Heterogeneity in Chronic Fatigue Syndrome – empirically defined subgroups from the PACE trial

Thomas E Williams[1], Trudie Chalder[2], Michael Sharpe[3], Peter D White[1]*

[1]Center for Psychiatry, Wolfson Institute of Preventive Medicine, Barts and the London School of Medicine, Queen Mary University of London, UK

[2]Academic Department of Psychological Medicine, King's College London, Weston Education Centre, London, UK

[3]Psychological Medicine Research, Department of Psychiatry, University of Oxford, Oxford, UK

* Corresponding Author - p.d.white@qmul.ac.uk

Original Funding: Medical Research Council, Department of Health (England), the Department for Work and Pensions, and the Chief Scientist Office (NHS Scotland)

Keywords: chronic fatigue syndrome, heterogeneity, nosology, subgroups, latent class analysis

Word count: 4,284

ABSTRACT

Background

Chronic fatigue syndrome is likely to be a heterogeneous condition. Previous studies have empirically defined subgroups using combinations of clinical and biological variables. We aimed to explore the heterogeneity of chronic fatigue syndrome.

Methods

We used baseline data from the PACE trial, which included 640 participants with chronic fatigue syndrome. Variable reduction, using a combination of clinical knowledge and principal component analyses, produced a final dataset of 26 variables for 541 patients. Latent class analysis was then used to empirically define subgroups.

Results

The most statistically significant and clinically recognisable model comprised five subgroups. The largest, "core" subgroup (33% of participants), had relatively low scores across all domains and good self-efficacy. A further three subgroups were defined by: the presence of mood disorders (21%); the presence of features of other functional somatic syndromes (such as fibromyalgia or irritable bowel syndrome) (21%); or by many symptoms – a

group which combined features of both of the above (14%). The smallest "avoidant-inactive" subgroup was characterised by physical inactivity, belief that symptoms were entirely physical in nature, and fear that they indicated harm (11%). Differences in the severity of fatigue and disability provided some discriminative validation of the subgroups.

Conclusions

In addition to providing further evidence for the heterogeneity of chronic fatigue syndrome, the subgroups identified may aid future research into the important aetiological factors of specific subtypes of CFS and the development of more personalised treatment approaches.

Introduction

Chronic fatigue syndrome (CFS) is a chronic disabling condition with an estimated population prevalence of 0.2-0.4% (National Institute for Health and Clinical Excellence 2007). Opinions differ as to whether myalgic encephalomyelitis (ME) and CFS are the same condition (Prins et al. 2006). In the absence of reliable biomarkers or clinical signs, diagnosis is based upon self-reported symptoms and the exclusion of alternative explanatory diagnoses. Several case definitions are commonly used for research (Sharpe et al. 1991; Fukuda et al. 1994; The National Task Force on Chronic Fatigue Syndrome 1994), although of the twenty published to date, all assessments of reproducibility and validation are limited by the absence of a gold standard. (Brurberg et al. 2014)

The reliance upon self-reported symptoms has led to doubts about the validity of CFS as an aetiologically homogenous diagnosis (Wakefield 2013; Komaroff 2015). As such, the last 20 years has seen much research examining the potential heterogeneity of CFS, with multiple attempts to empirically define cases and subgroups.

Initial work by Hickie *et al* used symptoms and demographics to empirically define a core group and a smaller polysymptomatic subgroup (Hickie et al. 1995; Wilson et al. 2001). Further studies have suggested empirical subgroups of CFS resembling fibromyalgia (Hadzi-Pavlovic et al. 2000), whilst others found that CFS and fibromyalgia could not be empirically distinguished (Sullivan et al. 2002).

The inclusion of biological variables in studies suggests these may also be important in defining illness subgroups. These include body-mass index (BMI); sleep-disordered breathing and insomnia; sympathetic nervous system activity; and endocrine features of metabolic strain (Vollmer-Conna et al. 2006; Aslakson et al. 2009). When such biological variables were combined with the variables of many symptoms and depression, they could empirically identify validated subgroups of CFS cases and separate them from well controls (Aslakson et al. 2006).

The PACE trial (White et al. 2011) found that both cognitive behaviour therapy (CBT) and graded exercise therapy (GET) improved both fatigue and physical function over 1 year when compared to standardised medical therapy alone. This study aimed to use the large database from the PACE trial baseline assessment to explore the heterogeneity of CFS, using not only self-reported measures of beliefs, behaviours, mood and co-morbidities, but also objective measures of activity, sleep and fitness.

Methods

Data collection

Methods of participant recruitment and data collection for the trial have been described elsewhere (White et al. 2007; White et al. 2011). Briefly, 3158 new patients attending specialist chronic fatigue syndrome services at six centres in the UK were assessed for suitability. CFS was defined using the Oxford criteria (Sharpe et al. 1991), requiring disabling fatigue to be the primary complaint, with exclusion of alternative medical or psychiatric diagnoses. 640 patients were recruited and gave informed consent to participate. Participants were subsequently randomly allocated to adaptive pacing therapy, cognitive behavioural therapy, or graded exercise therapy as supplements to specialist medical care, or to specialist medical care alone.

Only baseline data was used for this study. It included a maximum dataset of 961 original or derived variables for each participant (see table A1 for data sources), though the actual number of data-points for each patient was far fewer than this. The participant with the most data determined the number of variables within each category, hence, for example, medications were listed from *Medication 1* to *Medication 25,* since at least one participant was taking 25 medications. As separate variables documented dosage, frequency, duration, etc, of treatment, there were a total of 225 variables related to medication, most of which would not actually contain any data for any individual patient. Similar variable multiplication applied to healthcare utilisation, previous medical history and co-morbidities. In accordance with previous literature, variable reduction was initially undertaken

rationally, based upon previous research, clinical experience and logic (figure 1). For example, data on medication use was excluded in order to avoid defining subgroups by their treatment (Aslakson et al. 2006), as was healthcare utilisation data as this was felt to be influenced by many unrelated variables, such as availability of healthcare, geographical location, and previous personal experiences of healthcare.

Despite the large number of variables attributed to previous medical history and co-morbidities, the majority of individual diagnoses, if present, were rare within the cohort. When present as co-morbid conditions, they were invariably also reported in past medical history, and the few more common co-morbidities displayed considerable overlap with other variables (e.g. mood or anxiety disorders covered in Standard Clinical Interview for DSM (SCID) diagnoses, fibromyalgia diagnoses overlapping with responses to questions on joint, muscle or back pain). The resulting multicollinearity within the dataset prevented principal component analysis (PCA), and infrequent diagnoses could contribute little to the overall variability of the dataset. All items from past medical history were therefore removed, as were co-morbidities that we thought were better represented elsewhere within the data. Infrequent or absent SCID diagnostic variables were removed due to the same rationale. Many other variables were then removed, since they constituted single items from scales where we considered that the summary scale scores or subscale summary scores would be more useful (e.g. Hospital Anxiety and Depression Scale, Cognitive Behavioural Response Questionnaire), or because we felt the variables were either unlikely to relate to subgroups of CFS, or represented outcomes of the condition, rather than causal factors (e.g. EQ-5D, further economic and demographic data).

Substantial overlap also existed between many of the remaining questionnaires employed. The exclusion of overlapping variables is important in order to prevent such predictable correlations from producing artificial classes during latent class analysis (LCA). For example, *CDC CFS criteria Q7 Unrefreshing sleep* vs. *Jenkins Sleep Scale Q4 Wake up after your usual amount of sleep feeling tired*

or worn out? Similarly, some data were represented in multiple variables, such as the derived daytime activity sub-groups and mean daytime activity. In both cases, correlation matrices were utilised to assist in the identification of such overlaps, and a decision was made on which variable to retain based upon which was believed to be the most clinically useful or significant. The removal of correlated pairs was again also necessary to reduce the multicollinearity of the dataset, a requirement for principal component analysis.

Principal component analysis

PCA was employed to assist our decisions on variable reduction, as in previous studies (Hickie et al. 1995; Vollmer-Conna et al. 2006; Aslakson et al. 2009). Patients with missing data were excluded from the analysis, and non-binary variables for the remaining 541 patients were then dichotomised via median split to produce a binary categorical dataset. PCA was then performed, using SPSS version 22, on the orthogonally rotated correlation coefficient matrix of variables. PCA describes the total variability of a dataset whilst reducing it into a number of independent principal components. Each principal component was created by the weighted sum of multiple correlated individual variables. The first PCA described the maximum variability of the dataset. Each subsequent analysis did the same for the remainder of the variance, whilst remaining uncorrelated with all preceding principal components. Each component was typically defined by a few highly loaded variables, followed by many other variables with low loading values and hence minimal contribution. The utility of the technique is in identifying variables that achieve only very low loading across all identified principal components. Such variables were therefore removed from the dataset with minimal loss of the overall variance. By using PCA we were able to reduce substantially the number of variables needed in the subsequent latent class analysis.

Latent class analysis

In contrast to PCA, where the weighted variables are allocated to principal components, we used latent class analysis (LCA) (Latent Gold v5.0) to allocate each individual patient to a latent class. This

was achieved based upon their responses to the variables included in the analysis, which are described in detail in appendix A2. The assumption was that where there was an association between the responses of different patients to such observed variables, this was due to shared membership of an unobserved latent class. Class membership was then assigned to each patient in order to maximise the amount of variation of the dataset explained by class membership alone. The number of latent classes to be derived was specified prior to each analysis. In order to determine the optimum number of classes, the analysis was therefore repeated multiple times, each time requesting a higher number of classes. The derived models were then compared using statistical indices, together with an assessment of their perceived clinical utility. In keeping with previous work (Vollmer-Conna et al. 2006; Aslakson et al. 2009), we used the Akaike Information Criterion (with 3 as penalising factor) (AIC3) to determine the optimum number of classes; lower AIC3 values indicated a better fit.

Validation of latent classes

In the absence of a gold standard for comparison, initial assessments of validity included comparisons to previous heterogeneity studies. Additionally, we sought to demonstrate discriminative validity of the empirically derived subgroups by comparing them across pre-specified measures of functional and symptomatic severity, not included in either the PCA or LCA. These variables were chosen a priori, in keeping with the assessments of validity successfully utilised in previous studies (Aslakson et al. 2006; Aslakson et al. 2009). The Chalder fatigue questionnaire, Likert scoring (CFQ), the work and social adjustment scale (WSAS), and the 36 item short form health survey, physical function subscale (SF-36) were compared across subgroups using the Kruskal-Wallis test. (Ware & Sherbourne 1992; Mundt et al. 2002; Cella, Sharpe, et al. 2011) The SF-36 provides a measure of physical function, whilst the WSAS captures a more global measure of impairment or participation in life. In addition, we explored the relationship of our empirically derived subgroups with CDC and London criteria for CFS and ME, respectively, (The National Task Force on Chronic

Fatigue Syndrome 1994; Reeves et al. 2003) as well as age, sex and illness duration, using Chi-squared tests.

Results

Following our initial review of the dataset, 100 variables were chosen for analysis; 51 were removed as they represented correlations and overlaps (figure 1), leaving 49 binary, categorical variables in the PCA. Scree plots were used to determine the number of principal components to analyse based upon the point of inflection. PCA was used iteratively, repeatedly removing low-loading variables and re-running the analysis until a final dataset of 30 variables was achieved, each of which contributed substantially to the overall variance of the dataset (Table 1; see Table A2 for explanation of these variables).

Initial LCA modelling revealed that 4 of the 30 variables (mean total sleep time, age, Borg Scale / % of maximum heart rate achieved [following step test], and body mass index) were contributing insignificantly to the analysis, and these were removed (figure 1). Models were therefore derived using 26 variables. The minimum AIC3 value was reached for the 6 class solution, but was only marginally higher for 5 classes (16,372 and 16,376, respectively). Both were therefore evaluated for perceived clinical utility.

The 5 class solution is shown in table 2 and figure 2. The largest subgroup ("Core", 33%) was determined by less severe scores across the majority of the variables, with an added association with good self-efficacy. This class had an absence of the features associated with functional somatic syndromes (FSS), which we define here as including the features of irritable bowel syndrome (IBS – stomach pain, constipation/diarrhoea, indigestion) or the features of fibromyalgia (FM – back pain, joint pain). There were also low scores for anxiety and depression. The quality of sleep was both subjectively and objectively good, with high daytime activity and low scores on the cognitive behavioural response questionnaire (CBRQ).

The second largest group ("Mood affected", 21%), was defined by low to moderate scores on features of FSS, but very high scores on co-morbid major depressive disorder (MDD), co-morbid generalised anxiety disorder (GAD) and total HADS scores. CBRQ scores were moderate to high, particularly embarrassment and symptom focusing, and self-efficacy scores were moderate to low. Interestingly, the majority of patients in this class believed their symptoms to be the results of both physical and psychological factors (as opposed to purely or predominantly physical), and they had the highest levels of fitness.

The third group ("FSS", 21%), in contrast, had the highest scores across features of FSS, whilst MDD, GAD and HADS-total were low to moderate. Sleep was subjectively poor with low sleep efficiency, and CBRQ scores were moderately elevated.

The fourth group ("Polysymptomatic", 14%), had high scores on both features of FSS and mood. Physical symptoms of panic disorder, such as chest pain, palpitations and dyspnoea were also high, with a high proportion reporting new headache and dizziness. They had the highest responses to the CBRQ, very low self-efficacy, and largely believed their symptoms to be both physical and psychological in nature. Sleep was subjectively and objectively poor.

The smallest group ("Avoidant-Inactive", 11%), as for the core group, had low scores on features associated with FSS, anxiety and depression. In contrast, however, they had higher scores on the CBRQ, particularly for behaviour avoidance, fear avoidance and damage beliefs. In keeping with this, they were the most likely to believe their symptoms were physical in nature. Their sleep efficiency was low despite subjectively good quality sleep, and they had poor fitness and low daytime activity.

The 6 class solution (data not shown) produced broadly similar groups. The mood group, however, was split into two. The larger division (18%) had low scores for symptoms of panic disorder, moderate CBRQ scores and high sleep efficiency. The smaller division (12%) had moderate features of panic disorder, high CBRQ scores, low sleep efficiency and higher daytime activity.

We concluded that the subgroups of the 5 class solution were most clinically useful, and hence this solution was taken forward for validation.

Validation

The initial inspection of the subgroups led to the consensus that they were consistent with the existing literature on CFS, particularly our “Core”, “Polysymptomatic” and “Mood” subgroups described previously.

Comparison across the pre-specified measures of severity for the 5 class solution are shown in table 3 and summarised in figure 3. The differences were highly significant across the subgroups on all validating variables, providing support for discriminative validity. The polysymptomatic and FSS groups had the most fatigue, and the polysymptomatic group also reported the most severe functional impairment on both the WSAS and SF-36. Such functional impairment was matched by the FSS group when assessed by SF-36, though this was less severe on the WSAS. The core and avoidant-inactive groups demonstrated the least fatigue and the least functional impairment, although this differed between the SF-36 and WSAS: the core group had the least severe WSAS score, and avoidant-inactive the least severe SF-36.

Further validation was undertaken comparing the groups using age, sex, illness duration and whether they met the Centre for Disease Control (CDC) or London criteria for CFS (Reeves et al. 2003; The National Task Force on Chronic Fatigue Syndrome 1994) (table 4).

62% of patients met the CDC criteria for CFS, but differences across subgroups were not statistically significant ($p = 0.12$). In contrast, 53% of all patients met the London criteria for ME, and differences across subgroups could not be explained by chance alone ($p < 0.001$). This relationship, however, appeared to be driven by the low proportion of patients in the mood and polysymptomatic groups. This was to be expected, given that the high levels of co-morbid anxiety and depression that

characterised these subgroups would, by definition, result in exclusion from the London criteria (White et al. 2007).

There was a significant difference between subgroups for sex ($p = 0.014$), with the FSS group, and to a lesser extent avoidant-inactive, having a relatively greater proportion of women. Age and illness duration did not differ across subgroups (data not shown).

Discussion

Summary of classes and validation

This analysis provides further support for the heterogeneous nature of CFS in a large cohort of patients recruited from UK secondary care clinics. Five separate subgroups of CFS were defined based upon a combination of statistical likelihood and perceived clinical utility. The groups were predominantly defined by features of associated functional somatic syndromes, mood, self-efficacy ratings, cognitive behavioural responses questionnaire scores and assessments of sleep. These groups had face validity and were consistent with previous publications. As in previous studies (Aslakson et al. 2006; Aslakson et al. 2009), discriminative validity was demonstrated by showing significant differences across the classes in the independent, a priori measures of fatigue and disability. The polysymptomatic and FSS group had the greatest levels of fatigue and functional impairment, whilst both the mood and polysymptomatic groups tended to regard their symptoms as both physical and psychological in nature.

Comment on validation

Given the features used to define the polysymptomatic group, it is no surprise that they reported the greatest levels of fatigue and functional impairment. It is also logical that the functional somatic syndromes group, which suffered from the highest levels of back, joint and stomach pain, had a similarly high level of disability when assessed by the SF-36, which measures physical impairment.

In contrast, the avoidant-inactive group had the lowest impairment on the SF-36 physical function scale, but greater impairment on the WSAS. This may relate to the more global measure of impairment or participation in life captured by the WSAS. The avoidant-inactive group commonly reported features of behavioural avoidance, fear avoidance and concerns regarding the potential damaging effects of activity, which may therefore result in greater impairment of their ability to work, socialise and maintain relationships (represented by the WSAS), despite relatively preserved physical functioning (represented by the SF-36).

Comparison to previous studies

Despite the unique combination of variables used to define our groups, comparisons can be drawn with previously published studies of the heterogeneity of CFS. In particular, our polysymptomatic and FSS subgroups show similarities to those defined previously. Hickie *et al* found in multiple cohorts the presence of a polysymptomatic subgroup of CFS patients (Hickie et al. 1995; Wilson et al. 2001). The authors concluded that this subgroup was related to 'somatisation' (a tendency to report multiple symptoms that were not related to established disease), and found in post-hoc comparisons that such patients had higher levels of psychiatric comorbidity. This is consistent with our results. In further work an additional pain predominant subgroup was identified in addition to the polysymptomatic and core groups (Hadzi-Pavlovic et al. 2000), which may be more resistant to conventional treatments such as CBT (Cella, Chalder, et al. 2011). Subgroups defined by many symptoms and pain-predominance were also key features of a recent, large study including 8,433 CFS patients from UK and Dutch specialist centres (Collin et al. 2016).

More recent work has focused on population derived samples with increasing use of biological parameters. As such samples included patients with CFS (Centers for Disease Control criteria), uncategorised chronic fatigue and unfatigued controls, as well as different variables, the extent of available comparison is limited.

Vollmer-Conna and colleagues defined five classes in an American population derived cohort, based upon the key features of obesity, sleep disturbance, metabolic strain, multiplicity of symptoms and depressed mood(Vollmer-Conna et al. 2006). In contrast to this, we did not find that BMI contributed to the subgroup model. It should be noted, however, that BMI of our cohort was substantially lower than that of Vollmer-Conna's study (median BMI 24.7 vs. 29, respectively), that BMI may varying significantly between nations, and that the authors of the American study concluded that this raised BMI is likely to be contributing to the fatigue and relative sleep hypnoea identified in their sample. Despite such differences, however, Vollmer-Conna's description of two highly symptomatic and depressed groups, which was replicated in a separate American cohort (Aslakson et al. 2009), is reminiscent of the subgroups previously described and demonstrated again in this study.

There is support for the role of sleep in defining CFS subgroups, as identified by Vollmer-Conna *et* al and again found in this study. In a sample of 343 CFS patients, polysomnography identified 30% had a primary sleep disorder that could explain their fatigue, and the remaining patients could be grouped in 4 polysomnography subtypes (Gotts et al. 2013). As we have found, sleep efficiency varied between these subgroups, together with aspects of total sleep and sleep stage duration.

Our finding of a separate mood disorder group is in keeping with the well-recognised role that mood disorders play in CFS (Wessely et al. 1996; Harvey et al. 2009). Previous studies using latent class analysis have found that depression was an important contributing factor in their models (Aslakson et al. 2006; Aslakson et al. 2009). Similarly, it has been recognised that CFS, fibromyalgia and IBS, which have been collectively referred to as functional somatic syndromes , often cluster together (Aaron & Buchwald 2001; Mcbeth et al. 2015). It has been suggested that this may be due to shared predisposing factors, perhaps relating to mechanisms of chronic central nervous system sensitisation (Kato et al. 2009; Bourke et al. 2015) or comorbid mood disorders (Mcbeth et al. 2015), which,

together with anxiety disorders, are common in patients with functional somatic syndromes (Janssens et al 2015). Taken together, the delineation of a group characterised by predominant functional somatic syndrome features, a group with predominant mood features, and a polysymptomatic group with features of both is attractive in its simplicity, and could have clinical utility.

Critique

Strengths

The size and quality of the PACE trial database provides a number of strengths. The 640 patients were carefully screened to ensure accuracy of diagnosis, allowing 541 patients with complete datasets to be analysed. This large sample from six CFS clinics is representative of the CFS population attending secondary care facilities around the UK. The range of baseline variables incorporated into the analysis provided detailed characterisation of this group, including measures of fitness, daytime activity, sleep quality, and illness beliefs and coping which have not previously been included in CFS heterogeneity studies.

Limitations

As with all trial samples, it is possible that systematic differences existed between the trial sample and the UK CFS population, limiting the external validity of our results. The use of the broad and widely used Oxford criteria (Sharpe et al. 1991) and the inclusion of Action for ME representatives on the trial management committee (with the resulting addition of the adaptive pacing therapy arm) may have aided recruitment of a representative patient cohort. The participants, were still selected from secondary care facilities, had to be able to attend research and treatment sessions (and so were generally ambulant), and may be inherently more motivated to partake in treatment therapies than non-trial participants. The extent to which the results may be generalised to primary care patients, the severely disabled and patients with poor motivation to engage in therapies may therefore be limited.

Whilst we describe our subgroups as empirically derived, we acknowledge that the subgroups identified will be limited by the variables selected for inclusion, and that the majority of the decisions relating to variable reduction were undertaken post hoc. A large number of redundant variables, such as 25 separate variables for different medications, were removed as they rarely contained patient data (most patients were taking far less than 25 medicines). Reference to published literature and clinical experience, shaped further decisions to omit variables, aimed at maximising retention of what we perceived to be clinically useful data whilst removing the substantial overlap between many other variables. Whilst such decisions are liable to a selection bias, the use of PCA as a variable reduction tool once an initial list of plausible variables was compiled should have minimised this.

Similarly a post hoc decision was made when choosing the number of classes to include in the final latent class model. By using a combination of statistical indices and clinical experience, as in previous literature, however, we have been able to derive clinically recognisable and subsequently validated subgroups of chronic fatigue syndrome. Whereas a six class solution was minimally more statistically plausible, we feel the added complexity would have detracted from its future clinical utility.

We did not include haematological or biochemical data and were unable to investigate the endocrine, metabolic and immunological abnormalities that have previously been reported. We may also have inadvertently excluded other measures that may have revealed different sub-groups. Our assessments of sleep using actigraphy may also not be as detailed as the polysomnography that have been used in previous studies (Vollmer-Conna et al. 2006; Gotts et al. 2013). The validity of actigraphy in assessing sleep has, however, been shown to be acceptable (Tryon 2004). It is also clinically more practical, and has the advantage of being able to capture daytime activity. Clinicians might use alternative measures of activity, such as pedometers or Fitbit in clinical practice.

Conclusion

In conclusion, we have used a large, high quality dataset of variables from participants in the PACE trial to empirically identify and validate five subgroups of chronic fatigue syndrome. These provide evidence for the clinical heterogeneity of CFS. We have replicated previous findings of subgroups defined by mood disorders, functional somatic syndromes, and having many symptoms. A novel finding is the avoidant-inactive subgroup. Future work should focus on the aetiology and pathophysiologies of these subgroups, as well as their specific treatment responses. For example, one could hypothesise that the mood disorder subgroup may preferentially benefit from interventions targeted at depression and anxiety, especially given the known increased risk of suicide in patients with CFS (Roberts et al. 2016). Similarly, the avoidant-inactive group may particularly benefit from CBT addressing their negative cognitions related to the damaging effects of exercise and subsequent activity avoidance.

Future studies of aetiology and response to treatment could use these subgroups in order to obtain more homogeneous groups. The identification of these subgroups in clinical practice may allow translation of such research into potentially improved treatments.

Acknowledgements

We acknowledge the help of the PACE Trial Management Group, which consisted of the authors of this paper, excluding AP, plus (in alphabetical order): B Angus, H Baber, J Bavinton, M Burgess,

LV Clark, DL Cox, JC DeCesare, E Feldman, P McCrone, G Murphy, M Murphy, H O'Dowd, T Peto, L Potts, R Walwyn, and D Wilks.

Financial support

Funding for the PACE trial was provided by the Medical Research Council, Department for Health for England, The Scottish Chief Scientist Office, and the Department for Work and Pensions. TEW was funded by the North East Thames Foundation School.

Conflicts of interest

PDW is an appointed member of the Independent Medical Experts Group, which advises the UK Ministry of Defence regarding the Armed Forces Compensation Scheme. He has done voluntary consultancy work for the UK government and paid consultancy work for a reinsurance company. TC has received royalties from Sheldon Press and Constable and Robinson. MS has received royalties from Oxford University Press.

Ethical standards

The PACE trial was approved by the West Midlands Multicentre Research Ethics Committee; reference number MREC 02/7/89; ISRCTN 54285094.

References:

Aaron L.A. & Buchwald D., 2001. A review of the evidence for overlap among unexplained clinical conditions. *Annals of internal medicine*, 134(9 Pt 2), pp.868–81.

Aslakson E., Vollmer-Conna U., William C., White P.D., 2009. Replication of an empirical approach to delineate the heterogeneity of chronic unexplained fatigue. *Population health metrics*, 7, p.17.

Aslakson E., Vollmer-Conna U. & White P.D., 2006. The validity of an empirical delineation of heterogeneity in chronic unexplained fatigue. *Pharmacogenomics*, 7, pp.365–373.

Bourke J.H., Langford R.M. & White P.D., 2015. The common link between functional somatic syndromes may be central sensitisation. *Journal of Psychosomatic Research*.

Brurberg K.G., Fønhus M., Larun L., Flottorp S., Malterud, K., 2014. Case definitions for chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME): a systematic review. *BMJ open*, 4(2), p.e003973.

Cella M., Chalder T. & White P.D., 2011. Does the heterogeneity of chronic fatigue syndrome moderate the response to cognitive behaviour therapy? An exploratory study. *Psychotherapy and Psychosomatics*, 80(6), pp.353–358.

Cella, M., Sharpe, M. & Chalder, T., 2011. Measuring disability in patients with chronic fatigue syndrome: Reliability and validity of the Work and Social Adjustment Scale. *Journal of Psychosomatic Research*, 71(3), pp.124–128.

Chalder T., Berelowitz G., Pawlikowska T., Watts L., Wessely S., Wright D., Wallace, E.P., 1993. Development of a fatigue scale. *Journal of Psychosomatic Research*, 37(2), pp.147–153.

Collin S.M., Nikolaus S., Heron J., Knoop H., White P.D., Crawley E., 2016. Chronic fatigue syndrome (CFS) symptom-based phenotypes in two clinical cohorts of adult patients in the UK and The Netherlands. *Journal of Psychosomatic Research*, 81, pp.14–23.

Fukuda K., Straus S., Hickie I., Sharpe M., Dobbins J., Komaroff A., 1994. The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. *Annals of internal medicine*, 121(12), pp.953–959.

Gotts Z., Deary V., Newton J., Van der Dussen D., De Roy P., Ellis J., 2013. Are there sleep-specific phenotypes in patients with chronic fatigue syndrome? A cross-sectional polysomnography analysis. *BMJ open*, 3(6), p.e002999–.

Hadzi-Pavlovic D., Hickie I. B., Wilson A. J., Davenport T., Lloyd A.R., Wakefield D., 2000. Screening for prolonged fatigue syndromes: Validation of the SOFA scale. *Social Psychiatry and Psychiatric Epidemiology*, 35(10), pp.471–479.

Harvey S.B., Wessely, S., Kuh D., Hotopf M., 2009. The relationship between fatigue and psychiatric

disorders: evidence for the concept of neurasthenia. *Journal of psychosomatic research*, 66(5), pp.445–54.

Hickie I., Lloyd A., Hadzi-Pavlovic D., Parker G., Bird K., Wakefield D, 1995. Can the chronic fatigue syndrome be defined by distinct clinical features? *Psychological medicine*, 25(5), pp.925–935.

Janssens K, Zijlema W, Joustra M, Rosmalen J., 2015. Mood and anxiety disorders in chronic fatigue syndrome, fibromyalgia, and irritable bowel syndrome: Results from the LifeLines cohort study. Psychosomatic Medicine; 77(4), pp.449-57

Jenkins C., Stanton, B., Niemcryk S., Rose R., 1988. A scale for the estimation of sleep problems in clinical research. *Journal of Clinical Epidemiology*; 41(4), pp.313-321

Kato K., Sullivan P., Evengård B., Pedersen N., 2009. A population-based twin study of functional somatic syndromes. *Psychological medicine*, 39(3), pp.497–505.

Komaroff A., 2015. Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Real Illness. *Annals of internal medicine*, 162(12), pp.871–2.

Kroenke K., Spitzer R. & Williams J., 2002. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosomatic medicine*; 16(2), p258-266

Lorig K., Sobel D., Ritter P., Laurent, D., Hobbs M., 2001. Effect of a Self-Management Program on Patients with Chronic Disease. *Effective clinical practice*, 4(6), pp.256–262.

Mcbeth J., Tomenson B., Chew-Graham C., Macfarlane G., Jackson J., Littlewood A., Creed F., 2015. Common and unique associated factors for medically unexplained chronic widespread pain and chronic fatigue ☆. *Journal of Psychosomatic Research*, 79(6), pp.484–491.

Mundt J., Marks I., Shear M., Greist J., 2002. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *The British Journal of Psychiatry*.

National Institute for Health and Clinical Excellence, 2007. Chronic fatigue syndrome / myalgic

encephalomyelitis (or encephalopathy): diagnosis and management. *NICE guidelines [CG53]* .

Petrella R., Cunningham D., Paterson D., 2001. A self-paced step test to predict aerobic fitness in older adults in the primary care clinic. *Journal of the American Geriatrics Society*, 49(5), pp.632–638.

Prins J.B., Meer J.W., Bleijenberg, G., 2006. Review Chronic fatigue syndrome. *Review Literature And Arts Of The Americas*, 367, pp.346–355.

Reeves W. & Lloyd A., 2003. Identification of ambiguities in the 1994 chronic fatigue syndrome research case definition and recommendations for resolution. *BMC health services …*, 5(25).

Roberts E., Wessely S., Chalder T., Chang C., Hotopf M., 2016. Mortality of people with chronic fatigue syndrome: a retrospective cohort study in England and Wales from the South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLaM BRC) Clinical Record Interactive Search (CRIS) Register. *The Lancet*, 387, pp.1638–1643.

Sharpe M.C., Archard L. C., Banatvala J. E., Borysiewicz L. K., Clare A. W., David A., Edwards R. H., Hawton K. E., Lambert H. P., Lane R. J., 1991. A report--chronic fatigue syndrome: guidelines for research. *Journal of the Royal Society of Medicine*, 84(2), pp.118–121.

Skerrett T. & Moss-Morris R., 2006. Fatigue and social impairment in multiple sclerosis: the role of patients' cognitive and behavioral responses to their symptoms. *Journal of psychosomatic research*, 61(5), pp.587–93.

Sullivan P.F., Smith W. & Buchwald D., 2002. Latent class analysis of symptoms associated with chronic fatigue syndrome and fibromyalgia. *Psychological medicine*, 32(5), pp.881–888.

The National Task Force on Chronic Fatigue Syndrome, 1994. Report from the National Task Force on Chronic Fatigue Syndrome (CFS)/Post Viral Fatigue Syndrome (PVMS)/Myalgic Encephalomyelitis (ME). *Bristol: Westcare*.

Tryon W.W., 2004. Issues of validity in actigraphic sleep assessment. *Sleep*, 27(1), pp.158–165.

Vollmer-Conna U., Aslakson E. & White P.D., 2006. An empirical delineation of the heterogeneity of chronic unexplained fatigue in women. *Pharmacogenomics*, 7(3), pp.355–364.

Wakefield, J.C., 2013. The DSM-5 debate over the bereavement exclusion: Psychiatric diagnosis and the future of empirically supported treatment. *Clinical Psychology Review*, 33(7), pp.825–845.

Ware J. & Sherbourne C., 1992. The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical care*, 30(6), pp.473–483.

Wessely S., Chalder T., Hirsch S., Wallace P., Wright D., 1996. Psychological symptoms, somatic symptoms, and psychiatric disorder in chronic fatigue and chronic fatigue syndrome: a prospective study in the primary care setting. *Am J Psychiatry*, 153(8), pp.1050–1059.

White P.D., Goldsmith K., Johnson A.L., Potts L., Walwyn R., DeCesare J.C., Baber H.L., Burgess M., Clark L.V., Cox D.L., Bavinton J., Angus B.J., Murphy G., Murphy M., O'Dowd H., Wilks D., McCrone P., Chalder T., Sharpe M., 2011. Comparison of adaptive pacing therapy, cognitive behaviour therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): a randomised trial. *Lancet*, 377(9768), pp.823–836.

White P.D., Sharpe M., Chalder T., DeCesare J., Walwyn R., 2007. Protocol for the PACE trial: a randomised controlled trial of adaptive pacing, cognitive behaviour therapy, and graded exercise, as supplements to standardised specialist medical care versus standardised specialist medical care alone for patients with the. *BMC neurology*, 7, p.6.

Wilson A., Hickie I., Hadzi-Pavlovic D., Wakefield D., Parker G., Straus S.E., Dale J., McCluskey D., Hinds G., Brickman A., Goldenberg D., Demitrack M., Blakely T., Wessely S., Sharpe M., Lloyd A., 2001. What is chronic fatigue syndrome? Heterogeneity within an international multicentre study. *The Australian and New Zealand journal of psychiatry*, 35(4), pp.520–527.

Wolfe F., Smythe H., Yunus M., Bennett R., Bombardier C., Goldenberg D., Tugwell P., 1990. The american college of rheumatology 1990 criteria for the classification of fibromyalgia. *Arthritis & Rheumatism*. 33(2), p160-172

Zigmond S. & Snaith R., 1983. The hospital anxiety and depression scale. *Acta Psychiatr Scand*, 67(6), pp.361–370.

Figure Legends

Figure 1 – Summary of variable reduction

Footnote: Variable reduction was initially undertaken in accordance to clinical knowledge and previous literature. Correlation matrices were then used to remove overlap. Principal component analysis then identified variables minimally contributing to the overall variance, and four further variables were removed during latent component analysis due to a lack of statistical contribution to the derived models. PMHx = Past Medical History; SCID = Standardised Clinical Interview for DSM IV; HADS = Hospital Anxiety and Depression Scale; CBRQ = Cognitive Behavioural Responses Questionnaire; EQ5D = EuroQoL 5 items; PCA = Principal Components Analysis; LCA = Latent Class Analysis.

Figure 2 – Subgroups – 5 class solution

Footnote: Diagrammatic representation of the 5 class solution, highlighting the key features used to identify each subgroup. FSS, functional somatic syndromes; CBRQ, cognitive and behavioural response questionnaire.

Figure 3 – Subgroups – Validation of the 5 class solution

Footnote: Diagrammatic representation of the statistically significant comparisons used to provide support for discriminative validity.