HHS Public Access

Author manuscript

J Health Psychol. Author manuscript; available in PMC 2021 May 01.

Published in final edited form as:

J Health Psychol. 2020 ; 25(13-14): 2053–2071. doi:10.1177/1359105318784161.

Specificitet og sensitivitet af PEM

Meta-analysis investigating post-exertional malaise between patients and controls

Abigail Brown, **Leonard A Jason**

DePaul University, USA

Abstract

Post-exertional malaise is either required or included in many previously proposed case definitions of myalgic encephalomyelitis/chronic fatigue syndrome. A meta-analysis of odds ratios (ORs; association between patient status and post-exertional malaise status) and a number of potential moderators (i.e. study-level characteristics) of effect size were conducted. Post-exertional malaise was found to be 10.4 times more likely to be associated with a myalgic encephalomyelitis/chronic fatigue syndrome diagnosis than with control status. Significant moderators of effect size included patient recruitment strategy and control selection. These findings suggest that post-exertional malaise should be considered a cardinal symptom of myalgic encephalomyelitis/chronic fatigue syndrome.

Keywords

assessment; chronic fatigue syndrome; meta-analysis; myalgic encephalomyelitis; symptoms

Introduction

Myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) is a debilitating illness (Nacul et al., 2011a) characterized by profound fatigue, neurocognitive dysfunction, unrefreshing sleep, and a worsening of the symptom complex following mental or physical activity; secondary symptoms include pain and immune, autonomic, and neuroendocrine dysfunction (Carruthers et al., 2003). The illness has been referred to as chronic fatigue syndrome (CFS) (Fukuda et al., 1994), myalgic encephalomyelitis or encephalopathy (ME) (Ramsay, 1988), ME/CFS (Carruthers et al., 2003), and most recently as systemic exertion intolerance disease (SEID) (Institute of Medicine of the National Academies, Board on the Health of Select Populations, Committee on the Diagnostic Criteria for ME/CFS (IOM), 2015). For the purposes of this article, the term “ME/CFS” will be utilized.).

The heterogeneous patient samples used in ME/CFS research may be a product of the vague and poorly operationalized diagnostic criteria that have been established (Jason et al., 1999a). Since the illness became formally recognized as CFS in the late 1980s in the United

Corresponding author: Leonard A Jason, DePaul University, Chicago, IL 60614, USA., LJASON@depaul.edu.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

States following reports of cluster outbreaks in Nevada (Buchwald et al., 1992) and New York (Bell et al., 1994), consensus for a singular case definition has yet to be reached by researchers, practitioners, and patient advocates. Thus, the diagnosis of ME/CFS is an exclusionary process that relies heavily on self-reported symptom profiles (Afari and Buchwald, 2003). Therefore, selecting the cardinal or core symptoms of the illness and developing a standardized process for assessing these symptoms is vital (King and Jason, 2004). Many attempts have been made to clarify and define a case of ME/CFS since the late 1980s (Fukuda et al., 1994; Holmes et al., 1988; Ramsay et al., 1988). More recent efforts have occurred with the IOM's (2015) SEID, Carruthers et al.'s (2003, 2011) ME/CFS Clinical Canadian Criteria (CCC), and ME International Consensus criteria (ME-ICC).

While similar themes emerge across the case definitions for CFS, ME, ME/CFS, and SEID, they diverge substantially on which symptoms should be required for a diagnosis. Clarifying the "core" symptoms for a diagnosis of ME/CFS has become a focus for the field, as has the notion that case definitions should be arrived at empirically rather than be based upon expert, clinical consensus. It has been suggested that consistent inclusion of homogeneous patient groups into studies, as well as identification of phenotypical subtypes of patients, could assist in the pursuit of biomarkers for ME/CFS (Nacul et al., 2011b), which would ultimately allow for a more circumscribed investigation into potential treatments.

Core symptoms of ME and CFS

One common approach to establishing "core symptoms" of ME/CFS has been to examine which symptoms best distinguish between individuals with ME/CFS and control groups (e.g. healthy groups or groups with other illnesses). A number of statistical approaches have been utilized in the literature to address the question of what should be considered "core" to this illness. Hawk et al. (2006) employed stepwise discriminant function analysis to examine which of the eight Fukuda et al. (1994) symptoms could best distinguish individuals with ME/CFS from those with major depressive disorder (MDD). The authors found that when entering severity ratings for the eight Fukuda symptoms into the discriminant function analysis, post-exertional malaise (PEM), unrefreshing sleep, and impaired memory/concentration were the best predictors of group membership, correctly classifying 91.1 percent of cases. Using receiver operating characteristic curve (ROC) analysis, Jason et al. (2009b) found that items loading to a PEM factor on the ME/CFS Fatigue Types Questionnaire had good sensitivity (90%) and specificity (93%) in distinguishing between patients and controls.

Factor analysis has also been utilized to inform our understanding of "core" ME/CFS domains. Brown and Jason (2014) employed an exploratory factor analysis with a well-defined patient sample on a comprehensive list of 54 ME/CFS-related symptoms, and a three-factor solution was found to fit the data. Two of these factors were easily interpretable, and provided support for both PEM and neurocognitive impairment as core domains of the illness. A third factor encompassed items relating to symptom domains that have been considered secondary such as neuroendocrine, autonomic, and immune. These findings of a PEM factor were in line with other factor analytic studies that found PEM factors in other

symptom inventories (Arroll and Senior, 2009; Friedberg et al., 2000; Jason et al., 2007, 2015d; Nisenbaum et al., 2004).

Recently, more advanced statistical methods that utilize computer-learning techniques have been implemented to determine which symptoms best distinguish patients with ME/CFS from other groups using large data sets. Using a technique called data mining, Jason et al. (2011) found that the inability to concentrate, PEM, and unrefreshing sleep were the best symptom discriminators between patients with ME and CFS and controls. A more recent study that employed dating mining with a larger sample and empirically established severity thresholds, found that fatigue, PEM, neurocognitive dysfunction, and unrefreshing sleep differentiated ME and CFS patients from controls with good accuracy (Jason et al., 2015b). When the authors utilized those four-symptom criteria to categorize patients, they found that this identified group was significantly more functionally impaired than patients who did not meet these criteria. Interestingly, this empirically derived case definition has some similarities to the recent, consensus-based SEID criteria that called for PEM, unrefreshing sleep, and either cognitive dysfunction or orthostatic intolerance to be present for a diagnosis (IOM, 2015). Given the results of these previous studies and the move toward considering PEM a “core” symptom of this illness in the most recently proposed SEID case definition, this review and subsequent meta-analysis will focus solely on this symptom.

PEM

PEM, also referred to as post-exertional neuro-immune exhaustion (Carruthers et al., 2011), is included in most case definitions for ME/CFS, although the description of this symptom varies across criteria. The Fukuda et al. (1994) criteria simply refer to it as “postexertional malaise lasting more than 24 hours,” whereas the CCC (Carruthers et al., 2003) provide much greater specificity, “an inappropriate loss of physical and mental stamina, rapid muscular and cognitive fatigability … and a tendency for other associated symptoms to worsen.” The newest criteria for the illness, SEID (IOM, 2015) describe PEM as “prolonged exacerbation of a patient’s baseline symptoms after physical/cognitive/orthostatic stress; [it] may be delayed relative to the trigger.” PEM is considered cardinal or required for diagnosis under many case definitions (Carruthers et al., 2003, 2011; Ramsay et al., 1988) but is not required for diagnosis using the Fukuda et al. (1994) or Empiric criteria (Reeves et al., 2005). A recent article that examined 53 unique ME and CFS patient samples all meeting the Fukuda criteria, found that between 24.7 and 100 percent of these patient samples had PEM, with a mean of 85 percent (McManimen et al., 2015).

However described or defined, PEM is often referred to as the most debilitating aspect of the ME/CFS symptom complex by patients (U.S. Food and Drug Administration’s Patient-Focused Drug Development Initiative, 2013), leading to profound reductions in functioning (Davenport et al., 2011). Furthermore, PEM is often cited as a primary reason that treatment protocols based on vigorous, incremental exercise may be inappropriate for individuals with this illness (Nijs et al., 2008). Those researchers and clinicians who endorse a more psychogenic explanation for the illness consider PEM the result of deconditioning or a learned fear of activity and encourage patients to treat their illness with exercise or cognitive behavioral therapy to learn strategies for re-evaluating certain illness cognitions and

adopting recovery focused cognitions (Surawy et al., 1995; White et al., 2011). However, the majority of patients prefer pacing strategies (Shepherd, 2001), whereby they learn to assess and stay within their "energy envelope" (Jason et al., 2009a) to avoid PEM, rather than pushing themselves beyond their envelope as recommended by many exercise-based therapies. Learning to stay within one's energy envelope has been associated with improved physical functioning and less PEM for some patients (Brown et al., 2013a).

Subjective reporting of PEM

Given the varied case definitional descriptions of PEM, assessing and operationalizing PEM in both clinical and research settings has been a challenge for the field. A number of self-report measures have been developed and validated to assess for PEM in patients, including the ME/CFS Fatigue Types Questionnaire (Jason et al., 2009b), the Symptom Inventory (Wagner et al., 2005), the CFS Screening Questionnaire (Jason et al., 1997), the Medical Questionnaire (Komaroff and Buchwald, 1991), and the DePaul Symptom Questionnaire (Jason et al., 2010a). These questionnaires utilize varying symptom descriptions and question stems to elicit a patient's experience of PEM. For example, the DePaul Symptom Questionnaire asks respondents to rate five PEM-related items on frequency and severity Likert-type scales (e.g. "Dead, heavy feeling after starting to exercise"; "Next day soreness or fatigue after non-strenuous, everyday activities"; and "Mentally tired after the slightest effort,"), whereas the Symptom Inventory simply asks respondents about "unusual fatigue after exertion."

In a recent study, Jason et al., 2015a applied an item that is commonly used to define PEM according to the Fukuda criteria (taken from the CFS Screening Questionnaire), "Do you feel generally worse than usual or fatigued for 24 hours or more after you have exercised?" to a clinically evaluated sample of patients with ME and CFS. Approximately 25 percent of the patients responded "no" to this question. However, when this symptom was probed differently (e.g. by a physician or by the item: "Do you experience high levels of fatigue or weakness following normal daily activity?") all of the patients appeared to have PEM. Similarly, Jason et al. (1999a) found that within a clinically evaluated ME/CFS sample, the percentage of the sample endorsing PEM ranged from 40 to 93 percent, dependent upon how the symptom was operationalized. The results from these studies demonstrate the critical role symptom operationalization plays in ME/CFS diagnosis. Although self-reported PEM has been found to be a sensitive and specific discriminator between ME/CFS patients and healthy controls, as well as between ME/CFS patients and depressed individuals (Hawk et al., 2006), the varied approach to PEM assessment across studies makes it difficult to interpret the true occurrence of self-reported PEM in patients.

Considerable effort has focused on establishing a reliable and valid case definition for ME/CFS and on investigating potential diagnostic tests for the illness. However, these efforts have been complicated by an over-reliance on clinical consensus for establishing case definitions and inconsistent application of case definitions by researchers across study sites. This has resulted in the absence of an empirically-based case definition for ME/CFS, as well as failed replication studies on potential diagnostic tests and biomarkers. One step of empirically driven case definition development is establishing which symptoms might be

able to discriminate well between patients with ME/CFS and controls (healthy controls or other illness groups). As reviewed above, one symptom thought to be “core” or “cardinal” to this illness is PEM. However, to date, there have been no meta-analyses of the findings from studies that investigate PEM differences between patients and controls. Thus, a meta-analytic approach to synthesizing the data on PEM and an investigation of potential moderators of effect size in the literature are both logical next steps in case definition development. Only one other ME/CFS meta-analysis on another core area, cognitive functioning, has been conducted (Cockshell and Mathias, 2010). Similar to methods used by Cockshell and Mathias (2010), our study assessed PEM in ME/CFS samples (as contrasted to controls) in a comparable meta-analysis. This study extracted and pooled ORs from studies that report on occurrence of self-reported PEM in patients and controls. It is hypothesized that the presence of PEM will be associated with an increased odds of having ME/CFS as measured by a 95 percent confidence interval around the mean OR that does not contain the null value, log(OR) = 0.

Method

Overview of meta-analysis

Meta-analysis is a quantitative technique for summarizing results of studies that attempt to measure the same phenomenon (Card, 2011). The primary unit of interest in meta-analysis is the effect size, or the strength or practical importance of a study’s finding beyond its statistical significance. Meta-analysis also allows for measurement of effect-size heterogeneity in the literature, and if significant heterogeneity is detected, allows for an investigation of what observable, study-level characteristics might be driving this heterogeneity. Meta-analysis is a systematic and transparent process which is becoming increasingly common in the social, physical, and medical sciences. It consists of the following steps: establishing study inclusion and exclusion criteria; conducting a thorough and systematic review of the literature for appropriate studies; coding the subsequent sample of studies on key characteristics utilizing a standardized coding protocol; computing effect sizes for individual studies; calculating the overall mean effect size and confidence interval for the phenomenon of interest; investigating the presence of and contributors to heterogeneity of effect size in the sample of studies utilizing subgroup analysis; and finally, considering and addressing the potential impact of publication bias on the findings. The guidance and recommendations of Card (2011) primarily shaped the authors’ understanding of the stages of a rigorous meta-analysis.

Inclusion criteria

Studies were included in this study that met the following criteria: (1) they reported on the presence or occurrence of PEM in both patients with ME/CFS and controls, (2) they reported sufficient information for computing effect sizes, (3) they were published between January 1988 and December 2016, (4) they investigated an adult sample (18 years or older), (5) they presented data from independent samples, and (6) they were available in English.

Literature search

Eligible studies were identified through searches of two major databases, PsycINFO, and PubMed. Another published meta-analysis in the ME/CFS field (Cockshell and Mathias, 2010) relied upon the following search terms: “chronic fatigue syndrome”; “chronic fatigue and immune dysfunction syndrome”; “chronic fatigue disorder”; “chronic fatigue-fibromyalgia syndrome”; “chronic infectious mononucleosis-like syndrome”; “myalgic encephalomyelitis”; “myalgic encephalopathy”; “post viral fatigue syndrome”; and “royal free disease.” These terms and an additional term, “myalgic encephalomyelitis/chronic fatigue syndrome” were included. To avoid potential publication bias, the ProQuest Dissertation & Theses databases were also searched.

Study-level moderators

Given the substantial variability observed across studies of ME/CFS on a number of methodological design decisions, there are many potential study-level factors that may impact the outcome of a study beyond group membership (ME/CFS or control). That is, certain aspects of a study’s design may result in a larger or smaller observed difference between patients and controls on PEM outcomes.

Recruitment method.—ME and CFS patient samples are drawn from a number of sources, and this may result in substantial variability between studies. Patients may be identified for study participation from primary care, from tertiary (or specialized care) settings, through random community-based methods, or through convenience methods. Patients identified through tertiary care settings have been found to be more severely ill than patients from community-based samples (Jason et al., 2003). Furthermore, patients identified using randomized community approaches tend to be less severely ill and are more likely to be receiving a diagnosis for the first time compared to patients recruited from primary or tertiary care (Jason et al., 2009d). Community-based recruitment also results in more ethnically and socioeconomically diverse samples because these recruitment methods are not biased to only select for individuals with access to healthcare (Jason et al., 2000b). Finally, convenient methods such as online recruitment or recruiting through support groups will likely result in patient samples that are similar to tertiary care samples, as these are likely individuals that identify with this illness and are actively involved in ME/CFS communities (Jason et al., 2015c). Thus, the method of patient recruitment utilized in the studies to be included in this meta-analysis may be an important moderator of the observed differences between ME/CFS and controls on subjective PEM experience. It is hypothesized that studies that recruit from tertiary care settings or utilize convenience methods may select for patients with more severe symptomatology, whereas studies that recruit using randomized, community-based methods may have milder symptomatology. Thus, the effect of the PEM phenomenon may be significantly greater in studies that compare controls to patients recruited from tertiary care or through convenience sampling than in studies that utilize community-based methods.

Diagnosis.—Once a patient is recruited and brought into a study, the method of ME/CFS diagnostic confirmation may also vary across studies. Many studies employ thorough physical and psychiatric evaluations to diagnosis ME/CFS, while other studies rely upon

self-reported ME/CFS or documentation from an outside medical provider to confirm diagnosis. Clinically evaluated patient samples are more homogeneous with regards to symptomatology compared to non-clinically evaluated patient samples (Johnston et al., 2013). Furthermore, accepting self-reported diagnoses with no documentation may introduce significant bias into a study, and it has been suggested that prevalence studies based on self-reported ME/CFS should be interpreted cautiously (Johnston et al., 2013). It is hypothesized that studies that employ thorough evaluations may result in more profound differences on PEM outcomes between patients and controls, as these studies may avoid erroneous inclusion of non-patients or patients with other conditions (as might occur with self-reported ME/CFS). Documentation from outside physicians may also not be sufficient as they may not have the clinical expertise that specialists have, as has been demonstrated by many studies of physicians (Anderson et al., 2011; Bayliss et al., 2014).

Case definition.—The case definition adhered to for diagnosis would also be a desirable moderator to examine, given the breadth of findings from case definitional comparison studies, but the vast majority of studies employ the Fukuda et al. (1994) criteria, and this may make subgroup analyses difficult.

Symptom measurement.—Many different approaches to assessing symptomatology are represented in the literature. Some studies utilize a validated and accepted self-report tool to assess symptomatology such as the Centers for Disease Control and Prevention (CDC) Symptom Inventory (Wagner et al., 2005), the DePaul Symptom Questionnaire (Jason et al., 2010a), or the CFS Questionnaire (Komaroff et al., 1996), while other studies utilize non-validated tools to assess symptoms (e.g. a set of questions developed for a study that are not used by other researchers). Studies may also rely upon interviewing alone. Definitions of "presence" or "occurrence" of a symptom also vary across studies (Jason et al., 2000a). While the major case definitions (Carruthers et al., 2003; Fukuda et al., 1994) state that the symptom complex must be present for at least 6 months, it is unclear if a symptom has to occur at a certain severity and frequency to be considered present. Jason et al. (2010a) attempted to operationalize the CCC (Carruthers et al., 2003) by recommending that symptoms must be rated as occurring at least "half the time" and being of at least "moderate" severity to count as truly "occurring." However, many investigators do not offer this level of specificity and simply rely upon endorsement of a symptom at any intensity to count as present. For studies that provide additional information about symptom assessment, additional moderators of effect size can be investigated. For example, how PEM was assessed (e.g. through a validated tool, physician assessment, or a non-validated tool) will be treated as a moderator, as well as how "occurrence" of PEM was defined (e.g. utilizing intensity thresholds versus occurrence). It is hypothesized that studies that utilize a validated questionnaire to assess PEM and apply some sort of intensity thresholding may find a greater effect than studies that do not.

Control selection.—Many types of controls are represented in the literature as comparison samples, including physically and mentally "healthy" samples or sedentary but otherwise healthy samples. Another common approach is to utilize other illness groups as controls, such as samples with depressive disorders or other samples that may experience

some shared symptoms such as severe fatigue (e.g. lupus, multiple sclerosis, and cancer). It is hypothesized that studies that utilize healthy controls will find a greater effect than studies that utilize other illness groups.

Coding procedure

The first author identified articles for inclusion and further coding by reviewing the title and abstract. If necessary, the full article was scanned to determine eligibility. Relevant study information was recorded using a standardized coding protocol developed by the first author. This coding protocol was developed with the proposed effect size and moderator analyses in mind, as well as other potentially relevant information.

Analytic strategy

Computing effect size.—Effect sizes were computed as ORs from outcomes from the two independent groups (patients and controls). An OR describes the strength of association between two binary variables (Bland and Altman, 2000). For this study, the two binary variables were “presence of ME/CFS” (yes/no) and “presence of PEM” (yes/no). In order to account for the sample size of a study, the OR was transformed to the log scale and then weighted by the inverse variance as proposed by Lipsey and Wilson (2001) before being averaged. This sample size weighting was done because larger studies are thought to more precisely estimate the population effect size than smaller studies.

Statistical model.—A random effects model was used due to the assumed significant variability between the studies. This is a more conservative approach than utilizing a fixed-effects model, as a random effects model accounts for random error as well as study-level variability (e.g. research design, and sample characteristics) (Hunter and Schmidt, 2000). Given what is known about the heterogeneity of study design within the ME/CFS literature as discussed in the introduction, this approach is most appropriate. Furthermore, this random effects approach allows for a more valid generalization of the present findings to studies that are not included in the analysis (Hedges and Vevea, 1998). The statistical packages “metafor” version 1.9–9 (Viechtbauer, 2016) and “meta” version 4.8–1 (Schwarzer, 2017) for R were used for all analyses.

Heterogeneity analyses.—Variability in effect size across studies was statistically tested by investigating the Cochran (1954) Q statistic. The null hypothesis for the Q statistic states that variance in effect size is due to random error alone and is not due to true differences between studies. If the Q is statistically significant, this suggests that the variance of effect size is significantly greater than 0, and thus, the null hypothesis is rejected because at least some of this variability might be explained by known study-level characteristics. Moderator analyses may then be considered appropriate in order to investigate the potential factors contributing to the effect-size variability. However, it has been suggested that the Q statistic may not do well at detecting true heterogeneity due to power issues and that a failure to reject the null should not be taken as evidence of effect-size homogeneity (Higgins et al., 2003). An alternative statistic, I^2, developed by Higgins and Thompson (2002), measures the inconsistency of results across studies. This statistics provides the percentage of variation across studies included in the meta-analysis that is due to true heterogeneity rather than

random error (ranging from 0 to 100 percent). Both the Q and I^2 statistics were investigated and considered before moving forward with moderator analyses.

Moderator analyses.—Investigating moderators using subgroup analysis in a meta-analysis can be thought of as analogous to analysis of variance (ANOVA) in an individual study; groups defined by their level of some independent variable X (e.g. patient or control) are compared with the outcome Y (e.g. fatigue level). In subgroup analysis within a meta-analysis, groups are defined by their level of some observable study characteristic (e.g. patient recruitment method and type of control sample) and compared with the outcome of mean effect size. Initially, a meta-regression with a mixed-effects model and maximum likelihood estimation was utilized to see which potential moderators significantly contributed to effect-size variability (Van Houwelingen et al., 2002). Those moderators found to be significant were further investigated by comparing the resulting subgroups for significant differences within a fixed-effects model by computing the $Q_{\text{between statistic}}$ based on ANOVA (Borenstein et al., 2009). A Bonferroni correction was utilized based on the number of planned comparisons; subgroup contrasts had to be significant at $p < 0.001$. The within-group Cochran's Q statistic was also computed for each subgroup of studies just as it was computed for the total set of studies.

Investigation of publication bias.—Publication bias is said to occur when peer-reviewed, published articles in the literature (typically the basis for systematic reviews and meta-analyses) are not truly representative of the group of studies that have actually been conducted on a given phenomenon (Rothstein et al., 2005). The "file-drawer effect" refers to the tendency for studies with significant and positive results to be published more often than studies that fail to reject the null hypothesis or studies that result in findings in the opposite direction of what was hypothesized, and thus these non-significant or negative findings are "placed in the file-drawer" rather than being submitted for publication or disseminated (Rosenthal, 1979). Thus, systematic reviews and meta-analyses may be biased due to the fact that studies that are readily available for analysis likely show a stronger overall effect of a given phenomenon than if all conducted studies were included. One strategy to combat publication bias before beginning the analysis, as mentioned above, is the inclusion of unpublished (but accessible) dissertations and theses. Once the sample of studies to be included was established and the meta-analysis was conducted, a number of statistical approaches were used to investigate the potential impact of publication bias on the results. The following approaches, as described by Card (2011), were used.

Funnel plot.—A funnel plot allows for a graphical representation of potential publication bias and is a simple scatter plot. The effect sizes of all studies were plotted (on the x-axis) relative to a measure of study size (on the y-axis; standard error was utilized for this study), and the resulting scatter plot was evaluated for symmetry and a triangular shape.

Rank correlation test.—As developed by Begg and Mazumdar (1994), a more objective assessment of funnel plot symmetry involves the computation of an adjusted rank correlation between effect size and standard error for all included studies. For each study, the variance of the effect size from the mean effect size and the standardized effect size are both

computed and used to estimate Kendall's rank correlation. If power is adequate and the correlation is significant, this is indicative of funnel plot asymmetry and potential publication bias.

Egger's linear regression.—Another evaluation of funnel plot symmetry, as developed by Sterne and Egger (2005), involves regressing the standardized effect sizes onto the standard errors. In the resulting regression equation $z_i = B_0 + B_1 + e_1$; the slope (B_1) is the mean effect size, and the intercept (B_0) is the measure of bias. Thus, a non-zero intercept value is indicative of funnel plot asymmetry or potential publication bias.

Failsafe N.—Failsafe *N* refers to the number of excluded studies with an average effect size of zero that would have to be included in the meta-analysis to lower the observed mean effect size to a non-significant level. Rosenthal (1979) introduced this concept, and it can be thought of as the number of studies that found (on average) no effect that would have to have been "filed away" in order to make this meta-analysis meaningless. The larger the number, the more robust to publication bias the findings can be thought to be.

Results

Search outcome

The search of PubMed resulted in 6208 publications, 26 of which met inclusion criteria. The search of PsycInfo resulted in 788 additional, unique publications, only three of which met inclusion criteria. The search of ProQuest Dissertations & Theses Global resulted in 157 manuscripts. Of these, only two met inclusion criteria. In the case of duplicate samples, only the first study found that utilized the sample was included. Table 1 includes a description of all included studies ($N = 31$) on key study characteristics. The complete list of citations is included as Appendix A.

Effect size

The weighted mean effect size (log(OR)) with a 95 percent confidence interval for all studies was found to be 2.34 (1.81–2.87). Thus, the odds of the presence of PEM being associated with an ME/CFS diagnosis is roughly 10.4 times more likely than the presence of PEM being associated with a non-ME/CFS diagnosis. The forest plot is a visual representation of the effect size and 95 percent confidence interval of all studies included in the meta-analysis, with studies listed on the vertical axis and effect sizes on the horizontal axis of the figure (Card, 2011). The weighted mean effect size of all studies is indicated with a black diamond, as well as a dotted line indicating the null result (e.g. a log(OR) value of 0). The forest plot is included as Figure 1.

Tests of study heterogeneity

The Cochran (1954) *Q* statistic was significant, X^2 (30) = 145.48, $p < 0.001$. The I^2 (Higgins and Thompson, 2002) was 85.8 percent, suggesting that a considerable percentage of the variability in effect estimates is due to true heterogeneity. Together, both results suggest that the included studies had significant effect-size heterogeneity that is likely not accounted for by random error alone and thus moderator analyses were appropriate to investigate.

Moderator findings

A summary of subgroup mean effect-size comparisons are included as Table 2. All subgroups of studies including more than one study had a significant within-group Q statistic aside from one subgroup. This suggests that significant effect-size variability exists even within subgroups of studies that share certain moderators. The moderators with non-significant findings within the meta-regression included: publication status, method of diagnosis, case definition, mode of PEM assessment, and thresholding. That is, these were not found to be significant moderators of overall effect-size variability and thus were not further investigated.

Patient recruitment strategy and control type were found to be significant moderators within the meta-regression. Regarding patient recruitment strategy, in studies that utilized a convenience method for recruiting individuals with ME/CFS the effect was found to be significantly greater than in studies that recruited individuals with ME/CFS through tertiary care and primary care settings. Regarding control type, in studies that utilized healthy controls or a combination of healthy controls and individuals with MDD, the effect was found to be significantly greater than in studies that utilized chronically fatigued individuals, individuals with postural orthostatic tachycardia syndrome (POTS), or the “other” category which was composed of studies that utilized combinations of other illness groups such as multiple sclerosis, lupus, or Lyme disease.

Publication bias findings

The funnel plot is included as Figure 2. Although this is a subjective visual assessment, studies with small sample sizes appear to be more variable on effect size (representing the “base” of the triangle), and as sample sizes increase, the variability in effect size decreases (representing the “point” of the triangle). The shape suggests that publication bias may not be an issue. In addition, the failsafe N was found to be 4811 ($p < 0.001$) utilizing the Rosenthal approach. Thus, 4811 studies with an average effect size of zero would have to be included in the meta-analysis to lower the observed mean effect size to a non-significant level, suggesting that these results are quite robust to potential publication bias. Begg-Mazumdar’s rank correlation test, an objective assessment of funnel plot asymmetry (Kendall’s $tau = 0.19$, $p = 0.14$) suggests publication bias is not an issue, but power may be too low to detect significance. Egger’s linear regression approach found the intercept value (estimate of potential bias) to be = 1.07, $p = 0.02$, which suggests possible publication bias. However, Higgins and Green (2011) suggest that when utilizing ORs, both Begg-Mazumdar’s and Egger’s approaches may be problematic due to the natural correlation of ORs to their standard errors. Taking this into account, and the high failsafe N value, these results seem to be robust to publication bias.

Discussion

The major finding of this meta-analysis is that the presence of subjectively reported PEM is 10.4 times more likely to be associated with an ME/CFS diagnosis than with control status. This finding can reasonably be considered robust to publication bias and strongly suggests that self-reported PEM discriminates well between ME/CFS and controls and has meta-

analytic support as a cardinal symptom of the disease. Thus, case definitions that require PEM for a diagnosis may be most appropriate for use (e.g. Carruthers et al., 2003) and should be relied upon rather than the most commonly utilized polythetic Fukuda et al. (1994) criteria.

Implications of moderator analyses

The total sample of studies ($N = 31$) and all of the study subgroups defined by study-level characteristics evidenced significant within-group variability on effect size. Thus, the hypothesis that studies on ME/CFS and PEM are heterogeneous on effect size was supported. The overall estimate of effect size was significantly impacted by two study-level moderators: patient recruitment strategy and control selection. The hypothesis that studies that utilized a healthy control group would find a stronger effect was supported by the results. It should be noted that neither of the moderators changed the overall pattern of the effect (PEM is strongly associated with ME/CFS regardless), but significantly changed the strength of the effect. That is, studies that utilized healthy individuals or a combination of healthy individuals and individuals with MDD as comparison groups found significantly higher odds of ME/CFS being associated with PEM (40.4 times more likely when utilizing healthy individuals and 81.5 times more likely when utilizing the combination) than the studies that utilized other disease groups or chronically fatigued groups as comparisons. Surprisingly, the group of studies that utilized an MDD-only control group was not significantly different from the other disease study subgroups, but this is likely due to low power as just two studies were included in the MDD-only group. These findings suggest that PEM may discriminate ME/CFS from healthy or depressed individuals more strongly than it discriminates ME/CFS from other illness groups. This fits with previous literature suggesting that ME/CFS and MDD are distinct entities (Barnden et al., 2015; Christley et al., 2013).

When considering the impact of patient recruitment strategy, the effect was much stronger in the studies that utilized convenience methods (50.9 times more likely that PEM and ME/CFS are associated) than studies recruiting patients from primary or tertiary care settings. Thus, the hypothesis that studies that recruit from tertiary care settings or utilize convenience methods may select for patients with more severe symptomatology was only partially supported. The hypothesis that studies that recruit using randomized, community-based methods would have milder symptomatology was not supported. It is somewhat counterintuitive that the convenience sampling and tertiary care sampling subgroups were significantly different, given that these strategies tend to capture similar patient groups (Jason et al., 2015c). The strength of the phenomenon in the convenience sampling subgroup compared to the other subgroups may suggest that individuals with ME/CFS that are recruited from support groups or online are some of the most profoundly ill.

Many of the moderators that were hypothesized to be of import were not significant contributors to effect-size variability (diagnostic approach, method of PEM assessment, and case definition) and thus subgroup comparison hypotheses could not be investigated. Furthermore, the intersections of many of the proposed moderators may have provided more rich information (e.g. subgroups defined by their patient recruitment strategy and case

definition used), but the resulting subgroups would have been too small for comparison. Future meta-analytic studies of ME/CFS should investigate subgroups defined by a number of study-level characteristics if power allows.

Limitations

This study has a number of limitations. Most importantly, the number of studies that met inclusion criteria is relatively low (while still being appropriate for meta-analysis). The primary reasons studies were excluded was the lack of reporting on PEM. Many studies focused exclusively on the symptom of fatigue, missing the unique element of post-exertional sickness and symptom exacerbation that PEM describes. Other studies reported just one composite somatic symptom severity score that did not allow for the teasing out of unique symptom occurrence. Ideally, it might have been possible to reach out to lead authors about the latter issue to collect these data, but this was outside the scope of the present investigation.

Regarding publication bias, while an attempt was made to include dissertations and theses, many of the abstracts that seemed promising were inaccessible and thus could not be included in this study. After assessing for publication bias, a decision was made not to contact leaders in the field for unpublished data from the time-frame of interest due to the large result of the failsafe *N* analysis. However, this could also be considered a limitation.

The use of just one coder for the meta-analysis was both a limitation and a strength. This may mean that more bias was introduced than if multiple coders were used (Buscemi et al., 2006), but also allowed for more consistency in applying the inclusion criteria and in the subsequent coding process. That is, the introduction of bias may have been more systematic than if multiple coders had been utilized. In addition, while those studies that were included provided data on independent samples, it is not possible to be fully confident that individuals with ME and CFS were not represented in more than one study.

Future directions

This meta-analysis was only focused on subjective presence of PEM. While method of PEM assessment (thresholding for frequency and severity versus occurrence alone) was considered as a moderator, it would also be important to meta-analyze PEM severity outcomes in patients versus controls. However, this may be difficult until more researchers begin reporting on the intensity of specific symptom domains rather than just reporting composite somatic symptom scores. Future meta-analyses of PEM should also focus on studies that investigate objective performance on exercise testing, and how well this testing may distinguish between patients and controls. Cognitive functioning has already been investigated meta-analytically (Cockshell and Mathias, 2010), but other core symptoms of ME and CFS (sleep dysfunction, autonomic dysfunction, pain, etc.) could be investigated in a similar way.

Conclusion

As the field continues to move toward an empirical approach to ME/CFS case definition, it is key to utilize the tool of meta-analysis to quantitatively synthesize results. While

treatment trials have traditionally been the basis for meta-analyses in the ME/CFS field, more attention should be paid to the role of meta-analysis in empirical case definition development. Meta-analysis allows for a unique type of systematic communication between researchers and permits broader claims to be made about an understanding of phenomena. This study highlights the importance of considering not only the mean effect size of a sample of studies that purport to study the same outcome but also how that effect is moderated by the study design choices of researchers. Through increased collaboration, multi-site studies, and more consistent adherence to best practices (such as considering the minimum data elements for ME/CFS research reports recommended by Jason et al. (2012b)), the field can move closer to more comparable and replicable investigations. This study lends support for PEM as a core symptom of ME/CFS that is capable of distinguishing between individuals with and without this disease.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Funding was provided by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (grant no.: HD072 208).

Appendix A

List of studies included in the meta-analysis

References

Arpino C, Carrieri MP, Valesini G, et al. (1999) Idiopathic chronic fatigue and chronic fatigue syndrome: A comparison of two case-definitions. Annali Dell'istituto Superiore di Sanita 35(3): 435–441.

Bennett BK, Goldstein D, Chen M, et al. (2014) Characterization of fatigue states in medicine and psychiatry by structured interview. Psychosomatic Medicine 76(5): 379–388. [PubMed: 25076512]

Buchwald D and Garrity D (1994) Comparison of patients with chronic fatigue syndrome, fibromyalgia, and multiple chemical sensitivities. Archives of Internal Medicine 154(18): 2049–2053. [PubMed: 8092909]

Buchwald D, Umali P, Umali J, et al. (1995) Chronic fatigue and the chronic fatigue syndrome: Prevalence in a Pacific Northwest health care system. Annals of Internal Medicine 123(2): 81–88. [PubMed: 7778839]

Cockshell SJ and Mathias JL (2013) Cognitive deficits in chronic fatigue syndrome and their relationship to psychological status, symptomatology, and everyday functioning. Neuropsychology 27(2): 230–242. [PubMed: 23527651]

De Becker P, McGregor N and De Meirleir K (2001) A definition-based analysis of symptoms in a large cohort of patients with chronic fatigue syndrome. Journal of Internal Medicine 250(3): 234–2440. [PubMed: 11555128]

Davenport TE, Stevens SR, Baroni K, et al. (2011) Diagnostic accuracy of symptoms characterizing chronic fatigue syndrome. Disability and Rehabilitation 33(19–20): 1768–1775. [PubMed: 21208154]

Eymard D, Lebel F, Miller M, et al. (1993) Human herpesvirus 6 and chronic fatigue syndrome. Canadian Journal of Infectious Diseases 4(4): 199–202. [PubMed: 22346448]

Gaudino EA, Coyle PK and Krupp LB (1997) Post-Lyme syndrome and chronic fatigue syndrome. Neuropsychiatric similarities and differences. Archives of Neurology 54(11): 1372–1376. [PubMed: 9362985]

Gibson H, Carroll N, Clague JE, et al. (1993) Exercise performance and fatiguability in patients with chronic fatigue syndrome. Journal of Neurology and Neurosurgical Psychiatry 56(9): 993–998.

Hawk C, Jason LA and Torres-Harding S (2006) Differential diagnosis of chronic fatigue syndrome and major depressive disorder. International Journal of Behavioral Medicine 13(3): 244–251. [PubMed: 17078775]

Jason LA, Helgerson J, Torres-Harding SR, et al. (2003) Variability in diagnostic criteria for chronic fatigue syndrome may result in substantial differences in patterns of symptoms and disability. Evaluation & The Health Professions 26(1): 3–22. [PubMed: 12629919]

Jason LA, Sunnquist M, Brown A, et al. (2014) Examining case definition criteria for chronic fatigue syndrome and myalgic encephalomyelitis. Fatigue 2(1): 40–56. [PubMed: 24511456]

Jason LA, Sunnquist M, Kot B, et al. (2015) Unintended consequences of not specifying exclusionary illnesses for systemic exertion intolerance disease. Diagnostics 5(2): 272–286. [PubMed: 26854153]

Johnston S, Brenu EW, Staines D, et al. (2013) The prevalence of CFS/ME: a meta-analysis. Clinical Epidemiology 5: 105–110. [PubMed: 23576883]

Kaemingk K (1992) Cardiovascular reflex function, fatigue and depression in chronic fatigue syndrome. PhD Thesis, The University of Arizona, Tucson, AZ.

Keijmel SP, Saxe J, van der Meer JW, et al. (2015) A comparison of patients with Q fever fatigue syndrome and patients with chronic fatigue syndrome with a focus on inflammatory markers and possible fatigue perpetuating cognitions and behaviour. Journal of Psychosomatic Research 79(4): 295–302. [PubMed: 26272528]

Komaroff AL, Fagioli LR, Geiger AM, et al. (1996) An examination of the working case definition of CFS. The American Journal of Medicine 100(1): 56–64. [PubMed: 8579088]

Lane TJ, Manu P and Matthews DA (1991) Depression and somatization in the chronic fatigue syndrome. American Journal of Medicine 91(4): 335–344. [PubMed: 1951377]

McDonald C, Koshi S, Busner L, et al. (2014) Postural tachycardia syndrome is associated with significant symptoms and functional impairment predominantly affecting young women: A UK perspective. BMJ Open 4(6): e004127.

MacDonald KL, Osterholm MT, LeDell KH, et al. (1996) A case-control study to assess possible triggers and cofactors in chronic fatigue syndrome. American Journal of Medicine 100(5): 548–554. [PubMed: 8644768]

Nisenbaum R, Reyes M, Unger ER, et al. (2004) Factor analysis of symptoms among subjects with unexplained chronic fatigue: What can we learn about chronic fatigue syndrome? Journal of Psychosomatic Research 56(2): 171–178. [PubMed: 15016574]

Okamoto LE, Raj SR, Peltier A, et al. (2012) Neurohumoral and haemodynamic profile in postural tachycardia and chronic fatigue syndromes. Clinical Science 122(4): 183–192. [PubMed: 21906029]

Patrick DM, Miller RR, Gardy JL, et al. (2015) Lyme disease diagnosed by alternative methods: A phenotype similar to that of chronic fatigue syndrome. Clinical Infectious Disease 61(7): 1084–1091.

Strickland PS, Levine PH, Peterson DL, et al. (2001) Neuromyasthenia and chronic fatigue syndrome in Northern Nevada/California: A ten-year follow-up of an outbreak. Journal of Chronic Fatigue Syndrome 9(3–4): 3–14.

Suhadolnik RJ, Peterson DL, Reichenbach NL, et al. (2004) Clinical and biochemical characteristics differentiating chronic fatigue syndrome from major depression and healthy control populations: Relation to dysfunction in the RNase L pathway. Journal of Chronic Fatigue Syndrome 12(1): 5–35.

VanNess JM, Stevens SR, Bateman L, et al. (2010) Postexertional malaise in women with chronic fatigue syndrome. Journal of Women's Health 19(2): 239–244.

Vercoulen JH, Hommes OR, Swanink CM, et al. (1996) The measurement of fatigue in patients with multiple sclerosis. A multidimensional comparison with patients with chronic fatigue syndrome and healthy subjects. Archives of Neurology 53(7): 642–649. [PubMed: 8929171]

Wagner L (1997) Chronic fatigue syndrome, chronic fatigue, and psychiatric disorders: An examination of functional status among a national nursing sample. Unpublished Dissertation. DePaul University.

Wessely S, Chalder T, Hirsch S, et al. (1996) Psychological symptoms, somatic symptoms, and psychiatric disorder in chronic fatigue and chronic fatigue syndrome: A prospective study in the primary care setting. American Journal of Psychiatry 153(8): 1050–1059. [PubMed: 8678174]

References

Afari N and Buchwald D (2003) Chronic fatigue syndrome: A review. American Journal of Psychiatry 160: 221–236. [PubMed: 12562565]

Anderson VR, Jason LA, Hlavaty L, et al. (2011) A review and meta-synthesis of qualitative studies on ME/CFS. Patient Education and Counseling 86(2): 147–155. [PubMed: 21571484]

Arroll MA and Senior V (2009) Symptom typology and sub-grouping in chronic fatigue syndrome. Bulletin of the IACFS/ME 17(2): 39–52.

Barnden LR, Crouch B, Kwiatek R, et al. (2015) Evidence in chronic fatigue syndrome for severity-dependent upregulation of prefrontal myelination that is independent of anxiety and depression. NMR in Biomedicine 28(3): 404–413. [PubMed: 25702943]

Bayliss K, Goodall M, Chisholm A, et al. (2014) Overcoming the barriers to the diagnosis and management of chronic fatigue syndrome/ME in primary care: A meta synthesis of qualitative studies. BMC Family Practice 15(1): 44. [PubMed: 24606913]

Begg CB and Mazumdar M (1994) Operating characteristics of a rank correlation test for publication bias. Biometrics 50(4): 1088–1101. [PubMed: 7786990]

Bell DS, Bell KM and Cheney PR (1994) Primary juvenile fibromyalgia syndrome and chronic fatigue syndrome in adolescents. Clinical Infectious Diseases 18: S21–S23. [PubMed: 8148447]

Bland JM and Altman DG (2000) Statistics notes. The odds ratio. BMJ 320(7247): 1468. [PubMed: 10827061]

Borenstein M, Hedges LV, Higgins JPT, et al. (2009) Introduction to Meta-Analysis. London: John Wiley & Sons.

Brown AA and Jason LA (2014) Validating a measure of myalgic encephalomyelitis/chronic fatigue syndrome symptomatology. Fatigue: Biomedicine, Health, & Behavior 2: 132–152.

Brown AA, Evans M, Jones N, et al. (2013a) Examining the energy envelope and associated symptom patterns in ME/CFS: Does coping matter? Chronic Illness 9: 302–311.. [PubMed: 23585632]

Brown AA, Jason LA, Evans M, et al. (2013b) Comparing the ME International Consensus Criteria and the Fukuda criteria for ME/CFS. North American Journal of Psychology 15(1): 103–120. [PubMed: 25364305]

Buchwald D, Cheney PR, Peterson DL, et al. (1992) A chronic illness characterized by fatigue, neurologic and immunologic disorders, and active human herpesvirus type 6 infection. Annals of Internal Medicine 116: 103–113. [PubMed: 1309285]

Buscemi N, Hartling L, Vandermeer B, et al. (2006) Single data extraction generated more errors than double data extraction in systematic reviews. Journal of Clinical Epidemiology 59(7): 697–703. [PubMed: 16765272]

Card NA (2011) Applied Meta-Analysis for Social Science Research. New York: The Guilford Press.

Carruthers BM, Jain AK, De Meirleir KL, et al. (2003) Myalgic encephalomyelitis/chronic fatigue syndrome: Clinical working case defintion, diagnostic and treatment protocols. Journal of Chronic Fatigue Syndrome 11(1): 7–116.

Carruthers BM, van de Sande MI, De Meirleir KL, et al. (2011) Myalgic encephalomyelitis: International consensus criteria. Journal of Internal Medicine 270(4): 327–338. [PubMed: 21777306]

Christley Y, Duffy T, Everall IP, et al. (2013) The neuropsychiatric and neuropsychological features of chronic fatigue syndrome: Revisiting the enigma. Current Psychiatry Reports 15: 353. [PubMed: 23440559]

Cochran WG (1954) The combination of estimates from different experiments. Biometrics 10(1): 101–129.

Cockshell SJ and Mathias JL (2010) Cognitive functioning in chronic fatigue syndrome: A meta-analysis. Psychological Medicine 40(8): 1253–1267. [PubMed: 20047703]

Davenport TE, Stevens SR, Baroni K, et al. (2011) Diagnostic accuracy of symptoms characterizing chronic fatigue syndrome. Disability and Rehabilitation 33: 1768–1775. [PubMed: 21208154]

Friedberg F, Dechene L, McKenzie MJ, et al. (2000) Symptom patterns in long-duration chronic fatigue syndrome. Journal of Psychosomatic Research 48: 59–68. [PubMed: 10750631]

Fukuda K, Straus SE, Hickie I, et al. (1994) The chronic fatigue syndrome: A comprehensive approach to its definition and study. Annals of Internal Medicine 121: 953–959. [PubMed: 7978722]

Hawk C, Jason LA and Torres-Harding S (2006) Differential diagnosis of chronic fatigue syndrome and major depressive disorder. International Journal of Behavioral Medicine 13(3): 244–251. [PubMed: 17078775]

Hedges LV and Vevea JL (1998) Fixed- and random-effects models in meta-analysis. Psychological Methods 3(4): 486–504.

Higgins JP and Thompson SG (2002) Quantifying heterogeneity in a meta-analysis. Statistics in Medicine 21(11): 1539–1558. [PubMed: 12111919]

Higgins JP, Thompson SG, Deeks JJ, et al. (2003) Measuring inconsistency in meta-analyses. British Medical Journal 327(7414): 557–560. [PubMed: 12958120]

Higgins JPT and Green S (eds) (2011) Cochrane Handbook for Systematic Reviews of Interventions, version 5.1.0. London: John Wiley & Sons Available at: www.handbook.cochrane.org

Holmes GP, Kaplan JE, Gantz NM, et al. (1988) Chronic fatigue syndrome: A working case definition. Annals of Internal Medicine 108: 387–389. [PubMed: 2829679]

Hunter JE and Schmidt FL (2000) Fixed effects vs. random effects meta-analysis models: Implications for cumulative research knowledge. International Journal of Selection and Assessment 8(4): 275–292.

Institute of Medicine of the National Academies, Board on the Health of Select Populations, Committee on the Diagnostic Criteria for ME/CFS (IOM) (2015) Beyond ME/CFS: Redefining an Illness. Washington, D.C: The National Academies Press.

Jason LA, Benton M, Torres-Harding S, et al. (2009a) The impact of energy modulation on physical functioning and fatigue severity among patients with ME/CFS. Patient Education and Counseling 77: 237–241. [PubMed: 19356884]

Jason LA, Brown A, Clyne E, et al. (2012a) Contrasting case definitions for chronic fatigue syndrome, myalgic encephalomyelitis/chronic fatigue syndrome and myalgic encephalomyelitis. Evaluation and the Health Professions 35: 280–304. [PubMed: 22158691]

Jason LA, Corradi K and Torres-Harding S (2007) Toward an empirical case definition of CFS. Journal of Social Service Research 34(2): 43–54.

Jason LA, Evans M, Porter N, et al. (2010a) The development of a revised Canadian myalgic encephalomyelitis chronic fatigue syndrome case definition. American Journal of Biochemistry and Biotechnology 6(2): 120–135.

Jason LA, Evans M, So S, et al. (2015a) Problems in defining post-exertional malaise. Journal of Prevention and Intervention in the Community 43: 20–31. [PubMed: 25584525]

Jason LA, Jessen T, Porter N, et al. (2009b) Examining types of fatigue among individuals with ME/CFS. Disability Studies Quarterly 29(3).

Jason LA, King CP, Richman JA, et al. (1999a) U.S. case definition of chronic fatigue syndrome: Diagnostic and theoretical issues. Journal of Chronic Fatigue Syndrome 5(3–4): 3–33.

Jason LA, King CP, Taylor RR, et al. (2000a) Defining chronic fatigue syndrome: Methodological challenges. Journal of Chronic Fatigue Syndrome 7(3): 17–32.

Jason LA, Kot B, Sunnquist M, et al. (2015b) Comparing and contrasting consensus versus empirical domains. Fatigue: Biomedicine, Health & Behavior 3: 63–74.

Jason LA, Najar N, Porter N, et al. (2009c) Evaluating the centers for disease control's empirical chronic fatigue syndrome case definition. Journal of Disability Policy Studies 20(2): 93–100.

Jason LA, Paavola E, Porter N, et al. (2010b) Frequency and content analysis of chronic fatigue syndrome in medical textbooks. Australian Journal of Primary Health 16: 174–178. [PubMed: 21128580]

Jason LA, Plioplys AV, Torres-Harding S, et al. (2003) Comparing symptoms of chronic fatigue syndrome in a community-based versus tertiary care sample. Journal of Health Psychology 8(4): 459–464. [PubMed: 19127712]

Jason LA, Porter N, Brown M, et al. (2009d) CFS: A review of epidemiology and natural history studies. Bulletin of the IACFS/ME 17: 88–106. [PubMed: 21243091]

Jason LA, Richman JA, Rademaker AW, et al. (1999b) A commuity-based study of chronic fatigue syndrome. Archives of Internal Medicine 159(18): 2129–2137. [PubMed: 10527290]

Jason LA, Ropacki MT, Santoro NB, et al. (1997) A screening scale for CFS: Reliability and validity. Journal of Chronic Fatigue Syndrome 3: 39–59.

Jason LA, Skendrovic B, Furst J, et al. (2011) Data mining: Comparing the empiric CFS to the Canadian ME/CFS case definition. Journal of Clinical Psychology 68: 41–49. [PubMed: 21823124]

Jason LA, Sunnquist M, Brown A, et al. (2015c) Chronic fatigue syndrome versus systemic exertion intolerance disease. Fatigue: Biomedicine, Health & Behavior 3: 127–141.

Jason LA, Sunnquist M, Brown A, et al. (2015d) Factor analysis of the DePaul Symptom Questionnaire: Identifying core domains. Journal of Neurology and Neurobiology. Epub ahead of print 17 September. DOI: 10.16966/2379-7150.114.

Jason LA, Sunnquist M, Kot R, et al. (2015e) Unintended consequences of not specifying exclusionary illnesses for systemic exertion intolerance disease. Diagnostics 5: 272–286. [PubMed: 26854153]

Jason LA, Taylor RR, Kennedy CL, et al. (2000b) Chronic fatigue syndrome: Sociodemographic subtypes in a community-based sample. Evaluation & The Health Professions 23(3): 243–263. [PubMed: 11067190]

Jason LA, Unger ER, Dimitrakoff JD, et al. (2012b) Minimum data elements for research reports on CFS. Brain, Behavior, and Immunity 26(3): 401–406.

King CP and Jason LA (2004) Improving the diagnostic criteria and procedures for chronic fatigue syndrome. Biological Psychology 68: 87–106.

Komaroff AL and Buchwald D (1991) Symptoms and signs of chronic fatigue syndrome. Reviews of Infectious Diseases 13: S8–S11. [PubMed: 2020806]

Lipsey MW and Wilson DB (2001) Practical Meta-Analysis. Thousand Oaks, CA: SAGE.

McManimen SL, Jason LA and Williams YJ (2015) Variability in symptoms complicates utility of case definitions. Fatigue: Biomedicine, Health & Behavior 3: 167–172.

Nacul LC, Lacerda EM, Campion P, et al. (2011a) The functional status and well-being of people with myalgic encephalomyelitis/chronic fatigue syndrome and their carers. BMC Public Health 11: 402. [PubMed: 21619607]

Nacul LC, Lacerda EM, Pheby D, et al. (2011b) Prevalence of myalgic encephalomyelitis/chronic fatigue syndrome in three regions of England: A repeated cross-sectional study in primary care. BMC Medicine 9: 91. [PubMed: 21794183]

Nijs J, Paul L and Wallman K (2008) Chronic fatigue syndrome: An approach combining self-management with graded exercise to avoid exacerbations. Journal of Rehabilitation Medicine 40(4): 241–247. [PubMed: 18382818]

Nisenbaum R, Reyes M, Unger ER, et al. (2004) Factor analysis of symptoms among subjects with unexplained chronic fatigue: What can we learn about chronic fatigue syndrome. Journal of Psychosomatic Research 56(2): 171–178. [PubMed: 15016574]

Ramsay AM, David AS, Wessely S, et al. (1988) Myalgic encephalomyelitis, or what? Lancet 332(8602): 100–101.

Reeves WC, Wagner D, Nisenbaum R, et al. (2005) Chronic fatigue syndrome—A clinically empirical approach to its definition and study. BMC Medicine 3: 19. [PubMed: 16356178]

Rosenthal R (1979) The file drawer problem and tolerance for null results. Psychological Bulletin 86(3): 638–641.

Rothstein HR, Sutton AJ and Borenstein M (2005) Publication bias in meta-analysis In: Rothstein HR, Sutton AJ and Borenstein M (eds) Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments. West Sussex: John Wiley & Sons, pp. pp. 1–7.

Schwarzer G (2017) Meta: General package for meta-analysis version. R package version, pp. 48–41.

Shepherd C (2001) Pacing and exercise in chronic fatigue syndrome. Physiotherapy 87(8): 395–396.

Sterne JAC and Egger M (2005) Regression methods to detect publication and other bias in meta-analysis In: Rothstein HR, Sutton AJ and Borenstein M (eds) Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments. West Sussex: John Wiley & Sons, pp. 99–110.

Surawy C, Hackmann A, Hawton K, et al. (1995) Chronic fatigue syndrome: A cognitive approach. Behavior Research and Therapy 33(5): 535–544.

U.S. Food and Drug Administration's Patient-Focused Drug Development Initiative (2013) The Voice of the Patient: CFS and ME. Silver Spring, MD: Center for Drug Evaluation and Research.

Van Houwelingen HC, Arends LR and Stijnen T (2002) Advanced methods in meta-analysis: Multivariate approach and meta-regression. Statistics in Medicine 21: 589–624. [PubMed: 11836738]

Viechtbauer W (2016) Metafor: Meta-analysis package for R. R package version 19–9. DOI: 10.1002/sim.1040.

Wagner D, Nisenbaum R, Heim C, et al. (2005) Psychometric properties of the CDC Symptom Inventory for assessment of chronic fatigue syndrome. Population Health Metrics 3: 8. [PubMed: 16042777]

White PD, Goldsmith KA, Johnson AL, et al. (2011) Comparison of adaptive pacing therapy, cognitive behavior therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): A randomized trial. Lancet 377(9768): 823–836. [PubMed: 21334061]

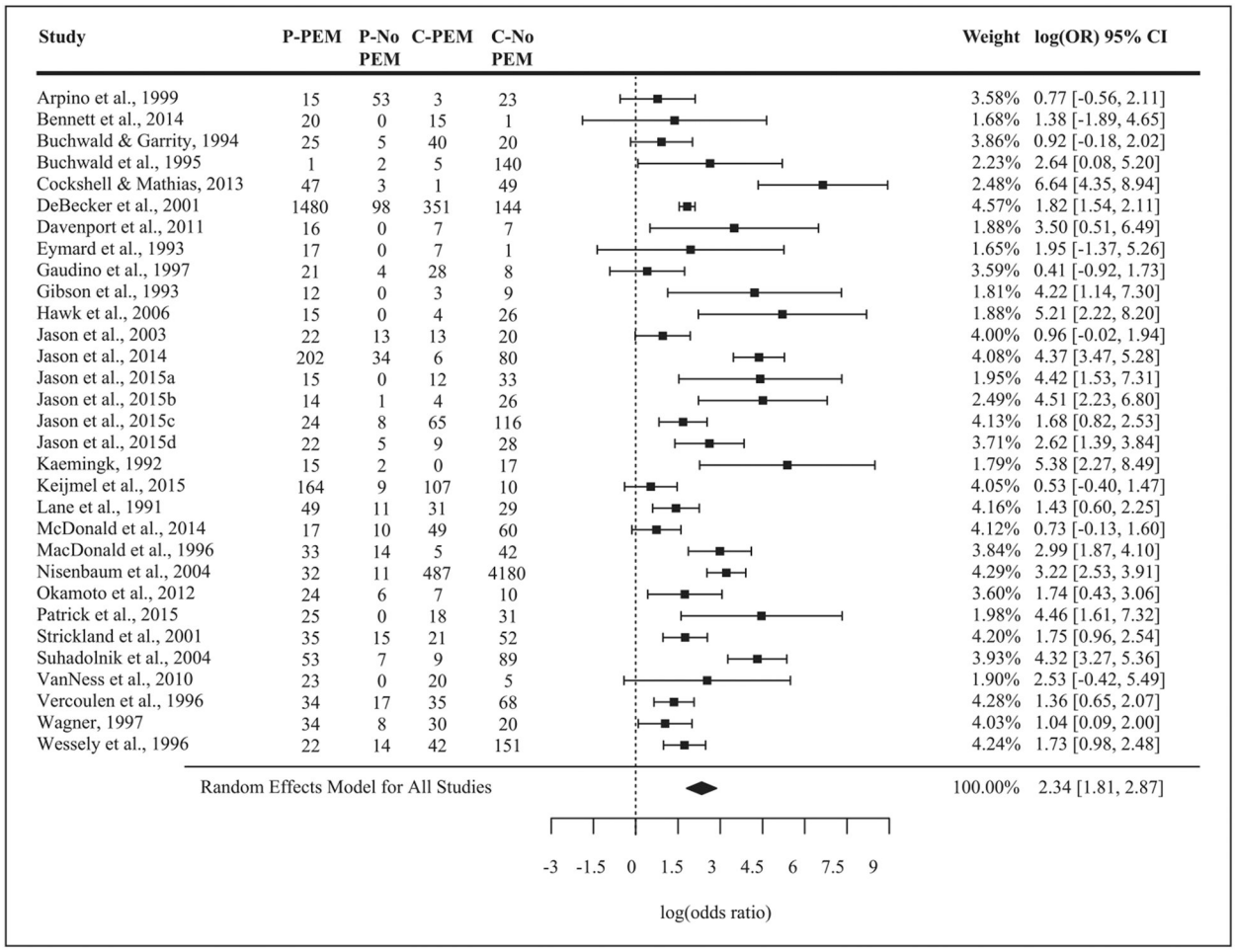

Figure 1. Forest plot, random effects model.

P-PEM: individuals with ME and CFS with post-exertional malaise; P-No PEM: individuals with ME and CFS without post-exertional malaise; C-PEM: controls with post-exertional malaise; C-No PEM: controls without post-exertional malaise; log(OR): log (odds ratio).

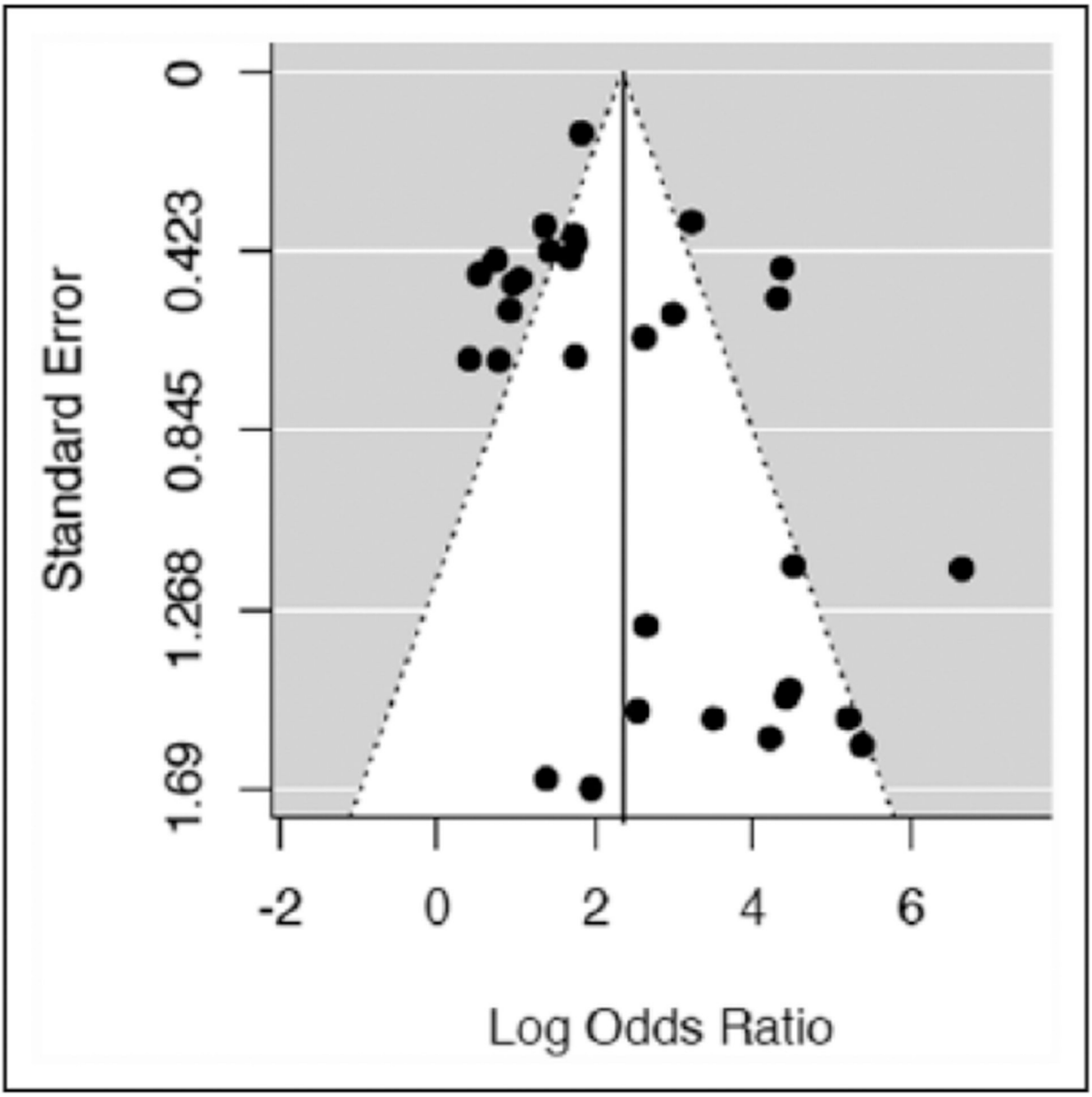

Figure 2.
Funnel plot.

Table 1.

Summary of Studies (N = 31).

Study	Patient recruitment	Diagnosis	Case Def.	PEM assessment	Controls
Arpino et al., 1993	Primary care	MD: Study	Fukuda and Holmes	Interview; Occur	CF
Bennett et al., 2014	Unknown	MD: Study	Fukuda	Interview; Occur	MDD
Buchwald and Garrity, 1994	Tertiary care	MD: Study	Holmes	NV Quest; Occur	FM and MCS
Buchwald et al., 1995	Primary care	MD: Study	Fukuda	NV Quest; Occur	CF and HC
Cockshell and Mathias, 2013	Primary and tertiary care	MD: Study	Fukuda	CDC-SI; Occur	HC
De Becker et al., 2001	Tertiary care	MD: Study	Fukuda	GSC; Sev	CF
Davenport et al., 2011	Convenience	Self-report	Fukuda	NV Quest; Occur	HC
Eymard et al., 1993	Tertiary care	MD: Study	Holmes	Interview; Occur	CF
Gaudino et al., 1997	Tertiary care	MD: Study	Fukuda	NV Quest; Freq	Lyme
Gibson et al., 1993	Tertiary care	MD: Outside	Oxford	Interview; Occur	HC
Hawk et al., 2006	Convenience	MD: Outside	Fukuda	NV Quest; Occur	MDD and HC
Jason et al., 2003	Random- CB	MD: Study	Fukuda and London	NV Quest; Occur	CF
Jason et al., 2014	Convenience	MD: Outside	Fukuda	DSQ; Freq and Sev	HC
Jason et al., 2015a	Convenience	MD: Study	Fukuda	Interview; Occur	MS, lupus, and HC
Jason et al., 2015b	Convenience	MD: Outside	Fukuda	NV Quest; Occur	MDD and HC
Jason et al., 2015c	Random- CB	MD: Study	Fukuda	Interview; Occur	CF and HC
Jason et al., 2015d	Convenience	MD: Outside	Fukuda	CDC-SI; Freq/Sev	MDD
Kaemingk, 1992 *	Tertiary care	MD: Outside	Unknown	NV Quest; Sev	HC
Keijmel et al., 2015	Tertiary care	MD: Outside	Fukuda	NV Quest; Occur.	Q-Fever
Lane et al., 1991	Tertiary care	MD: Study	Holmes	Interview; Occur	CF
McDonald et al., 2014	Tertiary care	Self-report	Unknown	NV Quest; Occur	POTS
MacDonald et al., 1996	Tertiary care	MD: Study	Holmes	Interview; Freq	HC
Nisenbaum et al., 2004	Random- CB	MD: Study	Fukuda	NV Quest; Occur	CF and HC
Okamoto et al., 2012	Tertiary care	MD: Study	Fukuda	Unknown	POTS
Patrick et al., 2015	Convenience	MD: Study	Canadian	NV Quest; Occur	Lyme, lupus, and HC
Strickland et al., 2001	Tertiary care	MD: Outside	Fukuda	NV Quest; Occur	CF
Suhadolnik et al., 2004	Tertiary care	MD: Study	Fukuda and Holmes	NV Quest; Occur	HC and MDD
VanNess et al., 2010	Tertiary care	MD: Outside	Fukuda	NV Quest; Occur	HC

Study	Patient recruitment	Diagnosis	Case Def.	PEM assessment	Controls
Vercoulen et al., 1996	Tertiary care	MD: Study	Unknown	NV Quest, Freq	HC and MS
Wagner, 1997 *	Other	MD: Review	Fukuda	NV Quest, Occur	CF
Wessely et al., 1996	Primary care	MD: Study	Fukuda	NV Quest, Occur	HC

Random-CB: randomized community-based sampling; Occur: reported occurrence of PEM only w/o thresholding; NV Quest: non-validated questionnaire used to assess PEM; CDC-SI: Centers for Disease Control Symptom Inventory; GSC: Goldstein Symptom Checklist; Sev: utilized some threshold of severity to assess PEM; Freq: utilized some threshold of frequency to assess PEM; DSQ: DePaul Symptom Questionnaire; CF: chronic fatigue or idiopathic chronic fatigue; MDD: major depressive disorder; FM: fibromyalgia; MCS: multiple chemical sensitivities; HC: healthy controls; MS: multiple sclerosis; POTS: postural orthostatic tachycardia syndrome.

*Dissertation.

Table 2.

Subgroup comparisons.

Moderators	log (odds ratio)	No. of studies	$Q_{within\ group}$
Publication status			
Published	2.33	29	137.32**
Dissertation	2.65	2	6.82
Recruitment approach[3]			
Tertiary care	1.85_{a}	15	55.8**
Convenience	3.93_{ab}	7	6.59**
Primary care	1.57_{b}	3	2.22**
Randomized community-based	2	3	15.94**
Other	2	3	19.54*
Diagnosis			
MD-Study	2.22	18	77.89**
MD-Outside	2.89	10	52.76**
Self-report	0.95	2	3.03*
Other	1.04	1	0
Case definition			
Fukuda	2.48	19	90.01**
Holmes	1.76	4	7.47**
Unknown	1.23	3	8.21**
Other	2.68	3	8.29**
Fukuda and Holmes	2.59	2	16.79*
PEM mode of assessment			
Non-validated questionnaire	2.14	18	85.56**
Interview	1.96	8	13.01**
Validated questionnaire	3.61	4	43.49**
Other	1.74	1	0

Moderators	log (odds ratio)	No. of studies	$Q_{within\ group}$
PEM thresholding			
No	2.33	23	99 97**
Yes	2.48	7	45.03**
Other	1.74	l	0
Control[a]			
HC	3.7_{abc}	8	32.23**
CF	$l.5_{ad}$	7	6.97**
Other	1.13_{be}	6	13.47**
CF and HC	2.51	3	7.52**
MDD and HC	4.43_{def}	3	0.31**
MDD	2.46	2	0.48**
POTS	$l.04_{cf}$	2	1.58**

MD–Study: physicians with ME/CFS expertise diagnosed cases; MD–Outside: physician without ME/CFS expertise diagnosed cases; HC: healthy controls; CF: chronic fatigue; POTS: postural orthostatic tachycardia syndrome; MDD: major depressive disorder; PEM: post-exertional malaise.

[a]Significant moderator category based on meta-regression.

Similar subscripts within a category for a given moderator reveal a significant contrast on $Q_{between\ group}$ at $p < 0.001$.

*$p < 0.05$,

**$p < 0.001$ on $Q_{within\ group}$.