12th Avenue, Vancouver, BC, V5Z 4R4, Canada. Phone: +1604 707 2488. Fax: +1604 707 2516. Email: jennifer.gardy@bccdc.ca

Shoshana M. Parker BSc[5]

Centre for Health Evaluation and Outcome Sciences, 588-1081 Burrard St., Vancouver, BC, V6Z 1Y6. Phone: +1604-682-2344 X 62492. Fax: +1604-806-8005. Email: sparker@cheos.ubc.ca

Alison C. Bested, MD, FRCPC[6]

1265 Morningside Avenue, Suite 107, Toronto ON, M1B 3V9. Phone: +1416-282-5773 Fax: +1416-282-9545. Email: acbes@bell.net

Patrick Tang, MD PhD[3,4]

British Columbia Centre for Disease Control, 655 West 12th Avenue, Vancouver, BC, V5Z 4R4, Canada. Phone: +1604 707 2616. Fax: +1 604 707 2675. Email: patrick.tang@bccdc.ca

David M. Patrick, MD, FRCPC, MHSc[1,2]

British Columbia Centre for Disease Control, 655 West 12th Avenue, Vancouver, BC, V5Z 4R4, Canada. Phone: +1604 707 2518. Fax: +1604 707 2516. Email: david.patrick@bccdc.ca

Affiliations:

1. School of Population and Public Health, University of British Columbia
2. British Columbia Centre for Disease Control
3. Department of Pathology and Laboratory Medicine, University of British Columbia
4. Laboratory Services, British Columbia Centre for Disease Control
5. Centre for Health Evaluation and Outcome Sciences
6. Clinical Associate Professor, Faculty of Medicine, University of British Columbia

Abstract

Performance scales measure the health of Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) sufferers, and are used to distinguish them from individuals with other illnesses, or to categorise sufferers into subgroups. Although the Karnofsky performance scale is commonly used, other performance scales may have greater resolution to distinguish between the narrow ranges of symptoms seen in ME/CFS. We recruited 25 participants with ME/CFS and 13 with alternatively diagnosed clinical Lyme syndrome (ADCLS), and compared them to 25 healthy controls and 11 controls with systemic lupus erythematosus (SLE) using seven performance scales. Scores for ME/CFS and ADCLS sufferers were significantly worse than healthy controls on all scales and 6/7 scales, respectively. Comparison to SLE controls showed less difference. Two scales (Short Form-36 physical component and Fatigue Severity Scale) distinguished completely between ME/CFS cases and healthy controls, whilst the Functional Capacity Scale (FCS) and Karnofsky also distinguished between the two groups at a level that may be diagnostically useful. When ME/CFS and ADCLS participants completed scales according to their best and worst days, the FCS appeared better at distinguishing the two groups than Karnofsky, but the difference was not statistically significant. Scales, such as the FCS, covering symptoms more appropriate for ME/CFS may be useful for future studies investigating ME/CFS.

Keywords: Performance scale, Karnofsky, Myalgic Encephalomyelitis/Chronic Fatigue Syndrome, systemic lupus erythematosus, symptom evaluation

Introduction

Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) is a multifactorial disease defined by unexplained fatigue. There is substantial heterogeneity of symptoms among ME/CFS sufferers, which are hard to characterise within a single metric. For this reason, performance scales are often used to measure the health of individuals with ME/CFS [1]. Such measures may be used to distinguish ME/CFS sufferers from healthy controls or from individuals with other illnesses, as well as to stratify ME/CFS patients into subgroups based on their degree of impairment.

The Karnofsky performance status scale [2] is a well-used and validated scale that allows patients to be classified on their functional impairment. Despite being commonly used in the ME/CFS literature, the Karnofsky scale includes extremes that are not appropriate for measurement in ME/CFS, from no evidence of disease to dead. Therefore, alternative scales are often utilised. For example, the Functional Capacity Scale (FCS) [3] which, while not formally validated against Karnofsky for use in assessing patients with ME/CFS, is commonly used clinically as it is designed to evaluate ME/CFS across a broader descriptive range of fatigue symptoms.

Other commonly used scales include the Short Form 36 (SF36) health survey [4], which reports overall physical and mental summaries based on eight scales covering concepts from bodily pain to social functioning. The SF36's individual scales can be useful for establishing which factors are causing most debility in ME/CFS patients. Additionally, the Fatigue Severity Scale (FSS) [5] and Pittsburgh Sleep Quality Index (PSQI) [6] are both

more specific measures related to fatigue and sleep quality, and the Centre for Epidemiologic Studies Depression scale (CES-D) [7] is a measure for depression.

Various studies have compared the utility of performance scales in ME/CFS, and often focus on SF36 and FSS [8-11]. Such studies have demonstrated that these two measures, among others, are able to identify patients with ME/CFS compared to healthy controls. However, to our knowledge, no studies have included the commonly used Karnofsky and FCS scales in their comparisons, particularly not to measure extremes of health within ME/CFS sufferers.

We sought to compare seven performance scales – Karnofsky, FCS, SF36 physical and mental FSS, PSQI, and CES-D – for their ability to distinguish between ME/CFS cases and healthy controls, as well as between individuals with alternatively diagnosed chronic Lyme syndrome (ADCLS), and a second control group with systemic lupus erythematosus (SLE), both patient groups also frequently experiencing fatigue [12]. We also compared the ability of the Karnofsky and FCS scales to describe changes in health for ME/CFS and ADCLS patients. Our results will aid researchers in identifying appropriate measures to use in future studies.

Methods

Subjects were recruited as part of the Chronic Complex Diseases (CCD) study at the University of British Columbia, a case-control study designed to compare participants with chronic complex diseases. Recruitment was through newspapers, websites and physician

referrals in southern British Columbia and patients were excluded if they were <19 years, unable to understand English, diagnosed with another medical condition explaining their symptoms, or on antibiotic therapy in the last month. ME/CFS participants had a physician-based diagnosis that fit the Canadian Case definition [13], and all also met the Fukuda (CDC) definition [14]. The Canadian Case definition was chosen for patient selection, since it has high specificity and thus selects a relatively homogeneous group of ME/CFS patients. ADCLS participants were diagnosed on clinical grounds supported only by alternative laboratory tests for Lyme disease [15], the validity of which is questioned by major reference laboratories and the CDC [16]. SLE participants had to fit standard diagnostic criteria for lupus [12]. ME/CFS and SLE participants were matched to healthy controls by gender and five-year age strata. The study protocol was approved by the University of British Columbia's IRB (H11-01998).

After collection of histories and medical data, subjects self-completed the Karnofsky [2], FCS, SF36 [4], FSS [5], PSQI [6], and CES-D scales [7]. Participants were asked to complete the Karnofsky, FCS, FSS and CES-D scales based on their experience over the preceding week, and the SF36, PSQI scales over the preceding four weeks/month in line with scale protocols. Participants also completed two additional Karnofsky and FCS assessments reflecting their best and worst days.

Statistical analyses were performed using STATA V12.1. Basic comparison of scales between pairs of groups was performed using non-parametric Mann-Whitney rank sum tests. *P* values <0.05 were considered significant. We also performed a Receiver Operating

Characteristic Curve (ROC) analysis [17] to determine how useful each scale was for distinguishing each study group from controls and other groups, and how Karnofsky and FCS distinguished individuals' best and worst days. Comparison between ROC curves of participants' best and worst days was performed using a modified chi-squared test [18].

Results

Twenty-five ME/CFS cases, 13 ADCLS cases, 25 healthy controls, and 11 SLE controls were recruited and all completed each questionnaire. Groups did not differ significantly in gender and age distribution, apart from ADCLS participants who were significantly younger than ME/CFS cases and healthy controls (P=0.02 for both comparisons) (Table 1).

We first examined the differences in scores across the scales for all pairwise comparisons of the four participant groups. Scores for ME/CFS participants were significantly worse than healthy controls on all seven scales (Table 1). Scores for ADCLS participants were also significantly worse than healthy controls on six of seven scales, with the SF36-mental component having a lower median score but at a non-significant P value (Table 1). SLE controls differed significantly from healthy controls on five scales – all but the SF36-mental and PSQI, however, SLE controls were closer to healthy controls in their distribution of scores (Table 1).

Scores for ME/CFS and ADCLS participants compared to SLE controls also showed significant differences between both groups in four and three of the seven scales, respectively (Table 1). Only the FSS scale differed significantly between CFS and ADCLS participants, with ME/CFS patients reporting more severe fatigue; however, there is a

broad overlap in score distribution between the two patient groups on this scale and the *P* value (*P*=0.02) is marginal given the many comparisons made (Table 1).

Value of Functional Scales in Distinguishing Between Study Groups

We next examined the utility of each of the seven scales in correctly classifying participants as belonging to one of the four groups. The Receiver Operating Characteristic Curve, or ROC, quantifies the predictive value of a classifier by measuring the area under the curve (AUC). An AUC of 1.0 represents perfect classification accuracy, while a score of 0.5 is equivalent to random guessing and indicates no useful predictive value. AUCs of 0.9 or above may be useful diagnostically, those of 0.75 to 0.9 moderately useful, and those below 0.75 of questionable value [17]. Several scales performed well in distinguishing participants with ME/CFS or ADCLS from healthy controls using ROC (Table 2). The SF36-physical, FCS, and Karnofsky scales all yielded an AUC greater than 0.9 for both groups compared to healthy controls. The FSS scale also performed well for both groups.

The SF36-mental health scale did not prove as useful, yielding an AUC of only 0.75 and 0.74 respectively for distinguishing between ME/CFS and ADCLS. Similarly, CES-D was not useful for distinguishing between any groups. Sub-component scores for the SF36 that focused on physical function, bodily pain, general health, vitality and social function performed better in distinguishing those with ME/CFS or ADCLS from controls than did sub-components based on emotional and mental health (Table 2). Most scales did not perform as well in distinguishing the SLE group from healthy controls or from the two

study groups. Notably, no scale proved useful in distinguishing subjects with ME/CFS from ADCLS (Table 2).

Value of two Functional Scales in Distinguishing between patients' best and worst days

We also used ROC analysis to determine which scales could optimally distinguish between best and worst days as reported by individual subjects (Figure 1). For patients with ME/CFS and ADCLS, the Functional Capacity Scale performed best at distinguishing between their best and worst day, with AUC of 0.987 (95% CI 0.968-1.000) and 0.982 (0.852-1.000) respectively, followed closely by the Karnofsky with AUCs of 0.925 (95% CI 0.848-1.000) and 0.947 (0.852-1.000) respectively. However, these differences were not statistically significant (P=0.10 and 0.41 for ME/CFS and ADCLS respectively) and had largely overlapping 95% confidence intervals.

Discussion

Out of seven performance scales investigated, the SF36-physical and FSS scales consistently and perfectly distinguished between ME/CFS cases and healthy controls in our dataset, while the FCS and Karnofsky scales had very high, but not perfect, discriminatory power. Such findings are in keeping with research from other studies, which have also found that SF36 and FSS good classifiers of ME/CFS [8, 10, 11].

The SF36-physical scale also exhibited high discriminatory power to distinguish ADCLS cases from healthy controls, with the Karnofsky and FCS scales also performing well. When comparing participants on their best and worst days, the Functional Capacity Scale

appeared to be a better discriminator of patients' good and bad days than Karnofsky for both patients with ME/CFS and ADCLS. The FCS demonstrated the increased clinical utility of a scale covering a narrower spectrum of symptoms. However, the differences between the Karnofsky and FCS scales were not statistically significant.

As none of the scales were useful for distinguishing between ME/CFS and ADCLS, this suggests that both syndromes have a similar phenotype with similar severity of symptoms, marked by profound functional impairment. Median values for the Karnofsky score of 60 and 65 for ME/CFS and ADCLS participants, respectively, suggest that the average patient with either of these conditions is unable to carry out normal activity or work. Further supporting this high level of disability, median FCS values of 4 and 5 suggest that patients with syndromes experience symptoms even at rest and have to limit daily activities. The SF36 physical and mental components are calibrated around a mean of 50.0 for the general population, and in this study ME/CFS and ADCLS sufferers had mean values of 26.6 and 31.0, respectively, for the physical component, again indicating a high level of disability.

Also in keeping with previous research, scales based primarily on physical symptoms (e.g., Karnofsky, SF36-physical) were able to more clearly distinguish between those with ME/CFS or ADCLS from healthy controls than those based on mental symptoms [11]. The SF6-mental scale and CES-D had less ability to discriminate between cases and controls, and the mean values for the mental component of SF36 were 45.0 and 44.2 for ME/CFS and ADCLS, respectively – much closer to the population average than the SF36-physical scores. This suggests the roots of these syndromes may be physical rather than mental; however,

22/25 ME/CFS cases and 11/13 ADCLS had CES-D scores indicating sub-threshold depression symptoms, whereas only 9/25 healthy controls had similar scores. Thus, mental health may play a role in exacerbating these syndromes, as seen in most chronic conditions, though perhaps as effect rather than cause. It is important to avoid a false dichotomy of physical and "mental" symptoms in these disorders, as fatigue, body pain, depression, and anxiety may all ultimately be determined by perturbations in neurochemistry.

This study was performed on a small number of participants, thus larger follow-up studies are needed to confirm our findings, particularly those related to the ADCLS participants, who differed in age from other groups. However, we would expect that the younger age of ADCLS participants would correlate with fewer symptoms, making the differences between ADCLS and control participants conservative. To help make the small number of ME/CFS participants selected for this study more homogeneous, we chose to define cases using the Canadian Case definition, and cases also happened to fit the Fukuda definition. However, other definitions of ME/CFS are available, such as the newly recommended US Institutes of Medicine case definition [19]. Further studies would be required for evaluation of these scales for distinguishing patients selected using other ME/CFS definitions. Furthermore, a fuller validation of scales for aiding in diagnosis might also require inclusion of people with other medical diagnoses. However, for measuring the performance of these scales for classifying ME/CFS and ADCLS patients, our study provides a good indication that scores such as FCS, FSS and SF36-physical provide good alternatives to the Karnofsky scale, and that the Functional Capacity Scale is highly appropriate for measuring change in symptoms within sufferers.

The widely used Karnofsky score runs a spectrum from completely moribund to extremely well, therefore, ME/CFS clinicians may wish to employ scales with a higher resolution into the spectrum of fatigue between these two extremes. We showed that the FCS compares well with the Karnofsky scale in distinguishing symptoms between patient groups, and also tends to do better for describing changes in symptoms experienced by the same patient. For this reason, the FCS, already in broad use in clinical settings, may prove to be highly useful in future studies in ME/CFS.

Acknowledgements:

In addition to the named authors, members of the Complex Chronic Disease Study Team include A. Mattman, S. Sirrs, W. D. Reid, P. Phillips, A. Avina-Zubieta, J. Reynolds, H. Wong, I. Hyams, B. Ng, G. Blaney, J. Spinelli, J. Isaac-Renton, L. Hoang, M. Krajden and M. McCabe, J. Singer.

.

References

1. Hardcastle, S.L., et al., *Severity Scales for Use in Primary Health Care to Assess Chronic Fatigue Syndrome/Myalgic Encephalomyelitis.* Health Care Women Int, 2014: p. 0.
2. Karnofsky DA, B.J., *The Clinical Evaluation of Chemotherapeutic Agents in Cancer*. Evaluation of Chemotherapeutic Agents, ed. M. CM. 1949: Columbia University Press.
3. *Functional Capacity Scale*. 2015 [cited 2015 2015-04-10]; Functional Capacity Scale]. Available from: http://ocfp.on.ca/docs/committee-documents/activity-log-and-functiona-capacity-scale.pdf?sfvrsn=3.
4. Ware J.E., J., Kosinski M., Bjorner J.B., Turner-Bowker, D.M., Gandek, B. and Maruish, M.E. , *User's Manual for the SF-36v2TM Health Survey (2nd Ed).* 2007, Lincoln, RI: QualityMetric Incorporated. .
5. Krupp, L.B., et al., *The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus.* Arch Neurol, 1989. **46**(10): p. 1121-3.
6. Buysse, D.J., et al., *The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research.* Psychiatry Res, 1989. **28**(2): p. 193-213.
7. Radloff, L., *The CES-D Scale: A self-report depression scale for research in the general population.* Applied Psychological Measurement, 1977. **1**: p. 385-401.
8. Davenport, T.E., et al., *Reliability and validity of Short Form 36 Version 2 to measure health perceptions in a sub-group of individuals with fatigue.* Disabil Rehabil, 2011. **33**(25-26): p. 2596-604.
9. Taylor, R.R., L.A. Jason, and A. Torres, *Fatigue rating scales: an empirical comparison.* Psychol Med, 2000. **30**(4): p. 849-56.
10. Jason, L.A., et al., *Fatigue Scales and Chronic Fatigue Syndrome: Issues of Sensitivity and Specificity.* Disabil Stud Q, 2011. **31**(1).
11. Jason, L., et al., *Measuring substantial reductions in functioning in patients with chronic fatigue syndrome.* Disabil Rehabil, 2011. **33**(7): p. 589-98.
12. Hochberg, M., *Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erhthematosus.* Arthritis Rheum, 1997. **40**: p. 1725.
13. Carruthers, B.M., et al., *Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Clinical Working Case Definition, Diagnostic and Treatment Protocols.* Journal of Chronic Fatigue Syndrome, 2003. **11**(1): p. 7-36.
14. Fukuda, K., et al., *The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group.* Ann Intern Med, 1994. **121**(12): p. 953-9.
15. Patrick, D.M., et al., *Lyme Disease Diagnosed by Alternative Methods: A Common Phenotype with Chronic Fatigue Syndrome.* Clin Infect Dis, 2015. **In press**.
16. *Caution while testing for Lyme disease.* MMWR, 2005. **54**(5): p. 125.
17. Zou, K.H., A.J. O'Malley, and L. Mauri, *Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models.* Circulation, 2007. **115**(5): p. 654-7.
18. DeLong, E.R., D.M. DeLong, and D.L. Clarke-Pearson, *Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach.* Biometrics, 1988. **44**(3): p. 837-45.

19. Clayton, E.W., *Beyond myalgic encephalomyelitis/chronic fatigue syndrome: an IOM report on redefining an illness.* JAMA, 2015. **313**(11): p. 1101-2.

Table 1: Details of patient age and gender, and scores for each performance scale

	Range (Best-worst)	Healthy Median (IQR)	ME/CFS Median (IQR)	ADCLS Median (IQR)	SLE Median (IQR)	P Healthy vs ME/CFS	P Healthy vs ADCLS	P Healthy vs SLE	P ME/CFS vs ADCLS	P ME/CFS vs SLE	P ADCLS vs SLE
Demographics											
Male* N (%)	NA	4 (16%)	4 (16%)	3 (23%)	0 (0%)	1.0	0.7	0.3	0.7	0.3	0.2
Age (years)	NA	53 (30;69)	54 (34;67)	45 (18;71)	51 (29;75)	0.9	**0.02**	0.5	**0.02**	0.6	0.4
Preformance Scales											
Karnofsky	100 - 0	100 (95;100)	60 (60;70)	65 (60;70)	90 (70;90)	**<0.0001**	**<0.0001**	**0.001**	0.4	**0.001**	**0.02**
Functional Capacity Scale	10 - 0	9 (8;9)	4 (3;5)	5 (4;6)	8 (5;8)	**<0.0001**	**<0.0001**	**0.001**	0.1	**0.001**	**0.05**
SF36 Physical	100 - 0	56 (52;60)	25 (22;33)	30 (25;36)	40 (32;55)	**<0.0001**	**<0.0001**	**0.001**	0.1	**0.0001**	**0.01**
SF36 Mental	100 - 0	55 (47;58)	48 (40;55)	47 (37;50)	51 (45;57)	**0.0029**	0.2	0.1	0.8	0.2	0.3
Fatigue Severity	0 - 63	18 (14;30)	62 (57.5;63)	54 (33;61)	41 (28;53)	**<0.0001**	**0.0006**	**0.0007**	**0.02**	**<0.0001**	0.2

Scale											
Pittsburgh Sleep Quality Index	0 - 12	6 (3;8)	11 (8;13)	10 (8;12)	7 (5;16)	**<0.0001**	**0.008**	0.2	0.3	0.06	0.4
Centre for Epidemiologic Studies Depression scale	0 - 60	13 (12;16)	21 (19;24)	20 (17;25)	19 (14;20)	**<0.0001**	**0.0001**	**0.01**	0.7	0.09	0.2

Values are Median (IQR) except for male gender which is N (%)

P values are Mann-Whitney rank sum tests for continuous variables and exact tests for male gender. *P* value <0.05 marked in bold.

ME/CFS= Myalgic Encephalomyelitis, ADCLS = Alternatively diagnosed clinical Lyme syndrome, SLE=systemic lupus erythematosus, IQR=Interquartile range.

Table 2: Area Under the Curve and 95% Confidence Intervals of a Receiver Operating Characteristic Curve for Distinguishing Patient Groups from Each Other

AUC	SLE vs. Healthy	ME/CFS vs. Healthy	ADCLS vs. Healthy	ME/CFS vs. SLE	ME/CFS vs. ADCLS	ADCLS vs. SLE
	0.840	**1.000**	**0.972**	0.898	0.654	0.811
SF36 Physical	(0.692;0.988)	**(1.000;1.000)**	**(0.921;1.000)**	(0.795;1.000)	(0.458;0.850)	(0.639;0.984)
	0.655	**0.747**	0.735	0.629	0.474	0.636
SF36 Mental	(0.455;0.854)	**(0.609;0.884)**	(0.555;0.916)	(0.426;0.832)	(0.265;0.683)	(0.402;0.871)
	0.826	**0.989**	**0.955**	0.843	0.678	0.731
FCS	(0.686;0.965)	**(0.971;1.000)**	**(0.898;1.000)**	(0.681;1.000)	(0.504;0.851)	(0.519;0.943)
	0.809	**0.980**	**0.962**	0.835	0.575	0.783
Karnofsky	(0.652;0.966)	**(0.944;1.000)**	**(0.910;1.000)**	(0.661;1.000)	(0.392;0.758)	(0.534;0.983)
	0.6473	0.833	0.763	0.697	0.593 (0.340;	0.608
PQSI	(0.436;0.859)	(0.713;0.952)	(0.607;0.919)	(0.447;0.947)	0.790)	(0.346;0.870)
	0.8582	**1.000**	0.843	**0.932**	0.721	0.657
FSS	(0.722;0.995)	**(1.000;1.000)**	(0.684;1.000)	**(0.846;1.000)**	(0.532;0.910)	(0.424;0.891)

	0.769	0.898	0.879	0.674	0.533	0.661
CES-D	(0.595;0.943)	(0.809;0.987)	(0.770;0.988)	(0.479;0.870)	(0.323;0.744)	(0.436;0.886)
SF-36 sub components						
Physical	0.838	**0.998**	**0.965**	0.869	0.631	0.717
Function	(0.689;0.987)	**(0.992;1.000)**	**(0.907;1.000)**	(0.746;0.992)	(0.403;0.858)	(0.503;0.931)
	0.678	**0.999**	**0.932**	0.896	0.602	0.818
Role Physical	(0.456;0.900)	**(0.997;1.000)**	**(0.817;1.000)**	(0.790;1.000)	(0.401;0.803)	(0.648;0.989)
	0.746	**0.937**	**0.928**	0.815	0.422	0.826
Bodily Pain	(0.585;0.906)	**(0.878;0.996)**	**(0.841;1.000)**	(0.667;0.963)	(0.213;0.630)	(0.658;0.992)
	0.915	**1.000**	**0.983**	0.780	0.546	0.738
General Health	(0.814;1.000)	**(1.000;1.000)**	**(0.948;1.000)**	(0.571;0.989)	(0.344;0.749)	(0.514;0.961)
	0.800	**0.952**	**0.908**	0.831	0.563	0.748
Vitailty	(0.653;0.947)	**(0.889;1.000)**	**(0.813;1.000)**	(0.672;0.990)	(0.357;0.769)	(0.547;0.949(
Social	0.756	**0.989**	**0.906**	0.871	0.671	0.734
Functioning	(0.581;0.932)	**(0.972;1.000)**	**(0.589;0.943)**	(0.739;1.000)	(0.486;0.855)	(0.631;0.937)
Role Emotional	0.586	0.641	0.766	0.544	0.446	0.640

	(0.372;0.799)	(0.486;0.796)	(0.589;0.943)	(0.351;0.736)	(0.262;0.631)	(0.401;0.871)
	0.562	0.691	0.780	0.609	0.420	0.682
Mental Health	(0.335;0.788)	(0.544;0.838)	(0.605;0.955)	(0.400;0.818)	(0.227;0.613)	(0.456;0.908)

Values are area under the curve (AUC) and (90% confidence intervals). Bold type highlights AUC for tests with good ability to discriminate between groups.

ME/CFS= Myalgic Encephalomyelitis, ADCLS = Alternatively diagnosed clinical Lyme syndrome, SLE=systemic lupus erythematosus, AUC= area under the curve, SF36= Short Form 36 health survey, FCS= Functional Capacity Scale, PSQI= Pittsburgh Sleep Quality Index, FSS= Fatigue Severity Scale, CES-D= Centre for Epidemiologic Studies Depression scale.

Figure legends:

Figure 1: Receiver Operating Characteristic Curve (ROC) for the Functional Capacity Scale and Karnofsky performance scale to compare ME/CFS and ADCLS participants on their best and worst days.

Reference line shows the curve if prediction was completely due to chance.