Scientific Research

Sensitivity and Specificity of the CDC Empirical Chronic Fatigue Syndrome Case Definition

Leonard A. Jason, Meredyth Evans, Abigail Brown, Molly Brown, Nicole Porter, Jessica Hunnell, Valerie Anderson, Athena Lerch

DePaul University, Chicago, American.
Email: Ljason@depaul.edu

ABSTRACT

In an effort to bring more standardization to the chronic fatigue syndrome (CFS) Fukuda et al. case definition [1], the Centers for Disease Control and Prevention (CDC) has developed an empirical case definition [2] that specifies criteria and instruments to diagnose CFS. The present study investigated the sensitivity and specificity of this CFS empirical case definition with diagnosed individuals with CFS from a community based study that were compared to non-CFS cases. All participants completed questionnaires measuring disability (Medical Outcome Survey Short-Form-36) [3], fatigue (the Multidimensional Fatigue Inventory) [4], and symptoms (CDC Symptom Inventory) [5]. Findings of the present study indicated sensitivity and specificity problems with the CDC empirical CFS case definition.

*Keywords***:** *Chronic Fatigue Syndrome, Empirical Case Definition, Centers for Disease Control and Prevention, Fukuda Criteria*

1. Sensitivity and Specificity of the CDC Empirical Case Definition

The Centers for Disease Control and Prevention (CDC) has developed an empirical case definition for chronic fatigue syndrome (CFS) that involves assessment of symptoms, disability, and fatigue [2]. The CDC empirical CFS case definition assesses three specific areas to determine whether a person meets criteria for this illness including: 1) disability, using the Medical Outcomes Survey Short Form-36 (SF-36) [3], 2) fatigue, using the Multidimensional Fatigue Inventory (MFI) [4], and 3) symptoms, using the CDC Symptom Inventory (SI) [5]. The authors of this empirical case definition feel that the specification of instruments and cut-off points will result in a more reliable and valid approach for the assessment of CFS.

The disability criterion for the Reeves *et al.* empirical CFS case definition [2] would be met by scoring below the 25th percentile on any one of the following four SF-36 sub-scales [3]: Physical Functioning (less than or equal to 70), Role Physical (less than or equal to 50), Social Functioning (less than or equal to 75), or Role Emotional (less than or equal to 66.7). Because a person could meet the disability criterion for the empirical CFS case definition by only showing impairment in one or more of these four areas, a person could meet the disabil-

ity CFS criterion by only having an impairment in role emotional areas (e.g., problems with work or other daily activities as a result of emotional problems). Ware *et al.* [3] found that the mean for Role Emotional for a clinical depression group was 38.9, indicating that almost all those with clinical depression would meet the CFS disability criterion, as they would be within the lower 25th percentile on this sub-scale.

To meet the fatigue criterion, the Reeves et al. empirical case definition [2] requires a score on the MFI [4] of greater than or equal to 13 on the General Fatigue sub-scale, or greater than or equal to 10 on the Reduced Activity sub-scale. In one study of three groups with CFS [6], the mean MFI General Fatigue scores ranged from 18.3 to 18.8 and these scores are clearly higher than the Reeves *et al.* cutoff of 13. In addition, Reduced Activity items refer to issues that a person with depression might easily endorse. If a person indicated that the following two items were entirely true: "I get little done," and "I think I do very little in a day"; they would meet criterion for fatigue on this sub-scale.

The SI [5] assesses information about the presence, frequency, and intensity of fatigue related symptoms during the past one month. The frequency and severity scores were multiplied for each of the eight critical Fukuda *et al.* [1] symptoms and were then summed. To meet the Reeves *et al.* [2] symptom criterion, a person

needed to have four or more symptoms and a total score greater or equal to 25 on the SI. This overall level of symptoms seems relatively low for patients with classic CFS symptoms (the criterion would be met if an individual rated only 2 core symptoms as occurring all the time, and if one was of moderate and the other of severe severity). In addition, the 8 case definition symptoms for the empirical case definition were based on a time period comprising the last month compared to what is specified in the Fukuda *et al.* criteria, which states that: "There needs to be the concurrent occurrence of 4 or more of the following symptoms, and all must be persistent or recurrent during 6 or more months of the illness and not predate the fatigue."

Jason, Najer, Porter, and Reh [7] recently investigated this CFS empirical case definition with 27 participants with a diagnosis of CFS and 37 participants with a diagnosis of a Major Depressive Disorder (MDD). All participants completed questionnaires measuring disability (SF-36), fatigue (MFI), and symptoms (SI). Jason et al. found that 38% of those with a diagnosis of MDD were misclassified as having CFS using the new CDC empirical case definition. Jason, Evans, *et al.* [8] later used this same sample to examine issues of sensitivity and specificity for the three instruments along with their cut-off points. Sensitivity is the probability that the test correctly classifies a person with CFS as positive, whereas specificity is the probability that a test correctly classifies a person without CFS as negative. When Jason, Evans, *et al.* used a Receiver Operating Characteristic (ROC) curve analysis with the Reeves *et al.* criteria [2], they found the disability, fatigue and symptom criteria had serious specificity and/or sensitivity problems. They concluded that the Reeves *et al.* criteria would not be considered a good diagnostic method for selecting CFS cases among a sample of CFS and MDD cases.

Reeves, Gurbaxani, Lin, and Unger [9] critiqued the study by Jason *et al.* [7] by stating that the study should have relied on better methods to diagnose the sample, including a medical and psychiatric examination. Another criticism brought up by Reeves et al. was the focus on MDD, particularly as some persons with CFS also suffer from MDD. Some individuals with CFS do have MDD, but the key issue is that MDD can be confused with CFS, as it has some overlapping symptoms with CFS. For example, it is possible that some patients with MDD also have chronic fatigue and four CFS Fukuda *et al.* [1] symptoms that can occur with depression (e.g., unrefreshing sleep, joint pain, muscle pain, impairment in concentration). Yet, CFS and MDD are different disorders, and they can be differentiated by use of appropriate assessment instruments [10].

Great care needs to be exercised when determining which scales, with which cut off points, should indicate that CFS criteria has been reached for CFS samples. For example, Jason, Brown, *et al.* [11] examined published studies using the SF-36 [3] which contrasted CFS with controls. The largest differences emerged for the Role Physical, Social Functioning, and Vitality SF-36 sub-scales. Rather than arbitrarily selecting the lower 25% for four SF-36 sub-scales, as was recommended by the authors of the empirical CDC CFS case definition [2], Jason, Brown, *et al.* used Receiver Operating Characteristics (ROC) to determine sub-scales that best discriminate CFS from Controls in two well defined samples, one involving a community data base collected in the mid 1990s, and the other a tertiary data base collected in the mid 2000s. Vitality, Social Functioning, and Role Physical had the highest AUCs, with good sensitivity and specificity.

Because the Jason, Brown, *et al.* study [11] only had data on the SF-36, these investigators were not able to examine the Reeves *et al.* [2] recommendations on fatigue or symptom criteria. In addition, the Jason *et al.* [7] sample, which had all three Reeves *et al.* measures, had been criticized as not having formal medical and psychiatric examinations to select cases. The present study includes the disability, fatigue, and symptom measures as recommended by Reeves *et al.* in a carefully defined sample. In this study, we employed an ROC to determine the sensitivity and specificity of the Reeves *et al.* criteria in a well characterized community-based CFS sample. This study included formal medical and psychiatric tests to determine CFS status.

2. Method

The present project was carried out in two stages. In Stage 1, we attempted to re-contact the 213 adults who were medically and psychiatrically evaluated from a community-based sample from 1995-1997. These adults were previously evaluated in our original Wave 1 CFS epidemiology project [12]. Stage 2 of the study encompassed a structured psychiatric assessment, a complete physical examination and a structured medical history.

The original Wave 1 sample collected from 1995-1997 is a stratified random sample of several neighborhoods in Chicago specifically selected to contain individuals from different ethnic and socioeconomic profiles. As a whole, Chicago, Illinois is an ethnically and socioeconomically diverse city. We sampled in eight Chicago community locations, including low socioeconomic areas such as West Garfield Park, middle-socioeconomic areas such as Bridgeport and Armour Park, gentrifying areas such as the near West Side, and high socioeconomic areas such as the Loop and the near North Side. Racial data indicate that the sample consists of 20.0% African-Americans, 52.6% Caucasians, 18.7% Latinos, 0.5% Native Americans, 5.5% Asian Americans, 1.4% multiracial individuals, and 1.3% individuals of other races [12]. The telephone numbers comprising the stratified random sample

were obtained from Survey Sampling, Incorporated. This company generated random telephone numbers using valid Chicago exchanges, resulting in a sample of both listed and unlisted numbers (as well as business and non-working numbers). In the first stage of data collection in the original study, procedures developed by Kish [13] were used to select one adult from each household for subsequent screening for CFS-like illness. Birth dates for each adult were gathered and the person with the most recent birthday was selected to be interviewed using the Stage 1 CFS Screening Questionnaire. The final sample of respondents consisted of 18,675 households.

2.1 Stage 1

The *CFS Screening Questionnaire* consists of two parts and was administered to all participants that could be located for this follow-up study. It assessed participants' sociodemographic characteristics and fatigue characteristics to determine whether any changes have occurred since the first wave of data collection in the original study. Basic demographic data included age, ethnicity, socioeconomic status, work status, marital status, parental status (including number of children) and gender. Consistent with the procedures followed in the original CFS epidemiology study [12] the *CFS Screening Questionnaire* contains questions measuring more specific aspects of fatigue and health status. In addition, questions assessed the level of impairment that fatigue and illness cause to daily activities, as well as the frequency and duration of the fatigue. Respondents were also asked if they have ever been diagnosed with any other medical or psychiatric conditions associated with chronic fatigue and what current treatments they were receiving. A version of the screening scale used in the present study was evaluated by Jason *et al.* [14]. They recruited four groups of subjects (*i.e.*, those diagnosed with CFS, lupus, and multiple sclerosis, and a healthy control group). All subjects were interviewed with a screening instrument twice over a two-week period of time. The screening scale exhibited high discriminant validity and excellent test-retest and inter-rater reliability. Hawk *et al.* [10] revised this CFS Screening Questionnaire, and administered the questionnaire to three groups (those with CFS, MDD, and healthy controls). The revised instrument, which was used in the present study, evidences good test-retest reliability and has good sensitivity and specificity.

2.2 Stage 2

In Stage 2, the *Structured Clinical Interview for the DSM-IV* (SCID) [15] was administered to assess current psychiatric diagnoses as defined on Axis I of the Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition (DSM-IV) [16]. The SCID is a valid and reliable semi-structured interview guide that approxi-

mates a traditional psychiatric interview [17]. It has been successfully used to assess psychiatric disorders in samples of people with CFS [18].

Following the structured psychiatric interview, participants were provided a medical history interview and complete medical examination. Prior to the physical examination, the interviewer who accompanied participants and provided transportation to the medical exam administered the Medical Questionnaire at the physician's office to assess current and past medical history. The Medical Questionnaire is a modified version of The Chronic Fatigue Questionnaire, a structured instrument developed by Komaroff and Buchwald [19] that was used in a study by Komaroff *et al.* [20]. This comprehensive instrument assesses symptoms related to CFS and chronic fatigue, as well as other medical and psychiatric symptoms, in order to help rule out exclusionary conditions such as HIV/AIDS, active malignancies, iatrogenic conditions resulting from the side effects of medication, unresolved cases of hepatitis, and active substance use. In addition, the Medical Questionnaire measures fatigue severity, fatigue-related social role impairment, psychosocial stressors, job satisfaction, toxic exposures prior to CFS onset, chemical sensitivities, presence of CFS or chronic fatigue in other network members, and family medical history. Because sleep disturbances are often reported by individuals with CFS and chronic fatigue, the Sleep Disturbance Questionnaire, which has been validated experimentally in a sleep laboratory [21], has been incorporated into the medical questionnaire to help identify participants with sleep disorders.

Participants also filled out the Medical Outcome Survey Short-Form-36 (SF-36) [3]. This 36-item instrument is composed of multi-item scales that assess functional impairment in eight areas: limits in physical activities (physical functioning), limits in one's usual role activities due to physical health (role physical), limits in one's usual role activities due to emotional health (role emotional), bodily pain, general health perceptions (general health), energy and fatigue (vitality), social functioning, and general mental health. Scores in each area reflect ability to function and higher values indicate better functioning. Reliability and validity studies have demonstrated high reliability and validity in a wide variety of patient populations for this instrument [22]. According to Reeves *et al.* [2] significant reductions in occupational, educational, social, or recreational activities were defined as scores lower than the 25th percentile on Physical Functioning (less than or equal to 70), *or* Role Physical (less than or equal to 50), *or* Social Functioning (less than or equal to 75), *or* Role Emotional (less than or equal to 66.7). A person would meet the disability criterion for the empirical CFS case definition by showing impairment in one or more of these four areas.

Participants also completed the CDC Symptom Inven-

tory (SI) [5]. The SI assesses information about the presence, frequency, and intensity of 19 fatigue related symptoms during the past one month. For each of the eight Fukuda *et al.* [1] symptoms, participants were asked to report the frequency (1 = a little of the time, 2 = some of the time, 3 = most of the time, 4 = all of the time) and severity (the ratings were transformed to the following scale: 0 = symptom not reported, 1 = mild, 2.5 = moderate, 4 = severe)[1]. The frequency and severity scores were multiplied for each of the eight critical Fukuda *et al.* symptoms and were then summed. Individuals having four or more symptoms and scoring greater or equal to 25 would meet symptom criterion on this instrument according to the CDC empirical case definition.

Additionally, the participants completed the Multidimensional Fatigue Inventory (MFI) [4]. This instrument is a 20-item self-report instrument consisting of five scales: general fatigue, physical fatigue, reduced activity, reduced motivation, and mental fatigue. Each scale contains four items rated from 1 to 5 with the scale score of 1= completely true and the scale score of 5 = no, not true. Reeves *et al.* [2] employed the MFI to measure severe fatigue, and to do this, they used only two of the five subscales; General Fatigue and Reduced Activity. Using the CDC empirical case definition standards, severe fatigue was defined as greater than or equal to 13 on General Fatigue *or* greater than or equal to ten on Reduced Activity.

Following the medical history interview, the physician conducted a detailed medical examination. This examination was carried out in order to rule out exclusionary medical conditions and detect evidence of diffuse adenopathy, hepatosplenomegaly, synovitis, neuropathy, myopathy, cardiac or pulmonary dysfunction, or any other medical disorder. An 18-tender-point examination was used to test for Fibromyalgia [23]. Laboratory tests administered to all participants included a chemistry screen (glucose, calcium, electrolytes, uric acid, liver function tests, and renal function tests), complete blood count with differential and platelet count, T4 and TSH, erythrocyte sedimentation rate, arthritic profile (which includes rheumatoid factor and antinuclear antibody), hepatitis B surface antigen, CPK, HIV screen, and urinalysis. An intra-dermal, intermediate-strength PPD skin test was applied, and a posterior-anterior chest x-ray was completed, if it was not already obtained by the participant within eight months of entering the study. At the time of evaluation, the examining physician was blinded to participants' status with respect to initial classification based upon the Stage 1 screen. Participants were reimbursed $100.00 for the time and effort involved in participation. Participants also signed the Human Subjects

Consent Form (See Jason, Porter, Hunnell, Rademaker, & Richman [24] for more details).

At the end of Stage 2, a team of physicians was responsible for making final diagnoses. Two physicians independently rated each file according to the current U.S. definition of CFS. Files that did not meet CFS criteria were rated as either idiopathic chronic fatigue (ICF), exclusionary for CFS due to medically/psychiatrically explained chronic fatigue [1], or control (participants with no exclusionary illness and less than 6 months of fatigue). Those with ICF had at least six months duration of fatigue, but with insufficient symptoms or fatigue to meet the case definition of CFS. The exclusionary group had chronic fatigue for at least six months duration, but with active medical conditions that explain chronic fatigue (e.g., untreated hypothyroidism), previously diagnosed medical disorders whose resolution has not been documented beyond reasonable clinical doubt, and whose continued activity may explain the chronic fatiguing illness (e.g., unresolved cases of hepatitis C). The exclusionary group also included those with chronic fatigue for at least six months duration, but with psychiatric explanations of the fatigue (e.g., delusional disorders, schizophrenia, etc). Controls had no exclusionary illnesses and less than 6 months of fatigue. Reviewing physicians had access to all information gathered on each participant during each of the phases of the study. The review panel was also provided with all results from the physical exam. If a disagreement occurred during the physician review process regarding whether a participant should receive a diagnosis of CFS, ICF, exclusionary due to medically/psychiatrically explained chronic fatigue, or control, the participant's file was rated by a third physician reviewer, and the diagnosis was determined by majority rule. We used refinements of the Fukuda *et al.* criteria as recommended by an International Research group and the CDC [25].

2.3 Sample Characteristics

In Wave 1, 213 adults were medically and psychiatrically evaluated from the community-based sample. For the follow-up study, data was available on 24 individuals diagnosed with CFS and 84 who did not have CFS. Wave 1 differences were examined between those we were able to versus those we were not able to re-evaluate at Wave 2, and we did not find any significant sociodemographic differences for age, gender, race, marital status, number of children, or education (See Jason, Porter, *et al.*, [24] for more details).

2.4 Statistical Analysis

The statistical software package used for data analysis was PASW (formerly SPSS) for Windows, version 17.0. A Receiver Operating Characteristic (ROC) curve analysis [26] was used to evaluate the ability of the scales to

---

[1]The scale we used had five choices, and we needed to convert the ratings to a four point scale in order to conform to Wagner *et al.*'s (2005) severity scaling system.

discriminate between patients with CFS in the community-based sample and those without this illness. The ROC curve graphically represents the probability of true positive results in diagnosis as a function of the probability of false positive results of this test. The area under the curve (AUC) is an indicator of the discriminatory ability of the scale: a straight line (area = 0.5) means that the scale is doing no better than chance in classifying CFS and non-CFS, while a perfect scale would have an ROC curve with an area of 1. The area under the ROC curve is a summary measure that essentially averages diagnostic accuracy across the spectrum of test values. The informative area under the ROC curve ranges from 0.5 to 1.0, and not from 0.0 to 1.0 as would the area under a probability distribution curve. An AUC of .99 means that 99% of the time a randomly selected individual from the CFS group will more adequately fulfill the fatigue criteria than a randomly selected individual from the control group. A test needs an AUC threshold of between 90-100% to have diagnostic meaning, and 95% or above to be considered a good diagnostic tool [27,28].

3. Results

3.1 ROC Analyses

Table 1 presents the ROC analyses for the CFS versus the non-CFS group. The MFI scales had AUCs that were low. When using the cutoff scores proposed by Reeves *et al.* [2], using either the General Fatigue or Reduced Activity criteria, 95% of those with CFS were identified, indicating good sensitivity, but the specificity was only .27, indicating that few of those without the illness would have been correctly identified. The AUC for the SI instrument was also low, and the sensitivity data (.59) suggests that this symptom scale has significant problems in identifying true cases of CFS. Finally, AUC findings for the SF-36 indicate low AUCs, and using Reeves *et al.*'s cutoff scores, that the sensitivity is acceptable at .96; however, specificity is inadequate at .17. When using all three criteria for fatigue, symptoms and disability, the sensitivity was at an unacceptably low level of .65. The sensitivity and specificity outcomes for the Reeves *et al.* criteria suggest that these recommended scales and cutoff points would not be considered a good diagnostic tool for selecting CFS cases from the general population.

4. Discussion

The present study investigated the sensitivity and specificity of the empirical CFS case definition [2] with diagnosed individuals with CFS from a community based study that were compared with non-CFS cases. Findings of the present study indicated sensitivity and specificity problems with the CDC empirical CFS case definition. When comparing the overall Reeves *et al.* criteria, only

Table 1. AUC values, standard errors and confidence intervals for CFS vs. other*

Scale	AUC	Std. Error	95% C.I.		Cut-offs	Sensitivity	Specificity
			LB	UB			
MFI							
Gen. Fatigue	0.69	0.07	0.56	0.82	≥ 13	0.74	0.39
Red. Activity	0.64	0.07	0.51	0.78	≥ 10	0.74	0.50
Meets[a]	0.61	0.07	0.47	0.74		0.95	0.27
SI							
Total[b]	0.69	0.07	0.55	0.84	≥ 25	0.59	0.73
SF-36							
Phys. Func.	0.60	0.06	0.48	0.72	≤ 70	0.68	0.51
Role Phy.	0.66	0.06	0.54	0.77	≤ 50	0.82	0.51
Soc. Func.	0.62	0.07	0.48	0.76	≤ 75	0.74	0.35
Role Emo.	0.57	0.07	0.43	0.70	≤ 66.7	0.73	0.44
Meets[c]	0.56	0.07	0.44	0.69		0.96	0.17
Meets Criteria[d]	0.70	0.08	0.56	0.85		0.65	0.76

[a]Meets Reeves *et al.* (2005) fatigue criteria.
[b]Meets Reeves *et al.* (2005) core symptoms criteria.
[c]Meets Reeves *et al.* (2005) substantial reductions criteria.
[d]Meets Reeves *et al.* (2005) CFS criteria.
*Some of the participants did not complete all three questionnaires, and were thus excluded from the overall sensitivity and specificity figures.

about 65% of true CFS cases were identified. In other words, these criteria are not able to identify an acceptable high percentage of individuals who have this illness.

If samples of CFS are not identified with sensitivity and specificity, it will be difficult to compare samples from different studies, and the search for biological markers will be compromised. Using the Reeves *et al.* criteria [2], the estimated rates of CFS have increased to 2.54% [29], rates that are about ten times higher than prior CDC estimates [30] and prevalence estimates of other investigators [31]. It is at least possible that the increases in the United States are due to a broadening of the case definition and possible inclusion of cases with primary psychiatric conditions. Chronic fatigue occurs in about 4-5% of the population [32]. If about 5% of the population has 6 or more months of fatigue, and about half of this is due to clear medical or psychiatric reasons [31], then the critical question is how many of the remaining 2.5% have CFS. The empirical CFS case definition estimates that 2.54% do have this illness, so that research group would suggest that almost all of the remaining 2.5% would fall within the CFS category. However, Jason *et al.* [7] believe that within this 2.54% are mood disorders, which are one of the most prevalent psychiatric disorders (one-month prevalence rate of major depressive episode is 2.2%) [33]. As an example, one mood disorder is MDD, which can be confused with CFS, as it has some overlapping symptoms with CFS. It is possible that some patients with MDD also have chronic fatigue and four CFS Fukuda *et al.* [1] symptoms that can occur with depression (e.g., unrefreshing sleep, joint pain, muscle pain, impairment in concentration). Fatigue and these four minor symptoms are also defining criteria for CFS, so it is possible that some patients with a primary affective disorder could be misdiagnosed as having CFS. Yet, these are distinct illnesses, as several CFS symptoms are not commonly found in depression, including prolonged fatigue after physical exertion, night sweats, sore throat, and swollen lymph nodes. Illness onset with CFS often occurs over a few hours or days, whereas primary depression generally shows a more gradual onset. Biological findings also differentiate the two conditions [34]. Including the latter type of patients in the current CFS case definition could confound the interpretation of epidemiologic and treatment studies, and complicate efforts to identify biological markers for this illness.

It is important for screening tests to have high sensitivity and specificity, particularly for disorders with low prevalence rates such as CFS (about 4.2 in a thousand) [31]. As an example, in a city of 1,000,000, with a true CFS rate of 4.2 per thousand, there would be 4,200 CFS cases. According to Bayes' theorem [35] if a diagnostic test had a 95% rate of sensitivity, the screening test would correctly identify 3,990 of these cases. However, if the test had 95% specificity, there would be 49,790

individuals who did not have CFS but were identified as having it using the test. Clearly, being able to identify true negatives with precision is of high importance in the diagnostic process.

We provide two case studies that illustrate several of the problems with the Reeves *et al.* [2] criteria. For example, one person who we diagnosed with CFS did not meet the Reeves *et al.* empirical case definition due to not meeting the frequency/severity requirement for the Symptom Inventory (SI). Yet, this person indicated that she had experienced a 95% decrease in daily activities over the past 6 months and an 80% decrease in daily energy level over the last 6 months. The person also reported having experienced 6 months of fatigue and more than 4 core symptoms. On a different scale from the medical questionnaire, using a 100 point scale, with higher scores indicating more problems, the person had a score of 80 on impaired memory and 85 on un-refreshing sleep. Our physician panel clearly felt that this person met all CFS Fukuda *et al.* [1] criteria, but the person was not included as a CFS case using the Reeves *et al.* criteria. In contrast, another person who we classified as ICF met the Reeves *et al.* empirical case definition. This person only had a 30% reduction in daily activity in the last 6 months and a 30% reduction of daily energy levels in the last 6 months. Our physician panel did not diagnose this participant has having CFS, yet the person was counted as a CFS case using the Reeves *et al.* criteria.

There are several limitations in this study. First, the community-based study of participants was relatively small. Clearly, these results need to be replicated by other investigators with larger samples. However, when the Reeves *et al.* [2] disability criteria were evaluated on a tertiary care setting [11], the findings also pointed to sensitivity and specificity problems. Another study using psychiatric controls also found the empirical case definition to be problematic due to specificity issues [7].

In summary, the scientific enterprise depends on reliable and valid ways of classifying patients into diagnostic categories, and this critical research activity can enable investigators to better understand etiology, pathophysiology, and treatment approaches for CFS and other disorders [36]. When diagnostic categories lack reliability and accuracy, the quality of treatment and clinical research can be significantly compromised. If CFS is to be diagnosed reliably across health care professionals, it is imperative to provide specific thresholds and scoring rules for the symptomatic criteria.

REFERENCES

[1]   K. Fukuda, S. E. Straus, I. Hickie, M. C. Sharpe, J. G. Donnibs and A. Komaroff, "The Chronic Fatigue Syndrome: A Comprehensive Approach to Its Definition and Study," *Annals of Internal Medicine*, Vol. 121, 1994, pp. 953-958.

[2] W. C. Reeves, D. Wagner, R. Nisenbaum, J. F. Jones, B. Gurbaxani, L. Solomon, *et al.*, "Chronic Fatigue Syndrome– A Clinically Empirical Approach to Its Definition and Study," *BMC Medicine*, Vol. 3, 2005. http://www.biomedcentral.com/1741-7015/3/19

[3] J. E. Ware, K. K. Snow and M. Kosinski, "SF-36 Health Survey: Manual and Interpretation Guide," *Quality Metric Incorporated*, 2000.

[4] E. M. Smets, B. J. Garssen, B. Bonke and J. C. DeHaes, "The Multidimensional Fatigue Inventory (MFI): Psychometric Properties of an Instrument to Assess Fatigue," *Journal of Psychosomatic Research*, Vol. 39, 1995, pp. 315-325.

[5] D. Wagner, R. Nisenbaum, C. Heim, J. F. Jones, E. R. Unger and W. C. Reeves, "Psychometric Properties of the CDC Symptom Inventory for Assessment of Chronic Fatigue Syndrome," *Population Health Metrics*, Vol. 3, 2005. http://www.pophealthmetrics.com/content/3/1/8

[6] L. A. Tiersky, R. J. Matheis, J. DeLuca, G. Lange and B. H. Natelson, "Functional Status, Neuropsychological Functioning, and Mood in Chronic Fatigue Syndrome," *Journal of Nervous and Mental Disease*, Vol. 191, 2003, pp. 324-331.

[7] L. A. Jason, N. Najar, N. Porter and C. Reh, "Evaluating the Centers for Disease Control's Empirical Chronic Fatigue Syndrome Case Definition," *Journal of Disability Policy Studies*, Vol. 20, 2009, pp. 93-100.

[8] L. A. Jason, M. Evans, M. Brown, N. Porter, A. Brown, J. Hunnell, *et al.*, "Fatigue Scales and Chronic Fatigue Syndrome: Issues of Sensitivity and Specificity," *Manuscript Submitted for Publication*, 2010.

[9] W. C. Reeves, B. M. Gurbaxani, J. S. Lin and E. R. Unger, "A Response to Jason *et al.*, 'Evaluating the Centers for Disease Control's Empirical Chronic Fatigue Syndrome Case Definition," *Journal of Disability Policy Studies Online First*, 2009. http://dps.sagepub.com/cgi/rapidpdf/1044207309346310v1.pdf

[10] C. Hawk, L. A. Jason and S. Torres-Harding, "Differential Diagnosis of Chronic Fatigue Syndrome and Major Depressive Disorder," *International Journal of Behavioral Medicine*, Vol. 13, 2006, pp. 244-251.

[11] L. A. Jason, M. Brown, M. Evans, V. Anderson, A. Lerch, A. Brown, *et al*., "Measuring Substantial Reduction in Functioning with CFS," *Manuscript Submitted for Publication*, 2010.

[12] L. A. Jason, K. M. Jordan, J. A. Richman, A. W. Rademaker, C. Huang, W. McCready, *et al.*, "A Community-Based Study of Prolonged and Chronic Fatigue," *Journal of Health Psychology*, Vol. 4, 1999, pp. 9-26.

[13] L. Kish, "Survey Sampling," *John Riley and Sons, Incorporated*, 1965.

[14] L. A. Jason, M. T. Ropacki, N. B. Santoro, J. A. Richman, W. Heatherly, R. Taylor, *et al.*, "A Screening Instrument for Chronic Fatigue Syndrome: Reliability and Validity," *Journal of Chronic Fatigue Syndrome*, Vol. 3, No. 1, 1997, pp. 39-59.

[15] R. L Spitzer, J. B. W. Williams, M. Gibbon and M. B. First, "Structured Clinical Interview for DSM-IV-Non-Patient Edition (SCID-NP, Version 2.0)," *American Psychiatric Press*, 1995.

[16] "Diagnostic and Statistical Manual of Mental Disorders," Forth Edition, *American Psychiatric Association*, 1994.

[17] M. B. First, R. L. Spitzer, M. Gibbon and J. B. W. Williams, "Structured Clinical Interview for DSM-IV Axis I Disorders-Patient Edition," *Biometrics Research Department*, 1995.

[18] R. R. Taylor and L. A. Jason, "Comparing the DIS with the SCID: Chronic Fatigue Syndrome and Psychiatric Comorbidity," *Psychology and Health: The International Review of Health Psychology*, Vol. 13, 1998, pp. 1087-1104.

[19] A. L. Komaroff and D. S. Buchwald, "Symptoms and Signs of Chronic Fatigue Syndrome," *Reviews of Infectious Diseases*, Vol. 13, 1991, pp. 8-11.

[20] A. L. Komaroff, L. R. Fagioli, A. M. Geiger, T. H. Doolittle, J. Lee, R. J. Kornish, *et al.*, "An Examination of the Working Case Definition of Chronic Fatigue Syndrome," *The American Journal of Medicine*, Vol. 100, 1996, pp. 56-64.

[21] D. Buchwald, R. Pascualy, C. Bombardier and P. Kith, "Sleep Disorders in Patients with Chronic Fatigue," *Clinical Infectious Diseases*, Vol. 18, 1994, pp. S68-S72.

[22] A. L. Stewart, S. Greenfield, R. D. Hayes, K. Wells, W. H. Rogers, S. D. Berry, *et al.*, "Functional Status and Well-Being of Patients with Chronic Conditions: Results from the Medical Outcomes Study," *Journal of the American Medical Association*, Vol. 262, No. 7, 1989, pp. 907-913.

[23] P. Goodnick and R. Sandoval, "Treatment of Chronic Fatigue Syndrome and Related Disorders: Immunological Approaches," In P. Goodnick and N. Klimas Eds., Chronic Fatigue and Related Immune Deficiency Syndromes, *American Psychiatric Press*, 1993, pp. 131-161.

[24] L. A. Jason, N. Porter, J. Hunnell, A. Rademaker and J. Richman, "CFS Prevalence and Risk Factors Over Time," *Manuscript Submitted for Publication*, 2010.

[25] W. C. Reeves, A. Lloyd, S. D. Vernon, N. Klimas, L. A. Jason, G. Bleijenberg, *et al.*, "Identification of Ambiguities in the 1994 Chronic Fatigue Syndrome Research Case Definition and Recommendations for Resolution," *BMC Health Services Research*, Vol. 3, 2003. http://www.biomedcentral.com/content/pdf/1472-6963-3-25.pdf

[26] J. A. Hanley and B. J. McNeil, "The Meaning and Use of the Area under a Receiver Operating Characteristics (ROC) Curve," *Radiology*, Vol. 143, 1982, pp. 29-36.

[27] K. H. Zou, A. J. O'Malley and L. Mauri, "Receiver-Operating Characteristic Analysis for Evaluating Diagnostic Tests and Predictive Models," *Circulation*, Vol. 115, 2007, pp. 654-657.

[28] M. H. Zweig and G. Campbell, "Receiver-Operating Characteristic (ROC) Plots: A Fundamental Evaluation Tool in Clinical Medicine," *Clinical Chemistry*, Vol. 39, 1993, pp. 561-577.

[29] W. C. Reeves, J. J. Jones, E. Maloney, C. Heim, D. C. Hoaglin, R. Boneva, *et al.*, "New Study on the Prevalence of CFS in Metro, Urban and Rural Georgia Populations," *Population Health Metrics*, Vol. 5, 2007. http://www.pophealthmetrics.com/content/5/1/5

[30] M. Reyes, R. Nisenbaum, D. C. Hoaglin, E. R. Unger, C. Emmons, B. Randall, *et al.*, "Prevalence and Incidence of Chronic Fatigue Syndrome in Wichita, Kansas," *Archives of Internal Medicine*, Vol. 163, 2003, pp. 1530-1536.

[31] L. A. Jason, J. A. Richman, A. W. Rademaker, K. M. Jordan, A. V. Plioplys, R. Taylor, *et al.*, "A Community-Based Study of Chronic Fatigue Syndrome," *Archives of Internal Medicine*, Vol. 159, 1999, pp. 2129-2137.

[32] L. A. Jason, M. Evans, M. Brown and N. Porter, "What is Fatigue? Pathological and Non-Pathological Fatigue," *Manuscript Submitted for Publication*, 2010.

[33] D. A. Regier, J. H. Boyd and J. D. Burke Jr., "One Month Prevalence of Mental Disorders in the United States: Based on Five Epidemiological Catchment Rea Sites," *Archives of General Psychiatry*, Vol. 45, 1988, pp. 977-986.

[34] L. A. Jason, K. Corradi, S. Torres-Harding, R. R. Taylor and C. King, "Chronic Fatigue Syndrome: The Need For Subtypes," *Neuropsychology Review*, Vol. 15, 2005, pp. 29-58.

[35] E. T. T. Jaynes, "Probability Theory: The Logic of Science," G. L. Bretthorst Ed., *Cambridge University Press*, 2003.

[36] C. King and L. A. Jason, "Improving the Diagnostic Criteria and Procedures for Chronic Fatigue Syndrome," *Biological Psychology*, Vol. 68, 2005, pp. 87-106.