Q J Med 2013; **106**:555–565
doi:10.1093/qjmed/hct061 Advance Access Publication 28 March 2013

Treatment outcome in adults with chronic fatigue syndrome: a prospective study in England based on the CFS/ME National Outcomes Database

E. CRAWLEY[1], S.M. COLLIN[1], P.D. WHITE[2], K. RIMES[3], J.A.C. STERNE[4], M.T. MAY[4] and CFS/ME NATIONAL OUTCOMES DATABASE

From the [1]Centre for Child & Adolescent Health, School of Social & Community Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, [2]Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, Charterhouse Square, London EC1M 6BQ, [3]Department of Psychology, University of Bath, Claverton Down, Bath BA2 7AY and [4]School of Social and Community Medicine, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol BS8 2PS, UK

Address correspondence to Dr E. Crawley, Centre for Child & Adolescent Health, School of Social & Community Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK. email: esther.crawley@bristol.ac.uk

Received 30 December 2012 and in revised form 11 February 2013

Summary

Background: Chronic fatigue syndrome (CFS) is relatively common and disabling. Over 8000 patients attend adult services each year, yet little is known about the outcome of patients attending NHS services.

Aim: Investigate the outcome of patients with CFS and what factors predict outcome.

Design: Longitudinal patient cohort.

Methods: We used data from six CFS/ME (myalgic encephalomyelitis) specialist services to measure changes in fatigue (Chalder Fatigue Scale), physical function (SF-36), anxiety and depression (Hospital Anxiety and Depression Scale) and pain (visual analogue pain rating scale) between clinical assessment and 8–20 months of follow-up. We used multivariable linear regression to investigate baseline factors associated with outcomes at follow-up.

Results: Baseline data obtained at clinical assessment were available for 1643 patients, of whom 834 (51%) had complete follow-up data. There were improvements in fatigue [mean difference from assessment to outcome: −6.8; 95% confidence interval (CI) −7.4 to −6.2; $P < 0.001$]; physical function (4.4; 95% CI 3.0–5.8; $P < 0.001$), anxiety (−0.6; 95% CI −0.9 to −0.3; $P < 0.001$), depression (−1.6; 95% CI −1.9 to −1.4; $P < 0.001$) and pain (−5.3; 95% CI −7.0 to −3.6; $P < 0.001$). Worse fatigue, physical function and pain at clinical assessment predicted a worse outcome for fatigue at follow-up. Older age, increased pain and physical function at assessment were associated with poorer physical function at follow-up.

Conclusions: Patients who attend NHS specialist CFS/ME services can expect similar improvements in fatigue, anxiety and depression to participants receiving cognitive behavioural therapy and graded exercise therapy in a recent trial, but are likely to experience less improvement in physical function. Outcomes were predicted by fatigue, disability and pain at assessment.

Introduction

Chronic fatigue syndrome (CFS), also known as myalgic encephalomyelitis (ME), is defined as persistent or recurrent debilitating fatigue that is not lifelong, not the result of ongoing exertion, not alleviated by rest, not explained by other conditions and results in a substantial reduction in function.[1–3] CFS is heterogeneous[4] and relatively common, with an estimated prevalence from population surveys of between 0.2% and 2.6% depending on case definition and recruitment methodology.[5–10] Adults with CFS can be very disabled[11] and are generally unwell for a long time, with a median duration of illness of 6.3 years.[12]

Over 8000 adults[13] are assessed and treated by specialist UK National Health Service (NHS) CFS/ME clinical teams each year and although we have trial evidence about which treatments are effective, little is known about what happens when patients are treated in the NHS setting or which factors are associated with treatment outcomes. One previous study showed that outcomes were better for patients with CFS enrolled in a trial of cognitive behavioural therapy (CBT; $N = 30$) when compared with those receiving CBT as part of everyday clinical practice.[14] The only other service evaluation we found showed that secondary care outpatients who received CBT improved their function and symptom count, although only 61% gave follow-up data.[15] There is clear trial evidence that CBT and graded exercise therapy (GET), are moderately effective for the treatment of CFS.[16,17] In a recent trial, CBT and GET were each associated with a mean improvement of ~7 points on the Chalder Fatigue Scale, comparing baseline with 52 weeks follow-up and ~19 points on the SF-36 physical function subscale[17].

NHS specialist CFS/ME services in England follow National Institute for Health and Clinical Excellence (NICE) guidelines, offering CBT or GET, or components of CBT or GET alone or with activity management and sleep management.[3] For severely affected patients, the guidelines recommend an activity management programme that draws on the principles of CBT and GET.[3]

In this study, we use data from the CFS/ME National Outcomes Database (NOD) to investigate: (i) what the prognosis is for patients accessing NHS specialist services; (ii) which routinely collected baseline measures predict post-treatment outcomes (fatigue, physical function, pain and mood) and (iii) whether outcomes are similar to those recorded in the Pacing, graded Activity and Cognitive Behavioural Therapy: a randomised Evaluation (PACE) trial for CBT and GET (the most effective treatments).

Methods

Study population

We included patients who had attended one of six NHS specialist CFS/ME services during the period 1 January 2005 to 31 December 2009. The six services were chosen because they had been collecting 12-month outcome data for at least 1 year during the study period and had achieved >40% 12-month follow-up. Patients were included if they were age 18 years or older and were diagnosed with CFS according to Centers for Disease Control criteria.[1,2] Services treated patients with CBT, GET, a combination of both or activity management, in group and/or individual treatment sessions of varying numbers and lengths.

Patient-level data

In each of the six services, patients routinely completed the following inventories prior to their initial clinical assessment (baseline): 11-item Chalder Fatigue Scale;[18] 10-item SF-36 physical function subscale;[19] 14-item Hospital Anxiety and Depression Scale (HADS);[20] and a Visual Analogue Pain Rating Scale (score of 0 for 'no pain' and 100 for 'pain as bad as possible'). The Chalder Fatigue Scale was scored using the 0–3 method for scoring each question (0 'Less than usual', 1 'No more than usual', 2 'More than usual' and 3 'Much more than usual'). On the physical function subscale of the SF-36 (RAND version), patients scored 0 ('Yes, limited a lot') 5 ('Yes, limited a little') or 10 ('No, not limited at all') for each question, (range 0–100 with 0 being most disabled). Inventory total scores (and each HADS subscale score) were coded as missing if more than one question was unanswered; if only one item was missing, an adjusted total score was calculated. The same set of inventories was sent to patients by post ~12 months after their initial clinical assessment. Patients were asked at assessment to report the duration of their illness (time elapsed, in months, between onset of symptoms and clinical assessment). Data on treatments received by each patient are not currently recorded in the NOD.

Service-level data

Information was collected from specialist services using a self-completed questionnaire which included questions on: numbers of patients assessed per annum; total number of staff employed (full-time equivalent); type of treatments offered (CBT, GET, activity management); number of treatment sessions the service usually offered/aimed to offer; ratio of

group/individual sessions and estimated average contact time with patients.

Follow-up interval

Each team sent out follow-up questionnaires at 12 months. Variation in when the questionnaires were sent and delays in return of 12-month follow-up questionnaires led to variation in the exact time of follow-up. Also, some teams obtained data at additional follow-up points (e.g. 6 and 24 months). To maximize the availability of follow-up data for our analysis, we determined a margin of follow-up either side of 12 months. We did this by fitting fractional polynomial generalized estimating equation (GEE) regression models of fatigue against time.[21,22] This method (implemented in Stata as *fracpoly* combined with *xtgee*) compares a linear GEE model with the best-fitting first and second degree models, each containing fractional polynomial terms (from a pre-defined set of integer, fractional and negative powers) for time. The differences in deviances between the linear and 1st-degree model and the first and second degree models are tested using a chi-squared test and the resulting *P*-values indicate whether the change in outcome over time is linear or whether it has a more complex shape. We inspected a plot of predicted values of fatigue against time from the model with the best fit to determine an appropriate follow-up interval in which observed fatigue scores could be assumed to be representative of the scores predicted at 12 months. In patients with more than one follow-up assessment, the closest to 12 months was used.

Statistical analysis

Our primary outcome measures were fatigue and physical function. Our secondary outcome measures were anxiety, depression and pain. We used chi-squared tests (10-year age group and sex) and Kruskal–Wallis tests (duration of illness, fatigue, physical function, anxiety, depression and pain) to compare the baseline characteristics of patients for whom we had/did not have complete follow-up data on primary and secondary outcomes. We defined clinically useful improvements within each person as having a difference of 2 points on the total Chalder Fatigue Scale and a difference of 11 points on the SF-36 physical function subscale between the baseline and follow-up measurements. We chose these cut-offs because they equated to 0.5 SD of the distribution of the baseline measurements. These characteristics were included in a random effects linear regression model, with service as a unit of random effect, to identify factors measured at baseline that were independently associated with outcomes measured in the follow-up interval. Scores from the different inventories were re-scaled, for this particular analysis, so that the range for each was ~0–10 so that a regression coefficient of 1 represents a 10% change in the score.

Ethical approval

The North Somerset and South Bristol Research Ethics Committee decided that the collection and analysis of CFS/ME patient data were part of service evaluation and as such did not require ethical review by a NHS Research Ethics Committee or approval by NHS Research and Development offices (REC reference number 07/Q2006/48).

Results

Patients and follow-up

Complete baseline data were available for 1643 patients assessed by the six services during the study period. Of these patients, 1269 (77.2%) were female, the mean age was 39.9 (12.6 SD) years, and patients had been ill for a median of 36 months [interquartile range (IQR) 16–84] before accessing a service. Follow-up data on fatigue were available for 53.7% (882 of 1643) patients up to 24 months after assessment. The median interval between assessment and follow-up was 375 (IQR: 357–402) days. We fitted a fractional polynomial model to these data to determine the interval that represented a reasonably steady state of fatigue around 12 months after assessment. The model with the best fit incorporated a term for $\sqrt{}$time. Fatigue was relatively constant between 8 and 20 months (240–600 days) after assessment (Figure 1). Hence, we used this interval to capture the maximum amount of follow-up data.

Of the 1643 patients with complete baseline data, 834 (50.8%) had complete (fatigue, physical function, anxiety, depression and pain) follow-up data between 8 and 20 months after assessment. We compared patients with and without complete follow-up data (Table 1). Patients without follow-up data were slightly younger than patients with follow-up data [mean age 38.5 (12.4 SD) years vs. 41.3 (12.7 SD) years, Student's *t*-test $P < 0.001$] but were similar in all other characteristics.

At follow-up, there were improvements in fatigue, SF-36 physical function, anxiety, depression and pain (Table 2). About 74% (620 of 834) of patients had a decreased Chalder Fatigue score at follow-up and 64% (534 of 834) had improved by >2 points (our definition of a clinically useful improvement). In

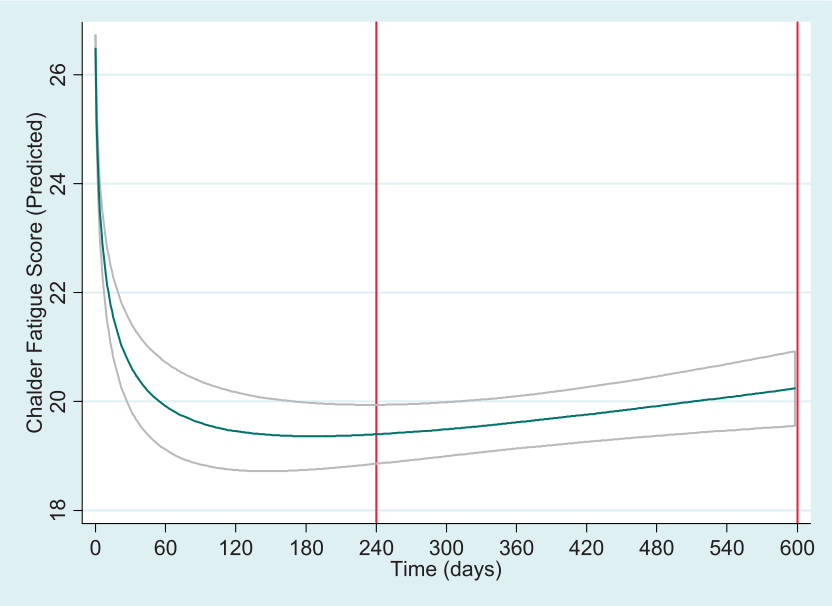

Figure 1. Predicted values of fatigue against time from generalized estimating equation regression model incorporating a fractional term for time (time$^{1/2}$) ($N = 882$). Vertical lines indicate follow-up interval adopted for this study (8–20 months).

Table 1 Characteristics of CFS patients with and without follow-up data

Characteristics	Without follow-up data ($N = 809$)	With follow-up data ($N = 834$)	P-value[a]
Age (years) median (IQR)	38 (28–47)	41 (32-51)	<0.001
Female N (%)	617 (76.3)	652 (78.2)	0.36
Duration of illness (months) median (IQR)	36 (17–96)	36 (15–84)	0.06
Chalder Fatigue (0–33) median (IQR)	27 (23–30)	28 (23–30)	0.22
SF-36 physical (scale 0–100) median (IQR)	40 (20–60)	40 (25–55)	0.46
Anxiety (scale 0–21) median (IQR)	10 (7–14)	10 (7–13)	0.30
Depression (scale 0–21) median (IQR)	9 (7–12)	9 (7–12)	0.92
VPain (scale 0–100) median (IQR)	52 (24–70)	51 (24–69)	0.52

[a]Chi-squared test for categorical variables; Kruskal–Wallis test for continuous variables

Table 2 Characteristics of CFS patients at baseline and at follow-up ($N = 834$)

Characteristics	Baseline mean (SD)	Follow-up mean (SD)	Mean Difference (95% CI)	P-value[a]
Chalder Fatigue (0–33)	26.5 (5.2)	19.7 (8.4)	−6.8 (−7.4 to −6.2)	<0.001
SF-36 physical function (scale 0–100)	40.6 (22.7)	45.0 (27.2)	4.4 (3.0 to 5.8)	<0.001
HADS anxiety (scale 0–21)	10.1 (4.6)	9.5 (4.6)	−0.6 (−0.9 to −0.3)	<0.001
HADS depression (scale 0–21)	9.6 (4.1)	7.9 (4.5)	−1.6 (−1.9 to −1.4)	<0.001
Visual Analogue Pain (scale 0–100)	47.3 (26.6)	42.0 (28.4)	−5.3 (−7.0 to −3.6)	<0.001

[a]Student's paired t-test

contrast, only 50% (416 of 834) of patients had an increased SF-36 physical function score at follow-up and only 16% (131 of 834) had improved by >22 points. In total, 14% (120 of 834) had clinically useful improvements on both scales.

Patient-level factors associated with treatment outcomes

There was strong evidence ($P < 0.001$) that each baseline measure (fatigue, physical function,

anxiety, depression and pain) was associated with fatigue at 8–20 months in models that were only adjusted for centre and year of assessment (Table 3). In a model that was adjusted for all baseline measures, fatigue, physical function and pain, but not anxiety or depression scores, were associated with fatigue at 8–20 months. Similarly, older age, physical function and pain, were associated with physical function at 8–20 months. There was little evidence that anxiety and depression were associated with physical function at follow-up (Table 4).

In fully adjusted regression models, each baseline measure was associated with its repeat measurement at follow-up (Table 5). Lower physical function at baseline was associated with higher levels of depression and pain at follow-up. Higher levels of pain at baseline were associated with higher levels of anxiety and depression at follow-up. Female sex was associated with lower HADS depression at follow-up. Higher levels of pain at follow-up were related to higher levels of baseline fatigue, physical function and depression (Table 5).

Duration of illness was inversely associated with baseline physical function [mean change in SF-36 score −0.03; 95% confidence interval (CI) −0.05 to −0.01; $P=0.01$] per additional month of illness) and associated with baseline pain (mean change in visual analogue pain score 0.05; 95% CI 0.02 to 0.07; $P=0.001$ per additional month of illness), but was not associated with baseline fatigue or mood or with any outcomes at 8–20 months in fully adjusted models.

Service-level factors associated with treatment outcomes

Including services as a categorical variable in multivariable model showed that treatment outcomes varied between services. Services reported that they treated patients with CBT, GET, a combination of both or activity management (Table A1). Comparing the three services that said they offered CBT/GET with the three services that only offered activity management, suggested that patients attending services offering CBT and GET had less improvement in fatigue at 12 months (fully adjusted coefficient 0.57; 95% CI 0.24 to 0.90; $P=0.001$,corresponding to a difference of ∼2 points on the Chalder Fatigue Scale). However, this effect mainly represented a comparison between two large services which assessed 372 (CBT/GET) and 501 (only activity management) patients, respectively, in 2010.

Discussion

This is the first study to investigate treatment outcomes and their predictors in patients treated by more than one NHS specialist CFS/ME service. Among the patients for whom follow-up data were available (51% of patients assessed), there were overall improvements in fatigue, physical function, anxiety, depression and pain at ∼12-month follow-up. As may be expected, the measurement of each of these symptoms at baseline predicted its subsequent value when repeated between 8 and 20 months. Patients who were less physically able at assessment had higher levels of fatigue, depression and pain at follow-up. Patients who were in more pain when assessed at baseline had worse scores for all outcomes at follow-up. There was little evidence that anxiety or depression was associated with either fatigue or physical function at follow-up. The size of improvement in fatigue was similar to that achieved in the PACE trial after CBT and GET, but this was not the case for physical function, where the improvement was considerably less in clinical services compared with the PACE trial.

Strengths and limitations

This is a large study with follow-up data on more than 800 patients with clinically characterized CFS. Six clinical services contributed to the outcome data, suggesting that the results are generalizable to other NHS specialist CFS/ME services which provide similar types of treatment. However, as the study was based on data from patients accessing specialist services, the results may not be generalizable to patients who receive similar treatments in primary care. The main limitation of this study is its relatively low follow-up rate. Compared with patients who did not complete follow-up, those who did had similar baseline levels of fatigue, physical function, pain and mood and similar time-to-assessment, but were slightly younger. However, we don't know whether patients lost to follow-up had better, the same or worse outcomes at 8–20 months, or how any such differential losses might affect results. We used 0.5 SD to define the clinically important difference. Although this is likely to be close to what patients define as clinically important,[23] further work needs to be done to define how much change is important to patients and what outcomes are relevant.

Services completed questionnaires about the type of treatment offered to patients, but patient-level treatment data were not collected by the NOD and therefore we cannot be certain what was provided. Similarly, services provided estimates of the

Table 3 Associations of baseline characteristics with fatigue (Chalder Fatigue Scale) at follow-up ($N = 834$)[a]

Characteristics	Mean change in fatigue (95% CI) adjusted for year	*P*-value	Mean change in fatigue (95% CI) adjusted for year and baseline fatigue	*P*-value	Mean change in fatigue (95% CI) adjusted for year and all variables in the table	*P*-value
10-year age group	0.02 (−0.12 to 0.16)	0.81	−0.01 (−0.15 to 0.13)	0.88	−0.08 (−0.22 to 0.06)	0.25
Sex (female vs. male)	0.15 (−0.27 to 0.56)	0.18	0.05 (−0.35 to 0.45)	0.79	−0.14 (−0.54 to 0.26)	0.49
Chalder Fatigue[b]	0.41 (0.31 to 0.52)	<0.001	0.41 (0.31 to 0.52)	<0.001	0.24 (0.12 to 0.35)	<0.001
SF-36 (physical function)[b]	−0.31 (−0.39 to −0.24)	<0.001	−0.24 (−0.32 to −0.16)	<0.001	−0.19 (−0.28 to −0.10)	<0.001
HADS anxiety[b]	0.13 (0.06 to 0.21)	<0.001	0.09 (0.01 to 0.16)	0.02	0.03 (−0.06 to 0.11)	0.54
HADS depression[b]	0.24 (0.16 to 0.33)	<0.001	0.15 (0.07 to 0.24)	0.001	0.06 (−0.04 to 0.16)	0.27
Visual Analogue Pain[b]	0.22 (0.16 to 0.28)	<0.001	0.17 (0.10 to 0.23)	<0.001	0.09 (0.02 to 0.16)	0.01
CBT/GET vs. Activity	0.48 (0.13 to 0.83)	0.007	0.49 (0.15 to 0.83)	0.004	0.57 (0.24 to 0.90)	0.001

[a]Random-effects regression model with centre as unit of effect.
[b]The measures were re-scaled to range from 0 to 10; hence, a regression coefficient of 1 represents a 10% change in the score. For SF-36, coefficient <1 indicates that higher levels of disability are positively associated with higher levels of fatigue at 8–20 months.

Table 4 Associations of baseline characteristics with physical function (SF-36 physical function subscale) at follow-up ($N = 834$)[a]

Characteristics	Mean change in physical function (95% CI) adjusted for year	*P*-value	Mean change in physical function (95% CI) adjusted for year and baseline physical function	*P*-value	Mean change in physical function (95% CI) adjusted for year and all variables in table	P-value
10-year age group	−0.40 (−0.55 to −0.25)	<0.001	−0.16 (−0.27 to −0.05)	0.005	−0.17 (−0.29 to −0.06)	0.003
Sex (female vs. male)	−0.67 (−1.12 to −0.23)	0.003	0.002 (−0.34 to 0.33)	0.99	−0.10 (−0.43 to 0.24)	0.57
Chalder Fatigue[b]	−0.47 (−0.58 to −0.36)	<0.001	−0.02 (−0.11 to 0.08)	0.69	0.01 (−0.09 to 0.11)	0.82
SF-36 (physical function)[b]	0.81 (0.75 to 0.87)	<0.001	0.81 (0.75 to 0.87)	<0.001	0.71 (0.64 to 0.79)	<0.001
HADS anxiety[b]	−0.17 (−0.25 to −0.09)	<0.001	−0.05 (−0.11 to 0.01)	0.14	0.001 (−0.07 to 0.07)	0.97
HADS depression[b]	−0.38 (−0.47 to −0.30)	<0.001	−0.07 (−0.14 to 0.00)	0.05	−0.05 (−0.13 to 0.04)	0.28
Visual Analogue Pain[b]	−0.41 (−0.47 to −0.34)	<0.001	−0.11 (−0.17 to −0.05)	<0.001	−0.11 (−0.17 to −0.05)	<0.001
CBT/GET vs. Activity	0.50 (0.13 to 0.88)	0.008	0.15 (−0.13 to 0.43)	0.28	0.15 (−0.13 to 0.42)	0.30

[a]Random-effects regression model with centre as unit of effect. [b]The measures were re-scaled to range from 0 to 10, hence, a regression coefficient of 1 represents a 10% change in the score.

Table 5 Associations of baseline characteristics with secondary outcomes (anxiety, depression and pain) at 8–20 months after initial clinical assessment ($N = 834$)[a]

Outcome at follow-up	HADS anxiety		HADS depression		Visual Analogue Pain	
Baseline characteristics	Mean change in fatigue (95% CI)	*P*-value	Mean change in fatigue (95% CI)	*P*-value	Mean change in fatigue (95% CI)	*P*-value
10-year age group	0.02 (−0.08 to 0.13)	0.64	0.08 (−0.02 to 0.18)	0.13	0.10 (−0.03 to 0.23)	0.15
Sex (female vs. male)	−0.10 (−0.40 to 0.21)	0.54	−0.32 (−0.63 to −0.02)	0.04	−0.26 (−0.63 to 0.13)	0.20
Chalder Fatigue[b]	−0.04 (−0.13 to 0.05)	0.42	−0.02 (−0.11 to 0.07)	0.65	−0.10 (−0.21 to 0.01)	0.06
SF-36 physical function[b]	−0.02 (−0.09 to 0.05)	0.56	−0.11 (−0.18 to −0.05)	0.001	−0.13 (−0.22 to −0.05)	0.002
HADS anxiety[b]	0.50 (0.43 to 0.56)	<0.001	−0.01 (−0.07 to 0.06)	0.88	0.04 (−0.04 to 0.12)	0.30
HADS depression[b]	0.10 (0.02 to 0.18)	0.01	0.58 (0.50 to 0.66)	<0.001	0.12 (0.03 to 0.22)	0.01
Visual Analogue Pain[b]	0.10 (0.05 to 0.16)	<0.001	0.07 (0.02 to 0.13)	0.006	0.56 (0.49 to 0.62)	<0.001
CBT/GET vs. Activity	0.09 (−0.16 to 0.35)	0.47	−0.23 (−0.95 to 0.49)	0.54	0.10 (−0.22 to 0.42)	0.53

[a]Random-effects regression models with centre as unit of effect, adjusted for year of assessment and all variables in table. [b]The measures were re-scaled to range from 0 to 10, hence, a regression coefficient of 1 represents a 10% change in the score. For SF-36, coefficient <1 indicates that higher levels of disability are positively associated with higher levels of anxiety, depression or pain at 8–20 months.

number of treatment hours received by patients, but this was not checked with hospital-level activity data. A meta-analysis of CBT and GET for CFS found that total treatment hours were one of the strongest predictors of outcome.[16] The lack of patient-level treatment data and a non-intervention group does not allow us to compare treatments or to determine whether observed improvements were due to the interventions themselves. Although our analyses suggested that patients who received activity management did better than those who received CBT and/or GET, this was based on a comparison between three services offering activity management only and three services offering CBT/GET only. As one service in each group was much larger than the others, our result is based on a comparison between two services. As treatment content and adherence were not assessed, it is not known whether CBT or GET conformed to existing protocols or whether activity management provided content consistent with CBT or GET principles. There are many other service-level factors that could affect outcome between two services, such as the number of sessions offered by the service, a known predictor of outcome,[16] the overall philosophy of the service, individual therapist effects and patient characteristics.[24]

We only have data on depression and anxiety from inventories. Patients did not have a psychiatric interview. The relationship between depression, anxiety and physical function could be explored further using a clinical assessment of depression rather than self-report questionnaire scores.

Comparison with previous literature

Baseline measures of fatigue and physical function in our study were very similar to the PACE trial. The mean change in fatigue in our study was 6.8 points (95% CI −7.4 to −6.2), which is consistent with results from the recently reported PACE trial, where the mean change in fatigue was 7.4 points in the CBT group.[25] This is in contrast with one previous study, which showed that improvement in fatigue was greater in trial participants.[14] However, in our study, the mean improvement on the SF-36 physical function subscale was much less [4.4 points (95% CI 3.0 to 5.8)] than in the PACE trial where the mean improvement was 19.2 points in the CBT arm and 21.0 in the GET arm. This remarkable discrepancy suggests that the poor outcome in physical function in our study may be due to differences in the delivery or content of treatments given in routine clinical practice. For example, most patients in NHS services appear to be offered five or six sessions (Table A1), whereas PACE trial participants attended 12–14 sessions. This needs urgent investigation.

Treatments in our study also appear to be less effective than in other randomized controlled trials comparing either CBT or GET with controls.[16,26]. Patients enrolled in randomized controlled trials can have better results than seen in clinical practice because trial clinicians follow a research protocol and may have more training, supervision, motivation and resources available to them.[14] Trials also increase expectation of efficacy and patients may be more motivated.[27] Trials usually have selection criteria exceeding those in routine care (e.g. excluding for significant risk of self-harm or co-morbid medical conditions and only including those willing and able to attend regular appointments and complete repeated questionnaires); it is not known whether such factors are associated with treatment outcomes.

Our study showed that older age predicted worse outcome at follow-up which is consistent with one previous study.[14] There was little evidence that duration of illness was associated with any outcomes in a fully adjusted model which is consistent with previous studies.[16,28] We found that poorer baseline physical functioning was associated with higher scores on the depression subscale of the Hospital Anxiety and Depression Scale at follow-up but not vice versa. This is in contrast with two previous follow-up studies of a clinical cohort ($n = 41$)[29] and a randomized controlled trial ($n = 114$)[28] which suggested that psychiatric co-morbid diagnoses were associated with a poorer outcome [30]. A previous cross-sectional study suggested that depression partially mediated the relation between fatigue and physical disability.[31] However, this study used a mediation analysis, which was inappropriate in situations where there are unmeasured confounders for the effect of both the exposure (fatigue) and the intermediate (depression) on the outcome (physical function).[32] In chronic fatigue, the relationship between depression and fatigue is complex,[33] and the relationship with disability is likely to have unmeasured confounders. This aspect of CFS/ME requires further investigation using longitudinal data that will allow causality to be determined.

Conclusions

Although NHS services are moderately effective in improving fatigue in patients with chronic fatigue syndrome, they are much less effective in improving physical function than similar treatments delivered in the PACE trial. This requires urgent investigation to determine whether it is due to differences in the

delivery or the content of treatments offered by services. Future research also needs to include patient-level treatment data to investigate variations in outcome that may be related to treatment data. We did not find that depression, anxiety or duration of illness at assessment predicted outcome. Clinicians providing assessments should not assume that co-morbid mood disorders or length of illness are predictors of outcome.

Acknowledgements

We thank the clinical leads and team members of all services participating in the National Outcomes Database and, in particular, the clinical leads of the services that contributed data to this study: Dr Hazel O'Dowd (North Bristol NHS Trust), Karen Butt and Dawn Dunn (Somerset Partnership NHS Foundation Trust), Sue Pemberton (Leeds and York Partnership NHS Foundation Trust), Dr Maurice Murphy (East London Foundation NHS Trust and Barts and the London NHS Foundation Trust) and Dr Yasmin Mullick and Dr Amolak Bansal (Epsom and St Helier University Hospitals NHS Trust). E.C. had the original idea for the study; S.C. performed all data management and statistical analyses; M.M. and J.S. provided statistical and methodological input; S.C. and E.C. wrote the first draft and incorporated comments from all authors; all authors contributed to the interpretation and writing of the paper and read and approved the final manuscript.

Funding

This work was supported by the National Institute for Health Research (NIHR) [grant number DHCS/2008/08/08/06; Clinician Scientist Award issued by the NIHR to E.C.] and by Action for M.E. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health. Action for M.E. is a member of the Steering Group for the NOD and was therefore involved in decisions on which data should be collected. Action for M.E. did not contribute to the study design, data analysis, interpretation of the results, or writing of the manuscript and was not involved in the decision to submit the manuscript for publication.

Conflict of interest: All authors have completed the Unified Competing Interest form and declare that they did not receive any financial support from companies for the submitted work. E.C., S.C., K.R., J.S. and M.M. have no relationships with any companies that might have an interest in the submitted work in the previous 3 years. P.D.W. has done consultancy work for the Departments of Health and Work and Pensions and a re-insurance company. S.C., P.D.W., K.R., J.S. and M.M. have no non-financial interest that may be relevant to the submitted work. E.C. is a medical advisor for the Association for Young people with ME (AYME) and the Sussex and Kent ME society.

References

1. Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, Komaroff A. The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. *Ann Intern Med* 1994; **121**:953–9.
2. Reeves WC, Lloyd A, Vernon SD, Klimas N, Jason LA, Bleijenberg G, *et al.* Identification of ambiguities in the 1994 chronic fatigue syndrome research case definition and recommendations for resolution. *BMC Health Serv Res* 2003; **3**:25.
3. *Chronic Fatigue Syndrome/Myalgic Encephalomyelitis (or Encephalopathy); Diagnosis and Management.* National Institute for Health and Clinical Excellence (NICE), Clinical Guideline 53. Issued August 2007. http://guidance.nice.org.uk/CG53 (14 March 2013, date last accessed).
4. Hickie I, Davenport T, Vernon SD, Nisenbaum R, Reeves WC, Hadzi-Pavlovic D, *et al.* Are chronic fatigue and chronic fatigue syndrome valid clinical entities across countries and health-care settings? *Aust N Z J Psychiatry* 2009; **43**:25–35.
5. Reeves WC, Jones JF, Maloney E, Heim C, Hoaglin DC, Boneva R, *et al.* Prevalence of chronic fatigue syndrome in metropolitan, urban, and rural Georgia. *Popul Health Metr* 2007; **5**:5.
6. Evengard B, Jacks A, Pedersen NL, Sullivan PF. The epidemiology of chronic fatigue in the Swedish Twin Registry. *Psychol Med* 2005; **35**:1317–26.
7. Buchwald D, Umali P, Umali J, Kith P, Pearlman T, Komaroff AL. Chronic fatigue and the chronic-fatigue-syndrome - prevalence in a Pacific-northwest health-care system. *Ann Intern Med* 1995; **123**:81–8.
8. Reyes M, Nisenbaum R, Hoaglin DC, Unger ER, Emmons C, Randall C, *et al.* Prevalence and incidence of chronic fatigue syndrome in Wichita, Kansas. *Arch Intern Med* 2003; **163**:1530–6.
9. Wessely S, Chalder T, Hirsch S, Wallace P, Wright D. The prevalence and morbidity of chronic fatigue and chronic fatigue syndrome: A prospective primary care study. *Am J Public Health* 1997; **87**:1449–55.
10. Skapinakis P, Lewis G, Mavreas V. Unexplained fatigue syndromes in a multinational primary care sample: specificity of definition and prevalence and distinctiveness from depression and generalized anxiety. *Am J Psychiatry* 2003; **160**:785–7.
11. Collin SM, Crawley E, May M, Sterne JA, Hollingworth W, UK CFS/ME National Outcomes Database ME. The impact of CFS/ME on employment and productivity in the UK: a cross-sectional study based on the CFS/ME National Outcomes Database. *BMC Health Serv Res* 2011; **11**:217.

12. Nisenbaum R, Jones A, Jones J, Reeves W. Longitudinal analysis of symptoms reported by patients with chronic fatigue syndrome. *Ann Epidemiol* 2000; **10**:458.

13. Collin SM, Sterne JA, Hollingworth W, May MT, Crawley E. Equity of access to specialist Chronic Fatigue Syndrome (CFS/ME. services in England (2008-2010): a national survey and cross-sectional study. *BMJ Open* 2012; **2**:e001417.

14. Quarmby L, Rimes KA, Deale A, Wessely S, Chalder T. Cognitive-behaviour therapy for chronic fatigue syndrome: comparison of outcomes within and outside the confines of a randomised controlled trial. *Behav Res Ther* 2007; **45**:1085–94.

15. Akagi H, Klimes I, Bass C. Cognitive behavioral therapy for chronic fatigue syndrome in a general hospital–feasible and effective. *Gen Hosp Psychiatry* 2001; **23**:254–60.

16. Castell BD, Kazantzis N, Moss-Morris RE. Cognitive behavioral therapy and graded exercise for chronic fatigue syndrome: a meta-analysis. *Clin Psychol Sci Prac* 2011; **18**:311–24.

17. White P, Goldsmith K, Johnson A, Potts L, Walwyn R, Decesare J, *et al*. Comparison of adaptive pacing therapy, cognitive behaviour therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): a randomised trial. *Lancet* 2011; **377**:823–36.

18. Chalder T, Berelowitz G, Pawlikowska T, Watts L, Wessely S, Wright D, *et al*. Development of a fatigue scale. *J Psychosom Res* 1993; **37**:147–53.

19. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; **30**:473–83.

20. Snaith RP. The Hospital Anxiety And Depression Scale. *Health Qual Life Outcomes* 2003; **1**:29.

21. Royston P, Altman DG. Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. *Appl Stat* 1994; **43**:429–67.

22. Royston P, Sauerbrei W. *Multivariable Modelling*. Chichester: Wiley, 2008.

23. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 2007; **7**:541–6.

24. Wiborg JF, Knoop H, Wensing M, Bleijenberg G. Therapist effects and the dissemination of cognitive behavior therapy for chronic fatigue syndrome in community-based mental health care. *Behav Res Ther* 2012; **50**:393–6.

25. White PD, Goldsmith KA, Johnson AL, Potts L, Walwyn R, DeCesare JC, *et al*. Comparison of adaptive pacing therapy, cognitive behaviour therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): a randomised trial. *Lancet* 2011; **377**:823–36.

26. Chambers D, Bagnall AM, Hempel S, Forbes C. Interventions for the treatment, management and rehabilitation of patients with chronic fatigue syndrome/myalgic encephalomyelitis: an updated systematic review. *J Roy Soc Med* 2006; **99**:506–20.

27. Everitt BS, Wessely S. *Clinical Trials in Psychiatry*. Oxford: Oxford University Press, 2004.

28. Bentall RP, Powell P, Nye FJ, Edwards RH. Predictors of response to treatment for chronic fatigue syndrome. *Br J Psychiatry* 2002; **181**:248–52.

29. Bonner D, Ron M, Chalder T, Butler S, Wessely S. Chronic fatigue syndrome: a follow up study. *J Neurol Neurosurg Psychiatry* 1994; **57**:617–21.

30. Bonner D, Ron M, Chalder T, Butler S, Wessely S. Chronic fatigue syndrome: a follow up study. *J Neurol Neurosurg Psychiatry* 1994; **57**:617–21.

31. Hadlandsmyth K, Vowles KE. Does depression mediate the relation between fatigue severity and disability in chronic fatigue syndrome sufferers? *J Psychosom Res* 2009; **66**:31–5.

32. Cole SR, Hernan MA. Fallibility in estimating direct effects. *Int J Epidemiol* 2002; **31**:163–5.

33. Skapinakis P, Lewis G, Mavreas V. Temporal relations between unexplained fatigue and depression: longitudinal data from an international study in primary care. *Psychosom Med* 2004; **66**:330–5.

Appendix

Table A1 Results of a survey of NHS specialist CFS/ME specialist clinical teams which contributed data to this study

Team	Number of patients assessed in 2010	Staff whole time equivalent (WTE)	Ratio of individual: group treatments[a]	Individual total contact time (number of sessions)[a]	Group total contact time (number of sessions)[a]	Treatments offered
1	372	3.1	60:40	5 h (5)	16 h (8)	CBT+/or GET
2[b]	76	1.7	100:0	12 h (12)		Activity management
3	501	14.4	90:10	6 h (6)	20 h (13)	Activity management
4	157	7.7	80:20	20 h (30)	16 h (8)	CBT+/or GET
5	160	2.3	40:60	4 h (4)	16 h (11)	CBT+pacing +/or GET
6	308	5.9	50:50	11 h (11)	10 h (4)	Activity management

[a]Patients who are not severely affected
[b]About 90% of patients are seen at home