Psychological Medicine, 2003, **33**, 263–281.
DOI: 10.1017/S0033291702007031 Printed in the United Kingdom

Fatigue in a community sample of twins

P. F. SULLIVAN,[1] P. KOVALENKO, T. P. YORK, C. A. PRESCOTT AND K. S. KENDLER

From the Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA

ABSTRACT

Background. Fatigue is a complex symptom associated with many physiological, psychological and pathological processes. Its correlates and typology remain inadequately understood.

Method. These data were from two large, longitudinal twin studies. Trained interviewers enquired as to the presence of a $\geqslant 5$ day period in the previous year of fatigue or tiredness that interfered with daily activities. A range of potential correlates was assessed in a structured interview: demography; health beliefs; the presence of nine physical disorders; mood, anxiety and addictive disorders; neuroticism and extraversion; recollections of parental rearing; and nine stressful life events. Statistical analyses included logistic regression, CART, MARS, latent class analysis and univariate twin modelling.

Results. Data were available for interfering fatigue (IF) on 7740 individual twins (prevalence 9·9% in the previous year). IF was significantly associated with 42 of 52 correlates (most strongly with major depression, generalized anxiety disorder, reported major health problems and neuroticism). Multivariate analyses demonstrated that IF is a highly complex construct with different sets of correlates in its subtypes. There were two broad clusters of correlates of IF: (*a*) major depression, generalized anxiety disorder and neuroticism; and (*b*) beliefs of ill health coexisting with alcoholism and stressful life events. Twin analyses were consistent with aetiological heterogeneity – genetic effects may be particularly important in women and shared environmental effects in men.

Conclusions. IF is a complex and common human symptom that is highly heterogeneous. More precise understanding of the determinants of IF may lead to a fuller understanding of more extreme conditions like chronic fatigue syndrome.

INTRODUCTION

Fatigue is a complex and enigmatic entity. To begin, it is a common human complaint (Lewis & Wessely, 1992; Hagnell *et al*. 1993; Pawlikowska *et al*. 1994; Loge *et al*. 1998; Addington *et al*. 2001) that appears to be continuously distributed in the population (David *et al*. 1990; Lewis & Wessely, 1992; Pawlikowska *et al*. 1994; Lawrie & Pelosi, 1995; Loge *et al*. 1998) with no clear ‘point of rarity’ that delineates an extreme segment of the distribution. Moreover, fatigue can be non-pathological and explicable by life circumstances (e.g. raising an infant with colic, training for endurance sports, or shift work).

Fatigue is associated with a wide range of physiological and psychological disorders. It is a common presenting symptom in primary care (David *et al*. 1990; Kirk *et al*. 1990; Cathébras *et al*. 1992; Bates *et al*. 1993; Fuhrer & Wessely, 1995) and a diagnostic clue for a broad range of physical illnesses. Fatigue is also an important source of morbidity in chronic illnesses such as HIV infection (Vogl *et al*. 1999), diabetes mellitus (Konen *et al*. 1996) or following treatment for cancer (Cella *et al*. 2001). It is also a diagnostic criterion for several common psychiatric disorders (e.g. major depression and generalized anxiety disorder) (World Health Organization, 1993; American Psychiatric Association, 1994) and a common consequence of other psychiatric disorders (e.g. alcohol withdrawal or bulimia nervosa).

[1] Address for correspondence: Dr Patrick F. Sullivan, Virginia Institute for Psychiatric and Behavioral Genetics, PO Box 980126, Richmond, VA 23298–0126, USA.

Fatigue is perhaps most perplexing in the context of several 'medically unexplained' syndromes (Barsky & Borus, 1999; Wessely *et al.* 1999; Aaron *et al.* 2000) that are associated with considerable morbidity and medical costs (Wolfe *et al.* 1995; Sharpe, 1996). Profound fatigue is the cardinal feature of chronic fatigue syndrome (Sharpe *et al.* 1991; Hickie & Wakefield, 1992; Fukuda *et al.* 1994), a prominent feature of fibromyalgia (Wolfe *et al.* 1990), and often associated with irritable bowel syndrome (Thompson *et al.* 1999).

It is relatively easy for physicians to ignore the complexities of fatigue, in part because medical training often emphasizes fatigue as an ephemeral symptom whose root cause is to be discovered as a diagnostic challenge. Physicians often regard fatigue as a relatively unimportant symptom whereas patients rank it as particularly important (Lewis & Wessely, 1992).

Fatigue is insufficiently studied given its clinical prevalence and importance. Moreover, a more thorough understanding of common fatigue would inform our understanding of rarer conditions like chronic fatigue syndrome. A number of important questions about the 'orphan' symptom of fatigue remain unanswered. First, what are its univariate and multivariate correlates? Secondly, as it is likely that fatigue is heterogeneous, what is its typology? Thirdly, what can the classical twin study reveal about the causes of fatigue? We address these questions using a number of statistical and conceptual approaches in the context of a large, population-based twin registry.

METHOD

Subjects

Subjects for the present investigation were drawn from two longitudinal studies conducted in a similar manner by the same research group. Both investigations were reviewed by several ethical review boards and all subjects provided written informed consent (or verbal consent for telephone interviews) prior to participation. Each sample was ascertained from the population-based Mid-Atlantic Twin Registry (formerly the Virginia Twin Registry). The first study was of female–female twin pairs (FF) and the second study of male–male and male–female twin pairs (MMMF). These studies are described at length elsewhere (Kendler & Prescott, 1999). Zygosity determination was based on questionnaire responses and DNA polymorphisms where required (Spence *et al.* 1988).

FF study

The first interview wave of the FF study (1987–9) was the source of the majority of data in this report although a few variables are from the third interview wave (1992–5). Wave one data collection included personal interviews with 2163 female Caucasian twins (92% of the eligible twins) whose mean age was 30·1 (S.D. = 7·6) years. The sample contained 1033 complete pairs (590 MZ, 440 DZ, and three pairs of unknown zygosity). Wave three data collection included telephone interviews with 1898 individuals (88% of the wave one sample) whose mean age was 35·1 years (S.D. = 7·5).

MMMF study

The data for this report are from the second interview wave of the MMMF study (1994–8). Of 9417 eligible individuals for the first wave, 6814 (72·4%) completed an interview. At least 1 year after the completion of the first wave interview, we successfully completed a wave two interview with 5629 individuals (82·6%) whose mean age was 37·0 (S.D. = 9·1) years.

Measures

The data analysed here were collected as part of a 1–3 h personal interview that assessed multiple domains (e.g. physical status, psychiatric disorders, personality and stressful life events). The interviews for both the FF and MMMF studies were highly homologous. All interviews were conducted by individuals with Master's degrees in Social Work or Psychology or a Bachelor's degree plus at least 2 years of clinical experience. All interviewers underwent rigorous training and all interviews were reviewed by a senior editor for consistency and accuracy. Interviewers were blinded to all prior data from the twin they interviewed as well as from data on the co-twin.

Interfering fatigue

The term 'fatigue' is complex and multi-faceted (Lewis & Wessely, 1992). Our dependent variable – interfering fatigue (IF) – was defined as follows. All subjects in both studies were asked 'in the last year, have you had a time lasting at least 5 days when you felt tired or fatigued most of the time?'. Interviewers were instructed to record an affirmative response only if the reported fatigue was unusual and worse than the subject's baseline energy level. Affirmative responses were followed by a query about the degree to which fatigue interfered with daily life. IF in the past year was considered present if subjects responded affirmatively to the first query and indicated that fatigue interfered 'completely' or 'a lot' with their daily life. Our use of unusual fatigue associated with a sizeable degree of interference as the dependent variable is consistent with prior studies of fatigue (Bates *et al.* 1993; Addington *et al.* 2001).

Correlates of IF

We used the wealth of data assessed on all subjects to explore the correlates of IF. Considerable efforts were made to ensure the comparability of data across the FF and MMMF studies. The potential correlates fall into six categories.

(*i*) *Demography*

These variables included age, gender, marital status, number of children $\leqslant 5$ years old, employment status, years of education and current residence in a city with a population $\geqslant 250\,000$.

(*ii*) *Physical status*

Body mass index (kg/m^2) was calculated from self-reported height and weight and obesity was defined as BMI $\geqslant 30$ kg/m^2. All subjects were asked whether they had any major health problem, their satisfaction with their health, whether they thought they were at risk of early death because of their health and whether their daily activities were limited by their health. Subjects were also asked about the presence in the prior year of nine specific illnesses (hypertension, blindness/serious visual impairment, cardiac or pulmonary disease, diabetes mellitus, cancer, severe arthritis, paralysis, or serious accident or injury). For women, the presence of current pregnancy was noted and a premenstrual symptom score was calculated (Kendler *et al.* 1990) (for women in the FF study, the premenstrual score was from the third interview wave).

(*iii*) *Lifetime psychiatric and drug use disorders*

These disorders expected to be prevalent in a community sample were assessed with a modified version of the Structured Clinical Interview for DSM-III-R (Spitzer *et al.* 1992; Williams *et al.* 1992). These disorders included major depression (Kendler *et al.* 1992*a*), generalized anxiety disorder (Kendler *et al.* 1992*b*), panic disorder (Kendler *et al.* 1993), bulimia nervosa (Kendler *et al.* 1991), anorexia nervosa (Walters & Kendler, 1995), alcohol dependence (Kendler *et al.* 1992*c*) and nicotine dependence (Kendler *et al.* 1999). Because of low prevalences, men were not asked about bulimia or anorexia nervosa. Specific phobias, agoraphobia and social phobia (Kendler *et al.* 1992*d*) were assessed via an adaptation of the Diagnostic Interview Schedule (Version III-A) (Robins & Helzer, 1985). For generalized anxiety disorder, $\geqslant 1$ month duration of illness was required and we considered the 'A' criterion satisfied if the anxiety affected two or more life circumstances. Nicotine dependence was assessed with a modified version of the Fagerstrom Tolerance Questionnaire and was defined as a score $\geqslant 7$ during the period of lifetime maximal cigarette use (Fagerstrom, 1978; Fagerstrom & Schneider, 1989). For the FF study, nicotine dependence was based on wave three data. All diagnoses were made without regard to diagnostic hierarchies.

(*iv*) *Personality*

Neuroticism and extraversion were assessed with an adapted version of the short form of the Eysenck Personality Questionnaire (Eysenck & Eysenck, 1975; Heath *et al.* 1992).

(*v*) *Parenting*

Parenting was assessed with the Parental Bonding Instrument (PBI) (Parker *et al.* 1979) with modifications that are detailed elsewhere (Kendler, 1996). Higher scores on these instruments are generally associated with better parenting.

(*vi*) *Stressful Life Events*

We assessed the occurrence of nine stressful life events in the previous year (Kendler *et al.* 1995). Some were 'personal' events that occurred primarily to the informant: assault, rape, or mugging; divorce, marital separation, broken engagement, or breakup of other romantic relationship; loss of confidant or separation from other loved one or close friend; legal problems; and insufficient money for family needs. Others were 'network' events that were focused on an individual in the respondent's social network: serious illness of a child and a serious personal crisis, death, or serious illness of someone in the network.

https://doi.org/10.1017/S0033291702007031 Published online by Cambridge University Press

Statistical analyses

The data being analysed consist of a single dependent variable and over 50 variables to be investigated as correlates and possible predictors. Analysis of data such as these is typically accomplished with logistic regression. However, the data set contains four complexities that complicate analysis.

First, typical applications of logistic regression assume that the observations are independently sampled. Given that the sampling unit for a twin study is a pair of related individuals, it was necessary to account for the non-independence or clustering of individual twins within twin pairs. We used generalized estimating equations (GEE) (Zeger & Liang, 1986, 1992; Zeger *et al.* 1988; Diggle *et al.* 1994), which treat the correlation between twins as a nuisance parameter to obtain robust and asymptotically correct regression parameter estimates. This was accomplished with PROC GENMOD in SAS (SAS Institute Inc., 1999). A potential weakness of the GEE approach is that it cannot readily accommodate multiple groups (i.e. differences across monozygotic and dizygotic twin pair types) that can be handled in a mixed model. We cross-checked the GEE results with a mixed model using the SAS macro GLIMMIX (Littell *et al.* 1999); in all instances, results from PROC GENMOD and GLIMMIX were nearly identical. We used the GEE approach because of its greater flexibility and computational speed.

Secondly, the inevitable presence of missing data can lead to erroneous conclusions if, for example, missingness is related to the dependent variable. It is plausible that IF might lead to incomplete self-report questionnaires or premature termination of an interview. It is notable that nearly all of the predictors had very few missing values (generally $<0{\cdot}25\%$) with the following three general exceptions. Some data were available only for women (anorexia, bulimia, premenstrual symptom score, and current pregnancy) – given the rarity of eating disorders in males, only women were queried about anorexia and bulimia to minimize subject burden. Although bulimia was missing for only three women and pregnancy for zero women, the diagnosis of anorexia was missing for 39% of women (because it was available only in women from FF but not the MMMF study) and the premenstrual symptom score for 17% of women. Despite our best efforts, participating subjects sometimes did not complete and return the self-report questionnaires that yielded scores for neuroticism, extraversion and the two parenting measures (one or more missing for 8·4% of the sample). Finally, assessment of nicotine dependence was available for nearly everyone in MMMF (missing in one subject) but was assessed in only the third FF interview wave. Because some subjects in the first FF wave did not participate in the third FF interview wave, this meant that nicotine dependence was missing for 3·4% of subjects.

Overall, complete data were available on 87·4% of the sample. Excluding the self-report variables and nicotine dependence, 98·4% of the sample had complete data. There was no association of IF with missing data on the self-report questionnaire data ($\chi^2=0{\cdot}83$, df$=1$, $P=0{\cdot}36$), nicotine dependence ($\chi^2=1{\cdot}42$, df$=1$, $P=0{\cdot}23$), or premenstrual symptom score ($\chi^2=0{\cdot}65$, df$=1$, $P=0{\cdot}42$). However, the typical listwise deletion of observations with missing values can clearly lead to erroneous conclusions. To deal with missingness, we used multiple imputation (Rubin, 1987); see http://www.multiple-imputation.com for an accessible introduction. This relatively flexible technique iteratively 'fills in' missing data as a function of non-missing variables. By doing this multiple times, the results can be pooled to arrive at statistically valid parameter estimates (Rubin, 1987). To accomplish multiple imputation, we used PROC MI (with its MCMC option) and PROC MIANALYZE in SAS (Yuan, 2000).

Thirdly, many of our predictor variables are known to be interrelated (e.g. female gender and major depression or nicotine dependence and lower educational attainment). Strong covariation within a set of predictor variables is an analytical challenge particularly when, as in the present circumstance, there is no strong theory to guide model selection. We used the following three different regression techniques to attempt to understand the patterns of covariation within our data. Logistic regression (Cox & Snell, 1989) was used to attempt to predict a discrete dependent variable (the presence or absence of IF) as a function of a set of predictors. In addition, we used two 'modern regression' techniques (Hastie *et al.* 2001). Classification and regression trees (CART) (Breiman *et al.* 1984; Venables & Ripley, 1999; Salford Systems, 2001*a*) is a tree-based method that evaluates combinations of predictors to formulate a set of binary rules that to classify IF. Trees have been used for a number of clinical problems (Hess *et al.* 1999; Grassi *et al.* 2001) and are appealing in that they parallel the clinical process of evaluating and classifying patients on the basis of binary rules. CART grows decision trees by recursive binary partitioning by selecting the best predictor and splitting value that appropriately separates outcome cases by a fit criterion. All possible splits are evaluated by an impurity function to identify the splits that improve classification. Each subsequent partition can then be further divided until optimal prediction accuracy is achieved while considering a penalty for tree complexity. CART deliberately grows an overly large tree which likely over-fits the data and then 'prunes' the tree back using multifold cross-validation (Stone, 1974; Efron & Tibshirani, 1993; Zhang, 1993) to select the optimal sub-tree that minimizes the misclassification cost. Multivariate adaptive regression splines (MARS) (Friedman, 1991; Salford Systems, 2001*b*) is an automated regression algorithm suited for the examination of a large number of predictor variables as they relate to either a quantitative or qualitative outcome. MARS considers all predictors in terms of piece-wise linear basis functions and defines potential knots at observed data points. Basis functions locally restrict the range of a variable and are advantageous for high dimensional model building in which the allocation of parameters must be conserved (Hastie *et al.* 2001). Models are developed in a forward growing stage by adding the basis function that reduces the training error greatest and is repeated until a preset number of terms are added. Interactions can be searched for by multiplication of a term already entered in the model with another candidate basis function. Similar to the CART tree growing stage, this will most likely result in a large, over-fit model that, while fitting the training data well, will usually perform poorly when applied to a replication sample. A backward deletion stage iteratively removes the basis functions that contribute least to model fit and is repeated, after refitting the model, until all terms have been removed. The optimal model from the set of pruned models is selected that minimizes a generalized cross-validation criterion.

Fourthly, it seems likely that IF is a particularly complex dependent variable. Heterogeneity is quite possible: there might be different 'types' of IF or different 'paths' to the same phenotype. For example, a sample with IF might consist of a set of individuals with depression and a set with chronic physical illness. Interactions among predictors is also possible: for example, the impact of chronic physical illness of IF might be different in individuals with low versus high neuroticism. The presence of strong theory can ease these difficulties, but, as we argue in the introduction, no such theory exists for IF.

Logistic regression is not an ideal technique in the presence of substantial heterogeneity and given a need to model interactions among a large number of predictors. CART and MARS are clearly better suited to attempting to understand a complex dependent variable. MARS appears to be particularly good at finding true positive interactions and avoiding false positives in high-order data (York & Eaves, 2001).

We applied an additional technique to search for heterogeneity. Latent class analysis (McCutcheon, 1987; Eaves *et al.* 1993; Yang & Becker, 1997) attempts to characterize the unobserved latent classes that give rise to the observed data and is analogous in intent to cluster analysis. We used a FORTRAN program (Eaves *et al.* 1993; Bucholz *et al.* 1996) with an efficient EM algorithm (Dempster *et al.* 1977) for maximum likelihood estimation. To determine the number of latent classes, we fit 1, 2, 3, …, 10 latent class models to the data (50 separate runs with randomized starting values were run to attempt to avoid the known problem of local minima). To determine the number of classes, we used two parsimony indices (Akaike's Information Criterion and the Schwarz Bayesian Criterion) (Akiake, 1987; Williams & Holahan, 1994). In general, models with more latent classes fit the observed data better but at a cost of increased complexity. Both of these indices of parsimony penalize a goodness-of-fit statistic for the number of parameters estimated in the model (p); AIC by two times p and SBC by p times the natural logarithm of the sample size. The intent is to determine the number of classes with a balance of goodness-of-fit and complexity.

Table 1. *Univariate associations with interfering fatigue*†

Variable	IF Present ($N=765$, 9·9%)	IF Absent ($N=6975$, 90·1%)	OR (95% CI)	χ^2 (df=1)	Univariate P
Demography					
Age	34·8 (9·3)	35·1 (9·2)	1·00 (0·99–1·01)	0·60	NS
Female gender	53·3% (408)	44·8% (3125)	1·40 (1·20–1·63)	18·49	****
Ever married	73·5% (562)	73·1% (5095)	1·03 (0·85–1·24)	0·09	NS
Currently married	55·8% (427)	60·4% (4210)	0·82 (0·70–0·96)	5·85	*
Any child ⩽5 years old	25·1% (192)	23·3% (1628)	1·11 (0·93–1·32)	1·28	NS
Currently employed/homemaker	94·6% (724)	96·9% (6757)	0·59 (0·42–0·84)	6·03	**
Education ⩽12 years	46·3% (354)	44·6% (3111)	1·07 (0·92–1·25)	0·69	NS
Residence in city ⩾250 000	26·9% (202)	27·6% (1906)	0·97 (0·82–1·15)	0·10	NS
Health status					
Body mass index (kg/m²)	−0·1 (1·0)	0·0 (1·0)	0·99 (0·92–1·07)	0·04	NS
Obesity (BMI⩾30)	14·2% (108)	12·7% (886)	1·18 (0·95–1·46)	2·03	NS
Major health problem	61·8% (473)	33·9% (2366)	3·12 (2·67–3·65)	169·68	****
Dissatisfied with health	23·7% (181)	7·1% (498)	4·11 (3·38–5·00)	95·60	****
At increased risk of early death	25·7% (196)	11·1% (775)	2·89 (2·41–3·48)	76·47	****
⩾15 sick days in bed in prior year	16·2% (124)	2·8% (196)	6·39 (5·02–8·12)	84·98	****
Daily activities limited by health	24·7% (189)	7·6% (526)	4·16 (3·42–5·05)	98·15	****
Current pregnancy‡	4·9% (20)	3·2% (101)	1·52 (0·93–2·48)	2·15	NS
Premenstrual symptom score‡	0·4 (1·0)	−0·1 (1·0)	0·89 (0·86–0·92)	48·71	****
Specific physical disorders					
Hypertension	12·5% (96)	6·8% (473)	2·06 (1·62–2·63)	22·07	****
Blindness/visual impairment	2·5% (19)	1·3% (93)	1·92 (1·15–3·21)	4·04	*
Severe pulmonary disease	13·5% (103)	5·0% (351)	2·94 (2·31–3·74)	41·23	****
Diabetes mellitus	3·5% (27)	1·9% (136)	1·84 (1·20–2·81)	5·30	*
Serious cardiac disease	3·4% (26)	1·1% (78)	3·27 (2·07–5·18)	12·12	***
Any form of cancer	1·6% (12)	0·7% (46)	2·32 (1·20–4·50)	3·29	NS
Any form of severe arthritis	9·7% (74)	3·0% (206)	3·64 (2·71–4·89)	35·00	****
Paralysis	2·2% (17)	0·3% (22)	7·25 (3·86–13·60)	12·67	***
Serious accident or injury	7·1% (54)	2·5% (176)	2·95 (2·13–4·08)	21·22	****
Psychiatric disorders – lifetime					
Major depression	65·9% (504)	32·4% (2262)	3·98 (3·39–4·66)	232·17	****
Generalized anxiety disorder	48·5% (371)	17·8% (12 443)	4·25 (3·63–4·98)	191·70	****
Panic disorder	6·5% (49)	2·0% (137)	3·07 (2·18–4·33)	20·60	****
Agoraphobia	15·7% (120)	5·0% (349)	3·35 (2·67–4·19)	55·26	****
Social phobia	16·9% (129)	7·5% (526)	2·39 (1·94–2·95)	39·60	****
Bulimia nervosa‡	9·8% (40)	4·9% (152)	2·06 (1·44–2·95)	9·95	****
Anorexia (broad definition) ‡§	6·4% (16)	3·3% (62)	1·84 (0·98–3·46)	2·36	NS
Alcohol abuse/dependence	43·8% (335)	29·2% (2036)	2·23 (1·89–2·62)	76·21	****
Nicotine dependence	29·5% (216)	19·1% (1290)	1·97 (1·64–2·36)	41·26	****
Psychiatric disorders – previous year					
Major depression	41·3% (316)	7·8% (545)	8·14 (6·86–9·66)	234·36	****
Generalized anxiety disorder	29·7% (227)	7·0% (488)	5·38 (4·48–6·46)	143·10	****
Nicotine dependence	22·5% (165)	13·6% (918)	2·01 (1·66–2·44)	35·04	****
Personality					
Neuroticism	0·6 (1·1)	−0·1 (1·0)	1·79 (1·67–1·93)	160·99	****
Extraversion	−0·1 (1·1)	0·0 (1·0)	0·89 (0·83–0·96)	8·49	**
Parenting					
Parental care	−0·2 (1·0)	0·0 (1·0)	0·79 (0·71–0·87)	21·24	****
Parental overprotection	−0·2 (1·0)	0·0 (1·0)	0·79 (0·71–0·87)	21·81	****
Stressful life events in previous year					
Assaulted, mugged, or raped	10·2% (78)	5·9% (412)	1·78 (1·38–2·31)	13·40	***
Divorce or relationship break-up	21·8% (167)	14·3% (994)	1·68 (1·39–2·03)	22·07	****
Serious illness of child	3·9% (30)	2·3% (161)	1·71 (1·16–2·53)	4·90	*
Serious personal crisis in network	56·5% (432)	41·8% (2917)	1·75 (1·51–2·03)	51·10	****
Death in network	32·4% (248)	27·3% (1905)	1·29 (1·09–1·51)	8·45	**
Serious illness in network	43·5% (333)	35·0% (2441)	1·42 (1·22–1·65)	18·52	****

Table 1. (*Cont.*)

Variable	IF Present ($N=765$, 9·9%)	IF Absent ($N=6975$, 90·1%)	OR (95% CI)	χ^2 (df=1)	Univariate P
Stressful life events in previous year (*cont.*)					
Loss of confidant	27·6% (211)	18·0% (1258)	1·67 (1·40–1·99)	26·03	****
Legal problems	14·1% (108)	5·6% (388)	2·74 (2·18–3·45)	40·10	****
Not enough money for family	24·1% (184)	13·7% (958)	1·98 (1·65–2·37)	38·33	****
Any stressful life event	88·0% (673)	79·1% (5521)	1·87 (1·50–2·35)	40·96	****

† IF (interfering fatigue), presence or absence of a period of 'interfering fatigue' in prior year that lasted at least 5 days and which interfered with daily life; OR, odds ratio; CI, confidence interval. The descriptive data are mean (S.D.) or per cent (number) as appropriate. All continuous data were standardized prior to analysis. The χ^2 tests are from logistic regression analyses with IF as the dependent variable (1 = present, 0 = absent) and the variable listed in the left hand column as independent variable while controlling for age and gender (except for age and gender). Generalized estimating equations were used to adjust for the non-independence or clustering of twins. Missing data were handled with multiple imputation.

‡ Females only.

§ The broad definition of the diagnosis of anorexia was available only for women in the FF1 sample.

* $P<0.05$; ** $P<0.01$; *** $P<0.001$; **** $P<0.0001$; NS, not significant.

Given that so little is known about IF, we conducted univariate twin analyses using Mx (Neale *et al.* 1999) as applied to contingency tables and with the computation of confidence intervals (Neale & Miller, 1997). Briefly, as this approach has been described at length elsewhere (Neale & Cardon, 1992; Kendler, 1993, 2001; Plomin *et al.* 1997), we used structural equation modeling to decompose the variance in monozygotic and dizygotic twin pairs into that due to additive genetic (a^2), common or shared environmental (c^2), and individual-specific environmental effects (e^2). Although it is customary to search for the 'best-fitting' sub-model, it is preferable to interpret parameter estimates and confidence intervals from the full model (Sullivan & Eaves, 2002).

RESULTS

Prevalence and correlates

Assessment of fatigue was available on 7740 individual twins. A period lasting at least 5 days in the year prior to interview during which the subject 'felt tired or fatigued most of the time' was reported by 36·2% of the sample. When we further required that fatigue was associated with interference with daily life ('completely' or 'a lot'), 9·9% of the sample reported IF, which was the dependent variable for this report.

The correlates of IF are depicted in Table 1. Of the 52 comparisons in Table 1, 42 (81%) are significant at the 0·05 level. IF was not significantly associated with age, lifetime marital status, the presence of young children, educational attainment, residence in a large city, body mass index/obesity, current pregnancy, a lifetime history of any form of cancer and a broad definition of anorexia. Most of the significant associations were in the expected directions – e.g. IF associated with female gender or the presence of a pathological risk factor like major depression or arthritis. IF was significantly associated with being unmarried, unemployed, lower extraversion and recollections of poorer parental rearing.

Ranking the findings would usually be accomplished via calculation of effect sizes. Such calculations are complex given the different variables (nominal and continuous) and the analyses (GEE with control for age and gender). A simple approach is to rank the variables by P values. The ten smallest P values in Table 1 were those associated with: (1) major depression in the prior year; (2) lifetime major depression; (3) lifetime generalized anxiety disorder; (4) reported major health problems; (5) increased neuroticism scores; (6) generalized anxiety disorder in the prior year; (7) the belief that daily activities were limited by health; (8) dissatisfaction with health; (9) $\geqslant 15$ sick days in bed in the prior year; and (10) the belief of risk of early death. Morbid obesity and lifetime diagnoses of anorexia or bulimia nervosa are exclusionary for CFS (Fukuda *et al.* 1994) and yet were in the bottom half of the list.

As elements of the definitions of both major depression and generalized anxiety disorder contain reference to fatigue (American Psychiatric Association 1987), we removed fatigue from these criteria and repeated the analyses: the significance levels were similar to those in Table 1 and both major depression and generalized anxiety disorder remained in the list of the smallest *P* values.

Multivariate analyses

Table 1 clearly suggests associations of IF with a wide set of predictor variables, many of which overlap. Moreover, it is possible that IF is aetiologically heterogeneous: if there are several different processes that result in IF, 'typical' regression analyses might not detect these complexities. Therefore, we applied three multivariate regression techniques to these data (logistic regression, CART, and MARS).

We included as many predictors as possible from Table 1 with several exceptions and modifications. We dichotomized all continuous variables (age, neuroticism, extraversion, and parental care and overprotection) at the sample medians so that all analyses would have similar input. We excluded predictors available only on females (premenstrual symptom score, anorexia and bulimia). We excluded predictors confounded with other predictors (ever married, and past year major depression, generalized anxiety disorder, and nicotine dependence). We created aggregate variables for any possible physical cause of fatigue (obesity, current pregnancy and any of nine conditions present in the previous year – hypertension, severe pulmonary disease, diabetes mellitus, serious cardiac disease, any form of cancer, any form of severe arthritis, paralysis, or serious accident or injury) and for the presence of any of seven stressful life events (assault, mugging, or rape; divorce or relationship break-up; serious illness of a child; death in interpersonal network; loss of confidant; legal problems; and not enough money for family needs). We excluded three variables that were highly prevalent (current employment and the network stressful life events of a serious crisis or serious illness). In the end, there were 24 predictors included from Table 1.

Logistic regression

The results are depicted in Table 2. Many variables significant in the univariate analyses in Table 1 were not significant in the multivariate analysis (gender, any physical cause of fatigue, panic disorder, agoraphobia, social phobia, nicotine dependence, extraversion and the two parental rearing measures). Several variables not significant in univariate analyses became marginally significant in the multivariate analysis (age, the presence of young children and lesser educational attainment). As might be expected, the odds ratios for nearly all predictors were attenuated in the multivariate analysis. The largest statistically independent odds ratios were associated with a report of $\geqslant 15$ sick days in bed in the prior year, lifetime major depression, generalized anxiety disorder, the presence of a reported major health problem, higher neuroticism, and a report of daily activities limited by health.

CART

Fig. 1 depicts the final CART model. The model was modestly successful in predicting the classification of IF (Cohen's $\kappa = 0{\cdot}17$, S.E. $= 0{\cdot}009$) suggesting that at least some of the variables important in understanding IF were present in the tree. The heterogeneity of IF was clearly suggested as terminal nodes with IF could be reached in five ways, via the presence of: (*i*) major depression and generalized anxiety disorder; (*ii*) major depression and major health problems; (*iii*) major depression, high neuroticism and high parental care; (*iv*) major depression, high neuroticism and living in a large city; and (*v*) major health problems and high neuroticism. Although the relatively low κ suggests that variables important in the classification of IF were not included in the CART model, the importance of major depression, major health problems, and neuroticism in the classification of IF seemed clear.

MARS

The final MARS model is shown in Table 3. Although the MARS algorithm is quite different, interpretation of the model is analogous to linear regression. MARS predicted the presence of IF (scored as 1) or its absence (scored as 0). Certain patterns of predictor variables contributed to an increase in the regression score from a baseline intercept (0·240). The magnitude of these coefficients fell into three rough groupings: (*a*) lifetime generalized anxiety disorder and having spent $\geqslant 15$ sick days in bed in the previous year

Table 2. *Multivariate logistic regression to predict interfering fatigue*†

Variable	OR (95% CI)	Multiple imputation *P*	GEE *P*
Age ⩾ sample median	0·83 (0·69–0·99)	*	NS
Female gender	1·04 (0·85–1·28)	NS	NS
Currently married	0·95 (0·79–1·14)	NS	NS
Any child ⩽5 years old	1·24 (1·01–1·52)	*	*
Education ⩽12 years	0·83 (0·70–0·99)	*	NS
Residence in city ⩾250 000	0·93 (0·77–1·12)	NS	NS
Major health problem	1·79 (1·46–2·20)	****	****
Dissatisfied with health	1·44 (1·13–1·83)	**	**
At increased risk of early death	1·36 (1·09–1·69)	**	**
⩾15 sick days in bed in previous year	2·82 (2·12–3·74)	****	****
Daily activities limited by health	1·53 (1·20–1·94)	***	***
Any physical cause‡	0·95 (0·78–1·16)	NS	NS
Major depression§	2·22 (1·85–2·67)	****	****
Generalized anxiety disorder§	1·80 (1·49–2·18)	****	****
Panic disorder§	1·08 (0·73–1·60)	NS	NS
Agoraphobia§	1·25 (0·95–1·65)	NS	NS
Social phobia§	1·17 (0·91–1·51)	NS	NS
Alcohol abuse/dependence§	1·40 (1·16–1·68)	**	**
Nicotine dependence§	1·15 (0·93–1·41)	NS	NS
Neuroticism ⩾ sample median	1·53 (1·23–1·89)	**	****
Extraversion ⩾ sample median	0·93 (0·78–1·10)	NS	NS
Parental care ⩾ sample median	0·88 (0·70–1·10)	NS	NS
Parental overprotection ⩾ sample median	0·84 (0·68–1·04)	NS	NS
Any stressful life event¶	1·23 (1·03–1·47)	*	NS

† Odds ratios, 95% confidence intervals, and associated *P* values calculated via multivariate logistic regression using multiple imputation on the entire data set of 7740 subjects with and without interfering fatigue. As it was not currently possible to use multiple imputation with generalized estimating equations in SAS, we ran this analysis using generalized estimating equations and deletion of all data on any subject with missing data (87·6% of the sample included) is shown in the righthand column.

‡ Included are self-reported physical conditions that have at least some potential causative role in IF. These conditions were obesity, current pregnancy, and any of the nine conditions present in the previous year (hypertension, severe pulmonary disease, diabetes mellitus, serious cardiac disease, any form of cancer, any form of severe arthritis, paralysis, or serious accident or injury).

§ Lifetime diagnosis.

¶ The occurrence of any of seven stressful life event in the previous year (assault, mugging, or rape; divorce or relationship break-up; serious illness of a child; death in interpersonal network; loss of confidant; legal problems; and insufficient money for family needs).

* $P<0·05$; ** $P<0·01$; *** $P<0·001$; **** $P<0·0001$; NS, not significant.

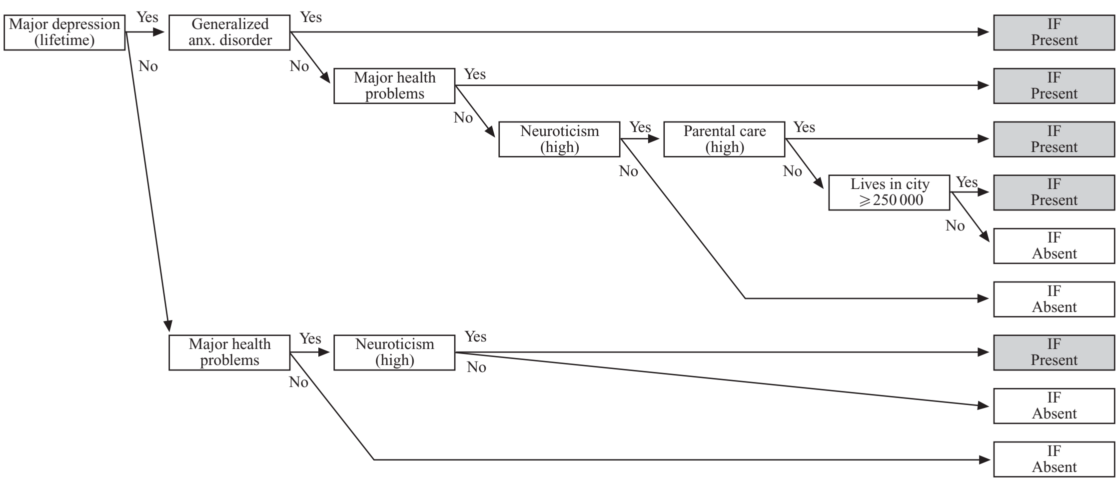

FIG. 1. Classification and regression tree analysis of interfering fatigue (IF): depiction of the final tree from CART analyses of 7740 individual twins. The dependent variable was IF and the 24 predictors were as in Table 2. This tree correctly classified 549 out of 765 individuals (71·8%) with IF and 4618 out of 6975 individuals (66·2%) without IF for total correct.

https://doi.org/10.1017/S0033291702007031 Published online by Cambridge University Press

Table 3. *Multivariate adaptive regression spline analysis of interfering fatigue**

Rank	Coefficient	Predictor	Interacting predictor
…	0·240	Intercept	Not applicable
1	0·307	Generalized anxiety disorder	None
2	0·273	⩾15 sick days in bed in previous year	None
3	0·176	Dissatisfied with health	**<15 sick days in bed in previous year**
4	0·135	Daily activities limited by health	None
5	0·127	Major depression	None
6	0·099	Agoraphobia	**Absence of generalized anxiety disorder**
7	0·067	Dissatisfied with health	**Absence of major depression**
8	0·059	Any stressful life event	**Absence of generalized anxiety disorder**
9	0·050	Major health problems	⩾15 sick days in bed in previous year
10	0·031	High neuroticism	None
11	0·030	Alcohol abuse/dependence	⩾15 sick days in bed in previous year

* On 7740 individual twins and allowing for a maximum of 45 basis functions and up to two-way interactions. All variables were dichotomous and are defined in Table 2. The model was set up to predict IF (1 = present, 0 = absent) from same 24 predictors used in Table 2. The final MARS model included five main effects and six interactions and accounted for 13·5% of the variance. Interpretation of the model is analogous to interpretation of a more typical regression model. For example, the first term indicates that the presence of generalized anxiety disorder would contribute 0·307 to the intercept of 0·240 and the sixth term that the presence of agoraphobia and the absence of GAD would contribute a further 0·099. Reverse-coded terms are shown in bold type.

contributed relatively large coefficients as main effects; (*b*) there were four terms with intermediate coefficients – dissatisfaction with health (in the presence of fewer sick days in bed), the belief that daily activities were limited by health, lifetime major depression, and agoraphobia (in the absence of generalized anxiety disorder); (*c*) the final five terms made more modest contributions and tended to include variables in the first two groupings – dissatisfaction with health (in the absence of major depression), any stressful life event (in the absence of generalized anxiety disorder), major health problems (in the presence of more sick days in bed), high neuroticism and the presence of lifetime alcohol abuse/dependence (in the presence of more sick days in bed).

The model R^2 was 13·5% suggesting modest prediction success but also that important predictors were not present in the MARS model. The model gives evidence for the presence of substantial heterogeneity in the predictors of IF as well as a complex pattern of interactions.

Latent class analysis

Unlike the regression analyses in Tables 1–3 and Fig. 1, LCA was applied only to subjects with IF as an alternative approach to delineating heterogeneity. A five class solution was suggested by the Schwarz Bayesian Criterion and an eight class solution by the Akaike Information Criterion. We chose to present the five class solution (Table 4) because the eight class solution was less interpretable and yet had similarities to the five class solution. The second column in Table 4 shows the prevalences of the predictors in all subjects with IF. LCA essentially assigned individuals with IF into one of five latent classes of the basis of their response profiles to 24 predictors.

Individuals in Class 1 tended to be older males with lesser education, poor physical health, and recollections of higher parental care and overprotection. Subjects assigned to Class 2 tended to be women who were married with young children and who had recollections of lower parental care and overprotection. Alcohol abuse/dependence was relatively uncommon. Individuals in Class 3 tended to be male, relatively infrequently reported major health problems and health dissatisfaction, and recalled higher parental care and overprotection. Although not highlighted in Table 4, Class 3 individuals had substantial prevalences of major depression and alcohol dependence. Those assigned to Class 4 reported considerable physical illness and adverse health

Table 4. *Latent class analysis in subjects reporting interfering fatigue**

Variable	Overall	Class 1	Class 2	Class 3	Class 4	Class 5
Proportion	1·00	0·12	0·25	0·32	0·19	0·13
Number of subjects	765	91	191	243	143	97
Age ⩾ sample median	0·50	**0·91**	0·40	0·49	0·69	**0·09**
Female gender	0·53	**0·32**	**0·94**	**0·15**	0·53	**0·91**
Currently married	0·56	0·72	**0·86**	0·51	0·49	**0·04**
Any child ⩽ 5 years old	0·25	0·12	**0·47**	0·24	0·20	**0·05**
Education ⩽ 12 years	0·46	**0·66**	0·41	0·36	0·63	0·39
Residence in city ⩾ 250 000	0·27	0·21	0·18	0·37	0·21	0·33
Major health problem	0·62	**1·00**	0·74	**0·37**	**0·91**	**0·22**
Dissatisfied with health	0·24	**0·61**	0·17	**0·04**	**0·55**	0·05
At increased risk of early death	0·26	**0·78**	0·10	0·13	**0·51**	**0·04**
⩾ 15 sick days in bed in previous year	0·16	**0·47**	0·13	0·01	0·35	0·04
Daily activities limited by health	0·25	**0·70**	0·13	0·06	**0·59**	**0·00**
Any physical cause	0·44	**0·88**	0·40	0·29	**0·73**	**0·07**
Major depression	0·66	0·50	0·53	0·64	**0·96**	0·68
Generalized anxiety disorder	0·48	0·33	0·44	0·30	**0·93**	0·52
Panic disorder	0·06	0·00	0·05	0·00	0·18	0·12
Agoraphobia	0·16	0·01	0·09	0·05	**0·50**	0·18
Social phobia	0·17	0·00	0·14	0·09	**0·45**	0·17
Alcohol abuse/dependence	0·44	0·34	**0·19**	0·52	**0·70**	0·43
Nicotine dependence	0·29	0·43	0·15	0·26	**0·58**	0·09
Neuroticism ⩾ sample median	0·73	0·70	0·68	0·59	**0·98**	0·81
Extraversion ⩾ sample median	0·48	0·51	0·44	0·54	0·37	0·59
Parental care ⩾ sample median	0·39	**0·70**	**0·10**	**0·72**	0·25	**0·06**
Parental overprotection ⩾ sample median	0·38	**0·65**	**0·08**	**0·72**	0·28	**0·05**
Any stressful life event	0·71	0·64	0·58	0·64	0·90	0·88

* Latent class analysis of 765 individuals reporting interfering fatigue. The second column notes the overall proportion of the 24 predictors in individuals with IF. All variables were dichotomous. The remaining columns describe the latent classes. For example, the age of 91 % of Class 1 was older than the sample median and 88 % of class 5 reported one of seven stressful life events in the previous year. To assist in interpretation, class-item prevalences > 20 % above or below the overall proportion are shown in bold type.

beliefs along with very high prevalences of major depression, anxiety disorder, alcohol abuse/dependence, nicotine dependence. Class 4 was also notable for high neuroticism and, although not highlighted, stressful life events. Class 5 contained younger, unmarried women who reported fewer health problems and had recollections of lower parental care and overprotection. In addition, although not highlighted, there was substantial major depression, high neuroticism, and stressful life events.

Twin Modelling

There were 3263 complete twin pairs. The relationships between members of twin pairs for IF are depicted in Fig. 2. Inspection of the tetrachoric correlations suggests marked differences between males and females in the patterns of correlations. Comparing monozygotic and dizygotic male pairs suggested the importance of shared and individual-specific environmental factors (as $r_{MZM} \approx r_{DZM}$ and $r_{MZM} < 1$) whereas comparison of monozygotic and dizygotic female pairs suggested the importance of genetic and individual-specific environmental factors (as $r_{MZF} > r_{DZF}$ and $r_{MZF} < 1$). The opposite-sex and female dizygotic correlations were similar and less than the male dizygotic correlation.

We then fit a univariate twin model to these five groups. The model fit was adequate (χ^2_4 GOF = 3·23, $P = 0·52$). As was suggested by the patterns of correlations, males and females were similar in the estimate of the proportion of variance due to individual-specific environmental effects. Males had small estimates of genetic effects (6 %) and rather considerable estimates of shared environmental effects (21 %) whereas these proportions were the opposite in females (26 % genetic and 1 % shared environment). However, as shown in Fig. 1, these estimates should be interpreted with caution as the 95 % confidence intervals are broad and contain zero (except for the estimates of individual-specific environmental effects).

Female monozygotic pairs

		Twin 2 No	Twin 2 Yes
Twin 1	No	478	48
	Yes	55	15

Tetrachoric: 0·290 (S.E. 0·098)

Female dizygotic pairs

		Twin 2 No	Twin 2 Yes
Twin 1	No	334	39
	Yes	52	8

Tetrachoric: 0·079 (S.E. 0·122)

Male monozygotic pairs

		Twin 2 No	Twin 2 Yes
Twin 1	No	604	44
	Yes	46	9

Tetrachoric: 0·260 (S.E. 0·110)

Male dizygotic pairs

		Twin 2 No	Twin 2 Yes
Twin 1	No	405	33
	Yes	39	8

Tetrachoric: 0·253 (S.E. 0·122)

Model fitting

	Male	Female
a^2	06 (0–46)	26 (0–44)
c^2	21 (0–41)	01 (0–30)
e^2	73 (54–90)	73 (56–92)
r_g	1·00 (–1 to +1)	

Male–female dizygotic pairs

		Female twin No	Female twin Yes
Male twin	No	844	105
	Yes	82	15

Tetrachoric: 0·106 (S.E. 0·085)

FIG. 2. Univariate twin analysis of interfering fatigue (IF): data from 3263 twin pairs. For each of the five types of twin pairs, the raw data and tetrachoric correlations are shown (with a summary in the lower left panel).

These results are consistent with the somewhat limited power of twin studies of our size for modestly familial discrete traits (Neale *et al.* 1994). Given that IF is clearly heterogeneous, we did not conduct additional twin modelling (e.g. multivariate modelling) as the currently available software does not allow us to account for heterogeneity and covariate effects suggested by the other analyses.

DISCUSSION

In this report, we investigated the prevalence and correlates of interfering fatigue (IF). It is important to review the components of our definition of IF: (*a*) feeling 'tired or fatigued'; (*b*) in the previous year; (*c*) most of the time for 5 days or longer; (*d*) unusual and worse than baseline; and (*e*) interfering with daily life ('completely' or 'a lot'). We believe that this is a reasonably firm definition with face validity and note that prior reports have used components (*d*) (Bates *et al.* 1993) and (*e*) (Bates *et al.* 1993; Addington *et al.* 2001). Because of our interest in the correlates of IF, we did not apply any exclusionary criteria that might have 'explained' IF. Removing individuals with, for example, a known physical illness would have seriously diminished our capacity to address the central aims of this report.

Univariate correlates

The past year prevalence of IF was reported by 765 of 7740 (9·9%) individual twins. Precise comparisons across studies in the literature are difficult owing to definitional differences (Lewis & Wessely, 1992), but the past year prevalence in our sample appears to be comparable to that of prior reports in community samples (Hagnell *et al.* 1993; Fukuda *et al.* 1997; Addington *et al.* 2001) and primary care settings (David *et al.* 1990; Cathébras *et al.* 1992; Bates *et al.* 1993; Fuhrer & Wessely, 1995).

As anticipated, univariate analyses showed that IF was strongly associated with a high percentage of the demographic, health status, physical and mental disorders, personality,

https://doi.org/10.1017/S0033291702007031 Published online by Cambridge University Press

parenting and stressful life events that we assessed. Nearly all of the 42 significant univariate associations were in the direction anticipated (e.g. IF associated with female gender, the presence of major depression and higher neuroticism). These findings are similar to prior studies, particularly the strong association of fatigue and both lifetime and past year major depression (Cathébras *et al.* 1992; Ridsdale *et al.* 1993; Pawlikowska *et al.* 1994; Wessely *et al.* 1996; Addington *et al.* 2001). There were weakly significant protective effect of current marital status and current employment on IF.

Many of the 10 non-significant comparisons showed trends in the expected direction of association with IF (i.e. the presence of young children, worse educational attainment, obesity, current pregnancy, any cancer, or anorexia nervosa) suggesting small effects that did not reach statistical significance with our sample.

Several non-significant findings are notable. The widely used criteria for chronic fatigue syndrome of the Centers for Disease Control (Fukuda *et al.* 1994) list both morbid obesity and any lifetime eating disorder as exclusionary. Although IF is clearly not synonymous with chronic fatigue syndrome, it is surprising that obesity, anorexia, and bulimia nervosa were not among the more strongly discriminating variables across groups. Given that obesity and eating disorders are often chronic conditions, one explanation may be that our requirement that IF be unusual and worse than baseline may have diminished the impact of these correlates. However, these counterintuitive findings suggest the need for further empirical evaluation of these elements of the CDC criteria.

When we ranked the comparisons, the strongest associations were with major depression, generalized anxiety disorder, several health beliefs and increased neuroticism. Given that fatigue is a component of the criteria for both major depression and generalized anxiety disorder, it is notable that the rankings were essentially unchanged when the fatigue criterion was removed.

Multivariate analyses

Analyses that attempt to understand further the data in Table 1 have three important complications. First, there is a high degree of covariation between many of the predictors. For example, in this sample, there are enormous associations between major depression and generalized anxiety disorder ($\chi^2_1 = 1297$, $P \sim 10^{-284}$) and between the presence of major health problems and having spent $\geqslant 15$ sick days in bed in the prior year ($\chi^2_1 = 432$, $P \sim 10^{-96}$). Secondly, it is possible that these predictors interact to influence risk of IF. Modelling interactions is feasible with many statistical procedures particularly when there is strong theory to guide modelling (or, unlike the present circumstance, if there are few predictors). We believe that the theory underlying IF is not strong and essentially all possible interactions must be considered. This poses a dilemma with 24 predictor variables given that there are 24 main effects and 276 two-way interactions: on one hand, there is a non-trivial risk of chance findings (i.e. Type 1 error) and, on the other hand, there is a real danger of 'over-fitting' or capitalizing on chance variation within a dataset to produce spurious improvements in model fit at the expense of generalizibility to other samples (Harrell *et al.* 1985, 1996). Thirdly, it is likely that IF is not a homogeneous construct. There may be several different 'types' of IF each of which with distinctive sets of covariates. For example, IF might have different correlates in men and women or in the presence of a history of major depression. Modelling these processes is feasible (e.g. with mixture models) if the underlying theory is strong. As above, without strong theory, the matter is far more complex.

To address these problems, we applied three relatively distinctive multivariate regression techniques in the hope that an interpretable pattern of results might emerge. We are aware that this report contains a considerable amount of information. To assist in interpretation and integration of the results, Table 5 summarizes the analyses.

The top portion of Table 5 compares and contrasts the univariate and three multivariate methods we applied. Each method had the same general goal: to attempt to explain a dependent variable (IF) as a function on one or more correlates. However, each approach was distinctive – logistic regression uses a linear maximum likelihood approach, CART a complex binary partitioning algorithm (Breiman *et al.* 1984), and MARS a complex algorithm involving the use of 'splines' (Friedman, 1991). As shown in Table 5, these methods have different characteristics

Table 5. *Overview and summary of regression analyses of interfering fatigue**

	Univariate logistic regression	Multivariate: Logistic regression	Multivariate: CART	Multivariate: MARS
Method accounts for				
Clustering of twins	Yes (GEE)	No (with MI)	No	No
Missing data	No	Yes (MI)	Yes	Yes
Covariation	No	Yes	Yes	Yes
Protection against over-fitting	n/a	No	Yes (CV)	Yes (CV)
Interactions	No	Possibly	Yes	Yes
Detecting heterogeneity	No	Possibly	Yes	Yes
Non-linear effects	Possibly	Possibly	Yes	Yes
Age (⩾median)		+		
Female gender	+++			
Currently married	+			
Any child ⩽5 years old		+		
Education ⩽12 years		+		
Residence in city ⩾250 000			+	
Major health problem	+++	+++	+++	
Dissatisfied with health	+++	++		++
At increased risk of early death	+++	++		
⩾15 sick days in bed in previous year	+++	+++		+++
Daily activities limited by health	+++	+++		++
Any physical cause	+++			
Major depression	+++	+++	+++	++
Generalized anxiety disorder	+++	+++	+++	+++
Panic disorder	+++			
Agoraphobia	+++			++
Social phobia	+++			
Alcohol abuse/dependence	+++	+++		+
Nicotine dependence	+++			
Neuroticism (⩾median)	+++	+++	+++	+
Extraversion (⩾median)	++			
Parental care (⩾median)	+++		+	
Parental overprotection (⩾median)	+++			
Any stressful life event (⩾median)	+++	+		+

* Summary of results from Tables 1–3 and Fig. 1. All variable are defined in the footnotes for Table 2. The qualitative importance of each of the 19 predictors is indicated by 0, 1, 2, or 3 plus signs. GEE, Generalized estimating equations; MI, multiple imputation; CV, multifold cross-validation; n/a, not applicable.

which seemed wise to use at the start of this project given the likely complexities of IF.

As it turned out, however, the multivariate methods tended to converge on the same answers. First, the logistic regression results tended to be similar whether or not we accounted for the non-independence or clustering of twins (Table 2). Secondly, the results were not strongly influenced by missing data (Table 2). Thirdly, the multivariate regression techniques tended to yield similar predictions about the probability of IF. The predicted values from the logistic regression model in Table 2 were highly correlated with the predictions from MARS ($r = 0{\cdot}93$, $P < 0{\cdot}0001$). Similarly, the predictions from CART were highly correlated with the logistic regression predictions ($P < 0{\cdot}00001$) and from MARS ($P < 0{\cdot}00001$).

Fourthly, of particular substantive importance, the multivariate results tended to implicate the same correlates or predictor variables. Based on the patterns of results shown in Table 5, we suggest that there are four classes of results. Three variables were significant in all three multivariate analyses (lifetime major depression, lifetime generalized anxiety disorder, and high neuroticism) suggesting that these are robust and potent correlates of IF. Six variables were significant in two multivariate analyses (a reported major health problem, dissatisfaction with health, more sick days in bed, the belief that daily activities were limited by health, alcohol abuse or dependence, and any stressful life event) indicating their status as potentially important correlates of IF. Seven variables were significant in one multivariate analysis. Eight variables were

https://doi.org/10.1017/S0033291702007031 Published online by Cambridge University Press

significant in none of the multivariate analyses. We would contend that the first two groups of variables are more likely to be involved with IF.

Moreover, we suggest that this pattern of results is consistent with a key conclusion regarding IF in a population sample of twins. There appear to be two broad clusters of correlates of IF defined by: (*a*) major depression, generalized anxiety disorder and the anxious-fearful personality trait of neuroticism; and (*b*) by a set of beliefs of ill health coexisting with alcoholism and stressful life events. These results are generally consistent with the correlates of fatigue in the literature (Cathébras *et al.* 1992; Lewis & Wessely, 1992; Hagnell *et al.* 1993; Ridsdale *et al.* 1993; Pawlikowska *et al.* 1994; Fuhrer & Wessely, 1995; Wessely *et al.* 1996; Loge *et al.* 1998; Addington *et al.* 2001).

Moreover, it is notable that female gender and the presence of any potential physical cause of IF were not strong predictors of IF. This does not necessarily mean that these variables are unimportant; rather, we suggest that their effects are accounted for in other ways. For example, the gender difference in IF could have been subsumed by the well-known gender differences for major depression (Weissman *et al.* 1993) and the impact of a physical cause of IF by one or more beliefs of ill health.

Heterogeneity

As a more direct assessment, we used latent class analysis to attempt to delineate heterogeneity within subjects with IF (Table 4). Consistent with prior studies (Hickie *et al.* 1995; Kirk *et al.* 1999; Nimnuan *et al.* 2001), there was strong evidence for the heterogeneity of IF (i.e. a five class rather than a homogeneous one class solution).

Several of the classes corresponded to prior expectations – e.g. a class of older males with lesser education and poor physical health (Class 1), a class with considerable burden of physical and mental illnesses (Class 4), and a class of women who were married with young children and who had recollections of lower parental care and overprotection (Class 2).

The remaining two classes were intriguing – both reported relatively few health problems and had a substantial prevalence of lifetime major depression. However, one class tended to consist of younger, unmarried women with recollections of worse parental rearing (Class 5) and the other of men with recollections of better parental rearing (Class 3).

Of note, the two variables concerning recollections of parental rearing were not prominent in the multivariate analyses of IF (Table 5), but were among the four most discriminating variables in the latent class analysis in Table 4. The lack of detection in the multivariate analyses could have been due to opposing effects in heterogeneous subgroups not detected by the regression methods. These results are consistent with prior indications that childhood rearing might have a direct impact on adult IF (Hotopf *et al.* 2000).

Twin analyses

Given that these data were collected in twins, it was logical to apply twin modelling methods to attempt to delineate the contributions to its variance. The results were intriguing but not conclusive. First, consistent with our prior expectations, individual-specific environmental effects were strong for women and men (73% of the variance in liability to IF). Secondly, the contributions to the familiality of IF were nearly opposite in men and women – men had small estimates of genetic effects (6%) and rather considerable estimates of shared environmental effects (21%) whereas these proportions were 26% genetic and 1% shared environment in females. Thirdly, due to the relatively small sample sizes (Neale *et al.* 1994), however, the confidence intervals for these estimates were large and overlapping and prohibit definitive conclusions about the relationship between gender and IF from the twin perspective.

We chose not to conduct further twin modelling (e.g. multivariate models) beyond the univariate results in Fig. 2 for two reasons. First, we believe that IF is a heterogeneous construct: given heterogeneity and the strong associations of IF with other variables, it is not clear that IF is the 'true' or most 'downstream' variable. For example, these results could reflect the prevalence and heritability of major depression for women or risk factors for poor health that function in the common environment for men. Secondly, rigorous analysis of IF awaits the development and proving of the next generation of

https://doi.org/10.1017/S0033291702007031 Published online by Cambridge University Press

twin methodology. For example, a number of groups are working on Bayesian approaches whose application to twin data could provide an appropriate analytical approach for IF.

Relevance to chronic fatigue syndrome (CFS)

One of the fundamental unanswered questions about CFS is definitional – given that fatigue is a common human complaint but CFS is quite rare (Reyes *et al*. 1997; Steele *et al*. 1998; Jason *et al*. 1999), can an uncommon syndrome be delineated from a common symptom? It seems reasonable to contend that a more complete understanding of IF is likely to inform discussions of CFS. By analogy, more detailed knowledge of endophenotypes like glucose homeostasis or adipocyte mass regulation are being investigated for their relevance to the pathological conditions of diabetes mellitus and obesity. It is impossible to prove that studying IF (or some other definition of common fatigue) will yield insights into CFS; however, this could prove to be a useful research strategy for CFS which continues to carry stigma, controversy and confusion. Moreover, all of the estimated 15 individuals in our sample with current CFS (0·2% × 7740) should have IF. A key question for future research is which of the different types of IF contain most of the individuals with CFS or if, in fact, individuals with CFS constitute a separate group not detected in the current analyses.

Conclusions

First, consistent with the prior literature, IF is a common human symptom with numerous associations with demographic, health status, specific physical disorders, psychopathology, personality, recollections of parental rearing, and stressful life events. The major multivariate correlates were with lifetime major depression, lifetime generalized anxiety disorder and high neuroticism along with major health problems, dissatisfaction with health, more sick days in bed, the belief that daily activities were limited by health, alcohol abuse or dependence, and any stressful life event.

Secondly, as anticipated, IF is a heterogeneous construct that probably consists of a number of different 'types'. IF is not a unitary symptom but rather a complex construct with different sets of correlates in various types.

Thirdly, consistent with the above, the twin analyses suggested etiological heterogeneity in men and women. Genetic effects may be particularly important in women and common or shared environmental effects in men. These conclusions are tempered by the low power of these twin models even with out relatively large sample size.

Finally, the above analyses suggest a preliminary typology of IF. The total number of 'types' of IF is unknown; however, we believe our data are consistent with at least four: (*i*) older, less educated men with physical health problems; (*ii*) women with major depression, generalized anxiety disorder, and stressful life events; (*iii*) healthy married individuals with young children; and (*iv*) healthy, unmarried, and younger individuals living in urban environments.

Limitations

These conclusions must be interpreted in light of a number of limitations. First, these data are cross-sectional and correlative in nature. These data may be consistent with causal pathways but clearly do not demonstrate causation between any of the predictors and IF. Secondly, there were several limitations of the dataset – the FF and MMMF interview waves were separated by as much as a decade, all participants were Caucasians from Virginia, we could not determine the timing or the length of the episodes of IF, we had no objective evaluation of an individual's complaint of IF, and some important variables (e.g. neuroticism) had a relatively large number (8%) of missing values. Thirdly, the low κ coefficient from CART and R^2 from MARS suggested that important correlates of IF were not measured in our study. Finally, we did not take into account the non-independence of twins within a pair in latent class analysis, CART, and MARS.

This study was funded by the National Institutes of Health (AI38429 and NS41483). P.F.S. is a consultant to Cypress Bioscience on the genetics of fibromyalgia. All authors reviewed and approved the final manuscript prior to submission. P.F.S. conceived the manuscript, performed some statistical analyses and wrote the manuscript. P.K. was the primary statistical analyst. T.Y. performed the CART and MARS analyses. K.S.K. was the principal investigator on the projects that collected the primary data.

REFERENCES

Aaron, L. A., Burke, M. M. & Buchwald, D. (2000). Overlapping conditions among patients with chronic fatigue syndrome, fibromyalgia, and temporomandibular disorder. *Archives of Internal Medicine* **160**, 221–227.

Addington, A. M., Gallo, J. J., Ford, D. E. & Eaton, W. W. (2001). Epidemiology of unexplained fatigue and major depression in the community: the Baltimore ECA follow-up, 1981–1994. *Psychological Medicine* **31**, 1037–1044.

Akiake, H. (1987). Factor analysis and AIC. *Psychometrika* **52**, 317–332.

American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association: Washington, DC.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association: Washington, DC.

Barsky, A. J. & Borus, J. F. (1999). Functional somatic syndromes. *Annals of Internal Medicine* **130**, 910–921.

Bates, D. W., Schmitt, W., Buchwald, D., Ware, N. C., Lee, J., Thoyer, E., Kornish, R. J. & Komaroff, A. L. (1993). Prevalence of fatigue and chronic fatigue syndrome in a primary care practice. *Archives of Internal Medicine* **153**, 2759–2765.

Breiman, L., Friedman, J. H., Olshen, R. & Stone, C. J. (1984). *Classification and Regression Trees*. Wadsworth: Belmont, CA.

Bucholz, K. K., Heath, A. C., Reich, T., Hesselbrock, V. M., Kramer, J. R., Numberger, J. I., Jr. & Schuckit, M. A. (1996). Can we subtype alcoholism? A latent class analysis of data from relatives of alcoholics in a multicenter family study of alcoholism. *Alcoholism Clinical Experimental Research* **20**, 1462–1471.

Cathébras, P. J., Robbins, J. M., Kirmayer, L. J. & Hayton, B. C. (1992). Fatigue in primary care: prevalence, psychiatric comorbidity, illness behavior, and outcome. *Journal of General Internal Medicine* **7**, 276–286.

Cella, D., Davis, K., Breitbart, W. & Curt, G. (2001). Cancer-related fatigue: prevalence of proposed diagnostic criteria in a United States sample of cancer survivors. *Journal of Clinical Oncology* **19**, 3385–3391.

Cox, D. R. & Snell, E. J. (1989). *The Analysis of Binary Data*. Chapman and Hall: London.

David, A., Pelosi, A., McDonald, E., Stephens, D., Ledger, D., Rathbone, R. & Mann, A. (1990). Tired, weak, or in need of rest: fatigue among general practice attenders. *British Medical Journal* **301**, 1199–1202.

Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society* **39**, 1–38.

Diggle, P. J., Liang, K.-Y. & Zeger, S. L. (1994). *The Analysis of Longitudinal Data*. Oxford University Press: New York.

Eaves, L. J., Silberg, J. L., Hewitt, J. K., Rutter, M., Meyer, J. M., Neale, M. C. & Pickles, A. (1993). Analyzing twin resemblance in multisymptom data: genetic applications of a latent class model for symptoms of conduct disorder in juvenile boys. *Behavior Genetics* **23**, 5–19.

Efron, B. & Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman and Hall: London.

Eysenck, H. J. & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire*. Hodder and Stoughton: London.

Fagerstrom, K.-O. (1978). Measuring degree of physical dependence to tobacco smoking with reference to individualization of treatment. *Addictive Behaviors* **3**, 235–241.

Fagerstrom, K.-O. & Schneider, N. G. (1989). Measuring nicotine dependence: a review of the Fagerstrom Tolerance Questionnaire. *Journal of Behavioral Medicine* **12**, 159–182.

Friedman, J. H. (1991). Multivariate adaptive regression splines. *Annals of Statistics* **19**, 1–141.

Fuhrer, R. & Wessely, S. (1995). The epidemiology of fatigue and depression: a French primary-care study. *Psychological Medicine* **25**, 895–905.

Fukuda, K., Strauss, S. E., Hickie, I., Sharpe, M. C., Dobbins, J. G. & Komaroff, A. (1994). The chronic fatigue syndrome: a comprehensive approach to its definition and study. *Annals of Internal Medicine* **121**, 953–959.

Fukuda, K., Dobbins, J. G., Wilson, L. J., Dunn, R. A., Wilcox, K. & Smallwood, D. (1997). An epidemiological study of fatigue with relevance for the chronic fatigue syndrome. *Journal of Psychiatric Research* **31**, 19–29.

Grassi, M., Villani, S. & Marinoni, A. (2001). Classification methods for the identification of 'case' in epidemiological diagnosis of asthma. *European Journal of Epidemiology* **17**, 19–29.

Hagnell, O., Gräsbeck, A., Öjesjö, L. & Otterbeck, L. (1993). Mental tiredness in the Lundby study. *Acta Psychiatrica Scandinavica* **88**, 316–321.

Harrell, F. E., Jr., Lee, K. L., Matchar, D. B. & Reichert, T. A. (1985). Regression models for prognostic prediction: advantages, problems, and suggested solutions. *Cancer Treatment Reports* **69**, 1071–1077.

Harrell, F. E., Jr., Lee, K. L. & Mark, D. B. (1996). Multivariate prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine* **15**, 361–87.

Hastie, T., Tibshirani, R. & Friedman, J. H. (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer: New York.

Heath, A. C., Neale, M. C., Kessler, R. C., Eaves, L. J. & Kendler, K. S. (1992). Evidence for genetic influences on personality from self-reports and from informant ratings. *Journal of Personality and Social Psychology* **63**, 85–96.

Hess, K. R., Abbruzzese, M. C., Lenzi, R., Raber, M. N. & Abbruzzese, J. L. (1999). Classification and regression tree analysis of 1000 consecutive patients with unknown primary carcinoma. *Clinical Cancer Research* **5**, 3403–3410.

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**, 925–935.

Hickie, I. & Wakefield, D. (1992). Diagnosing CFS. *CHIDS Chronicle*.

Hotopf, M., Wilson-Jones, C., Mayou, R., Wadsworth, M. & Wessely, S. (2000). Childhood predictors of adult medically unexplained hospitalisations. Results from a national birth cohort study. *British Journal of Psychiatry* **176**, 273–280.

Jason, L. A., Richman, J. A., Rademaker, A. W., Jordan, K. M., Plioplys, A. V., Taylor, R. R., Mccready, W., Huang, C. F. & Plioplys, S. (1999). A community-based study of chronic fatigue syndrome. *Archives of Internal Medicine* **159**, 2129–2137.

Kendler, K. S. (1993). Twin studies of psychiatric illness: current status and future directions. *Archives of General Psychiatry* **50**, 905–915.

Kendler, K. S. (1996). Parenting: a genetic-epidemiologic perspective. *American Journal of Psychiatry* **153**, 11–20.

Kendler, K. S. (2001). Twin studies of psychiatric illness: an update. *Archives of General Psychiatry* **58**, 1005–1014.

Kendler, K. S. & Prescott, C. A. (1999). A population-based twin study of lifetime major depression in men and women. *Archives of General Psychiatry* **56**, 39–44, erratum (2000) **57**, 94–95.

Kendler, K. S., Silberg, J. L., Neale, M. C., Kessler, R. C., Heath, A. C. & Eaves, L. J. (1990). Genetic and environmental factors in the aetiology of menstrual, premenstrual and neurotic symptoms: a population-based twin study. *Psychological Medicine* **22**, 85–100.

Kendler, K. S., Maclean, C., Neale, M., Kessler, R., Heath, A. & Eaves, L. (1991). The genetic epidemiology of bulimia nervosa. *American Journal of Psychiatry* **148**, 1627–1637.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C. & Eaves, L. J. (1992*a*). A population-based twin study of major depression in women: the impact of varying definitions of illness. *Archives of General Psychiatry* **49**, 257–266.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C. & Eaves, L. J. (1992*b*). Generalized anxiety disorder in women: a

https://doi.org/10.1017/S0033291702007031 Published online by Cambridge University Press

population-based twin study. *Archives of General Psychiatry* **49**, 267–272.

Kendler, K. S., Heath, A. C., Neale, M. C., Kessler, R. C. & Eaves, L. J. (1992*c*). A population-based twin study of alcoholism in women. *Journal of the American Medical Association* **268**, 1877–1882.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C. & Eaves, L. J. (1992*d*). The genetic epidemiology of phobias in women: the interrelationship of agoraphobia, social phobia, situational phobia, and simple phobia. *Archives of General Psychiatry* **49**, 273–281.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C. & Eaves, L. J. (1993). Panic disorder in women: a population-based twin study. *Psychological Medicine* **23**, 397–406.

Kendler, K. S., Kessler, R. C., Walters, E. E., Maclean, C. J., Neale, M. C., Heath, A. C. & Eaves, L. J. (1995). Stressful life events, genetic liability, and onset of an episode of major depression in women. *American Journal of Psychiatry* **152**, 833–842.

Kendler, K. S., Neale, M. C., Sullivan, P. F., Gardner, C. O. & Prescott, C. A. (1999). A population-based twin study in women of smoking initiation and nicotine dependence. *Psychological Medicine* **29**, 299–308.

Kirk, J., Douglass, R., Nelson, E., Jaffe, J., Lopez, A., Ohler, J., Blanchard, C., Chapman, R., McHugo, G. & Stone, K. (1990). Chief complaint of fatigue: a prospective study. *Journal of Family Practice* **30**, 33–39.

Kirk, K. M., Hickie, I. B. & Martin, N. G. (1999). Fatigue as related to anxiety and depression in a community-based sample of twins aged over 50. *Social Psychiatry and Psychiatric Epidemiology* **34**, 85–90.

Konen, J. C., Curtis, L. G. & Summerson, J. H. (1996). Symptoms and complications of adult diabetic patients in a family practice. *Archives of Family Medicine* **5**, 135–145.

Lawrie, S. M. & Pelosi, A. J. (1995). Chronic fatigue syndrome in the community: prevalence and associations. *British Journal of Psychiatry* **166**, 793–797.

Lewis, G. & Wessely, S. (1992). The epidemiology of fatigue: more questions than answers. *Journal of Epidemiology and Community Health* **46**, 92–97.

Littell, R. C., Milliken, G. A., Stroup, W. W. & Wolfinger, R. D. (1999). *SAS System for Mixed Models*. SAS Publishing: Cary, NC.

Loge, J. H., Ekeberg, Ø. & Kaasa, S. (1998). Fatigue in the general Norwegian population. *Journal of Psychosomatic Research* **45**, 53–65.

McCutcheon, A. L. (1987). *Latent Class Analysis*. Sage Publications: Beverly Hills.

Neale, M. C. & Cardon, L. R. (1992). *Methodology for the Study of Twins and Families*. Kluwer Academic Publisher Group: Dordrecht, the Netherlands.

Neale, M. C. & Miller, M. B. (1997). The use of likelihood-based confidence intervals in genetic models. *Behavior Genetics* **27**, 113–120.

Neale, M. C., Eaves, L. J. & Kendler, K. S. (1994). The power of the classical twin study to resolve variation in threshold traits. *Behavior Genetics* **24**, 239–258.

Neale, M. C., Boker, S., Xie, G. & Maes, H. (1999). *Mx: Statistical Modeling*. Medical College of Virginia, Department of Psychiatry: Richmond, VA.

Nimnuan, C., Rabe-Hesketh, S., Wessely, S. & Hotopf, M. (2001). How many functional somatic syndromes? *Journal of Psychosomatic Research* **51**, 549–557.

Parker, G., Tupling, H. & Brown, L. (1979). A parental bonding instrument. *British Journal of Medical Psychology* **52**, 1–10.

Pawlikowska, T., Chalder, T., Hirsch, S. R., Wallace, P., Wright, D. J. M. & Wessely, S. C. (1994). Population based study of fatigue and psychological distress. *British Medical Journal* **308**, 763–766.

Plomin, R., Defries, J. C., McClearn, G. E. & Rutter, M. (1997). *Behavioral Genetics*. W H Freeman and Company: New York.

Reyes, M., Gary, H. E., Dobbins, J. G., Randall, B., Steele, L., Fukuda, K., Holmes, G. P., Connell, D. G., Mawle, A. C., Schmid, D. S., Stewart, J. A., Schonberger, L. B., Gunn, W. J. & Reeves, W. C. (1997). Surveillance for chronic fatigue syndrome. *Morbidity and Mortality Weekly Report* **46**, 1–15.

Ridsdale, L., Evans, A., Jerrett, W., Mandalia, S., Osler, K. & Vora, H. (1993). Patients with fatigue in general practice: a prospective study. *British Medical Journal* **307**, 103–106.

Robins, L. N. & Helzer, J. E. (1985). *Diagnostic Interview Schedule (DIS): Version III-A*. Washington University School of Medicine: St Louis, MO.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley and Sons, Inc.: New York.

Salford Systems (2001*a*). *CART™ User's Guide*. Salford Systems: San Diego, CA.

Salford Systems (2001*b*). *MARS™ User's Guide*. Salford Systems: San Diego, CA.

SAS Institute Inc. (1999). *SAS/STAT® Software: Version 8*. SAS Institute, Inc.: Cary, NC.

Sharpe, M. (1996). Chronic fatigue syndrome. *Psychiatric Clinics of North America* **19**, 549–573.

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. *et al.* (1991). A report – chronic fatigue syndrome: guidelines for research. *Journal of the Royal Society of Medicine* **84**, 118–121.

Spence, J. E., Corey, L. A., Nance, W. E., Marazita, M. L., Kendler, K. S. & Schieken, R. M. (1988). Molecular analysis of twin zygosity using VNTR DNA probes. *American Journal of Human Genetics* **43**, A159.

Spitzer, R. L., Williams, J. B. W., Gibbon, M. & First, M. B. (1992). The Structured Clinical Interview for DSM-III-R (SCID) I: history, rationale, and description. *Archives of General Psychiatry* **49**, 624–629.

Steele, L., Dobbins, J. G., Fukuda, K., Reyes, M., Randall, B., Koppelman, M. & Reeves, W. C. (1998). The epidemiology of chronic fatigue in San Francisco. *American Journal of Medicine* **105**(3A), 83S–90S.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictors. *Journal of the Royal Statistical Society* **36**, 111–147.

Sullivan, P. F. & Eaves, L. J. (2002). Evaluation of analyses of univariate discrete twin data. *Behavior Genetics* **32**, 221–227.

Thompson, W. G., Longstreth, G. F., Drossman, D. A., Heaton, K. W., Irvine, E. J. & Muller-Lissner, S. A. (1999). Functional bowel disorders and functional abdominal pain. *Gut* **45** (Suppl. 2), 1143–1147.

Venables, W. N. & Ripley, B. D. (1999). *Modern Applied Statistics With S-Plus*. Springer-Verlag: New York.

Vogl, D., Rosenfield, B., Breitbart, W., Thaler, H., Passik, S., Mcdonald, M. & Portenoy, R. K. (1999). Symptom prevalence, characteristics, and distress in AIDS outpatients. *Journal of Pain Symptom Management* **18**, 253–262.

Walters, E. E. & Kendler, K. S. (1995). Anorexia nervosa and anorexic-like syndromes in a population-based female twin sample. *American Journal of Psychiatry* **152**, 64–71.

Weissman, M. M., Bland, R., Joyce, P. R., Newman, S., Wells, J. E. & Wittchen, H.-U. (1993). Sex differences in rates of depression: cross-national perspectives. *Journal of Affective Disorders* **29**, 77–84.

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. *American Journal of Psychiatry* **153**, 1050–1059.

Wessely, S., Nimnuan, C. & Sharpe, M. (1999). Functional somatic syndromes: one or many? *Lancet* **354**, 936–939.

Williams, J. B. W., Gibbon, M., First, M. B., Spitzer, R. L., Davies, M., Borus, J., Howes, M. J., Kane, J., Pope, H. G., Rounsaville, B. & Wittchen, H.-U. (1992). The Structured Clinical Interview for DSM-III-R (SCID) II: multisite test-retest reliability. *Archives of General Psychiatry* **49**, 630–636.

Williams, L. J. & Holahan, P. J. (1994). Parsimony-based fit indices for multiple-indicator models: do they work? *Structural Equation Modeling* **1**, 161–189.

Wolfe, F., Smythe, H. A., Yunus, M. B., Bennett, R. M., Bombardier, C., Goldenberg, D. L., Tugwell, P., Campbell, S. M., Abeles, M., Clark, P., Fam, A. G., Farber, S. J., Fiechtner, J. L., Franklin, C. M., Gatter, R. A., Hamaty, D., Lessard, J., Lichtbroun, A. S., Masi, A. T., Mccain, G. A., Reynolds, W. J., Romano, T. J., Russell, I. J. & Sheon, R. P. (1990). The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia. Report of the Multicenter Criteria Committee. *Arthritis and Rheumatism* **33**, 160–172.

Wolfe, F., Ross, K., Anderson, J., Russell, I. J. & Hebert, L. (1995). The prevalence and characteristics of fibromyalgia in the general population. *Arthritis and Rheumatism* **38**, 19–28.

World Health Organization (1993). *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. World Health Organization: Geneva.

Yang, I. & Becker, M. P. (1997). Latent variable modeling of diagnostic accuracy. *Biometrics* **53**, 948–958.

York, T. P. & Eaves, L. J. (2001). Common disease analysis using Multivariate Adaptive Regression Splines (MARS). *Genetic Epidemiology* **21** (Suppl. 1). S649–S654.

Yuan, Y. C. (2000). P267-25: Multiple imputation for missing data: concepts and new development. In *SAS White Papers* (ed. SAS Institute), pp. 1–11. SAS Institute: Rockville, MD.

Zeger, S. L. & Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* **42**, 121–130.

Zeger, S. L. & Liang, K. Y. (1992). An overview of methods for the analysis of longitudinal data. *Statistics in Medicine* **11**, 1825–1839.

Zeger, S. L., Liang, K. Y. & Albert, P. S. (1988). Models for longitudinal data: a generalized estimating equation approach. *Biometrics* **44**, 1049–1060 (erratum (1989) **45**, 347).

Zhang, P. (1993). Model selection via multifold cross validation. *Annals of Statistics* **21**, 299–313.