REVIEW

Annals of Internal Medicine

Diagnostic Methods for Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Systematic Review for a National Institutes of Health Pathways to Prevention Workshop

Elizabeth Haney, MD; M.E. Beth Smith, DO; Marian McDonagh, PharmD; Miranda Pappas, MA; Monica Daeges, BA; Ngoc Wasson, MPH; and Heidi D. Nelson, MD, MPH

Background: The diagnosis of myalgic encephalomyelitis (ME)/chronic fatigue syndrome (CFS) is based on clinical criteria, yet there has been no consensus regarding which set of criteria best identifies patients with the condition. The Institute of Medicine has recently proposed a new case definition and diagnostic algorithm.

Purpose: To review methods to diagnose ME/CFS in adults and identify research gaps and needs for future research.

Data Sources: MEDLINE, PsycINFO, and Cochrane databases (January 1988 to September 2014); clinical trial registries; and reference lists.

Study Selection: English-language studies describing methods of diagnosis of ME/CFS and their accuracy.

Data Extraction: Data on participants, study design, analysis, follow-up, and results were extracted and confirmed. Study quality was dual-rated by using prespecified criteria, and discrepancies were resolved through consensus.

Data Synthesis: Forty-four studies met inclusion criteria. Eight case definitions have been used to define ME/CFS; a ninth, recently proposed by the Institute of Medicine, includes principal elements of previous definitions. Patients meeting criteria for ME represent a more symptomatic subset of the broader ME/CFS population. Scales rating self-reported symptoms differentiate patients with ME/CFS from healthy controls under study conditions but have not been evaluated in clinically undiagnosed patients to determine validity and generalizability.

Limitations: Studies were heterogeneous and were limited by size, number, applicability, and methodological quality. Most methods were tested in highly selected patient populations.

Conclusion: Nine sets of clinical criteria are available to define ME/CFS, yet none of the current diagnostic methods have been adequately tested to identify patients with ME/CFS when diagnostic uncertainty exists. More definitive studies in broader populations are needed to address these research gaps.

Primary Funding Source: Agency for Healthcare Research and Quality. (PROSPERO: CRD42014009779)

Ann Intern Med. 2015;162:834-840. doi:10.7326/M15-0443 **www.annals.org**
For author affiliations, see end of text.

The terms *myalgic encephalomyelitis* (ME) and *chronic fatigue syndrome* (CFS) have been used to describe a debilitating multisystemic condition characterized by chronic, disabling fatigue and various other symptoms. The term *CFS* was introduced in the 1980s after research failed to identify a clear viral association with what was previously labeled chronic Epstein–Barr virus syndrome (1–4). Other terms, such as *postviral fatigue syndrome* and *chronic fatigue immune dysfunction syndrome*, were also used in attempts to associate the syndrome with possible underlying causes (1, 2, 5, 6). Although the most recent international consensus report advocates moving away from the term *CFS* in favor of the term *ME* to better reflect an underlying disease process involving widespread inflammation and neuropathology (7, 8), experts do not agree about these mechanisms and the cause of CFS remains unclear.

A recent Institute of Medicine (IOM) report proposes a name, *systemic exertion intolerance disease* (*SEID*), that describes the central elements of the disease. The report focuses on the adverse effect that physical, cognitive, or emotional exertion can have on patients with this condition and acknowledges that this is a complex and severe disorder for which specific causes are not yet proven (9).

The diagnosis of ME/CFS is based on clinical criteria that attempt to distinguish it from other conditions that also present with fatigue. Eight published case definitions have been used since the first one established by the Centers for Disease Control and Prevention (CDC) in 1988 (2), and the IOM proposed a ninth in February 2015 (9). All include persistent fatigue not attributable to a known underlying medical condition, as well as additional clinical signs and symptoms that do not all need to be present to establish the diagnosis (10). However, there has been no consensus about which, if any, of these clinical criteria should be considered the reference standard. The variations in case definitions imply that they may describe different conditions and lead to different diagnoses, complicating ME/CFS research and clinical care. For example, depending on the case definition, prevalence rates of ME/CFS in the United States range from 0.3% to 2.5% (1, 11, 12).

This systematic review is part of a larger report to inform a research agenda for the National Institutes of

See also:

Related articles 841, 860
Editorial comment 871

Web-Only
CME quiz

Health (NIH) 2014 Pathways to Prevention Workshop, an evidence-based methodology workshop (13). The purpose of this systematic review was to evaluate and compare studies of methods to diagnose ME/CFS, identify limitations of current studies, and determine needs for future research.

METHODS

Key questions guiding this review were developed in collaboration with the NIH ME/CFS Working Group following a standard protocol, including input from key informants and a technical expert panel, registration in the PROSPERO database for systematic reviews (14), and posting on an Agency for Healthcare Research and Quality (AHRQ) public Web site. Key questions concerned describing clinical methods for diagnosing ME/CFS and evaluating their concordance and accuracy, describing variations in diagnostic methods by patient subgroups, and identifying consequences of diagnosis for patients. This article focuses on the published case definitions and on the concordance and accuracy of methods for diagnosis of ME/CFS. A technical report details the methods and includes an analytic framework, search strategies, and additional evidence tables (13).

Data Sources and Searches

A research librarian searched electronic databases to identify relevant articles published between January 1988 (year of the first case definition) and September 2014: MEDLINE (Ovid), PsycINFO, the Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, and the National Health Sciences Economic Evaluation. Searches were supplemented by references identified from additional sources, including reference lists and experts.

Study Selection

English-language studies of adults with ME/CFS as defined by any of the established case definitions, and those for whom ME/CFS was a diagnostic consideration, were eligible for inclusion. For this review, we use the combined term "ME/CFS" when referring to the condition in general, and we use the individual terms to represent study populations fulfilling specific sets of clinical criteria defined as ME or CFS. Studies of diagnostic tests or case definitions were included if they were conducted in clinical settings or settings applicable to clinical practice settings; we excluded studies of inpatients or institutionalized individuals. We also excluded studies of disease cause and studies that reported the diagnosis of specific symptoms of ME/CFS (for example, postexertional malaise).

We included studies that 1) compared case definitions (for example, Fukuda/CDC, Canadian, International) and provided measures of agreement or 2) tested the ability of the method to identify patients with ME/CFS by using 1 of the case definitions as a reference standard and reported at least 1 of the specified outcomes. Because there is no single accepted definition for ME/CFS and therefore no "gold standard," any of the case definitions published since 1988 were accepted as reference standards. Included outcomes of diagnostic accuracy were sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, negative likelihood ratio, c-statistic, receiver-operating characteristic curve and area under the receiver-operating characteristic curve, net reclassification index, and concordance. Studies of any design were included if they described potential harms from diagnosis, such as psychological harms, labeling, risk from diagnostic tests, and misdiagnosis. These studies are included in the full report (13).

Two investigators independently evaluated each study to determine inclusion eligibility. Disagreement was resolved by consensus, with a third investigator making the final decision as needed.

Data Extraction and Quality Assessment

An investigator abstracted details of the patient population, study design, setting, inclusion and exclusion criteria, population characteristics, sample size, case definition for diagnosis, and results. A second investigator reviewed extracted data for accuracy and completeness. Investigators rated the quality (risk of bias) of the individual studies on the basis of criteria adapted from the Agency for Healthcare Research and Quality (AHRQ) *Methods Guide for Medical Test Reviews* (15). A second investigator reviewed ratings, and disagreements were resolved by consensus with a third investigator as needed. Quality and strength of evidence ratings were assessed for all studies of diagnostic test accuracy (comparison of a diagnostic test to a reference standard) but could not be assessed for other studies with descriptive, cross-sectional, and case series designs.

Data Synthesis and Analysis

Studies of diagnostic tests could not be combined in a quantitative meta-analysis because of heterogeneity of patient populations, study designs, reported outcomes, and reference standards. Therefore, data were synthesized qualitatively with attention to such factors as patient characteristics and risk of bias.

Role of the Funding Source

The AHRQ funded the review, and a working group convened by the NIH helped develop the review's scope and key questions. Neither had a role in study selection, quality assessment, or synthesis. The investigators are solely responsible for the content.

RESULTS

Among the 6175 abstracts identified by searches and additional papers identified through other sources, 44 studies met inclusion criteria (**Appendix Figure**, available at www.annals.org). These included 8 studies describing case definitions (**Table**) (2, 5–7, 16–19), 22 evaluating diagnostic tests (**Appendix Tables 1** and **2**, available at www.annals.org) (1, 10, 20–39), and 14 describing consequences of diagnosis (**Appendix Table 3**, available at www.annals.org) (11, 40–52). The new IOM

Table. Comparisons of Symptoms Using Different Case Definitions

Symptoms	SEID	ME With or Without CFS				CFS			
	IOM (SEID), 2015 (9)	London (ME): Dowsett et al, 1994 (16)	Canadian (ME/CFS): Carruthers et al, 2003 (5)	Revised Canadian (ME/CFS): Jason et al, 2010 (17)	International (ME): Carruthers et al, 2011 (7)	CDC: Holmes et al, 1988 (2)	Oxford: Sharpe et al, 1991 (19)	CDC: Fukuda et al, 1994 (6)	CDC: Reeves et al, 2005* (18)
General physical									
Fatigue	X†‡	X	≥6 mo	≥6 mo		≥6 mo	≥6 mo or >50%	≥6 mo	≥6 mo
Sudden or new onset	X†‡		X§			X‖	X		X¶
Impairment of daily function	≥6 mo X†‡					≥50%	X		
Neurologic/neurocognitive									
Muscle weakness				X		X‖			
Muscle pain			X		X**	X‖		X¶	X¶
Postexertional malaise	X†	X	X	X	X	X‖		X¶	X¶
New headaches			X			X‖		X¶	X¶
Arthralgias (migratory)			X	X	X**	X‖		X¶	X¶
Sleep disturbances	X†	X††	X	X	X**	X‖		X¶	X¶
Neurologic/ neuropsychiatric		X††	X‡‡		X**	X‖			
Memory or cognitive	X§§	X††	X‡‡	X‖‖	X**		X	X¶	
Dysequilibrium		X							
Temperature dysregulation					X¶¶				
Neuroendocrine/immune									
Autonomic dysfunction			X***	X***					
Fever or chills						X‖			
Sore throat						X‖		X¶	X¶
Lymph node pain						X‖		X¶	X¶
Neuroendocrine dysfunction			X***	X***	X†††				
Immune manifestations			X***	X***					
Impairment of other systems									
Cardiovascular	X†/‡‡‡				X¶¶				
Pulmonary					X¶¶				
Gastrointestinal					X†††				
Genitourinary					X†††				

CDC = Centers for Disease Control and Prevention; CFS = chronic fatigue syndrome; IOM = Institute of Medicine; ME = myalgic encephalomyelitis; SEID = systemic exertion intolerance disease.
* Defined functional impairment by Short-Form Health Survey (SF-36) scores, fatigue by Multidimensional Fatigue Inventory (MFI), and symptoms by Symptom Inventory Case Definition subscale.
† All 3 required.
‡ Impairment in function, fatigue, and new onset included as 1 of the 3 required symptoms.
§ Onset may be gradual.
‖ Eight of 11 minor symptoms.
¶ Four or more symptoms present concurrently for ≥6 months.
** At least 1 from 3 of the 4 symptom categories (neurocognitive, pain, sleep, neurosensory/motor).
†† At least 1 of 3 symptoms.
‡‡ At least 2 neurologic/cognitive manifestations.
§§ At least 1 of 2 required.
‖‖ At least 2 cognitive manifestations.
¶¶ At least 1 energy production/transportation impairment (cardiovascular, pulmonary, thermostatic, temperature).
*** At least 1 symptom from 2 of the categories of autonomic, neuroendocrine, and immune manifestations.
††† At least 1 symptom for ≥3 categories of immune, gastrointestinal, and genitourinary impairments.
‡‡‡ Orthostatic intolerance.

case definition was also included for completeness (bringing the total to 9 available case definitions), even though it was published after the literature search dates.

Methods for Diagnosing ME/CFS

Nine case definitions using clinical criteria have been developed to identify patients with ME/CFS and help clinicians distinguish ME/CFS from other conditions that present with fatigue (Table) (2, 5–7, 9, 16–19). Although most case definitions require that other conditions be excluded before ME/CFS is diagnosed, no studies compared strategies for ruling out alternative diagnoses or specifically defined which conditions should be ruled out. The IOM case definition, published in February 2015, incorporates required elements of fatigue, postexertional malaise, and sleep disturbance, along with cognitive impairment or orthostatic hypotension (9). The Oxford case definition incorporates the fewest symptoms (new onset of fatigue with impairment of physical and mental function), suggesting that it includes patients who would not meet other criteria for ME/CFS (19).

Concordance of Methods for Diagnosing ME/CFS

Seven studies compared symptoms of patients with ME/CFS diagnosed by using different case definitions and found that symptoms varied depending on the clinical criteria used (Appendix Table 1) (1, 10, 20–24). In general, populations defined by ME or ME/CFS criteria had more severe symptoms or more functional impairment than those defined by CFS criteria alone (1, 10, 20–24).

Three studies enrolling a total of 6087 patients compared symptoms of patients with CFS identified by the 1994 CDC criteria with symptoms of patients without CFS (healthy controls; other fatigued patients; and patients with psychiatric, rheumatologic, and other chronic diseases) (25–27). In general, patients without CFS were less impaired than those with CFS, although results varied. In 1 study, patients with CFS and multiple sclerosis had similar scores on the 36-item Short-Form Survey (SF-36) on physical function, vitality, and social function scales (27).

Accuracy of Measures for Diagnosing ME/CFS

Nine studies evaluated methods to discriminate ME/CFS from other conditions by using 1 of the published case definitions as a reference standard (Appendix Tables 2 and 4, available at www.annals.org) (29–37). One study met criteria for good quality (30), 7 for fair quality (31–37), and 1 for poor quality (29). Several studies used the same or very similar study populations to report different outcomes, most commonly recruiting from CFS self-help groups (34–36) or community samples (32, 33). Major limitations of studies included small size (<50 participants) (29, 34–36), recruitment from specialty clinics only (30), lack of blinding to the reference standard result (29–36), and comparing cases with primarily healthy or nonfatigued controls (29, 31, 33–36).

By using computerized modeling to identify key symptoms, 3 studies found that symptom-based instruments had high sensitivity and specificity for identifying patients who meet 1 of the ME/CFS case definitions (Appendix Table 2) compared with healthy controls (30, 31, 37).

Another study randomly assigned a broad spectrum of 198 participants with fatigue (including patients with systemic lupus erythematosus, fibromyalgia, and CFS defined by Oxford criteria) to derivation or validation cohorts (30). Participants completed symptom questionnaires, and the symptoms with the highest sensitivity and specificity for CFS were selected to develop and evaluate computer-generated classification criteria to distinguish patients with CFS from the other patients. Four methods of classification were tested in the derivation cohort, and for each algorithm, the sensitivity, specificity, and accuracy were determined in the validation cohort. A strategy that included 24 symptoms, the artificial neural network, had good discriminative ability (sensitivity, 0.95; specificity, 0.85; accuracy, 0.90) (30). This study met criteria for good quality because it included a broad spectrum of patients with conditions considered to be competing diagnoses for ME/CFS and included a validation cohort.

An evaluation of responses to the DePaul Symptom Questionnaire from 515 patients with CFS and 176 controls used K-means clustering to distinguish patients with fewer symptoms from those with more symptoms, who presumably had CFS (37). After testing of 4 methods of clustering, the unsupervised thresholding model was used to assign a diagnostic label to each participant, and the diagnosis assigned by each of 3 different clinical criteria (1994 CDC [CFS], Canadian [ME/CFS], and 2011 International [ME]) was compared with the assigned diagnostic label. Then, the individual symptoms were ranked by predictive value and compared with the 3 case definitions and the use of all 54 DePaul Symptom Questionnaire symptoms.

Results indicated that model accuracy obtained by using the top 11 ranked symptoms was better than that obtained with all 54 DePaul Symptom Questionnaire symptoms or the 1994 CDC (CFS), Canadian (ME/CFS), and 2011 International (ME) criteria (90.2%, 82.3%, 83.8%, 84.1%, and 78.7%, respectively). The top-ranked symptoms corresponded to fatigue, exertional malaise, sleep disturbance, cognitive impairment, and myalgias. This study met criteria for fair quality because it lacked a validation group, but it included a relatively large, broad spectrum of participants.

In another study of 368 patients and 452 controls, the Schedule of Fatigue and Anergia for CFS Scale was developed by using a composite set of criteria as a reference standard and specific symptoms from 4 symptom checklists (31). Latent class analysis was used to select 10 symptoms having the highest correlation to CFS-like fatigue; then, a composite score was tested to determine sensitivity and specificity. The 10 symptoms included fatigue, exertional malaise, myalgias, cognitive difficulties (including poor concentration, poor memory, speech difficulties), poor sleep, and headaches. A total score of 3 to 4 out of 4 had a sensitivity of 81% for the 3-class solution and a specificity of 100%. This study met criteria for fair quality because patients were recruited from specialty clinics rather than from a broader population and because it lacked a validation cohort.

Variation in Diagnostic Testing According to Subgroups

Three studies evaluated diagnostic tests in subgroups of patients with ME/CFS (28, 38, 39). Compared with patients younger than 25 years, patients older than 50 were more impaired, had lower self-efficacy, and had worse scores on the Fatigue Impact Scale, Chalder Fatigue Scale, Hospital Anxiety and Depression Scale–Depression subscale, and SF-36 (28). Likewise, older patients had lower resting heart rates, higher left ventricular ejection time, and lower baroreflex sensitivity (ability to maintain blood pressure) than younger patients.

Two studies of the same population evaluated the ability of self-reported function scales to predict recovery from cardiopulmonary exercise testing in patients

with CFS defined by 1994 CDC (CFS) criteria and nondisabled sedentary controls (38, 39). The SF-36 subscales of physical function, role-physical, bodily pain, general health, vitality, and social functioning identified patients with failure to recover at 1 day; the subscales role-emotional, vitality, and bodily pain identified those with failure to recover at 1 week (38). Having 3 or more symptoms of postexertional malaise optimally distinguished between patients with CFS and controls (39).

DISCUSSION

Of the 8 previously published sets of clinical criteria for ME and/or CFS, case definitions for ME and ME/CFS identify patients with more impairment, lower functioning, and more severe symptoms than the CFS-alone case definitions. The new IOM case definition incorporates principal elements of previous definitions, and the association of these elements to ME/CFS is supported by modeling studies (31, 37). None of the case definitions or other diagnostic methods has been adequately tested to determine how well they differentiate patients with ME/CFS from patients with other conditions. Although some symptom-based instruments discriminate patients with ME/CFS from healthy controls, their utility in differentiating patients with diagnostic uncertainty remains inconclusive because they have not been widely tested in broad spectrums of patients. The few studies that evaluated how diagnostic tests vary by patient subgroups were inconclusive.

The clinical applicability of current research on diagnostic methods for ME/CFS is limited in several ways. All case definitions require the exclusion of competing diagnoses before assigning a ME/CFS diagnosis, yet no studies evaluated strategies for the evaluation and assignment of alternative diagnoses. In addition, most studies were designed as descriptive studies and enrolled healthy or nonfatigued participants as controls. Studies evaluated whether tests distinguished ME/CFS from these types of controls, but not the essential clinical question of whether the test could distinguish ME/CFS from other fatiguing illnesses. Only 1 study included participants with overlapping symptoms and tested a strategy for diagnosis in both a derivation and a validation cohort; and only 2 studies evaluated a diagnostic test by using control groups of fatigued or other chronically ill patients (30). In addition, studies used varying case definitions as the reference standard precluding comparisons across studies. Finally, many studies recruited participants from specialty clinics, potentially reflecting more severe disease, or site-dependent or local practices, limiting generalizability to other patients with ME/CFS. Consistent with a prior systematic review (53), no studies identified specific patients with identifiable causes.

Future research should be based on a standard case definition, or a set of reference standards, to allow comparison of results across studies. The IOM has provided a consensus case definition that could serve this purpose. Consensus groups and researchers should consider retiring the Oxford case definition because it differs from the other case definitions and is the least restrictive, probably including individuals with other overlapping conditions. The new IOM case definition and algorithm provide a starting place for future studies of diagnostic testing.

Future studies evaluating the capability of diagnostic methods for ME/CFS should include a broad range of patients with conditions that require clinical distinction from ME/CFS, such as fibromyalgia and depression. Moreover, studies should report how well a particular method distinguishes ME/CFS from other conditions by using standard performance measures, such as concordance, sensitivity, and specificity. Studies should report findings according to important features of ME/CFS, such as postexertional malaise, neurocognitive status, and autonomic function, to identify subgroups that may respond differently to specific treatments. Collaborative groups could consider establishing an international ME/CFS registry that would track the natural history of patients to determine which set of clinical criteria best identifies patients for whom no alternative diagnosis will be found with subsequent testing, and for whom the diagnosis of ME/CFS will continue to be appropriate over time. Given the devastating effect of this condition on patients and families, researchers should involve patients and advocates in trial planning and development so that future research is relevant and meaningful to those affected by ME/CFS.

In conclusion, 9 sets of clinical criteria are used to define ME/CFS, yet none of the current diagnostic methods have been adequately tested to identify patients with ME/CFS when diagnostic uncertainty exists. More definitive studies in broader populations are needed to address these research gaps.

From Oregon Health & Science University and Providence Cancer Center, Providence Health and Services Oregon, Portland, Oregon.

Acknowledgment: The authors thank the following individuals for their contributions to this project: Richard Bryant, MD, for providing expert consultation throughout the report; Andrew Hamilton, MLS, MS, for conducting literature searches; and Spencer Dandy, BS, for assistance with preparing this report (all are located at the Oregon Health & Science University). They also thank Suchitra Iyer, PhD, Task Order Officer at the Agency for Healthcare Research and Quality; Carmen Green, MD, National Institutes of Health (NIH) Working Group Chair; the NIH; the Technical Expert Panel; and reviewers of the draft report.

Financial Support: By the Agency for Healthcare Research and Quality (contract 290-2012-00014-i, task order 4), Rockville, Maryland.

Disclosures: Dr. Haney reports grants from the Agency for Healthcare Research and Quality during the conduct of the study. Ms. Daeges reports grants from the Agency for Healthcare Research and Quality during the conduct of this study. Authors not named here have disclosed no conflicts of interest. Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M15-0443.

Requests for Single Reprints: Elizabeth Haney, MD, 3181 SW Sam Jackson Park Road, Mail Code BICC, Portland, OR 97239; e-mail, haneye@ohsu.edu.

Current author addresses and author contributions are available at www.annals.org.

References

1. **Jason LA, Brown A, Clyne E, Bartgis L, Evans M, Brown M.** Contrasting case definitions for chronic fatigue syndrome, myalgic encephalomyelitis/chronic fatigue syndrome and myalgic encephalomyelitis. Eval Health Prof. 2012;35:280-304. [PMID: 22158691] doi: 10.1177/0163278711424281

2. **Holmes GP, Kaplan JE, Gantz NM, Komaroff AL, Schonberger LB, Straus SE, et al.** Chronic fatigue syndrome: a working case definition. Ann Intern Med. 1988;108:387-9. [PMID: 2829679]

3. **Ramsay M.** Myalgic Encephalomyelitis and Postviral Fatigue States: The Saga of Royal Free Disease. London: Gower Medical Publishing; 1986.

4. A new clinical entity? Lancet. 1956;270:789-90. [PMID: 13320887]

5. **Carruthers BM, Jain AK, de Meirleir KL, Peterson DL, Klimas NG, Lerner A, et al.** Myalgic encephalomyelitis/chronic fatigue syndrome: clinical working case definition, diagnostic and treatment protocols. J Chronic Fatigue Syndr. 2003;11:7-115.

6. **Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, Komaroff A.** The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. Ann Intern Med. 1994;121:953-9. [PMID: 7978722]

7. **Carruthers BM, van de Sande MI, De Meirleir KL, Klimas NG, Broderick G, Mitchell T, et al.** Myalgic encephalomyelitis: International Consensus Criteria. J Intern Med. 2011;270:327-38. [PMID: 21777306] doi:10.1111/j.1365-2796.2011.02428.x

8. **Carruthers BM, van de Sande MI, De Meirleir K, Klimas N, Broderick G, Mitchell T, et al.** Myalgic encephalomyelitis–adult & pediatric: international consensus primer for medical practitioners. 2012. Accessed at http://sacfs.asn.au/download/me_international_consensus_primer_for_medical_practitioners.pdf on 11 February 2015.

9.Institute . **Institute of Medicine.** Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome Redefining an Illness. Washington, DC: National Academies Pr; 2015.

10. **Jason LA, Brown A, Evans M, Sunnquist M, Newton JL.** Contrasting chronic fatigue syndrome versus myalgic encephalomyelitis/chronic fatigue syndrome. Fatigue. 2013;1:168-183. [PMID: 23914329]

11. **Reyes M, Nisenbaum R, Hoaglin DC, Unger ER, Emmons C, Randall B, et al.** Prevalence and incidence of chronic fatigue syndrome in Wichita, Kansas. Arch Intern Med. 2003;163:1530-6. [PMID: 12860574]

12. **Reeves WC, Jones JF, Maloney E, Heim C, Hoaglin DC, Boneva RS, et al.** Prevalence of chronic fatigue syndrome in metropolitan, urban, and rural Georgia. Popul Health Metr. 2007;5:5. [PMID: 17559660]

13. **Smith M, Nelson H, Haney E, Pappas M, Daeges M, Wasson N, et al.** Diagnosis and treatment of myalgic encephalomyelitis/chronic fatigue syndrome. Evidence Report/Technology Assessment No. 219. (Prepared by the Pacific Northwest Evidence-based Practice Center under contract 290-2012-00014-I.) AHRQ Publication No. 15-E001-EF. Rockville, MD: Agency for Healthcare Research and Quality; December 2014. Accessed at http://effectivehealthcare.ahrq.gov/ehc/products/586/2005/chronic-fatigue-executive-141211.pdf on 11 February 2015.

14. **Smith MEB, Nelson N, Haney E, McDonagh M, Pappas M, Daeges M, et al.** Diagnosis and treatment of myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). Accessed at www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014009779 on 11 Februrary 2015.

15. **Agency for Healthcare Research and Quality.** Methods Guide for Medical Test Reviews. AHRQ Publication No. 12-EC017. Rockville, MD: Agency for Healthcare Research and Quality; June 2012. Accessed at www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=558 on 11 February 2015.

16. **Dowsett E, Goudsmit E, Macintyre A, Shepherd C.** London Criteria for myalgic encephalomyelitis. In: Report from The National Task Force on Chronic Fatigue Syndrome (CFS), Post Viral Fatigue Syndrome (PVFS), Myalgic Encephalomyelitis (ME). Westcare. 1994. 96-98. Accessed at: www.actionforme.org.uk/Resources/Action%20for%20ME/Documents/get-informed/national%20task%20force.pdf on 26 March 2015

17. **Jason L, Evans M, Porter N, Brown M, Brown A, Hunnell J, et al.** The development of a revised Canadian myalgic encephalomyelitis chronic fatigue syndrome case definition. Am J Biochem Biotechnol. 2010;6:120-35.

18. **Reeves WC, Wagner D, Nisenbaum R, Jones JF, Gurbaxani B, Solomon L, et al.** Chronic fatigue syndrome–a clinically empirical approach to its definition and study. BMC Med. 2005;3:19. [PMID: 16356178]

19. **Sharpe MC, Archard LC, Banatvala JE, Borysiewicz LK, Clare AW, David A, et al.** A report–chronic fatigue syndrome: guidelines for research. J R Soc Med. 1991;84:118-21. [PMID: 1999813]

20. **Jason LA, Torres-Harding SR, Taylor RR, Carrico AW.** A comparison of the 1988 and 1994 diagnostic criteria for chronic fatigue syndrome. J Clin Psychol Med Settings. 2001;8:337-43.

21. **Brown AA, Jason LA, Evans MA, Flores S.** Contrasting Case Definitions: The ME International Consensus Criteria vs. the Fukuda et al. CFS Criteria. N Am J Psychol. 2013;15:103-120. [PMID: 25364305]

22. **Van Hoof E, De Meirleir K.** chronic fatigue syndrome and myalgic encephalomyletis: are both conditions on the same continuum? N Am J Psychol. 2005;7:189-204.

23. **Jason LA, Sunnquist M, Brown A, Evans M, Newton JL.** Are Myalgic Encephalomyelitis and chronic fatigue syndrome different illnesses? A preliminary analysis. J Health Psychol. 2014. [PMID: 24510231]

24. **Jason LA, Torres-Harding SR, Jurgens A, Helgerson J.** Comparing the Fukuda et al. criteria and the Canadian case definition for chronic fatigue syndrome. J Chronic Fatigue Syndr. 2004;12:37-52.

25. **Aslakson E, Vollmer-Conna U, White PD.** The validity of an empirical delineation of heterogeneity in chronic unexplained fatigue. Pharmacogenomics. 2006;7:365-73. [PMID: 16610947]

26. **Katon WJ, Buchwald DS, Simon GE, Russo JE, Mease PJ.** Psychiatric illness in patients with chronic fatigue and those with rheumatoid arthritis. J Gen Intern Med. 1991;6:277-85. [PMID: 1890495]

27. **Komaroff AL, Fagioli LR, Doolittle TH, Gandek B, Gleit MA, Guerriero RT, et al.** Health status in patients with chronic fatigue syndrome and in general population and disease comparison groups. Am J Med. 1996;101:281-90. [PMID: 8873490]

28. **Lewis I, Pairman J, Spickett G, Newton JL.** Is chronic fatigue syndrome in older patients a different disease? – a clinical cohort study. Eur J Clin Invest. 2013;43:302-8. [PMID: 23397955] doi:10.1111/eci.12046

29. **Tiev KP, Demettre E, Ercolano P, Bastide L, Lebleu B, Cabane J.** RNase L levels in peripheral blood mononuclear cells: 37-kilodalton/83-kilodalton isoform ratio is a potential test for chronic fatigue syndrome. Clin Diagn Lab Immunol. 2003;10:315-6. [PMID: 12626460]

30. **Linder R, Dinser R, Wagner M, Krueger GR, Hoffmann A.** Generation of classification criteria for chronic fatigue syndrome using an artificial neural network and traditional criteria set. In Vivo. 2002;16:37-43. [PMID: 11980359]

31. **Hadzi-Pavlovic D, Hickie IB, Wilson AJ, Davenport TA, Lloyd AR, Wakefield D.** Screening for prolonged fatigue syndromes: validation of the SOFA scale. Soc Psychiatry Psychiatr Epidemiol. 2000;35:471-9. [PMID: 11127722]
32. **Jason LA, Evans M, Brown A, Brown M, Porter N, Hunnell J, et al.** Sensitivity and specificity of the CDC empirical chronic fatigue syndrome case definition. Psychology. 2010;1:9-16.
33. **Jason L, Brown M, Evans M, Anderson V, Lerch A, Brown A, et al.** Measuring substantial reductions in functioning in patients with chronic fatigue syndrome. Disabil Rehabil. 2011;33:589-98. [PMID: 20617920] doi:10.3109/09638288.2010.503256
34. **Gaab J, Engert V, Heitz V, Schad T, Schürmeyer TH, Ehlert U.** Associations between neuroendocrine responses to the Insulin Tolerance Test and patient characteristics in chronic fatigue syndrome. J Psychosom Res. 2004;56:419-24. [PMID: 15094026]
35. **Gaab J, Hüster D, Peisen R, Engert V, Schad T, Schürmeyer TH, et al.** Low-dose dexamethasone suppression test in chronic fatigue syndrome and health. Psychosom Med. 2002;64:311-8. [PMID: 11914448]
36. **Gaab J, Rohleder N, Heitz V, Engert V, Schad T, Schürmeyer TH, et al.** Stress-induced changes in LPS-induced pro-inflammatory cytokine production in chronic fatigue syndrome. Psychoneuroendocrinology. 2005;30:188-98. [PMID: 15471616]
37. **Watson SP, Ruskin AS, Simonis V, Jason LA, Sunnquist M, Furst JD.** Identifying defining aspects of chronic fatigue syndrome via unsupervised machine learning and feature selection. International Journal of Machine Learning & Computing. 2014;4:133-8.
38. **Davenport TE, Stevens SR, Baroni K, Van Ness JM, Snell CR.** Reliability and validity of Short Form 36 Version 2 to measure health perceptions in a sub-group of individuals with fatigue. Disabil Rehabil. 2011;33:2596-604. [PMID: 21682669] doi:10.3109/09638288.2011.582925
39. **Davenport TE, Stevens SR, Baroni K, Van Ness M, Snell CR.** Diagnostic accuracy of symptoms characterising chronic fatigue syndrome. Disabil Rehabil. 2011;33:1768-75. [PMID: 21208154] doi:10.3109/09638288.2010.546936
40. **Jason LA, Taylor RR.** Measuring attributions about chronic fatigue syndrome. J Chronic Fatigue Syndr. 2001;8:31-40.
41. **Asbring P, Närvänen AL.** Women's experiences of stigma in relation to chronic fatigue syndrome and fibromyalgia. Qual Health Res. 2002;12:148-60. [PMID: 11837367]
42. **Assefi NP, Coy TV, Uslan D, Smith WR, Buchwald D.** Financial, occupational, and personal consequences of disability in patients with chronic fatigue syndrome and fibromyalgia compared to other fatiguing conditions. J Rheumatol. 2003;30:804-8. [PMID: 12672203]
43. **Deale A, Wessely S.** Diagnosis of psychiatric disorder in clinical evaluation of chronic fatigue syndrome. J R Soc Med. 2000;93:310-2. [PMID: 10911826]
44. **Dickson A, Knussen C, Flowers P.** Stigma and the delegitimation experience: an interpretative phenomenological analysis of people living with chronic fatigue syndrome. Psychol Health. 2007;22:851-67.
45. **Green J, Romei J, Natelson BH.** Stigma and chronic fatigue syndrome. J Chronic Fatigue Syndr. 1999;5:63-95.
46. **Guise J, McVittie C, McKinlay A.** A discourse analytic study of ME/CFS (Chronic Fatigue Syndrome) sufferers' experiences of interactions with doctors. J Health Psychol. 2010;15:426-35. [PMID: 20348363] doi:10.1177/1359105309350515
47. **Jason LA, Taylor RR, Stepanek Z, Plioplys S.** Attitudes regarding chronic fatigue syndrome: the importance of a name. J Health Psychol. 2001;6:61-71. [PMID: 22049238] doi:10.1177/135910530100600105
48. **Brimmer DJ, Maloney E, Devlin R, Jones JF, Boneva R, Nagler C, et al.** A pilot registry of unexplained fatiguing illnesses and chronic fatigue syndrome. BMC Res Notes. 2013;6:309. [PMID: 23915640] doi:10.1186/1756-0500-6-309
49. **Devasahayam A, Lawn T, Murphy M, White PD.** Alternative diagnoses to chronic fatigue syndrome in referrals to a specialist service: service evaluation survey. JRSM Short Rep. 2012;3:4. [PMID: 22299071] doi:10.1258/shorts.2011.011127
50. **Lawn T, Kumar P, Knight B, Sharpe M, White PD.** Psychiatric misdiagnoses in patients with chronic fatigue syndrome. JRSM Short Rep. 2010;1:28. [PMID: 21103120] doi:10.1258/shorts.2010.010042
51. **Newton JL, Mabillard H, Scott A, Hoad A, Spickett G.** The Newcastle NHS Chronic Fatigue Syndrome Service: not all fatigue is the same. J R Coll Physicians Edinb. 2010;40:304-7. [PMID: 21132135] doi:10.4997/JRCPE.2010.404
52. **Woodward RV, Broom DH, Legge DG.** Diagnosis in chronic illness: disabling or enabling–the case of chronic fatigue syndrome. J R Soc Med. 1995;88:325-9. [PMID: 7629762]
53. **Brurberg KG, Fønhus MS, Larun L, Flottorp S, Malterud K.** Case definitions for chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME): a systematic review. BMJ Open. 2014;4:e003973. [PMID: 24508851] doi:10.1136/bmjopen-2013-003973

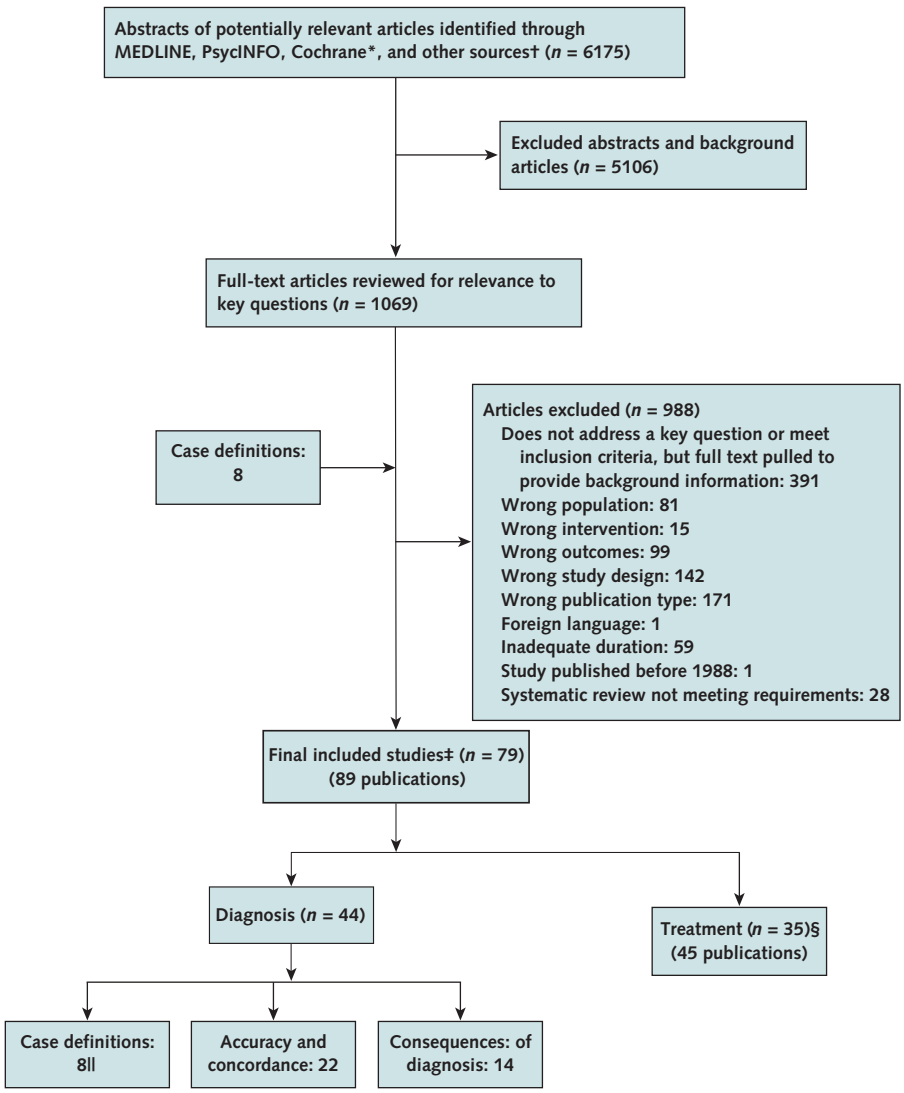
Current Author Addresses: Drs. Haney, Smith, McDonagh, and Nelson; Ms. Pappas; Ms. Daeges; and Ms. Wasson: 3181 SW Sam Jackson Park Road, Mail Code BICC, Portland, OR 97239.

Author Contributions: Conception and design: E. Haney, M.E.B. Smith, M. McDonagh, H.D. Nelson.
Analysis and interpretation of the data: E. Haney, M.E.B. Smith, M. McDonagh, M. Pappas, N. Wasson, H.D. Nelson.
Drafting of the article: E. Haney, M.E.B. Smith, M. McDonagh, M. Pappas, H.D. Nelson.
Critical revision of the article for important intellectual content: E. Haney, M.E.B. Smith, M. McDonagh, H.D. Nelson.
Final approval of the article: E. Haney, M.E.B. Smith, M. McDonagh, M. Pappas, N. Wasson, H.D. Nelson.
Obtaining of funding: M.E.B. Smith, M. McDonagh, H.D. Nelson.
Administrative, technical, or logistic support: M. Pappas, N. Wasson.
Collection and assembly of data: E. Haney, M.E.B. Smith, M. Pappas, N. Wasson, H.D. Nelson.

Web-Only References

54. **Beck AT, Steer RA, Carbin MG.** Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. Clin Psychol Rev. 1988;8:77-100.
55. **Brown M, Kaplan C, Jason L.** Factor analysis of the Beck Depression Inventory-II with patients with chronic fatigue syndrome. J Health Psychol. 2012;17:799-808. [PMID: 22104663] doi:10.1177/1359105311424470
56. **Chalder T, Berelowitz G, Pawlikowska T, Watts L, Wessely S, Wright D, et al.** Development of a fatigue scale. J Psychosom Res. 1993;37:147-53. [PMID: 8463991]
57. **Morriss RK, Wearden AJ, Mullis R.** Exploring the validity of the Chalder Fatigue scale in chronic fatigue syndrome. J Psychosom Res. 1998;45:411-7. [PMID: 9835234]
58. **Cella M, Chalder T.** Measuring fatigue in clinical and community settings. J Psychosom Res. 2010;69:17-22. [PMID: 20630259] doi:10.1016/j.jpsychores.2009.10.007
59. **Hickie I, Lloyd A, Hadzi-Pavlovic D, Parker G, Bird K, Wakefield D.** Can the chronic fatigue syndrome be defined by distinct clinical features? Psychol Med. 1995;25:925-35. [PMID: 8588011]
60. **Hickie I, Lloyd A, Wakefield D, Parker G.** The psychiatric status of patients with the chronic fatigue syndrome. Br J Psychiatry. 1990;156:534-40. [PMID: 2386862]
61. **Broadbent DE, Cooper PF, FitzGerald P, Parkes KR.** The Cognitive Failures Questionnaire (CFQ) and its correlates. Br J Clin Psychol. 1982;21 (Pt 1):1-16. [PMID: 7126941]
62. **Hawk C, Jason LA, Torres-Harding S.** Reliability of a chronic fatigue syndrome questionnaire. J Chronic Fatigue Syndr. 2007;13:41-66.
63. **Jason LA, Ropacki MT, Santoro NB, Richman JA, Heatherly W, Taylor R, et al.** A screening instrument for chronic fatigue syndrome: reliability and validity. J Chronic Fatigue Syndr. 1997;3:39-59.
64. **Jackson CA, Earl L.** Prevalence of fatigue among commercial pilots. Occup Med (Lond). 2006;56:263-8. [PMID: 16733255]]
65. **Bjelland I, Dahl AA, Haug TT, Neckelmann D.** The validity of the Hospital Anxiety and Depression Scale. An updated literature review. J Psychosom Res. 2002;52:69-77. [PMID: 11832252]
66. **McCue P, Buchanan T, Martin CR.** Screening for psychological distress using internet administration of the Hospital Anxiety and Depression Scale (HADS) in individuals with chronic fatigue syndrome. Br J Clin Psychol. 2006;45:483-98. [PMID: 17076959]
67. **Karnofsky DA, Abelmann WH, Craver LF, Burchenal JH.** The use of the nitrogen mustards in the palliative treatment of carcinoma. With particular reference to bronchogenic carcinoma. Cancer. 1948;1:634-56.
68. **Grieco A, Long CJ.** Investigation of the Karnofsky Performance Status as a measure of quality of life. Health Psychol. 1984;3:129-42. [PMID: 6536486]
69. **Smets EM, Garssen B, Bonke B, De Haes JC.** The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. J Psychosom Res. 1995;39:315-25. [PMID: 7636775]
70. **Lin JM, Brimmer DJ, Maloney EM, Nyarko E, Belue R, Reeves WC.** Further validation of the Multidimensional Fatigue Inventory in a US adult population sample. Popul Health Metr. 2009;7:18. [PMID: 20003524] doi:10.1186/1478-7954-7-18
71. **Stewart AL, Hays RD, Ware JE Jr.** The MOS short-form general health survey. Reliability and validity in a patient population. Med Care. 1988;26:724-35. [PMID: 3393032]
72. **Stewart AL, Greenfield S, Hays RD, Wells K, Rogers WH, Berry SD, et al.** Functional status and well-being of patients with chronic conditions. Results from the Medical Outcomes Study. JAMA. 1989;262:907-13. [PMID: 2754790]
73. **Morriss RK, Wearden AJ.** Screening instruments for psychiatric morbidity in chronic fatigue syndrome. J R Soc Med. 1998;91:365-8. [PMID: 9771495]
74. **Main CJ.** The Modified Somatic Perception Questionnaire (MSPQ). J Psychosom Res. 1983;27:503-14. [PMID: 6229628]
75. **Schrezenmaier C, Gehrking JA, Hines SM, Low PA, Benrud-Larson LM, Sandroni P.** Evaluation of orthostatic hypotension: relationship of a new self-report instrument to laboratory-based measures. Mayo Clin Proc. 2005;80:330-4. [PMID: 15757013]
76. **Pennebaker JW.** The Psychology of Physical Symptoms. New York: Springer; 1982.
77. **Vercoulen JH, Swanink CM, Fennis JF, Galama JM, van der Meer JW, Bleijenberg G.** Dimensional assessment of chronic fatigue syndrome. J Psychosom Res. 1994;38:383-92. [PMID: 7965927]
78. **Bergner M, Bobbitt RA, Carter WB, Gilson BS.** The Sickness Impact Profile: development and final revision of a health status measure. Med Care. 1981;19:787-805. [PMID: 7278416]
79. **Ware JE Jr, Sherbourne CD.** The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992;30:473-83. [PMID: 1593914]
80. **Buchwald D, Pearlman T, Umali J, Schmaling K, Katon W.** Functional status in patients with chronic fatigue syndrome, other fatiguing illnesses, and healthy individuals. Am J Med. 1996;101:364-70. [PMID: 8873506]
81. **Schmitz N, Hartkamp N, Kiuse J, Franke GH, Reister G, Tress W.** The Symptom Check-List-90-R (SCL-90-R): a German validation study. Qual Life Res. 2000;9:185-93. [PMID: 10983482]
82. **Zung WW.** A self-rating depression scale. Arch Gen Psychiatry. 1965;12:63-70. [PMID: 14221692]

Appendix Figure. Summary of evidence search and selection.



* Cochrane Central Register of Controlled Trials, Database of Abstracts of Reviews of Effects, Health Technology Assessment, National Health Sciences Economic Evaluation Database, and Cochrane Database of Systematic Reviews.
† Identified from such sources as reference lists, hand searches, and suggestions by experts.
‡ Studies that provided data and contributed to the body of evidence were considered "included."
§ Studies included for the treatment key questions are reported elsewhere (13).
|| The Institute of Medicine case definition (9) is an additional case definition, which was released subsequent to the search.

Appendix Table 1. Included Studies of Methods Used to Diagnose Myalgic Encephalomyelitis/Chronic Fatigue Syndrome

Study, Year (Reference)	Study Design/N Case definition	Diagnostic Test	Significant Findings
Gaab et al, 2005 (36)	Descriptive observational study. 21 CFS patients; 20 controls. CDC (Fukuda, 1994); Oxford (Sharpe, 1991)	ACTH, plasma cortisol, salivary cortisol, differential blood count, IL-6 and TNF-alpha (baseline, and 10, 60 minutes after the TSST) German translation of the Fatigue Scale (Chalder 1993), SIP-8, SCL-90R, HADS All subjects underwent the TSST: after basal blood and saliva samples were taken they were told to prepare for a fake job interview, then given a mental arithmetic task in front of an audience and told they would be videotaped for further analysis of their behavior.	The HADS, SCL-90R and SIP-8 scores were all significantly higher in the CFS group AUC for IL-6 and TNF-alpha vs. Chalder fatigue scale total score, mental fatigue and physical fatigue scores NS.
Jason et al, 2011 (33)	Observational descriptive study. 2 populations: 1) 114 recruited from tertiary care and 2) 32 community based sample with 47 in a non-fatigued control group. CDC (Fukuda, 1994)	SF-36 subscales for differentiating CFS patients.	Community-based sample (cases vs. controls) - AUC (SE) by subscale of SF-36: Vitality: 0.88 (0.04) Social functioning: 0.87 (0.04); Role-physical: 0.86 (0.04); Bodily pain: 0.85 (0.04); Physical Functioning: 0.84 (0.05); General Health: 0.80 (0.05); Mental Health: 0.75 (0.06); Role-Emotional: 0.67 (0.07). Tertiary care-based sample (cases vs. community controls) - AUC (SE) by subscale of SF-36: Vitality: 0.91 (0.03); Social functioning: 0.87 (0.04); Role-physical: 0.91 (0.03); Bodily pain: 0.86 (0.04); Physical Functioning: 0.87 (0.04); General Health: 0.91 (0.35); Mental Health: 0.71 (0.05); Role-Emotional: 0.63 (0.05).
Jason et al, 2010 (32)	Observational descriptive study. 213 adults from community based sample from neighborhoods in Chicago (see above). Final sample n = 108. 24 who had CFS and 84 who did not. Fukuda, 1994 refined as recommended by International Research group and the CDC (Reeves, Lloyd et al BMC health services research vol 3, 2003).	MFI-20 SF-36 To evaluate the CDC Empiric CFS definition (Reeves et al, BMC Medicine 2005) which assesses 3 areas: disability SF-36), fatigue (MFI-20) and symptoms (CDC symptom inventory). Aim to determine specific instruments and cutoffs to facilitate a more reliable approach to assessment of CFS.	AUC, sensitivity, specificity MFI-20 subscale General fatigue: 0.69, 74%, 39% Reduced activity: 0.64, 74%, 50% Meeting Reeves fatigue criteria: 0.61, 95%, 27% CDC Symptom Inventory Meeting Reeves core symptoms criteria (total): 0.69, 59%, 73%; SF-36 subscales: Physical functioning: 0.60, 68%, 51%; Role physical: 0.66, 82%, 51%; Social functioning: 0.62, 74%, 35%; Role emotional: 0.57, 73%, 44%; Meeting Reeves substantial reductions criteria: 0.56, 96%, 17%; Meeting Reeves CFS criteria: 0.70, 65%, 76%
Hadzi-Pavlovic et al, 2000 (31)	Observational descriptive study. 2,669 adults consisting of 613 patients referred to CFS clinic over 5 years, plus 452 controls recruited by the CFS patients, and 1,593 primary care attenders. Oxford (Sharpe, 1991)	To develop and evaluate the SOFA/CFS instrument for identifying CFS. SOFA developed from General Health Questionnaire 5 items from the Zung depression Scale Chronic Fatigue Symptoms Checklist Somatization Checklist (39 physical symptoms)	Initial phase: clinical sample and their selected controls. 10 items with highest loadings on the first factor - total score of these 10 items. Sensitivity, specificity A cut-off score of 1/2 classified 341/368 CFS cases and 409/430 control subjects correctly: 93%, 95% Kraemer's QROC: 87%, 89% Including the 69 CFS subjects who had a diagnosis other than CFS or for whom there was low confidence in the diagnosis as "non-cases" did not change the sensitivity, but reduced the specificity to 83% QROC: 86%, 65% LCA performed on 368 CFS subjects only Sensitivity, specificity Cut-off of ≥2: 3 class: 100%, 90%; 4 class: 97%, 98% Cut-off of ≥3: 3 class: 81%, 100%; 4 class: 66%, 100%

Continued on following page

Appendix Table 1–Continued

Study, Year (Reference)	Study Design/N Case definition	Diagnostic Test	Significant Findings
Linder et al, 2002 (30)	Observational descriptive study. 99 CFS, 41 SLE, 58 FMS. Randomly assigned to development (n = 158) or validation (n = 40) cohort. Oxford (Sharpe, 1991)	Applied self-learning artificial neural network to general diagnostic criteria sets for CFS compared to traditional classification criteria.	Sensitivity, specificity, accuracy Applied traditional CDC (Holmes, 1988) definition (group A): 62.6%, 93.9%, 78.3% Traditional format unweighted classification criteria in validation cohort (group B): 90.0%, 65.0%, 77.5%. Three symptoms: sudden onset of fatigue, sore throat, and impaired vision have the greatest discriminatory power in differentiating CFS from systemic lupus erythematosus and fibromyalgia. Weighting of classification criteria with regression coefficients in validation cohort (group B): 90.0%, 75.0%, 82.5% (optimum accuracy is obtained using sudden onset of fatigue, sore throat, and irritability (positive associations); negative associations with GI disturbances, allergies and dyspnea) Regression tree analysis in the validation cohort (group B): 95.0%, 80.0%, 87.5% (at most, 5 symptoms need to be ascertained before a classification can be made) Artificial neural network in the validation cohort (group B): 95.0%, 85.0%, 90.0% (uses 24 of the 26 symptoms)
Tiev et al, 2003 (29)	Observational descriptive study. 11 with CFS and 14 healthy volunteer controls. CDC (Fukuda, 1994)	Ratio of Rnase L isoforms measured from PBMCs. MFI-20	Using 0.4 as the cutoff for RNase L isoform ratio. Sensitivity: 91%, Specificity: 71%.
Watson et al, 2014 (37)	Observational descriptive study of 3 cohorts: 1) DePaul sample:187 CFS and 96 controls; 2) Biobank sample: 233 CFS and 80 controls; 3) Newcastle sample: 95 suspected CFS. ME/CFS as identified by any of 3 different case definitions: CDC (Fukuda, 1994), Canadian, 2003 and ME-ICC	DePaul Symptom Questionnaire; computerized thresholding using a k-means clustering approach.	Sensitivity, specificity, accuracy for case definition Unsupervised thresholding (UT): CDC (Fukuda, 1994): 83.1, 85.8, 83.8; Canadian (Carruthers, 2003): 82.9, 87.5, 84.1; ME-ICC (Carruthers, 2011): 74.4, 91.5, 78.7. Supervised thresholding: CDC (Fukuda, 1994): 80.8, 86.4, 82.2; Canadian (Carruthers, 2003): 85.8, 87.5, 86.3; ME-ICC (Carruthers, 2011): 89.9, 81.3, 87.7. Two-two static threshold: CDC (Fukuda, 1994): 80.8, 85.8, 82.1; Canadian (Carruthers, 2003): 77.9, 89.8, 80.9; ME-ICC (Carruthers, 2011) 67.4, 91.5, 73.5 (sensitivity and accuracy for Fukuda and ME-ICC p = 0.01 vs. UT). One-one static threshold: CDC (Fukuda, 1994): 98.1, 42.0, 83.8 (p = 0.01 vs. UT for sensitivity and specificity); Canadian (Carruthers, 2003): 97.3, 50.0, 85.2 (p = 0.01 vs. UT for sensitivity and specificity); ME-ICC (Carruthers, 2011) 93.4, 52.8, 83.1 (p = 0.01 vs. UT for sensitivity and specificity).

ACTH = adrenocorticotropic hormone; AUC = area under the curve; BMC = BioMed Central; CDC = Centers for Disease Control and Prevention; CFS = chronic fatigue syndrome; FMS = fibromyalgia syndrome; HADS = Hospital Anxiety and Depression Scale; IL = interleukin; LCA = latent class analysis; ME = myalgic encephalomyelitis; ME-ICC = myalgic encephalomyelitis–international consensus criteria; MFI-20 = Multidimensional fatigue inventory; NS = not significant; PBMC = peripheral blood derived mononuclear cell; QROC = quality receiver operating characteristic; Rnase L = latent Ribonuclease; SCL-90R = symptom checklist 90-revised; SF-36 = 36-item Short Form Survey; SIP-8 = Sickness Impact Profile 8-item; SLE = systemic lupus erythematosus; SOFA = schedule of fatigue and anergia; TNF = tumor necrosis factor; TSST = Trier social stress test; UT = unsupervised threshold.

Appendix Table 2. Included Studies Evaluating the Concordance of Different Diagnostic Criteria and Comparisons Between Populations

Study, Year (Reference)	Populations Case Definition Measures	Findings
Aslakson et al, 2006 (25)	N = 159 women: 51 CFS, 55 chronic fatigue (not CFS), 53 nonfatigued controls Reeves, 1994 case definition of ICF/CFS and CDC (Fukuda, 1994) criteria Measures: SF-36, Zung depression scale Methods: Used latent class analysis to compare empiric classification to the CDC (Fukuda, 1994) categories (CFS, idiopathic chronic fatigue, and nonfatigued controls)	Empirically derived latent class solution compares favorably against established research criteria for CFS and idiopathic chronic fatigue.
Brown et al, 2013 (21)	N = 113: 74 CDC (Fukuda, 1994) criteria, 39 ME International Consensus (Carruthers, 2011) criteria Measures: International Consensus criteria, Fukuda CFS questionnaire, DSM-IV SCID interview and medical appointment to rule out other reason for symptoms, SF-36, Cognitive Trailmaking Tests A and B from Halstead-Reitan Battery	CDC (Fukuda, 1994) vs. International ME (Carruthers, 2011) Demographics differences Concurrent psychiatric diagnosis: 27% (20/74) vs. 62% (24/39); $p < 0.001$ Sudden onset of illness (<1 month): 26% (19/74) vs. 44% (16/39); $p = 0.05$ Mean (SD) SF-36 subscales (0-100 scale, higher scores indicate better health); only significant outcomes are reported here Physical functioning: 51.0 (22.63) vs. 36.64 (23.32); $p = 0.001$ Bodily pain: 46.65 (21.42) vs. 27.28 (19.45); $p < 0.001$ Vitality: 19.86 (15.26) vs. 13.85 (13.15); $p = 0.04$ Social functioning: 45.25 (24.22) vs. 30.45 (21.99); $p = 0.002$ Symptom complaints more common in International ME vs. CDC PEM: $p = 0.004$ Neurological: memory/concentration ($p = 0.01$), slowness of thought ($p = 0.001$), absent mindedness ($p = 0.02$), confusion/disorientation ($p = 0001$), difficulty reasoning ($p = 0.01$), forgetting what you're trying to say ($p = 0.001$), difficulty finding the right word ($p = 0.002$), need to focus on one thing at a time ($p < 0.001$), frequently lose train of thought ($p = 0.001$), trouble expressing thoughts ($p>0.001$), difficulty retaining information ($p < 0.001$), difficulty recalling information ($p < 0.001$), put words/numbers in wrong order ($p = 0.04$), slow to react ($p < 0.001$), attention deficit ($p = 0.05$), poor hand-eye coordination ($p = 0.02$). Pain: muscle pain ($p < 0.001$), pain in multiple joints ($p < 0.001$), headaches ($p = 0.02$)
Jason et al, 2001 (20)	N = 55: 14 CDC (Holmes, 1988) criteria; 18 CDC (Fukuda, 1994) criteria Measures: Comparison of symptom frequency; and SF-36	CDC (Holmes, 1988) criteria vs. CDC (Fukuda, 1994) criteria vs. chronically fatigued psychiatric group % symptom frequency Sore throat: 85.7 vs. 44.4 vs. 51.5; $p < 0.05$ Lymph node pain 85.7 vs. 27.8 vs. 27.3; $p < 0.01$ for Fukuda vs. psychiatric group All others symptoms p = NS Mean SF-36 sub-scales scores (0-100 scale, higher scores indicate better health) Bodily pain: 33.3 vs. 44.5 vs. 53.7; $p < 0.05$ General health: 34.9 vs. 55.5 vs. 49.9; $p < 0.05$ Physical health composite: 30.9 vs. 37.0 vs. 39.9; $p < 0.05$ for Fukuda vs. psychiatric group All other subscales and composite scales p = NS Mean degree of impairment (0-100 scale, lower scores indicate better health) 64.1 vs. 46.5 vs. 65.6; $p < 0.05$ for Fukuda vs. psychiatric group

Continued on following page

Appendix Table 2–Continued

Study, Year (Reference)	Populations Case Definition Measures	Findings
Jason et al, 2013 (10)	N = 489: 189 DePaul sample; 242 BioBank sample; 96 Newcastle sample CDC (Fukuda, 1994) and Canadian (Carruthers, 2003) Measures: DePaul Symptom Questionnaire, SF-36	CDC (Fukuda, 1994) vs. Canadian (Carruthers, 2003) Mean (SD) SF-36 subscales (0-100 scale, higher scores indicate better health); only significant outcomes are reported here DePaul sample Physical functioning: 35.6 (19.6) vs. 28.1 (17.9); p < 0.05 Bodily pain: 59.3 (24.3) vs. 36.6 (19.7); p < 0.001 BioBank sample Physical functioning: 46.8 (22.9) vs. 33.2 (21.6); p < 0.001 Bodily pain: 60.0 (24.8) vs. 41.1 (21.0); p < 0.001 General health: 29.8 (17.8) vs. 22.8 (14.2); p < 0.01 Social functioning: 42.7 (28.8) vs. 24.0 (21.6); p < 0.001 Mental health: 72.2 (13.7) vs. 66.0 (19.6); p < 0.05 Vitality: 20.6 (13.7) vs. 12.0 (12.3); p < 0.001 Newcastle sample Physical functioning: 49.1 (25.8) vs. 29.6 (25.4); p < 0.05 Bodily pain: 45.2 (25.0) vs. 29.5 (21.3); p < 0.05 General health: 35.3 (18.9) vs. 20.7 (12.5); p < 0.01 Social functioning: 39.4 (20.9) vs. 25.0 (20.5); p < 0.05 Symptom complaints more common in Canadian (Carruthers, 2003) vs. CDC (Fukuda, 1994); p < 0.05 for those noted below. PEM: 3/5 subcategories in all 3 samples; 4/5 in DePaul and Solve samples Sleep parameters (unrefreshing sleep): 1/6 in all 3 samples; 3/6 other sleep parameters in DePaul and Solve samples only Pain: 5/7 subcategories in all 3 samples, 7/7 in DePaul and Solve samples Neurocognitive: 4/13 in all 3 samples; 15/15 in DePaul and Solve samples Autonomic: 4/7 in all 3 samples, 7/7 in DePaul and Solve samples Neuroendocrine: 5 /10 in all 3 samples; 10/10 in DePaul and Solve samples Immune: 4/5 in all 3 samples; 5/5 in DePaul and Solve samples
Jason et al, 2012 (1)	N = 114 meeting CDC (Fukuda, 1994) criteria CDC (Fukuda, 1994), Canadian (Carruthers, 2003), and Revised Ramsay, 1988 Measures: CFS questionnaire (validated by Jason 1997) to assess symptoms, with modified scoring system ranging from 0-100 with higher scores indicating more impairment; DSM-IV SCID interview, medical, and neurological history and exam, other explanation for CFS-like symptoms; CFS Questionnaire (Komaroff 1996) to rule out other disorders; MOS-SF; Cognitive Trailmaking Test Parts A and B Heart rate lying down, 2 minutes after standing, and 10 minutes after standing Methods: Used symptom counts, chi-square and MANOVA to assess differences between group	Of 114 people meeting Fukuda CFS criteria, 56 did not meet the ME/CFS criteria and 97 did not meet the ME criteria (56 were classified as ME/CFS and 27 as ME). 1 person was unable to be categorized. ME/CFS vs. CFS not ME/CFS Demographics differences Disability: 32% (18/57) vs. 16% (9/56); p = 0.06 Current psychiatric diagnoses: 58% (33/57) vs. 20% (11/56); p = 0.05 Sudden illness onset (<1 month): 41% (22/57) vs. 24% (13/56); p = 0.0 Physical cause of fatigue: 64% (36/57) vs. 65% (35/56); p = 0.04 Mean (SD) SF-36 subscales (0-100 scale, higher scores indicate better health); only significant outcomes are reported here Physical functioning: 38.0 (21.9) vs. 53.8 (23.4); p = 0.00 Bodily pain: 32.2 (20.0) vs. 48.0 (22.1); p = 0.00 General health: 28.5 (16.0) vs. 36.5 (18.3); p = 0.02 Vitality: 14.8 (12.0) vs. 20.9 (16.6); p = 0.02 Social functioning: 34.0 (22.7) vs. 46.6 (24.2); p = 0.01 Symptom complaints more common among ME/CFS vs. CFS not ME/CFS Fatigue: p = 0.00; PEM: p = 0.00; unrefreshing sleep: p = 0.00; need to nap each day: p = 0.05; difficulty falling asleep: p = 0.01; all pain parameters (muscle pain, pain in multiple joints, headaches, chest pain, abdomen pain, eye pain): all p < 0.02; all neurological parameters (impaired memory and concentration, abnormal sensitivity to light, slowness of thought, confusion/disorientation, difficulty finding the right work, difficulty comprehending information, need to have focus on one thing at a time): p = 0.00; all autonomic parameters (racing heart, shortness of breast, dizziness, feel unsteady on feet): p < 0.01; and tender/sore lymph nodes: all p = 0.00 Symptom complaints more common among ME vs. CFS not ME/CFS Headaches: p = 0.05; chest pain: p = 0.04; abdomen pain: p = 0.00; eye pain: p = 0.00; difficulty finding the right word: p = 0.05; need to have focus on one thing at a time: p = 0.02; all autonomic parameters (racing heart, shortness of breast, dizziness, feel unsteady on feet): all p < 0.02; tender/sore lymph nodes: p = 0.02; and hot/cold spells: p = 0.05 ME/CFS vs. CFS not ME/CFS; ME vs. CFS not ME Mean (SD) heart rate (bpm) Lying down: 80.7 (14.8) vs. 74.5 (11.1); p = 0.02; 84.4 (16.4) vs. 75.4 (11.4); p = 0.00 Standing 2 minutes: 94.2 (17.1) vs. 85.7 (14.6); p = 0.00; 96.9 (18.9) vs. 87.7 (14.9); p = 0.00 Standing 10 minutes: 94.6 (14.5) vs. 86.2 (13.6); p = 0.00; 97.8 (14.4) vs. 88.1 (13.9); p = 0.00 Mean (SD) Trailmaking test scores A-time: 32.9 (13.6) vs. 26.8 (9.9); p = 0.02; 35.3 (15.8) vs. 28.2 (10.3); p = 0.02 B-time: 56.1 (25.1) vs. 46.8 (14.9); p = 0.03; 61.2 (28.3) vs. 48.5 (17.3); p = 0.00 Symptoms and Psychiatric Comorbidity: ME/CFS group had 7.3 of the 13 Kroenke (2003) symptoms vs 5.1 for Fukuda CFS (p < 0.05); ME group had 8.1 of the 13 Kroenke (2003) symptoms vs 5.6 for Fukuda CFS (p < 0.01).

Continued on following page

Appendix Table 2–Continued

Study, Year (Reference)	Populations Case Definition Measures	Findings
Jason et al, 2004 (24)	N = 780 reported fatigue from random telephone survey CDC (Fukuda, 1994), Canadian (Carruthers, 2003), and Revised Ramsay, 1988 Measures: Work status, psychiatric comorbidity, symptoms, functional impairment as measured by medical outcomes study (MOS)	Canadian vs. CFS Fukuda vs. Chronic Fatigue-Psych No differences between groups on the Fatigue Scale or the Mental composite score of the MOS. Physical composite score: 32.5 vs. 37.8 vs. 39.9 No different in psychiatric status Rates of current psychiatric diagnoses: 47.8% vs. 75.0% vs. 87.9% ($p < 0.01$) Rates of lifetime psychiatric diagnoses: 78.3% vs. 83.3% vs. 100% ($p < 0.050$) Symptoms (all significant at $p < 0.05$): Fatigue General muscle weakness: 82.6% vs. 66.7% vs. 54.5% Neck weak: 52.2% vs. 25.0% vs. 24.2% Shoulders weak: 52.2% vs. 25.0% vs.24.2% Back weak: 47.8% vs. 33.3% vs. 18.2% Disturbed Sleep Trouble staying asleep: 30.4% vs. 66.7% vs. 39.4% Neuropsychiatric Confusion or Disorientation: 39.1% vs. 8.3% vs. 12.1% Difficulty retaining information: 56.5% vs. 41.7 % vs. 27.3% Need to focus on one thing at a time: 65.2% vs. 25.0% vs. 24.2% Slow to process visual and auditory information: 30.4% vs. 8.3% vs. 6.1% Disturbances in eyesight: 43.5% vs. 33.3% vs. 18.2% Infectious Lymph node pain: 34.8% vs. 25.0% vs.12.1% Rheumatologically Neck muscles ache: 65.2%vs. 75.0% vs. 36.4% Back muscles ache: 65.2% vs. 66.7% vs. 36.4% Stiff after sitting: 39.1% vs. 58.3% vs. 21.2% Sinus infection: 4.3% vs. 41.7% vs. 12.1% Sinus congestion: 26.1% vs. 50.0% vs. 15.2% Cardiopulmonary chest pains: 34.8% vs. 33.3% vs. 9.1% Gastrointestinal Bloating: 26.1% vs. 50.0% vs.15.2% Lower abdominal pain: 26.1% vs. 41.7% vs. 9.1% Neurological Feel weak or dizzy after standing: 43.5% vs. 41.7% vs. 18.2% Dizziness when move head suddenly: 47.8% vs. 16.7% vs. 18.2% Alcohol intolerance: 47.8% vs. 33.3% vs. 15.2% Reproductive Decreased sexual interest/function: 30.4% vs. 58.3% vs. 18.2%
Jason et al, 2014 (23)	N = 270: 73 CFS and 112 ME in DePaul sample; 27 CFS and 58 ME in Newcastle sample CDC (Fukuda, 1994) and ME-ICC (Carruthers, 2011) Measures: DePaul Symptom Questionnaire and SF-36	CFS vs. ME General health: 28.6 vs. 22.6 for the DePaul sample; 32.3 vs.19.1 for the Newcastle sample ($p = 0.01$) Bodily pain 50.0 vs. 25.6 for the DePaul sample ($p < 0.001$); no difference for the Newcastle sample Physical functioning 34.1 vs. 26.9 for the DePaul sample ($p < 0.01$); no difference for the Newcastle sample Role physical 7.9 vs. 2.5 ($p < 0.05$) for the DePaul sample; no difference for the Newcastle sample Vitality 15.4 vs. 11.2 ($p < 0.05$); no difference for the Newcastle sample
Katon et al, 1991 (26)	N = 129: 19 CFS; 79 chronic fatigue; 32 rheumatoid arthritis CDC (Holmes, 1988) Measures: General Health Questionnaire total score, MOS-SF, Modified Somatic Perception Questionnaire, Pennebaker inventory of Limbic Languidness	CFS vs. RA GHQ scores Mean (SD) total score: 12.5 (8.0) vs. 5.1 (4.6); $p < 0.001$ Score of ≥11: 53% (47/98) vs. 13% (3/31); $p < 0.001$ Mean (SD) MOS-SF (1-100 scale, higher score indicates better health); significant results only reported here Mental health: 17.7 (5.5) vs. 23.0 (5.4); $p < 0.01$ Health perception: 3.4 (1.4) vs. 5.3 (2.1); $p < 0.001$ No significant difference for SF-36 physical function and role functional, Modified Symptoms Perception Questionnaire, or the Pennebaker Inventory of Limbic Languidness.

Continued on following page

Appendix Table 2–Continued

Study, Year (Reference)	Populations Case Definition Measures	Findings
Komaroff et al, 1996 (27)	N = 5,881: 223 CFS; 2,474 controls; 5,881 chronic disease CDC (Fukuda, 1994) Measures: SF-36	Significant p values for means on SF-36 subscales: comparisons vs. CFS Physical functioning: p < 0.00001 general population, HTN, DM, AMI, and depression; p = 0.00004 CHF Role physical: p < 0.00001 all Bodily pain: p < 0.00001 all General health: p < 0.00001 all Vitality: p < 0.00001 all but MS which was NS (p = 0.1369) Social functioning: p < 0.00001 Role emotional: p < 0.00001 general population, HTN, DM, and depression; p = 0.3918 CHF; p = 0.1077 MS Mental health: p < 0.00001 all but MS which p = 0.0005
Lewis et al, 2013 (28)	N = 50: 25 CFS ages 16-29; 25 CFS ages >50 CDC (Fukuda, 1994) Measures: Heart rate variability, Baroreceptor sensitivity, FIS, CFQ, HADS, HADS-A and HADS-D, SF-36, Chalder fatigue scale, ESS, OGS - 5 items, each graded 0-4, t-tests statistics	Age 16-29 years vs. ≥50 years; only significant results reported here Mean (SD) BMI (kg/m^2): 22 (3) vs. 26 (3); p = 0.002 Mean (SD) FIS: 85 (33) vs. 107 (27); p = 0.02 Mean (SD) Chalder Fatigue severity scale (0-56 scale, lower score indicates better health): 9 (3) vs. 11 (1); p = 0.002 Mean (SD) HADS-D: 7 (3) vs. 10 (4); p = 0.005 Mean (SD) total SF-36 score (0-100, higher scores indicate better health): 20 (5) vs. 16 (5); p = 0.03 Mean (SD) self-efficacy scores: 31 (12) vs. 22 (14); p = 0.02 Mean (SD) heart rate (bpm): 80 (15) vs. 71 (8); p = 0.007 Mean (SD) LVET (ms): 274.6 (16) vs. 285.8 (9); p = 0.004 Mean (SD) LFnu: 51.5 (17) vs. 63.8 (18); p = 0.01 Mean (SD) HFnu: 49.1 (18) vs. 36.2 (18); p = 0.0 1 Mean (SD) LF/HF: 1.5 (0.9) vs. 2.2 (1.4); p = 0.04 Mean (SD) BRS: 19.7 (12) vs. 9.9 (5); p = 0.0004 Autonomic and hemodynamic differences: higher LVET (p = 0.004), higher LFnu (p = 0.01), higher HFnu (p = 0.01), higher LF/HF (p = 0.04), lower BRS (p = 0.0004) for the subjects > 50 vs. those age 16-26. No difference in HR, systolic BP, diastolic BP, mean BP, total HRV, BEI, or systolic BP with active stand.
Van Hoof and De Meirleir, 2005 (22)	N = 67: 41 CFS and 26 ME CDC (Fukuda, 1994) and London criteria for ME (National Task Force, 1994) Measures: SF-36, MFI-20, KPS, exercise	CFS vs. ME Demographic differences; only significant differences reported here Mean age (SD): 43 (10) vs. 34 (7) years; p = 0.001 Mean (SD) SF-36 subscale scores (0-100 scale, higher scores indicate better health) Role emotional: 62 (44.05) vs. 83 (31.05); p = 0.024 Mental health: 60 (17.90) vs. 69 (13.41); p = 0.049 Mean (SD) MFI-20 (4-20 scale, lower score indicates better health) General fatigue: 18 (2.73) vs. 17 (2.88); p = 0.029 Physical parameters; only significant differences reported here Mean (SD) age predicted heart rate (bpm): 178.04 (10.67) vs. 185.57 (6.64); p = 0.049 Mean (SD) VO_2 predicted: 26.81 (3.66) vs. 29.39 (2.28); p = 0.049 Note: Only the Role Emotional SF-36 subscale seemed able to discriminate ME patients from CFS patients. The analysis correctly classified 59.7% of the cases. 73% of the ME cases were correctly classified, and 51% of the CFS patients.

AMI = acute myocardial infarction; BEI = baroreflex effective index; BMI = body mass index; BP = blood pressure; BRS = baroreflex sensitivity; CDC = Centers for Disease Control and Prevention; CFQ = cognitive failures questionnaire; CFS = chronic fatigue syndrome; CHF = congestive heart failure; DM = depressed mood; DSM-IV = Diagnostic and Statistical Manual, fourth edition; ESS = Epworth sleepiness scale; FIS = fatigue impact scale; GHQ = general health questionnaire; HADS = Hospital Anxiety and Depression Scale; HADS-A = anxiety subscale of HADS; HADS-D = depression subscale of HADS; HF = high frequency; HFnu = high frequency normalized units; HR = heart rate; HRV = heart rate variability; HTN = hypertension; ICF = idiopathic chronic fatigue; KPS = Karnofsy Performance Scale; LF = low frequency; LFnu = low frequency normalized units; LVET = left ventricular ejection time; MANOVA = multivariate analysis of variance; ME = myalgic encephalomyelitis; ME = myalgic encephalomyelitis–international consensus criteria; MFI-20 = multidimensional fatigue inventory; MOS-SF = Medical Outcomes Study Short Form; MS = multiple sclerosis; NS = not significant; OGS = orthostatic grading scale; PEM = post exertional malaise; RA = rheumatoid arthritis; SCID = structured clinical interview for DSM-IV; SF-36 = 36-item Sort Form Survey; VO2 = volume oxygen.

Appendix Table 3. Included Studies of Harms of Diagnosis

Study, Year (Reference)	Study Design	N/Population	Findings
Åsbring et al, 2002 (41)	Qualitative study	N = 25 women (12 CFS, 13 fibromyalgia) were interviewed to the point of saturation of themes regarding stigma.	Two main aspects of stigmatization were reported 1) Women experienced their moral character being called into question. 2) They experienced distress from being psychologized by others, especially doctors (who decided in advance that problems were fictitious or psychological); and that this experience was deeply violating.
Assefi et al, 2003 (42)	Descriptive observational study of survey data	N = 555 (207 CFS, 76 fibromyalgia, 87 CFS+fibromyalgia, 31 syndromal fatigue, 154 medical conditions) of 630 (88%) patients from a university CFS clinic responded to a survey about financial, occupational, and personal consequences of their illness.	Disability outcomes reported by >20% of CFS (n = 207) group: Lower standard of living: 44% (92/207) Significant decrease in social life: 84% (174/207) Lost friends: 38% (79/207) Significant decrease in recreational activities: 90% (186/207) Of those CFS patients employed (n = 119) Taking a new job requiring fewer skills: 25% (30/119) Took a substantial pay cut: 30% (35/119)
Brimmer et al, 2013 (48)	Prospective cohort; descriptive study of patients referred to registry by provider or recruited from CFS support group.	N = 93 patients referred to CFS registry over the course of 1 year.	Review of the CFS registry referrals: 33 patients were classified as having CFS, 13 as insufficient fatigue or symptoms and 47 patients as having an exclusionary condition. 24 (65%) of the provider-referred patients and 13 (35%) of the support group referral patients met criteria for CFS.
Devasahayam et al, 2012 (49)	Descriptive observational study of survey data assessing referral letter rejections and case notes on consecutive referrals to a specialist CFS clinic.	N = 418 referrals received to CFS service.	Analysis of referral rejection letters: 52 (36%) of the reasons for rejected referrals were likely alternative psychiatric diagnosis and 67 (35%) were likely alternative medical diagnosis.
Deale and Wessely, 2000 (43)	Descriptive observational study of questionnaire data	N = 68 patients met Oxford criteria (Sharpe, 1991) for CFS completed a questionnaire asking about psychiatric diagnoses or labels given during their illness and then underwent interview to assess for those psychiatric disorders with the DSM III-R.	Reported psychiatric diagnosis 46% (31/68) given psychiatric diagnosis (usually depression) 68% (21/31) given depression diagnosis were misdiagnosed 35% (13/37) not given psychiatric diagnosis met DSM III-R criteria for treatable psychiatric disorder, present for ≥6 months
Dickson et al, 2007 (44)	Qualitative study	N = 14 people with self-reported CFS were interviewed about living with CFS.	Reported difficulties about living with CFS 71% (10/14) experienced delay in getting CFS diagnosis 57% (8/14) were prescribed antidepressants for depression diagnosis instead of CFS diagnosis Descriptive results Participants reported that they perceived many medical practitioners to hold stereotypical views of patients with CFS, namely that disease was either psychological or indicative of an affective disorder. Problems with friends and partners centered on the fact that the patient is not visibly ill, and that the symptoms are inconsistent or variable.
Green et al, 1999 (45)	Observation al descriptive study of survey data	N = 45 of 67 (67%) initially recruited patients with CFS reported perceptions of stigma.	Reported perceptions of stigma 95% reported feeling estranged 70% thought others attribute their symptoms to psychological or personality 40% felt need to be secretive about their symptoms in some circumstances
Guise et al, 2010 (46)	Qualitative study of interview data	N = 38 members of an internet-based ME/CFS support group were asked to comment on how they felt about the way medical people treated them.	Descriptive results Patients with CFS reported that health professionals lack clinical expertise and empathy; and that they encountered professionals who lacked expectation of treatability, described themselves as fortunate in terms of experiences with medical professionals, and described themselves as able to cope and actively seeking out information and treatment.

Continued on following page

Appendix Table 3–Continued

Study, Year (Reference)	Study Design	N/Population	Findings
Jason and Taylor, 2001 (40)	Randomized controlled trial, survey of perceptions	N = 105 medical trainees (Study 1) N = 141 undergraduate psychology students (Study 2) Randomly assigned to being told the case presented to them had CFS, Florence Nightingale Disease, or ME. The case studies were identical. N = 93 mental health practitioners (Study 3) Randomly assigned to 1/3 treatments for CFS, and given identical case studies of a woman with prototypic CFS symptoms, diagnosed by a physician; treatments were 1) Ampligen - IV immune modulator, 2) CBT with graded activity, or 3) cognitive coping skills therapy.	Studies 1 and 2: told case was CFS vs. Florence Nightingale Disease vs. ME Correctly diagnosed: 54% vs. 19% vs. 28%; $p < 0.01$ Disease result of as-yet-undiscovered cancer, infection or other illness: 22% vs. 47% vs. 28%; $p < 0.05$ Reported patient was likely to improve: 41% vs. 42% vs. 16%; $p < 0.05$ Study 3: Data not shown Participants assigned to Ampligen were more likely to think that the patient was correctly diagnosed as having CFS ($p < 0.05$) and also thought the patient was significantly more disabled than did individuals in the CBT with graded activity condition ($p < 0.05$)
Jason et al, 2001 (47)	Randomized controlled trial, survey of perceptions	N = 105 medical trainees (Study 1) N = 141 undergraduate psychology students (Study 2) Randomly assigned to being told the case presented to them had CFS, Florence Nightingale Disease, or ME. The case studies were identical.	Told case was CFS vs. Florence Nightingale Disease vs. ME Mean score of whether correct diagnosis (1-6 scale; 1 = not at all and 6 = very likely): 4.5 vs. 3.9 vs. 4.0; $p < 0.01$ Proportion that associated "causal factors" with diagnosis: 28% vs. 31% vs. 49%; $p < 0.01$ Mean score of whether diagnosis was associated "organ donor ship" (1-6 scale; 1 = not at all and 6 = very likely): 3.7 vs. 3.5 vs. 3.1; $p < 0.05$
Lawn et al, 2010 (50)	Case series from a specialist CFS clinic.	N = 135 patients participating in the PACE trial.	Psychiatric interview using the Structured Clinical Interview for DSM-IV Disorders 102 patients (76%) had a comorbid psychiatric diagnosis; 31% depression, 11% dysthymia, 35% anxiety, 11% social phobia, 15% specific phobia, 6% post-traumatic stress disorder and 2% obsessive compulsive disorder.
Newton et al, 2010 (51)	Case series from specialist CFS clinic	N = 260 patients referred to CFS specialist service between 2008 and 2009.	Reviewed medical notes of patients referred to CFS specialist service Of those referred, 60% were diagnosed with CFS; 40% had alternative diagnosis including other chronic disease (47%), sleep disorder (20%), psychological (15%), idiopathic fatigue (13%), cardiovascular (4%) and other (1%).
Reyes et al, 2003 (11)	Prospective cohort; random digit-dialing survey and clinical examination with 1-year follow-up telephone interview and clinical examination.	N = 3,528 subjects with fatigue 1 month duration (2762 with fatigue 6 months). 3 physicians and 2 psychiatrists independently reviewed each subject's clinical and laboratory data and classified the individual according to the CDC (Fukuda, 1994) criteria.	Descriptive results of exclusionary diagnosis identified in the telephone interview Among 1,155 subjects who had fatigue 6 months, not relieved by rest with 4 of 8 CFS symptoms, 600 had a medical or psychiatric diagnosis. Of 299 subjects without a medical/psychiatric diagnosis who underwent a clinical examination, 43 had CFS, 112 had insufficient symptoms or fatigue, 141 (47.2%) had a medical or psychiatric diagnosis that had not previously been identified and 3 were not classified.
Woodward et al, 1995 (52)	Qualitative study	N = 20 general practitioners (Study 1) and N = 50 patients with diagnosis of CFS (Study 2).	Descriptive results of interviews 14/20 physicians reluctant to diagnosis CFS (scientific uncertainties about condition, beliefs about appropriate professional practice and uncertainty about impact of diagnosis on patient's lives). 45/50 patients stated that diagnosis was the single most helpful event over the course of their illness. Described harms from not having a diagnosis (fear, anxiety, confusion, self-doubt, bitterness). Subjects in this study did not appear to endorse harm from labeling, but helpful

CBT = cognitive behavioral therapy; CDC = Centers for Disease Control and Prevention; CFS = chronic fatigue syndrome; DSM-III-R = Diagnostic and Statistical Manual third edition revised; ME = myalgic encephalopathy; PACE = Pacing, grade Activity and Cognitive behavior therapy: a randomized Evaluation.

Appendix Table 4. Measures Used as Diagnostic Tests for ME/CFS

Measure	Abbreviation	Description	Validation Studies in ME/CFS Population	Studies Using the Measure
Beck Depression Inventory (54)	BDI	Self-reported multiple-choice inventory of 21 questions for measuring the severity of depression. Scores of 0-9 indicate minimal depression, 10-18 mild depression, 19-29 moderate depression, 30-63 severe depression.	Validated in population receiving treatment for CFS (55)	Gaab et al, 2002 (35)
Chalder Fatigue Scale (28, 56)	None	Self-reported, 14- or 11-item fatigue scale. Items scored dichotomously on a 4-point scale (0,0,1,1), lower scores indicate better outcomes, total scores ≥4 designate clinically significant levels of fatigue. Note: Several different scoring methods are used for this scale.	Validated in those identified using Oxford (Sharpe, 1991) criteria (57) Validated in CFS patient meeting either Oxford (Sharpe, 1991) or CDC (Fukuda, 1994) criteria (58)	Lewis et al, 2013(28); Gaab et al, 2004 (34); Gaab et al, 2005 (36)
Chronic Fatigue Symptoms Checklist (59, 60)	CFSC	Self-reported set of 40 symptoms, 30 thought to be typical of CFS symptoms and 10 considered atypical. Each item is scored 0-4, with 0 = never suffer from it; 1 = mild or rare symptoms during the last month causing minor disruption; 2 = moderate or frequent symptoms during the last month causing major disruption; 3 = severe or very frequent symptoms during the last month unable to perform usual activities; and 4 = suffered from it previously for ≥1 month but not now.	Designed for CFS patients	Hadzi-Pavlovic et al, 2000 (31)
Cognitive Failures Questionnaire (61)	CFQ	The CFQ measure self-reported failures in perception, memory and motor function over the previous 6-months. It consists of 25 items, each graded on a scale of 5 point Likert-scale, total scores are calculated by adding the individual item scores. Final scores range from 0-100, lower scores indicate better health.	None	Lewis et al, 2013 (28)
DePaul Symptom Questionnaire (62)	DSQ	Single item of questionnaire: rate the severity of your post-exertional malaise over the past 6 months using a 5-point Likert scale with lower scores indicating less severity.	Developed for CFS population, based off of a validated measure for CFS (63)	Watson et al, 2014 (37); Jason et al, 2014 (23); Jason et al, 2012 (1)
General Health Questionnaire (64)	GHQ	A 60-item questionnaire to screen individuals for psychiatric disorders, scores are given as means and scores above 3 indicate disorders; a 30-item version of the same questionnaire uses a threshold of 6 to indicate general psychological distress.	None	Katon et al, 1991 (26); Hadzi-Pavlovic et al, 2000 (31)
Hospital Anxiety and Depression Scale (65)	HADS	Self-reported scale of 14-items for the detection of depression and anxiety in hospitalized patients. Scores range from 1-21 interpreted as: normal (0-7), mild (8-10), moderate (11-14), severe (15-21). Subscales for anxiety (HADS-A) and depression (HADS-D).	Validated in patients identified using CDC (Fukuda, 1994) criteria (66)	Lewis et al, 2013 (28); Gaab et al, 2002 (35); Gaab et al, 2005 (36)
Karnofsky Performance Scale (67)	KPS	Descriptive ordinal scale that measures the patient's ability to carry on normal activities/the degree of assistance required. The scale range is comprised of 10-point intervals from 0-100, where 0 = dead and 100 = normal, no complaints or evidence of disease. Score thresholds: 80-100 = normal health; 50-80 = an inability to work, with a varying amount of assistance needed at home; 10-40 = an inability for self care requiring the equivalence of institutional care	Validated in patients with chronic pain, but not specifically CFS (68)	Van Hoof and De Meirleir, 2005(22)
Multidimensional Fatigue Inventory (69)	MFI-20	Self-reported instrument used to measure fatigue consisting of 5 subscales: general fatigue, physical fatigue, mental fatigue, reduced motivation, and reduced activity. Each subscale has 4 statements regarding levels of fatigue experienced in the previous days (20 total) rated on a Likert-type scale from 1-5 for a final subscale score of 4-20, lower scores indicate less fatigue.	Validated in those with >12 months of fatigue (69) Validated in population self-reporting symptoms meeting CDC (Fukuda, 1994) criteria(70)	Van Hoof and De Meirleir, 2005 (22); Davenport et al, 2011 (38); Gaab et al, 2002 (35); Jason et al, 2010 (32); Tiev et al, 2003 (29)
Medical Outcome Study Short Form (71)	MOS-SF	Measures functioning and well being of 6 health concepts: physical functioning, social functioning role functioning, mental health, health perceptions, and bodily pain. Each area has varying numbers of items and are scored on scales from 1-100, with higher scores indicating better health.	Validated in patients with chronic conditions (72) Validated in those identified using Oxford (Sharpe, 1991) criteria (73)	Jason et al, 2012 (1); Jason et al, 2004 (24); Katon et al, 1991 (26)

Continued on following page

Appendix Table 4–Continued

Measure	Abbreviation	Description	Validation Studies in ME/CFS Population	Studies Using the Measure
Modified Somatic Perception Questionnaire (74)	MSPQ	Self-reported 13-item scale for patients with chronic pain or disabilities, it is used to identify somatic complaints that may be associated with psychological responses such as anxiety or depression. Each item is scored 0-3 (0 = not at all and 3 = extremely could not have been worse) for a total score of 0-39 with lower scores indicated lower general somatic symptoms.	None	Katon et al, 1991 (26)
Orthostatic Grading Scale (75)	OGS	Self-reported 5-item scale assessing for symptoms of orthostatic intolerance because of orthostatic hypotension. Each item is scored 0-4, with total score of 0-20, with lower scores indicated better health.	None	Lewis et al, 2013 (28)
Pennebaker Inventory of Limbic Languidness (76)	PILL	Self-reported 54-item questionnaire measures the tendency for someone to notice and report a broad array of physical symptoms and sensations. Each item scored from 0-4 (0 = never or almost never experienced and 4 = more than once a week) for a total score of 0-216 interpreted as: 0-21 below normal range; 22-66 well within normal range; 67-84 slightly above average, within normal range; and ≥85 top 25%.	None	Katon et al, 1991 (26)
Sickness Impact Profile 8-items (77, 78)	SIP-8	Self-reported measure of perceived impact of illness or disease on physical and psychosocial functioning, it can be self or interviewer administered. The 8 subscales used are home management, mobility, alertness behavior, sleep/rest, ambulation, social interactions, work and recreation and pastimes. A total score is calculated by addition of the weights of items (range 0–5,799). Lower scores indicate better health.	None	Gaab et al, 2002 (35); Gaab et al, 2005 (36)
36-item Short Form survey (79)	SF-36	Self-reported survey of 36 questions of patient health on 8 subscales: vitality, physical functioning, bodily pain, general health perceptions, physical role functioning, emotional role functioning, social role functioning, and mental health. The scale has a range from 0-100, with higher scores indicating better health.	Validated in those identified using CDC (Holmes, 1988) criteria (27, 80)	Aslakson et al, 2006 (25); Brown et al, 2013 (21); Jason et al, 2001 (20); Jason et al, 2013 (10); Komaroff et al, 1996 (27); Van Hoof and De Meirleir, 2005 (22); Davenport et al, 2011 (38); Jason et al, 2010 (32)
Somatization Checklist (31)	None	Self-reported set of 39 physical symptoms derived from diagnostic interview schedule for making a DSM-III/III-R diagnosis of somatization disorder. Items were answered yes or no for current and lifetime symptoms.	None	Hadzi-Pavlovic et al, 2000 (31)
Symptom Checklist-90 (81)	SCL-90	Self-reported checklist of 90 questions to assess psychological status in the following categories: somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, psychoticism.	None	Gaab et al, 2002 (35); Gaab et al, 2005 (36)
Zung Self-Rating Depression Scale (82)	ZDS	Self-reported questionnaire of 20-items that rate affective, psychological, and somatic symptoms associated with depression. Each item is rated from 1 (a little of the time) to 4 (most of the time) with final scores ranging from 20-80, interpreted as: 20-44 normal, 45-59 mildly depressed, 60-69 moderately depressed, ≥70 severely depressed.	None	Aslakson et al, 2006 (25); Hadzi-Pavlovic et al, 2000 (31)

CDC = Centers for Disease Control and Prevention; CFS = chronic fatigue syndrome; DSM III/III-R = Diagnostic and Statistical Manual third edition/third edition revised; ME = myalgic encephalomyelitis.