Fatigue among employees : treatment in primary care and associations with chronic fatigue syndrome

Citation for published version (APA):

Huibers, M. J. H. (2003). *Fatigue among employees : treatment in primary care and associations with chronic fatigue syndrome*. [Doctoral Thesis, Maastricht University]. Universiteit Maastricht. <https://doi.org/10.26481/dis.20031217mh>

Document status and date:

Published: 01/01/2003

DOI:

[10.26481/dis.20031217mh](https://doi.org/10.26481/dis.20031217mh)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 26 Apr. 2023

FATIGUE AMONG EMPLOYEES

FATIGUE AMONG EMPLOYEES

Treatment in primary care and associations with chronic fatigue syndrome

M.J.H. Huibers

ISBN: 90-90174-77-X

COVER Mikon van Gastel | Imaginary Forces New York

LAY-OUT Suzanne Mol | MOL grafische vormgeving

PRODUCTION Drukkerij Wilko, Amersfoort

© 2003, M.J.H. Huibers

All rights reserved. No part of this thesis may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage or retrieval system, without permission in writing from the author, or, when appropriate, from the publishers of the publications.

FATIGUE AMONG EMPLOYEES

Treatment in primary care and associations with chronic fatigue syndrome

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan
de Universiteit Maastricht,
op gezag van de Rector Magnificus,
Prof. dr. A.C. Nieuwenhuijzen Kruseman,
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op woensdag 17 december 2003 om 16.00 uur precies

door

Marcus J.H. Huibers

PROMOTORES

Prof. dr. C.P. van Schayck

Prof. dr. G. Bleijenberg (UMC St. Radboud, Nijmegen)

Prof. dr. J.A. Knottnerus

CO-PROMOTOR

Dr. A.J.H.M. Beurskens

BEOORDELINGSCOMMISSIE

Prof. dr. ir. P.A. van den Brandt (voorzitter)

Prof. dr. A. Arntz

Prof. dr. G.J. Dinant

Prof. dr. P. Knipschild

Prof. dr. J.W.M. van der Meer (UMC St. Radboud, Nijmegen)

The studies presented in this thesis were conducted at the Departments of Epidemiology and General Practice, Care and Public Health Research Institute (Caphri), Maastricht University, and the Department of Medical Psychology, UMC St. Radboud Nijmegen.

Studies were funded by the Health Research and Development Council (ZorgOnderzoek Nederland), The Netherlands (grant nr. 2830180). The recruitment of patients was made possible by the donation of services by 'Arbo Unie Limburg' (Occupational Health Service Limburg). The Maastricht Cohort Study is part of The Netherlands concerted action on 'Fatigue at Work' granted by The Netherlands Organization for Scientific Research (NWO grant nr. 580-02.201).

Financial support for the printing of this thesis was kindly provided by the Departments of Epidemiology and Medical, Clinical and Experimental Psychology (Maastricht University), the Section Occupational Epidemiology (Maastricht University), the 'Fatigue at Work' programme (NWO), the Dutch Cochrane Centre and Arbo Unie Maastricht/Hoensbroek.

Voor Marieke

Contents

CHAPTER 1	
Introduction	9
CHAPTER 2	
The effectiveness of psychosocial interventions delivered by general practitioners: a Cochrane Review	15
CHAPTER 3	
An alternative trial design to overcome validity and recruitment problems in primary care intervention studies	51
CHAPTER 4	
Efficacy of cognitive behaviour therapy by general practitioners for unexplained fatigue among employees: a randomised controlled trial	67
CHAPTER 5	
Why did cognitive behaviour therapy by general practitioners for unexplained fatigue among employees not work? A qualitative analysis of the delivery of the intervention	83
CHAPTER 6	
Predictors of outcome in fatigued employees on sick leave: results from a randomised trial	99
CHAPTER 7	
Predicting the two-year course of fatigue and the onset of long-term sickness absence in fatigued employees: results from the Maastricht Cohort Study	115
CHAPTER 8	
Fatigue, burnout and chronic fatigue syndrome among employees on sick leave: do attributions make the difference?	131

CHAPTER 9

Prevalence of chronic fatigue syndrome-like caseness in the working population: results from the Maastricht Cohort Study	147
---	-----

CHAPTER 10

The development of the chronic fatigue syndrome in fatigued employees: predictors of outcome in the Maastricht Cohort Study	157
--	-----

CHAPTER 11

General discussion	175
--------------------	-----

<i>Samenvatting (Dutch summary)</i>	189
-------------------------------------	-----

<i>Dankwoord</i>	199
------------------	-----

<i>About the author</i>	201
-------------------------	-----

<i>List of publications</i>	203
-----------------------------	-----

Introduction

Fatigue among employees in the working population is widely recognized as a considerable health problem. The reported prevalence of fatigue as a complaint is estimated to be 22% in the working population.¹ When fatigue becomes severe and persistent, it can have serious unfavourable consequences such as reduced performance, functional impairment, sickness absence² and even work disability.³ In the Netherlands, recognition of fatigue among employees as a policy problem has led to the large-scale research program 'Fatigue at Work' in 1996, a national initiative in which over sixty researchers from many different backgrounds participate.

DEFINING THE CONCEPT OF FATIGUE

Fatigue is a difficult concept. The aetiology of fatigue is anything but clear, medical explanations do not apply in most cases despite the physical aspects of fatigue, and the associations with other health complaints and factors are far-reaching and confounding.⁴ These observations have led some authors to regard fatigue as an expression of a functional syndrome.^{5, 6} In order to study fatigue, it is essential to choose a conceptual approach. We endorse the view of Lewis and Wessely who, in 1992, proposed fatigue as a continuum of severity, as opposed to a discrete disorder.⁷ The continuum approach accounts for the many different manifestations of fatigue that can be observed, ranging from the mild complaints of fatigue that everyone experiences once in a while to severe, persisting, disabling fatigue disorders such as the chronic fatigue syndrome (CFS).

The subject of the present thesis is severe, persistent, medically unexplained fatigue among employees. When placed on the fatigue continuum that runs from left (mild complaints) to right (severe complaints), we study an advanced stage of fatigue that has passed the centre on the right. The proposed position of our subject is displayed in the figure below. It should be noted that fatigue among employees does not overlap conceptually with CFS.

PROPOSED POSITION OF FATIGUE AMONG EMPLOYEES ON THE FATIGUE CONTINUUM

We do not refer to our subject as work-related fatigue. We feel this term is confusing since it implies that fatigue is directly caused by factors at work or that fatigue only acts up in the working situation. Rather, we refer to our subject as severe, persistent, unexplained fatigue among employees, in short, *fatigue among employees*, an expression that does not refer to the aetiology or situational aspects of fatigue.

RESEARCH THEMES

In this thesis, we focus on two main research themes. The first theme pertains the treatment options in primary care for fatigued employees who are absent from work on sick leave. The second theme deals with possible relations between fatigue among employees and the chronic fatigue syndrome, a similar but conceptually different disorder. We take up a multidisciplinary approach, combining insights and expertise from the fields of Epidemiology, General Practice and Medical Psychology. In the remainder of this chapter, I will describe the background and rationale of both themes and subsequently introduce the chapters in which we address the underlying main research questions.

TREATMENT OF FATIGUE IN PRIMARY CARE

A major part of the present thesis relates to a randomised controlled trial in which we evaluate the efficacy of cognitive behaviour therapy delivered by general practitioners to fatigued employees on sick leave. As stated, our subject is fatigue among employees, but one might ask, why general practitioners, why cognitive behaviour therapy, and why employees on sick leave?

Most individuals in the Netherlands with health problems, whether they are employees or not, come to their general practitioner (GP) first. Approximately five to ten percent of the patients in primary care present with fatigue as their main complaint.⁴ In most cases, a medical explanation cannot be found.^{8, 9} There are no evidence-based treatments available for this group of patients and referral to secondary health care often does not take place due to the general, indistinct nature of the complaint. As a result, fatigued patients rely heavily on the care of their GP, placing time-consuming demands on the already overloaded general practice. Therefore, GPs might benefit from tools to help these patients efficiently and effectively.

In recent years, several studies have demonstrated that cognitive behaviour therapy (CBT) by professional therapists is effective in the treatment of chronic fatigue syndrome.¹⁰ One of these studies was conducted by members of our research group.¹¹ A central assumption in CBT is that cognitive and behavioural factors play a role in the perpetuation of the complaint. Given its success in secondary care, we hypothesized that CBT for fatigue might be a promising tool for GPs in primary care as well, a hypothesis that was later on proposed by others as well.⁴

Our hypothesis raises two main research questions. Firstly, are GPs able to deliver CBT effectively?

And secondly, is a short CBT programme modified for use in primary care effective in the treatment of fatigued patients who are presumably less advanced than chronic fatigue syndrome patients?

We chose to target our CBT intervention at fatigued employees on sick leave for some time. With the continuum of fatigue in mind, we aimed to select study participants who would be less advanced than CFS patients, but whose fatigue complaints and functional impairment would be severe enough to justify a short but intensive treatment in primary care. When severe, persistent fatigue becomes associated with sick leave, return to work becomes an important objective, from a societal as well as a health perspective. Consequently, a third main research question emerges: does CBT delivered by GPs establish the resumption of work?

To address the first main research question, we conducted a systematic review of the effectiveness of CBT and other psychosocial interventions that are delivered by GPs in primary care. The results of our review are described in *chapter 2*. In *chapter 3*, we present the considerations for the alternative design of the randomised controlled trial we applied. Main research questions two and three are addressed in *chapter 4* in which we present the efficacy of the intervention. We have evaluated the delivery of the intervention in relation to the effects of CBT, and the results are described in *chapter 5*. In *chapter 6*, predictors of fatigue and work resumption in our trial population are presented. Finally, in *chapter 7* we report on the predictors of fatigue and sickness absence in the Maastricht Cohort Study, an affiliated prospective study in which over 12.000 employees were followed for three years.¹²

FATIGUE AMONG EMPLOYEES IN RELATION TO CHRONIC FATIGUE SYNDROME

In this second part, we focus on associations of fatigue among employees and (the development of) the chronic fatigue syndrome. Although fatigue among employees is considered to be a serious problem, little research is available on the clinical features of the condition. In that respect, the chronic fatigue syndrome has attracted a wider range of attention in recent years.¹³ The chronic fatigue syndrome (CFS) is a condition characterized by persistent medically unexplained fatigue lasting for at least six months, several other unexplained symptoms (e.g. sore throat, headache, painful joints) and severe functional impairment.¹⁴ The reported prevalence ranges from 0.006% to 3.0%.¹⁵ The cause of CFS is unknown and most patients do not recover spontaneously.¹⁵ On the fatigue continuum, CFS occupies the most extreme position on the right (see figure).

Why is it so important then to explore the possible relations between fatigue among employees and CFS? First of all, the two conditions share several important characteristics, such as severe, persistent, unexplained fatigue and functional impairment, in some instances leading to sickness absence and work disability. But more importantly, in concept the two conditions are two different stages on the same continuum of fatigue severity. As stated, the major advantage of the concept of fatigue as a continuum is to account for different manifestations of fatigue. Two important inferences can be drawn

from this understanding. Firstly, depending on changes in their fatigue condition, individuals can actually shift back and forth on the continuum. Secondly, if stages on the continuum are related in some way, one condition might lead to another, more severe condition. The latter inference is of particular importance since it refers to a shared aetiology and the possibility of prevention in previous stages. Considering the above, the main research questions underlying the possible associations of fatigue among employees and CFS are twofold. First, are fatigue among employees and CFS different but related stages on the continuum of fatigue severity? And if so, can fatigue among employees eventually lead to the development of CFS?

In *chapter 6*, we investigate the predictors of both fatigue and chronic fatigue syndrome at follow-up in our trial population. In *chapter 8*, we address the first main research question by investigating the baseline characteristics of our trial population of fatigued employees and comparing them to diagnosed CFS patients from another study. The prevalence of CFS in the working population is estimated in *chapter 9*, based on data from the Maastricht Cohort Study. Finally, we address the second main research question in *chapter 10* when we investigate the development of CFS, the persistence or recurrence of fatigue, and the recovery from fatigue in fatigued employees of the Maastricht Cohort study.

References

1. Bültmann U, Kant IJ, Kasl SV, Beurskens AJHM, Van den Brandt PA. Fatigue and psychological distress in the working population: psychometrics, prevalence and correlates. *J Psychosom Res* 2002; 52:445-52.
2. Janssen N, Kant IJ, Swaen GMH, Janssen PPM, Schröer CAP. Fatigue as a predictor of sickness absence: results from the Maastricht cohort study on fatigue at work. *Occ Environ Med* 2003; 60 (Suppl 1):i71-i6.
3. Amelsvoort van LGPM, Kant IJ, Beurskens AJHM, Schröer CAP, Swaen GMH. Fatigue as a predictor of work disability. *Occup Environ Med* 2002; 59:712-713.
4. Sharpe M, Wilks D. ABC of psychological medicine: Fatigue. *BMJ* 2002; 325:480-3.
5. Wessely S, Nimnuan C, Sharpe M. Functional somatic syndromes: one or many? *Lancet* 1999; 354:936-9.
6. Mayou R, Farmer A. ABC of psychological medicine: Functional somatic symptoms and syndromes. *BMJ* 2002; 325:265-8.
7. Lewis G, Wessely S. The epidemiology of fatigue: more questions than answers. *J Epidemiol Community Health* 1992; 46:92-7.
8. Knottnerus JA, Knipschild PG, Wersch JWW van, Sijstermanns AHJ. Unexplained fatigue and hemoglobin. *Can Fam Physician* 1986; 32:1601-4.
9. Knottnerus JA, Starmans JA, Vissers A. Diagnostische conclusies naar aanleiding van onverklaarde vermoeidheid. *Huisarts Wet* 1987; 30:9-12.
10. Whiting P, Bagnall AM, Sowden AJ, Cornell JE, Mulrow CD, Ramirez G. Interventions for the treatment and management of chronic fatigue syndrome: a systematic review. *JAMA* 2001; 286:1360-8.
11. Prins JB, Bleijenberg G, Bazelmans E, et al. Cognitive behaviour therapy for chronic fatigue syndrome: a multicentre randomised controlled trial. *Lancet* 2001; 357:841-7.
12. Kant IJ, Bültmann U, Schroer CAP, Beurskens AJHM, Van Amelsvoort LGPM, Swaen GMH. An epidemiological approach to study fatigue in the working population: The Maastricht Cohort Study. *Occup Environ Med* 2003; 60 (Suppl 1):i32-i9.
13. Wessely S, Hotopf M, Sharpe M. Treatment of CFS: the evidence. *Chronic Fatigue and its Syndromes*. Oxford New York Tokyo: Oxford University Press, 1998:357-98.
14. 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.
15. Afari N, Buchwald D. Chronic Fatigue Syndrome: a Review. *American Journal of Psychiatry* 2003; 160:221-36.

The effectiveness of psychosocial interventions delivered by general practitioners: a Cochrane Review

Marcus J.H. Huibers ^{1,2}

Anna J.H.M. Beurskens ¹

Gijs Bleijenberg ³

Constant P. van Schayck ²

THE COCHRANE LIBRARY, ISSUE 2, 2003

¹ DEPARTMENT OF EPIDEMIOLOGY, MAASTRICHT UNIVERSITY, THE NETHERLANDS

² DEPARTMENT OF GENERAL PRACTICE, MAASTRICHT UNIVERSITY, THE NETHERLANDS

³ DEPARTMENT OF MEDICAL PSYCHOLOGY, UMC ST RADBOUD NIJMEGEN, THE NETHERLANDS

Abstract

BACKGROUND

Many patients visit their general practitioner (GP) because of problems that are psychosocial in origin. However, for many of these problems there is no evidence-based treatment available in primary care, and these patients place time-consuming demands on their GP. Therefore, GPs could benefit from tools to help these patients more effectively and efficiently. In this light, it is important to assess whether structured psychosocial interventions might be an appropriate tool for GPs. Previous reviews have shown that psychosocial interventions in primary care seem more effective than usual care. However, these interventions were mostly performed by health professionals other than the GP.

OBJECTIVES

To present a systematic review of the literature addressing the effectiveness of psychosocial interventions by general practitioners by assessing the clinical outcomes and the methodological quality of selected studies.

SEARCH STRATEGY

The literature search was conducted using the CCDAN Trials Register, the Cochrane Library and reference lists of relevant studies for citation tracking. Also, personal communication with experts took place.

SELECTION CRITERIA

Randomised controlled trials, controlled clinical trials and controlled patient preference trials addressing the effectiveness of psychosocial interventions by GPs for any problem or disorder. Studies published before January 2002 were eligible for entry.

DATA COLLECTION AND ANALYSIS

Methodological quality was independently be assessed by two reviewers using the Maastricht-Amsterdam Criteria List and the CCDAN Quality Rating Scale. The qualitative and quantitative characteristics of selected trials were independently extracted by two reviewers using a standardised data extraction form. Levels of evidence were used to determine the strength of the evidence available. Results from studies that reported similar interventions and outcome measures were meta-analysed.

MAIN RESULTS

Eight studies were included in the review. Selected studies addressed different psychosocial interventions for four distinct disorders or health complaints. There is good evidence that problem-solving treatment by general practitioners is effective for major depression. The evidence concerning the remaining interventions for other health complaints (reattribution or cognitive behavioural group therapy for somatisation, counselling for smoking cessation, behavioural interventions to reduce alcohol reduction) is either limited or conflicting.

REVIEWERS' CONCLUSIONS

In general, there is little available evidence on the use of psychosocial interventions by general practitioners. Of the psychosocial interventions reviewed, problem-solving treatment for depression seems the most promising tool for GPs, although a stronger evidence-base is required and the effectiveness in routine practice remains to be demonstrated. More research is required to improve the evidence-base on this subject.

Background

Many patients visit their general practitioner (GP) because of problems that are psychosocial in origin. The reported prevalence of psychological or psychosocial disorders (e.g. depression, anxiety, stress, somatisation, unexplained or functional symptoms)¹ in primary care ranges from 30% to 70%.^{2, 3} However, the extent of mental health disorders in general practice is difficult to estimate due to problems of classification.⁴

There is an increasing awareness among GPs that responsibility for managing these patients falls mainly to them.⁵ However, for many of these complaints, there is no evidence-based treatment available. Immediate referral to specialist care is often not an option, because of waiting lists or the vague nature of complaints. These patients often visit their GP frequently without being offered an appropriate treatment and place time-consuming demands on already overloaded primary care services. Therefore, GPs need the tools to help these patients in an effective and efficient way.

Previously, several reviews have addressed the efficacy or effectiveness of psychosocial interventions (e.g. counselling, problem solving therapy, cognitive behavioural interventions, psychotherapy) in primary care.⁶⁻¹⁰ A general characteristic of psychosocial interventions is that some sort of psychological process is the central dynamic in the treatment of the patient. The overall conclusion in these reviews was that psychosocial interventions seem more effective than the usual care that is given by the GP, but that clear clinical evidence is lacking. Rowland⁹ points out in her systematic review (an updated Cochrane Library version was headed by Bower) on counselling for depression that the four studies she was able to include had many methodological deficiencies, which makes the effectiveness of these studies harder to interpret. Friedli⁸ states that despite some supporting evidence for the effectiveness of counselling in primary care, it is still unknown which type of psychosocial intervention is helpful for which patient in primary care. These reviews also demonstrated how difficult it is to define 'psychosocial interventions'.

However, the psychosocial interventions addressed in the aforementioned reviews were mostly performed by primary care workers other than the GP (e.g. nurses, counsellors, psychologists, psychiatrists, internists), leaving the question of whether psychosocial interventions performed by GPs might be effective. In a narrative review of psychological management approaches by general practitioners, Cape et al. came to the conclusion that the preliminary evidence for the clinical effectiveness of GP psychological management in routine consultations is scarce but encouraging.^{11, 12}

There are several reasons why GPs could use the knowledge and skills to perform psychosocial interventions in primary care. As patients come to their GP first with their health concerns, it is desirable that all potential treatment options in primary care are considered before a patient is referred to specialist care. This approach is known as the 'stepped-care principle'.^{13, 14} Secondly, many GPs already take the time to support their distressed patients without having adequate tools to structure these extend-

ed visits. Although these visits can be helpful, the contribution to the overall improvement often seems small against the investment of time. Therefore, many GPs would benefit from psychosocial skills that enable them to use their time more efficiently. Thirdly, the degree of success of any kind of psychosocial intervention depends largely on the trust one places in the care-provider. Since patients and GPs already maintain a relationship, it can be assumed that the familiarity of the doctor's office is preferred to visiting an unknown non-GP primary care worker. Finally, it is recommendable to endorse a holistic approach of patient care ¹⁵ in which physicians care not only about the somatic aspects of care but rather try to improve the well-being of patients in all aspects related to health.

Objectives

In this review, we aimed to present a systematic review of all the available literature addressing the effectiveness of psychosocial interventions by general practitioners. An important underlying objective was to assess whether it is wise to have GPs delivering psychosocial interventions in the first place, for the reasons mentioned in the Background section. In most instances, psychosocial interventions are performed by psychologists and psychiatrists, not by general practitioners. We might have started with assessing the effects of training GPs in applying psychosocial interventions and focus on the feasibility. Instead, we chose to investigate the effects of psychosocial interventions by GPs on patients' well-being, regardless of the anticipated diversity (and therefore incomparability) in type of interventions, participants and outcomes among studies eligible for inclusion.

The objectives of this review were:

1. To identify all randomised controlled trials (RCTs) controlled clinical trials (CCTs) and controlled patient preference trials (CPPTs) of psychosocial interventions delivered by general practitioners.
2. To assess the methodological quality and the relevant characteristics of the selected studies.
3. To assess the effectiveness of psychosocial interventions by general practitioners compared to the reference treatment (whether 'usual care' or another experimental intervention) by reviewing the clinical outcomes of the selected studies.

Criteria for considering studies for this review

TYPES OF STUDIES

Studies published before January 2002 were eligible for entry in the review. Identified trials in all languages were considered eligible. If necessary, assistance was going to be sought from the Cochrane Collaboration Depression Anxiety and Neurosis group (CCDAN) for translation purposes.

All RCTs, CCTs and CPPTs addressing the effectiveness of psychosocial interventions by GPs were considered eligible for inclusion in the study. Patient preference for a certain treatment option is a strong predictor for patient compliance and thus for the effectiveness of the intervention. CPPTs, in which patients with a strong preference for a particular treatment are not randomised, but are allocated to the treatment of their choice, would therefore be included in the study.

Bower discussed the different objectives and designs of pragmatic versus explanatory trials.¹⁶ Explanatory trials aim to isolate the active elements of an intervention in order to determine the relative potential of the treatment under highly controlled circumstances, whereas pragmatic trials aim to determine the relative value of the treatment in the routine practice settings. Although explanatory studies serve the objective of this review best, pragmatic trials were also eligible for entry in the review as long as the psychosocial intervention was standardised to some degree (see section 'types of intervention').

TYPES OF PARTICIPANTS

As was stated in the objective, there were no restrictions on the type of participants in studies to be selected.

TYPES OF INTERVENTIONS

All trials reporting the effectiveness of psychosocial interventions delivered by GPs were eligible for inclusion. A methodological difficulty was choosing a clear and objective definition for a psychosocial intervention. Rowland¹⁷ and Bower¹⁶ present the British Association for Counselling definition of counselling as appropriate for use in their systematic reviews. However, counselling is "a broadly non-directive approach"¹⁷ while we chose to also include other, more structured approaches, like cognitive behavioural interventions or problem solving therapy.

In this review, a GP-administered psychosocial intervention should meet the following criteria:

1. the intervention is explicitly delivered by a GP (or family physician or family doctor), although the GP intervention may be compared with a similar intervention administered by a different health professional.

2. the intervention is a systematic treatment in which a psychological process is the central dynamic.
3. the intervention consists of a standardised number of at least two face-to-face contacts between patient and GP.

Studies were excluded if:

1. the intervention is a (psycho)pharmacological intervention.
2. the GP is only one of many interventionists and the results were not presented for every discipline (e.g. GP, nurse, counsellor) separately.
3. psychosocial interventions were only delivered in combination with other types of treatment (e.g. (placebo) pharmacotherapy, physiotherapy) and the effectiveness of the psychosocial intervention alone could not be evaluated.

TYPES OF OUTCOME MEASURES

We did not restrict ourselves to one disorder, say depression, since this would result in a scope too narrow for our purposes. Since there were no restrictions on type of participant, type of disorder, problem or complaint, or type of psychosocial intervention, all outcome measures were reviewed and evaluated on their relevance. To be included in the review, studies had to report at least one of two types of outcome:

1. (clinical) improvement rates or numbers (e.g. % patients who are clinically recovered)
2. scores on relevant validated scales (e.g. BDI, SCL-90)

Search strategy for identification of studies

The search strategy included the following sources:

1. The Cochrane Collaboration Depression Anxiety and Neurosis group Controlled Trials Register (CCDAN-CTR). This is a specialised register that contains more than 13,800 records on trials comparing treatment options that are within the scope of the CCDAN. The register is updated quarterly, adding the results on searches of The Cochrane Library, CINAHL, EMBASE, LILACS, MEDLINE, NRR, PSYCLIT, PSYCINFO, PSYINDEX and SIGLE. Also, quarterly systematic screening of relevant journals and conference proceedings takes place. For information on the full search strategies, visit www.iop.kcl.ac.uk/iop/ccdan.index.htm.
The coordinator of the CCDAN-CTR, Hugh McGuire, performed several searches (CCDAN-CTR issues 2001-2 to 2001-4) (for the full search strategy, retrieve the Cochrane Library version of this review).
2. The Cochrane Library. It became apparent that some studies identified earlier through screening of journals did not appear in the first CCDAN-CTR search. Trials coordinator Hugh McGuire therefore performed additional searches in The Cochrane Library (issues 2001-2 to 2002-1) (for the full search strategy, retrieve the Cochrane Library version of this review).
3. Reference lists of relevant studies were scanned to retrieve additional studies that were not identified in the aforementioned database searches. This process is known as citation tracking.
4. Personal communication with experts in the field (CCDAN and otherwise) took place.

Methods of the review

SELECTION OF STUDIES

MH and AB independently screened the abstracts of the studies that were retrieved by the computer assisted searches. Studies that were identified by citation tracking or personal communication were also screened by both MH and AB. Eligibility criteria were used to select relevant studies. Disagreement between MH and AB about the selection of a trial was resolved by discussion between MH, AB and a third reviewer (GB). Thus, a final selection of studies was established. Since the reviewers were familiar with some of the studies beforehand, it was decided that studies were not blinded for assessment.

METHODOLOGICAL QUALITY

Methodological quality of the studies was independently assessed by two reviewers, AB and GB, using the Maastricht-Amsterdam Criteria List (MACL)¹⁸ and the CCDAN Quality Rating Scale (QRS).¹⁹ A third reviewer (MH) was involved in rating methodological quality if disagreement occurred between AB and GB. Reviewers were thus allowed to adjust their score assignments. If additional information on study characteristics was provided by authors on request, this information was used in the scoring of the items of the MACL and QRS.

The Maastricht-Amsterdam Criteria List (MACL) was originally developed in the field of musculoskeletal disorders, but is known to produce disease-nonspecific quality ratings. Also, the MACL includes all criteria of other prominent quality scales like the Jadad List²⁰ and the Delphi List.²¹ The MACL contains 17 items to assess internal validity (e.g. selection bias, performance bias, attrition bias and detection bias, 10 items), external validity (descriptive criteria, five items) and statistical aspects (two items). The total score on the MACL can range from 0-17. To prevent different interpretations between reviewers of study characteristics, each item of the MACL is explained in a separate appendix that provides uniform operationalisations of criteria. Despite this appendix, two items on the acceptability of compliance and withdrawal still raised questions about the appropriate interpretation among the reviewers. Therefore, we defined acceptable compliance as 75% of the patients in all treatment groups attending all treatment sessions and acceptable withdrawal/dropout as 80% of all patients in all study groups completing all study assessments, in addition to the original appendix.

The CCDAN Quality Rating Scale was used as a field-specific quality rating scale. The QRS consists of 23 items, including items on sample size, allocation, use of diagnostic criteria, compliance, attrition and statistical analysis. Scores on the QRS can range from 0-46. An explanatory appendix of the QRS is not available. Bower adapted the QRS for scoring trials on active non-drug interventions performed in primary care by deleting three items: blinding of subjects, blinding of assessors and details on side effects.¹⁶ It was argued that these items represent aspects less relevant for pragmatic stud-

ies of counselling in primary care. We chose to use the original, unadapted QRS, since scores on this scale will also reflect the methodological limitations that come 'a priori' with the field of research addressed in this review.

DATA EXTRACTION

The study characteristics (i.e. qualitative and quantitative characteristics) of selected trials were extracted with the use of a specially designed data extraction form by two reviewers independently (AB and GB). Disagreement between AB and GB was resolved by discussion between AB, GB and MH. Missing information was obtained from the investigators when possible. The following study characteristics were extracted: study size (total, group), number of active GPs, number of practices, inclusion criteria, type of intervention (content, number of sessions, duration of sessions), type of control condition, training of GPs (content, number of sessions, duration of sessions), supervision of GPs (content, number of sessions, duration of sessions), integrity check (content), study outcome (main effects, compliance with treatment by patients, withdrawal from study assessments) and overall conclusion of authors.

STATISTICAL ANALYSES

Although selected studies were very heterogeneous in types of interventions, types of complaints, study population and outcome measures, meta-analyses were performed if at least two studies reported data on a particular outcome of a clinically comparable intervention and a similar follow-up period. Analyses were performed using Review Manager 4.1. Relative risks were calculated for dichotomous outcomes and weighted mean differences for continuous outcomes. If standard deviations were not available we used the average standard deviation reported by other studies for that outcome.²² The random-effect model was used as default method of analysis because of the potential statistical heterogeneity among studies. Heterogeneity was investigated by evaluating whether confidence intervals around the individual effect sizes of pooled studies overlapped sufficiently and described accordingly in the results.

In addition to meta-analyses, numbers-needed-to-treat (NNT) were calculated for the last available endpoint if outcomes were presented as dichotomous data. We used the denotation as suggested by Altman²³: number of patients needed to be treated for one additional patient to benefit or be harmed are denoted as NNTB (benefit, positive score) or NNTH (harm, negative score), with confidence intervals that include 'infinity' if necessary (for example, NNTB 10, CI=NNTH 20 to infinity to NNTB 4). If the confidence interval is limited to positive scores (e.g. CI=NNTB 5 to NNTB 15) or negative scores and therefore does not include infinity, the number-needed-to-treat is statistically significant.

LEVELS OF EVIDENCE

For a more qualitative approach, so-called 'levels of evidence' were used.^{18, 24, 25} This rating scale enables reviewers to summarize the strength of scientific evidence by classifying results of studies with comparable interventions and more or less comparable outcome measures under one of four levels:

1. *good evidence*: provided by generally consistent findings in two or more high-quality studies.
2. *moderate evidence*: provided by generally consistent findings in one high-quality study and one or more low-quality studies or by generally consistent findings in two or more low-quality studies.
3. *limited or conflicting evidence*: only one study (either high- or low-quality) or inconsistent findings in two or more studies.
4. *no evidence*: no studies.

High quality studies were defined as studies that fulfilled five or more of the 10 MACL items on internal validity (range 0-10). Generally consistent findings were defined as 75% or more the studies having statistically significant findings in the same direction.

Description of studies

SELECTION OF STUDIES

Our computer-assisted searches of databases yielded 179 references in the CCDAN TR and 436 references in the Cochrane Library. Of these, 42 full-text articles were retrieved based on title and abstract and screened for eligibility. In addition, nine articles were identified through citation tracking or personal communication and retrieved for further screening. Of these 51 full-text papers, 19 articles were selected for additional reading. Eight authors were contacted for additional information or unpublished data and seven authors responded to our requests. This process resulted in eight studies that met eligibility criteria and that were included in the review (table 1). The remaining 36 studies (43 articles) were excluded from the review. In one included study²⁶, therapists were two research GPs and one psychiatrist. Since the effects were no different for the two GPs compared to the psychiatrist, we decided to include the study in the review, despite the fact that effects were not presented for GPs separately. In a second included study²⁷, the therapist performing all interventions was a physician trained in family medicine as well as internal and social medicine who worked in a preventive medicine unit in primary care, which raised our doubts whether this therapist could be classified as a typical general practitioner. However, since the study formally met our inclusion criteria, we decided to include the study in the review.

EXCLUDED STUDIES

Of the 36 studies that were initially retrieved but finally excluded from the review, 11 studies were excluded because the intervention was not performed by a GP.²⁸⁻³⁸ Seven studies were excluded because the intervention was no psychosocial intervention to our definition.³⁹⁻⁴⁵ Another seven studies were excluded because the intervention was administered by a group of GPs and other health professionals (e.g. nurse practitioners, general internists, psychiatrists), without separate analyses for each discipline being presented.⁴⁶⁻⁵² In one of these studies, the intervention consisted of two parts administered by a GP and a nurse.⁵⁰ In six studies, the selection criterion of at least two standardised intervention sessions was not met: in four studies, the intervention studies consisted of only one session⁵³⁻⁵⁶ while in two other studies, no required number of standardised intervention sessions was formulated.⁵⁷⁻⁵⁹ In these last two studies, the training of GPs in applying a psychosocial intervention was the focus of attention, rather than the application of the intervention itself. As a consequence, number of intervention sessions were not explicitly monitored, and it was often unknown if GPs even applied their new skills, as one of the authors stated. Three studies were excluded because the study design was no RCT, CCT or CPPT.⁶⁰⁻⁶² In two studies, the effects of 'isolated' psychosocial intervention could not be assessed since the intervention was only administered in combination with another treatment (nicotine gum and (placebo-)antidepressant treatment).^{63, 64}

TABLE 1

CHARACTERISTICS OF INCLUDED STUDIES

Study	Blankenstein 2001
Methods	RCT to compare two conditions. Randomisation performed on practice level. Assessments on three occasions: baseline, 1 and 2 years.
Participants	162 patients (aged 20-45) who were true or part somatisers (frequent GP attenders, 15 visits or more in 3 years and 5 or more somatisation symptoms) recruited from 20 GPs in 17 practices. Therapists: 10 GPs.
Interventions	T (n:75) = tailored and modified reattribution intervention (dealing with illness worry, feeling understood, broadening the agenda, making the link) and care as usual, 2-3 10-30-min sessions. C (n:87) = care as usual. Therapists received 20-hr training programme (application intervention, feedback on audiotaped consultations) and booster sessions in first year. Supervision of therapists unknown.
Outcomes	Two year follow-up: complete attribution achieved in 33 of 51 patients. T superior to C on all primary outcomes: medical consumption, subjective health (VAS), sick leave. No complete recovery in either groups (not supported by data).
Notes	non-compliance: T=24/75. withdrawal: T=7/75, C=6/87. no integrity check.
Study	Lidbeck 1997
Methods	RCT to compare two conditions. Randomisation biased towards more patients in experimental group. Assessments on three occasions: baseline, 2 and 6 months.
Participants	50 patients (aged 30-60) with somatisation disorder (unexplained somatic symptoms and 1 or more symptoms meeting ICHPPC criteria for specific functional disorders) recruited from GPs and other physicians in an out-patient clinic. Therapist: 1 research GP.
Interventions	T (n:33) = group cognitive behavioural therapy (focus on reducing dread, physical examination, patient education, cognitive restructuring, relaxation training), 8 3-hr sessions in 2 months. C (n:17) = waiting list. Therapist received training in stress relaxation. Supervision of therapist unknown.
Outcomes	6 month follow-up: T superior to C on illness behaviour (IBQ), hypochondriasis (Whitley index), and medication use. No sign. differences on social problems (SPQ), anxiety (HAD), depression (HAD), sleep (SDI).
Notes	non-compliance: T=1/32. withdrawal: T=1/32, C=unknown. no integrity check.
Study	McIntosh 1997
Methods	RCT to compare three conditions. Assessments on four occasions: baseline, 3, 6 and 12 months.
Participants	159 patients (aged 15 or older) with high alcohol consumption (one or more CAGE items or 4 or more standard drinks each day in 28 days) recruited by nurse in family practice centre. Therapists: 1 research GP (and 1 nurse practitioner).
Interventions	T1 (n:40) = cognitive behavioural intervention (CBI) by research GP (basic information, help understanding function alcohol, plan of action, moderation strategies, use of daily drinking records and self-help manual), 2 30-min sessions over 2 weeks. T2 (n:66) = CBI by research nurse practitioner, 2 30-min sessions over 2 weeks. C (n:53) = brief advice by own GP (1 of 12) on drinking, 1 5-min session. Training or supervision of therapists unclear.
Outcomes	12 month follow-up: no sign. differences between groups (overall reduction in drinking and alcohol-related problems in all groups).
Notes	non-compliance: T1&T2=27/106. withdrawal: T1&T2&C=16/159. no integrity check.

TABLE 1 - CONTINUED

CHARACTERISTICS OF INCLUDED STUDIES

Study	Mynors-Wallis 1995
Methods	RCT to compare three conditions. Randomisation is stratified (severity of depression). Assessments on three occasions: baseline, 6 and 12 weeks.
Participants	91 patients (aged 18-65) with major depression (research criteria for major depression and Hamilton rating scale score 13 or more) recruited from 26 GPs in 15 practices. Therapists: 2 research GPs (and 1 psychiatrist)
Interventions	T1 (n:30) = problem solving treatment (PST) (explanation rationale, emotional symptoms are caused by problems in living that can be dealt with; identification problems; stages of PST explained). T2 (n:31) = amitriptyline (50 mg-150 mg) and general support. C (n:30) = placebo and general support. All groups: 6 30-60-min sessions in 12 weeks. Therapists received training in PST (theory, role playing, treating five patients under supervision) and drug administration. Supervision of therapists unknown.
Outcomes	12 week follow-up: T1 superior to C on depression (BDI and Hamilton) and social functioning (SAS). No sign. difference on psychological symptoms (PSE). No sign. differences between T1 and T2. Recovered cases (Hamilton <8): T1=18/30, T2=16/31, C=8/30 (difference T1 and C significant). Patient satisfaction ("T is (very) helpful"): T1=28/30, T2=21/31.
Notes	non-compliance: T1=2/30, T2=6/31, C=18/30. withdrawal (82 patients included in analysis): T1=1/29, T2=2/27, C=14/26. no integrity check.
Study	Mynors-Wallis 2000
Methods	RCT to compare four conditions. Randomisation is stratified (severity of depression). Assessments on four occasions: baseline, 6, 12 and 52 weeks.
Participants	151 patients (aged 18-65) with major depression (research criteria for major depression, Hamilton rating scale score 13 or more, minimum illness duration 4 weeks) referred by 24 GPs. Therapists: 3 research GPs (and 2 practice nurses).
Interventions	T1 (n:39) = problem solving treatment (PST) by GP (stages: clarification of problems, choice of goals, generation of solutions, choice of solutions, implementation of solutions, evaluation). T2 (n:41) = PST by practice nurse. T3 (n:36) = fluvoxamine (100 mg) or paroxetine (20 mg) and general support. T4 (n:35) = combination treatment (medication by GP and PST by nurse). All groups: 6 30-60-min sessions in 12 weeks (except T4=12 sessions). Therapists (GPs) received training in PST (theory, treating five patients under supervision), training manual and were supervised throughout trial.
Outcomes	52 week follow-up: no sign. differences between groups on depression (BDI, Hamilton), psychological symptoms (clinical interview schedule) or social functioning (SAS). Recovered cases (Hamilton <8): T1=24/39, T2=23/41, T3=20/36, T4=23/35 (no significant differences).
Notes	non-compliance: T1=14/39, T2=9/41, T3=6/36, T4=6/35. withdrawal: T1=14/39, T2=13/41, T3=6/36, T4=5/35. no integrity check.

TABLE 1 - CONTINUED

CHARACTERISTICS OF INCLUDED STUDIES

Study	Richmond 1985
Methods	CCT to compare two conditions. Group allocation according to day of entry in study. Assessments on two occasions: baseline and 6 months.
Participants	200 patients (aged 16-65) who smoked recruited from 4 GPs in one practice. Therapists: three GPs.
Interventions	T (n:100) = smoking cessation counselling (use of 1 wk diary of smoking habits and self-help manual, information on effects of smoking, strategy for maintaining abstinence, discussing alternatives, withdrawal symptoms, weight management, risks of smoking, benefits of abstinence, problems encountered), 6 sessions (duration?) in 6 months. C (n:100) = usual care and use of 1 wk diary of smoking habits, 2 sessions (duration?) in 6 months. Training or supervision of therapists unknown.
Outcomes	6 month follow-up: abstinence T=33/100, C=3/100
Notes	non-compliance: T=25/100, C=23/100. withdrawal (non-compliers analysed as smokers): T=25/100, C=23/100. no integrity check.
Study	Richmond 1995
Methods	CCT to compare four conditions. Group allocation according to weekly blocks. Assessments on three occasions: baseline, 6 and 12 months.
Participants	378 patients (aged 18-70) with high weekly alcohol consumption (men > 35 drinks, women > 21) recruited from 119 GPs in 40 practices. Therapists: 119 GPs.
Interventions	T1 (n:96) = behavioural change programme 'Alcoholscreen' (use of self-help manual and diary, education, counselling, advice on changing drinking behaviour, supporting new drinking habits), 5 5-25-min sessions in 5 months. T2 (n:96) = brief advice to reduce drinking, 1 5-min session. C1 (n:93) = assessment of drinking behaviour and follow-up only. C2 (n:93) = follow-up only. All GPs received training, supervision unknown.
Outcomes	12 month follow-up (not available for C2): No sign. differences between groups in % patients reporting drinking above predefined weekly level (men < 28 drinks, women < 14) or alcohol-related problems (MAST). T1 superior to other treatments on alcohol consumption below level for patients who attended 2 sessions or more.
Notes	non-compliance: T1=49/96 2nd visit, 92/96 5th visit. withdrawal: T1=30/96, T2=26/96, C1=32/93 (C2 unavailable). no integrity check.
Study	Segnan 1991
Methods	RCT to compare four conditions. Assessments on three occasions: baseline, 6 and 12 months.
Participants	923 patients (aged 20-60) who were smokers free of a life-threatening disease recruited by 44 GPs. Therapists: 44 GPs.
Interventions	T1 (n:275) = repeated counselling (RC) (use of brochure, reinforcement antismoking message), 5 sessions (duration?) in 9 months. T2 (n:294) = RC plus prescription of nicotine gum. T3 (n:292) = RC plus spirometric test, results discussed by GP. C (n:62) = minimal intervention (counselling and use of brochure), 1 session. Therapists received 2 3-hr training sessions (introduction counselling techniques), 40% of GPs attended both sessions. Supervision of therapists unknown
Outcomes	12 month follow-up: No sign. differences between groups on % biochemically verified quitters, overall low quit rates (4.8-7.5%).
Notes	non-compliance: approx. 66% attended less than 4 sessions. withdrawal: total N=129/923. integrity check: GP compliance with treatment protocol: RC=67.3%; RC+gum=92.5%; RC+spiro=83.7%; MI=75%.

STUDY CHARACTERISTICS

The most important characteristics and outcomes of included studies were summarized in table 1. In the section 'notes' of this table, additional information is presented on non-compliance of patients with treatment (e.g. session attendance), withdrawal by patients from study assessments (i.e. drop-out, loss-to-follow-up) and the use of integrity checks (assessment of GP compliance with the treatment protocol).

The eight included studies reported the effectiveness of psychosocial interventions by general practitioners for the following health complaints or disorders: depression ^{26, 65}, somatisation ^{27, 66}, smoking addiction ^{67, 68} and excessive alcohol consumption.^{69, 70}

METHODOLOGICAL QUALITY OF INCLUDED STUDIES

For the qualification of methodological quality (high or low), the 10 MACL items on internal validity were used. Scores on each item are shown in table 2. Each item was scored as 'positive' (+), 'negative' (-) or 'unclear' (?). A positive score on item 2 to 10 was accredited one point, while a negative or unclear point was accredited no points. Item 1 consisted of 1a and 1b: positive scores on these items were accredited 0.5 points each. In total six studies ^{26, 27, 64-66, 68, 69} had a high methodological quality (five points or more) and two studies ^{67, 70} had a low methodological quality (less than five points).

Item 1a addresses the use of a randomisation method and item 1b addresses the concealment of treatment allocation. Six studies scored positive on item 1a while only five studies scored positive on item 1b. Items 2, 5 and 6 concern the blinding of patients (item 2), care providers (item 5) and outcome assessors (item 6). Not surprisingly, none of the studies scored positive on items 2 and 5. However, three studies did score positive on item 6. Items 7 and 9 concern the relevance of outcome measures (item 7) and the comparability of the timing of outcome assessments (item 9). All studies scored positive on these items. Item 3 addresses the avoidance or comparability of co-interventions. Five studies did not provide any information on this subject and scored unclear on this item. Acceptable compliance of patients was a problem in many studies: six studies scored negative on item 4 that concerns this matter while two studies scored positive. Item 8 concerns the acceptability of withdrawal/drop-out by patients. Four studies scored positive and four studies scored negative on this item. Finally, seven studies reported an intention-to-treat analysis. These studies scored positive on item 10.

In table 2, the total scores on the MACL and the QRS are presented for each study. For means of comparison, we defined four categories (1 to 4) based on the total range of each scale to indicate lowest quality (1) and highest quality (4) for each scale. Category values for the MACL were 0-4 (1), 5-8 (2), 9-12 (3) and 13-17 (4). Category values for the QRS were 0-11 (1), 12-23 (2), 24-35 (3), 36-46 (4). All studies had MACL and QRS total scores in similar categories which indicates that the MACL and QRS produced fairly consistent findings.

TABLE 2

QUALITY ASSESSMENT: INTERNAL VALIDITY ITEMS (MACL) AND SCORES ON MACL AND QRS

Items and scales	Mynors-Wal. 95	Mynors-Wal. 00	Blankenstein 01	Lidbeck 97	Richmond 85	Segnan 91	McIntosh 97	Richmond 95
1a randomisation	+	+	+	+	-	+	+	-
1b concealment of allocation	+	+	+	+	-	+	?	-
2 blinded care provider	-	-	-	-	-	-	-	-
3 co-interv. avoided	+	+	?	?	?	-	?	?
4 acceptable compliance	-	-	-	+	-	-	+	-
5 blinded patient	-	-	-	-	-	-	-	-
6 blinded outcome assessor	+	+	-	-	-	?	?	+
7 relevance measures	+	+	+	+	+	+	+	+
8 acceptable withdrawal	-	-	+	+	-	+	+	-
9 timing assessments	+	+	+	+	+	+	+	+
10 intention-to-treat	-	+	+	+	+	+	+	+
MACL total score	10.5	13	10.5	10.5	6.5	11	9.5	9
QRS total score	34	37	32	26	19	30	26	29

Results

DEPRESSION STUDIES

In two high-quality RCTs by the same research group ^{26, 65}, the effects of six-session problem solving treatment (PST) by a GP on major depression were compared to PST by a practice nurse, antidepressant treatment, placebo treatment or combination treatment (PST and antidepressant medication).

In the 1995 study ²⁶, PST administered by one of two GPs (and a psychiatrist) was superior to placebo treatment combined with general support on depression and social functioning, but not on psychological symptoms at 12 week follow-up. The NNTB of PST compared to placebo treatment was 3 (95% CI NNTB 1.75 to NNTB 10.29). There were no differences in effect between PST and antidepressant treatment (amitriptyline) combined with general support at 12 week follow-up. The NNTB of PST compared to antidepressant treatment was 11.92 (95% CI NNTH 6.08 to infinity to NNTB 3.01).

In the 2000 study ⁶⁵, the effects of PST administered by one of three GPs on depression, psychological symptoms or social functioning were no different compared to the effects of PST by a practice nurse, antidepressant treatment or combination treatment at 52 week follow-up. The NNTB of PST by a GP was 18.38 (95% CI NNTH 6.21 to infinity to NNTB 3.71) compared to PST by a practice nurse and 16.71 (95% CI NNTH 6.13 to infinity to NNTB 3.53) compared to antidepressant medication. The NNTH of PST by a GP compared to combination treatment was 23.95 (95% CI NNTH 3.83 to infinity to NNTB 5.63).

Both studies reported equivalent outcome measures for PST by a GP and antidepressant treatment at 12 weeks and several data were meta-analysed. There were no differences between PST by a GP and antidepressant treatment in recovery (Hamilton rating scale score of 7 or lower), depression (Hamilton rating scale and beck depression inventory) or social functioning (social adjustment scale). The pooled relative risk for recovery of PST by a GP compared to antidepressant treatment was 0.93 (95% CI 0.62 to 1.39). The weighted mean difference between PST by a GP and antidepressant treatment was 0.77 (95% CI -2.45 to 4) on the Hamilton rating scale, -1.12 (95% CI -4.75 to 2.51) on the beck depression inventory. Evaluation of the confidence intervals around the weighted mean differences on the social adjustment scale revealed statistical heterogeneity between the two studies. It was therefore decided that scores on this scale should not be pooled.

There is good evidence (level 1) that problem-solving treatment by a GP is no less effective than antidepressant treatment on depression, psychological symptoms and social functioning. There is limited evidence (level 3) that problem-solving treatment by a GP is more effective than placebo treatment on depression and social functioning. Finally, there is limited evidence (level 3) that problem-solving by a GP is no less effective than problem-solving by a practice nurse or combination treatment on depression, psychological symptoms or social functioning.

SOMATISATION STUDIES

In two high-quality RCTs, the effectiveness of a psychosocial intervention by a GP on somatisation was assessed.

In one study ⁶⁶, the effects of a tailored and modified reattribution intervention of two to three sessions by one of 10 GPs were compared to the effects of care as usual. At two year follow-up, reattribution intervention was superior to usual care on all primary outcomes (consumption of medical resources, subjective health and sick leave) and on secondary outcome somatisation, although none of the patients had recovered completely (no data presented).

In a second study ²⁷, the effects of cognitive behavioural group therapy of eight sessions administered by the author were compared to a waiting list condition. At 6-month follow-up, group therapy was superior to the waiting list condition on illness behaviour, hypochondriasis and medication use but not on social problems, anxiety, depression or sleep.

No data concerning recovery were presented. Calculations of NNT were therefore not possible. Since interventions, outcome measures and timing of assessment were incomparable for the two studies, meta-analyses were not performed.

There is limited evidence (level 3) that reattribution intervention by a GP is more effective than usual care on consumption of medical resources, subjective health, sick leave and somatisation. There is also limited evidence (level 3) that cognitive behavioural group therapy by a GP is more effective on illness behaviour, hypochondriasis and medication use.

SMOKING CESSATION STUDIES

In one high-quality RCT and one low-quality CCT the effects of counselling by a GP on smoking cessation were assessed.

In the high-quality study ⁶⁸, the effects of five-session 'repeated counselling' (RC) delivered by one of 44 GPs were no different compared to the effects of a one-session minimal intervention (MI), repeated counselling plus nicotine gum (RC+gum) or repeated counselling plus spirometry (RC+spiro): biochemically validated smoking abstinence rates at 12 month follow-up were respectively 4.8%, 5.5%, 7.5%, and 6.5%. The NNTB of RC compared to MI was 162.38 (95% CI NNTH 18.64 to infinity to NNTB 15.16). The NNTH of RC was 49.29 (95% CI NNTH 16.51 to infinity to NNTB 49.93) compared to RC+gum and 95.02 (95% CI NNTH 20.19 to infinity to NNTB 35.11) compared to RC+spiro.

In the low-quality study ⁶⁷, six-session smoking cessation counselling delivered by one of three GPs was superior to a minimal intervention (usual care and use of a diary) consisting of two sessions: at 6 month follow-up, 33% of the patients in the counselling group were biochemically validated as abstinent from smoking versus 3% in the minimal intervention group. The NNTB of counselling compared to minimal intervention was 3.33 (95% CI NNTB 2.51 to NNTB 4.95).

Evaluation of the confidence intervals around the effect sizes revealed statistical heterogeneity

between the two studies, possibly due to differences in methodological quality. It was therefore decided that abstinence rates from both studies should not be pooled.

There is conflicting evidence (level 3) that counselling by a GP is more or no less effective than minimal intervention on smoking behaviour. There is limited evidence (level 3) that counselling by a GP is no less effective than counselling plus nicotine gum or counselling plus spirometry by a GP on smoking behaviour.

ALCOHOL REDUCTION STUDIES

In one high-quality RCT and one low-quality CCT the effectiveness of a psychosocial intervention by a GP on alcohol consumption was assessed.

In the high-quality study ⁶⁹, the effects of a two-session cognitive behavioural intervention (CBI) administered by one research GP were compared to a CBI by a nurse practitioner and one-session brief advice by one of 12 regular GPs. At 12 month follow-up, there were no differences between the groups in alcohol consumption (quantity-frequency) or alcohol-related problems, although there was an overall reduction on these outcomes in all groups.

In the low-quality study ⁷⁰, the effects of a five-session behavioural change programme (Alcoholscreen) by one of 119 GPs were compared to one-session brief advice to stop drinking, assessment of drinking behaviour only and follow-up measurement only. At 12 month follow-up, there were no differences between the groups in alcohol consumption (% patients drinking above predefined consumption level) or alcohol-related problems, although Alcoholscreen was superior to other treatment conditions if only those patients were analysed who had attended two sessions or more. No data concerning recovery were presented. Calculations of NNT were therefore not possible. Since interventions and outcome measures were (partially) incomparable for the two studies, meta-analyses were not performed.

There is limited evidence (level 3) that a cognitive behavioural intervention by a GP is no more effective than a cognitive behavioural intervention by a nurse practitioner or brief advice on alcohol consumption or alcohol-related problems. Furthermore, there is limited evidence (level 3) that a behavioural change programme is no more effective than brief advice, assessment of drinking behaviour only or follow-up measurement only on alcohol consumption or alcohol-related problems.

EFFECTIVENESS IN RELATION TO STUDY CHARACTERISTICS

In table 3, we present relevant characteristics of studies in order to investigate possible associations with outcomes of studies.

The first two items (O1 and O2) concern the outcomes of studies. Item O1 refers to the comparison between a psychosocial intervention (interv) and a less intensive condition (lesser treatm, e.g.

TABLE 3

OUTCOMES (O) AND CHARACTERISTICS (C) OF STUDIES: QUALITATIVE DESCRIPTION

	Mynors-Wal. 95	Mynors-Wal. 00	Blankenstein 01	Lidbeck 97	Richmond 85	Segnan 91	McIntosh 97	Richmond 95
O1 = interv vs lesser treatm	+	n.a.	+	+	+	-	-	-
O2 = interv vs comp treatm	+	+	n.a.	n.a.	n.a.	+	+	n.a.
C1 = method. quality (high=+, low=-)	+	+	+	+	-	+	+	-
C2 = training GPs (yes=+, no=-)	+	+	+	+	-	+	-	+
C3 = supervision GPs (yes=+, no=-)	-	+	-	-	-	-	-	-
C4 = use treatment protocol (yes=+, no=-)	-	-	-	-	-	-	-	-
C5 = reg GP (+) vs res GP (-)	-	-	+	-	+	+	-	+
C6 = number GPs (= <10=+, >10+)	+	+	+	+	+	-	+	-
C7 = integrity check (yes=+, no=-)	-	-	-	-	-	+	-	-
C8 = number sessions (= >5=+, <5=-)	+	+	-	+	+	+	-	+
C9 = duration sessions (= >30m=+, <30m=-)	+	+	-	+	-	-	+	-
C10 = compliance (= >75%=+, <75%=-)	+	-	-	+	+	-	+	-
C11 = withdrawal (= <20%=+, >20%=-)	-	-	+	+	-	+	+	-

usual care, minimal intervention). If the psychosocial intervention was superior to the less intensive condition, this item was scored positive (+) and negative (-) if the psychosocial intervention was not superior. If such a comparison was not made, 'not applicable' (n.a.) was denoted. Item O2 refers to the comparison between a psychosocial intervention (interv) and other interventions that were comparable in intensity or more intensive (comp treatm, e.g. antidepressant treatment, counselling plus gum). If the psychosocial intervention was no less effective or superior to the comparable intervention(s), this item was scored positive (+) and negative (-) if the psychosocial intervention was less effective. If such a comparison was not made, 'not applicable' (n.a.) was denoted.

The remaining items concern the characteristics of studies. Item C1 refers to the methodological quality of studies and was scored positive (+) if the quality was high and negative (-) if the quality was low. Item C2 was scored positive (+) if the article explicitly mentioned the training of GPs and negative (-) if there was no mention of any GP training. Item C3 was scored positive (+) if the article explicitly mentioned the supervision of GPs during the trial and negative (-) if there was no mention of any GP supervision. Item C4 was scored positive (+) if the article explicitly mentioned the use of a treatment protocol and negative (-) if there was no mention of the use of a treatment protocol. Item C5 was scored positive (+) if patients were treated by their own, regular GP (reg GP) and negative (-) if patients were treated by an unknown research GP (res GP) assigned to them. Item C6 was scored positive (+) if 10 GPs or less were administering the intervention and negative (-) if more than 10 GPs administered the intervention. Item C7 was scored positive (+) if the article explicitly mentioned the use of an integrity check of the performance of GPs and negative (-) if there was no mention of the use of an integrity check. Item C8 was scored positive (+) if the intervention consisted of five or more sessions and negative (-) if the intervention consisted of less than five sessions. Item C9 was scored positive (+) if each intervention session lasted 30 minutes or more and negative (-) if each intervention session lasted less than 30 minutes. Item C10 was scored positive (+) if the compliance of patients in all treatment groups was 75% or more and negative (-) if the compliance of patients in all treatment groups was less than 75%. Finally, item C11 was scored positive (+) if the withdrawal rate of patients in all study groups was 20% or less and negative (-) if the withdrawal rate of patients in all study groups was more than 20%. If information on any of these characteristics was unclear, the item was scored negative (-). Type of psychosocial intervention was not scored as a relevant characteristic since interventions in the selected studies were largely incomparable.

In general, these characteristics of studies were not associated with the outcomes of studies. A modest association with outcome was observed for the number of GPs administering the intervention. In the four studies in which the psychosocial intervention was superior to the less intensive condition^{26, 27, 66, 67}, the number of GPs was small (10 or less). In the remaining three studies in which the psychosocial intervention was not superior to the less intensive condition, the number of GPs was high (more than 10) in two studies^{68, 70} and small in only one study.⁶⁹

Discussion

EFFECTIVENESS OF PSYCHOSOCIAL INTERVENTIONS BY GENERAL PRACTITIONERS

In this review, the main objective was to provide a systematic overview of the available evidence on the effects of psychosocial interventions that are delivered by general practitioners in primary care. Not surprisingly, we found that it is impossible to draw an overall conclusion concerning the effectiveness of "psychosocial interventions by general practitioners" since the eight selected studies were not comparable in numerous aspects (e.g. intervention, outcome, patient population). Under these circumstances, the classification of the available results into a single 'level of evidence' is inappropriate. Rather, we found that the available evidence addressed different psychosocial interventions for four distinct disorders or health complaints (depression, somatisation, smoking addiction and excessive alcohol consumption). There is some evidence that problem-solving treatment (PST) by a GP is effective in the treatment of major depression: there is good evidence (level 1) provided by two studies that PST is no less effective than antidepressant treatment, while there is limited evidence (level 3) that PST is more effective than placebo treatment and no less effective than PST by a nurse practitioner or combination therapy. These findings should be interpreted with considerable caution: the two studies on PST were conducted by the same research team and groups consisting of only 30 to 40 patients were treated by a small number of experienced research GPs.

As for the treatment of somatisation, there is limited evidence (level 3) that a reattribution intervention by a GP and cognitive behavioural group therapy by a GP are more effective than usual care by a GP. There is conflicting evidence (level 3) provided by two studies that counselling by a GP is more or no less effective than a minimal intervention in helping patients to stop smoking, while there is limited evidence (level 3) that counselling by a GP is no less effective than counselling plus the prescription of nicotine gum or the use of spirometry. It is noted that the findings from another review indicate a small benefit obtained through brief advice from physicians, with an additional small benefit from more intensive advice.⁷¹ Finally, behavioural interventions by a GP to reduce alcohol consumption seem no more effective than other, more simple interventions: there is limited evidence (level 3) that a cognitive behavioural intervention (CBI) by a GP is no more effective than a CBI by a nurse practitioner or brief advice and there is limited evidence that a behavioural change programme is no more effective than brief advice, assessment of drinking or follow-up assessments only.

In summary, despite fairly good methodological quality and positive findings of some studies, evidence for the effectiveness of psychosocial interventions by general practitioners does not exceed level 3 (limited or conflicting evidence), except for good evidence that problem-solving treatment is no less effective than antidepressant treatment for depression.

EFFECTIVENESS IN RELATION TO STUDY CHARACTERISTICS

The effectiveness of any intervention is influenced by the triad of the intervention-receiver (e.g. the patient), the intervention-giver (e.g. the GP) and the intervention itself. These three factors are inevitably linked, and any disturbance in this relation at any point can result in negative effects. We determined several characteristics that might be associated with the outcome of studies: methodological quality, training of GPs, supervision of GPs, use of a treatment protocol, treatment by the regular GP versus treatment by an unfamiliar research GP, number of active GPs, performance of integrity checks, number of intervention sessions, duration of intervention sessions, compliance of patients and withdrawal of patients. We observed that none of these characteristics seem associated with either a positive or a negative outcome of studies. Only one observation stands out: in all of the four studies that report a positive effect of the psychosocial intervention compared to a less intensive condition (usual care, minimal intervention), interventions are administered by a small number of GPs (less than ten), as opposed to the three 'negative' studies on psychosocial interventions versus a less intensive condition of which only one study reports a small number GPs as interventionists.

This finding exposes an interesting problem in primary care and general practice research: in order to have a large sample size, investigators are often forced to recruit a large number of GPs who can select patients for the trial and perform the intervention. However, there are many objections to having a large number of GPs actively participating in the trial: patient recruitment by GPs is usually slow, accurate training and supervision of a large number of GPs is time-consuming and costly and GPs do not build experience if they only apply the intervention to a small number of patients.⁷² A solution to this problem is the deployment of a small number of research GPs who have (unfamiliar) patients assigned to them for the purpose of the trial.⁷³ These research GPs tend to be highly motivated, highly trained and highly experienced. Consequently, findings from studies in which a small sample of research GPs performs the intervention should be interpreted with caution, since research GPs might not represent the typical GP and treatment effects might therefore be an overestimation of the effects in daily practice. In this review, of the six studies that have a small number of GPs delivering the intervention, four studies use research GPs instead of the regular GP of patients. Although treatment by the regular GP versus treatment by an unfamiliar research GP was not associated with outcome, these circumstances make the available evidence even harder to interpret.

Of course, this highly qualitative analysis of comparing study characteristics in relation to outcome lacks a clear validity basis due to the small number of selected studies. Nevertheless, the potential influence of factors like these, especially the influence GP-factors (number, type, training, experience), should be explored in future research.

METHODOLOGICAL QUALITY

In line with previous studies^{18, 24, 25}, we used internal validity criteria of the MACL to classify methodological quality since internal validity best reflects the most essential methodological aspects that are required for good quality (and lack of bias). The MACL is accompanied by an explanatory appendix that proved to be helpful but not always sufficient. Especially the items on the acceptability of compliance and withdrawal posed reviewers for a problem since 'acceptability' was not defined in any quantum or measure. We therefore chose to define acceptable compliance as 75% of the patients in all treatment groups attending all treatment sessions and acceptable withdrawal/drop-out as 80% of all patients in all study groups completing all study assessments. Uniform interpretations of rating scale criteria were even more problematic when items on the QRS were scored. This rating scale has no explanatory appendix (a glossary is under construction) and in many instances this led to inconsistency among reviewers. Despite these problems, scores on the field-unspecific MACL and the field-specific QRS produced consistent findings for each study. This consistency stems partly from the fact that we did not modify the MACL or QRS items to make the scales more applicable for this particular selection of studies. The potential methodological quality of studies may be highly determined by the limitations of the field of research (e.g. the issue of blinding), and we feel that methodological quality should be an objective and universal classification, assessed independent of the circumstances in which a particular study is performed.

Six out of eight selected studies had a satisfactory methodological quality. Although controlled patient preference trials were eligible as well, only randomised controlled trials (six high-quality studies) and controlled clinical trials (two low-quality studies) met eligibility criteria for the review. As was to be expected, none of the selected studies reported blinding of patients or GPs. Information on the avoidance or comparability of co-interventions was poor in all studies: five studies did not even mention this subject^{27, 66, 67, 69, 70} and the remaining three^{26, 65, 68} were very scarce with information. Acceptable compliance of patients with treatment, which is an indication of the acceptability of the treatment to patients, was achieved in only two studies^{27, 69} while withdrawal rates were acceptable in only half of the studies^{27, 66, 68, 69}. The latter is of particular concern since unacceptable withdrawal/dropout rates, especially due to selective withdrawal, represent a major threat to the validity of results.

In general, it should be noted that many studies reported poorly on the supervision and training of GPs during the trial, the use of treatment protocols, the performance of integrity checks, the compliance of patients apart from the attendance of sessions, the nature of usual care (what care exactly did patients receive?), co-interventions received, subgroup analyses or prognostic analyses (which patients benefit most?). Only one study reported that GPs were supervised throughout the trial.⁶⁵ In six studies it was mentioned that GPs were trained^{26, 27, 65, 66, 68, 70}, but only three studies elaborated to some extent on the specific content of the training.^{26, 65, 66} None of the studies reported the use of a treatment protocol. Only one study reported the performance of integrity checks.⁶⁸ Details

on other aspects (compliance, nature of usual care, co-interventions, subgroup/prognostic analyses) were generally insufficient, if mentioned at all. This lack of vital information makes it difficult to interpret the results of studies. A similar conclusion was drawn in the reviews by Rowland and Bower.^{9, 10, 16, 17} Especially in this field of research, in which the blinding of patients and caregivers is virtually impossible, a thorough description of all factors that might introduce bias is of paramount importance.

METHODOLOGICAL ASPECTS OF THIS REVIEW

We stated before that it is difficult to define 'psychosocial interventions'. We deliberately chose a rather conservative definition, and this has affected the inclusion of studies. For example, several studies we identified initially were not included because interventions did not consist of at least two standardised treatment sessions. In other excluded studies, psychosocial interventions were delivered by general practitioners and other health professionals (in the United States, it is customary that both 'family physicians' and general internists work in primary care). We are confident though that these restrictions enabled us to make clear assessments of the available evidence: a broader inclusion of studies would have resulted in even more incomparability among studies, adding no information or biased information to the evidence gathered in this review.

Another important issue has been addressed by Van Tulder²⁴ when he argued that "it is questionable whether the criteria list [...] evaluates the quality of the study or the quality of the publication". We used any additional information that was provided by authors on request to score quality criteria. Since we mostly addressed authors in case additional information on the intervention was needed to determine eligibility, the use of unpublished information might have been in favour of those studies in which descriptions of the intervention were less clear. This illustrates the need for clear guidelines as proposed in the CONSORT statement. Or, as members of the CONSORT group have stated it: "Despite several decades of educational efforts, the reporting of RCTs needs improvement".²⁴

IN CONCLUSION

The results of this review do not allow an overall conclusion concerning the effectiveness of "psychosocial interventions by general practitioners". The available evidence addressed different psychosocial interventions for four distinct disorders or health complaints. The results of problem-solving treatment by general practitioners are promising for major depression, although the effectiveness in routine practice settings remains to be seen. The evidence for the remaining interventions (retribution or cognitive behavioural group therapy for somatisation, counselling for smoking cessation, behavioural interventions to reduce alcohol reduction) is either limited or conflicting. This finding can partly be explained by the incomparability of interventions and the small number of included stud-

ies. Therefore, the evidence has to be interpreted with caution: there is no strong evidence for the effectiveness of these interventions nor for the lack of effectiveness. Except the number of GPs delivering the intervention, none of the characteristics of studies seemed to be associated with study outcome. The methodological quality of studies was fairly satisfactory, although most studies reported poorly on relevant study characteristics that are likely to influence the validity of results.

Reviewers' conclusions

IMPLICATIONS FOR PRACTICE

The research question in this review was: "are psychosocial interventions delivered by general practitioners effective?". The underlying question for daily practice was: "is it wise to have general practitioners deliver psychosocial interventions?". The answer to both questions is that it remains to be seen. There is a void of evidence-based knowledge: due to a small number of available studies and incomparability among psychosocial interventions, the evidence for the use of such interventions is mostly limited and does not allow general conclusions concerning the effectiveness, or ineffectiveness, of psychosocial interventions by general practitioners. Problem-solving treatment by highly experienced GPs seems a promising tool in the treatment of depressed patients, although the effectiveness of this intervention by regular GPs in routine care remains to be demonstrated.

Most psychological or psychosocial problems in primary care are vague and diffuse by nature, as opposed to the well-defined and often diagnosed problems that are treated in specialist care. The moderate or limited evidence for psychosocial interventions presented here is restricted to a handful of well-defined health complaints. Evidence on the effectiveness of psychosocial tools that are more suitable for the broad range of psychosocial complaints typically observed in primary care was not identified in this review. This too should be considered a void of knowledge.

In our view, the lack of evidence does not imply that GPs should be discouraged from delivering psychosocial interventions. One way to improve 'medicine-based-evidence'⁷⁵ is to promote the use of psychosocial interventions by general practitioners and collect the experiences from daily practice. However, the time consumption and related costs of these interventions might be a problem in general practice, especially if interventions turn about to be ineffective. Also, little is known about the potential harm GP-delivered psychosocial interventions might cause. From a common sense standpoint, the answer to whether or not GPs should experiment with psychosocial interventions depends largely on the time available in the daily practice of GPs, the personal interest and competence of GPs and the prevalence of psychosocial complaints in primary care.

IMPLICATIONS FOR RESEARCH

Based on the findings in this review, several recommendations for future research can be made. A possible explanation for the limited number of available studies in this review could be that psychosocial interventions are not considered to be helpful tools in general practice. A first step in future research should be to assess the need for GPs to have the use of psychosocial interventions at their disposal in the first place. If so, more empirical studies are needed. Secondly, results from a recent systematic review suggested that nurse practitioners can provide care equivalent to the care of GPs.²² Since nurse practitioners emerged as possible counterparts of GPs in this review, future studies could

investigate who is most suited to be the provider of psychosocial interventions in primary care. Although a comparison between studies in this review is unjustified, it seems that psychosocial interventions are most promising in the treatment of psychological problems like depression and somatisation. Future research that will contribute to a higher level of evidence for the effectiveness of problem-solving treatment for depression and cognitive behavioural interventions for somatisation is helpful in that respect.

It is the reviewers' opinion that future research should also concentrate on more general psychosocial interventions or "tools" that are applicable for a wide range of problems observed in primary care (e.g. depressive symptoms, generalised anxiety, functional or unexplained symptoms). A prerequisite for the building of evidence for any intervention is that articles reporting the effectiveness of studies contain high-quality descriptions of all aspects that are relevant for the interpretation of results. As has been stated earlier, this prerequisite applies to this particular field of research especially. Future RCTs should especially assess and describe the supervision, training and monitoring of GPs, all aspects concerning patient compliance and the actual care received in all conditions (content of usual care, co-interventions). Furthermore, future RCTs should ideally contain study samples large enough to include several subgroups of primary care patients.

Acknowledgments

The reviewers would like to thank the following persons:

The CCDAN Editorial Team, Nancy Rowland (UK), Peter Bower (UK) and Raymond Ostelo (Netherlands) for their support, information and advice on how to conduct a review properly.

Nettie Blankenstein (Netherlands), Michael King (UK), Marta Buszewicz (UK), Oliver Davidson (UK), Ingrid Arnold (Netherlands), Ulrik Malt (Norway) and Jose Catalan (UK) for responding to our requests for (additional) information and study data.

This study was funded by the Health Research and Development Council (ZorgOnderzoek Nederland), The Netherlands.

References

1. Wessely S, Nimnuan C, Sharpe M. Functional somatic syndromes: one or many? *Lancet* 1999; 354:936-9.
2. Chocron Bentata L, Vilalta Franch J, Legazpi Rodriguez I, Auquer K, Franch L. Prevalence of psychopathology at a primary care center. *Atencion Primaria* 1995; 16:586-90, 592-3.
3. Tiemens BG, Ormel J, Simon GE. Occurrence, recognition and outcome of psychological disorders in primary care. *Am J Psychiatry* 1996; 153:636-44.
4. Gray DP. Counsellors in general practice. *J Royal Coll Gen Practitioners* 1988; 38:50-1.
5. Terluin B, Meijman FJ. What do Dutch GPs do when patients have psychosocial problems? [Wat doen Nederlandse huisartsen bij psychosociale problemen van hun patiënten?]. *Huisarts en Wetenschap* 1999; 42:151-2.
6. Corney R. The effectiveness of counselling in general practice. *Int Rev Psychiatry* 1992; 4:331-8.
7. Brown C, Schulberg HC. The efficacy of psychosocial treatments in primary care: a review of randomised clinical trials. *Gen Hosp Psychiatry* 1995; 17:414-24.
8. Friedli K, King M. Counselling in general practice, a review. *Prim Care Psychiatry* 1996; 2:205-16.
9. Rowland N, Godfrey C, Bower P, Mellor Clark J, Heywood P, Hardy R. Counselling for depression in primary care: a systematic review of the research evidence. *Br J Guid Couns* 2000; 28:215-31.
10. Bower P, Rowland N, Hardy R. The clinical effectiveness of counselling in primary care: a systematic review and meta-analysis. *Psychol Med* 2003; 33:203-15.
11. Cape J, Barker C, Buszewicz M, Pistrang N. General practitioner psychological management of common emotional problems (I): definitions and literature review. *Br J Gen Pract* 2000; 50:313-8.
12. Cape J, Barker C, Buszewicz M, Pistrang N. General practitioner psychological management of common emotional problems (II): a research agenda for the development of evidence-based practice. *Br J Gen Pract* 2000; 50:396-400.
13. Davison GC. Stepped care: doing more with less? *J Consult Clin Psychol* 2000; 68:580-5.
14. Von Korff M, Tiemens B. Individualized stepped care of chronic illness. *West J Med* 2000; 172:133-7.
15. Richardson JD. Sexual difficulties. A general practice specialty. *Aust Fam Physician* 1989; 18:200-4.
16. Bower P, Rowland N, Mellor Clark J, Heywood P, Godfrey C, Hardy R. Effectiveness and cost effectiveness of counselling in primary care. *The Cochrane Library*. Oxford: Update Software, 2002.
17. Rowland N, Bower P, Mellor Clark J, Heywood P, Godfrey C. Counselling for depression in primary care. *The Cochrane Library*. Oxford: Update Software, 2001.
18. Tulder MW van, Assendelft WJ, Koes BW, Bouter LM. Method guidelines for systematic reviews in the cochrane collaboration back review group for spinal disorders. *Spine* 1997; 22:2323-30.
19. Moncreiff J. Quality rating scale for RCT. *Int J Methods Psychia Res* 2001; 10:3.
20. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Cont Clin Trials* 1996; 17:1-12.
21. Verhagen AP, de Vet HC, de Bie RA, et al. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *J Clin Epidemiol* 1998; 51:1235-41.

22. Horrocks S, Anderson E, Salisbury C. Systematic review of whether nurse practitioners working in primary care can provide equivalent care to doctors. *BMJ* 2002; 324:819-23.
23. Altman D. Confidence intervals for the number needed to treat. *BMJ* 1998; 317:1309-12.
24. Tulder MW van, Ostelo RWJG, Vlaeyen JWS, Linton SJ, Morley SJ, Assendelft WJv. Behavioral treatment for chronic low back pain. *Spine* 2001; 26:270-81.
25. Ostelo RWJG, Vet HCWd, Waddell G, Kerckhoffs MR, Leffers P, Tulder MW van. Rehabilitation following lumbar disc surgery. *The Cochrane Library*. Oxford: Update Software, 2002.
26. Mynors Wallis LM, Gath DH, Lloyd Thomas AR, Tomlinson D. Randomised controlled trial comparing problem solving treatment with amitriptyline and placebo for major depression in primary care. *BMJ* 1995; 310:441-5.
27. Lidbeck J. Group therapy somatization disorders in general practice: effectiveness of a short cognitive-behavioural treatment model. *Acta Psychia Scan* 1997; 96:14-24.
28. Barkham P. Brief prescriptive therapy in two-plus-one sessions: initial cases from the clinic. *Behav Psychotherapy* 1989; 17:161-75.
29. Issakidis C, Sanderson K, Teesson M, Johnston S, Buhrich N. Intensive case management in Australia: a randomized controlled trial. *Acta Psychia Scan* 1999; 99:360-7.
30. McLeod C, Budd M, McClelland. Treatment of somatization in primary care. *Gen Hosp Psychiatry* 1997; 19:251-8.
31. Ockene J, Kristeller J, Goldberg R, et al. Increasing the efficacy of physician-delivered smoking interventions: a randomized clinical trial. *J Gen Int Med* 1991; 6:1-8.
32. Ockene I, Hebert J, Ockene J, Saperia G, Stanek E, Nicolosi R. Effect of physician-delivered nutrition counseling training and an office-support program on saturated fat intake, weight, serum lipid measurements in a hyperlipidemic population: Worcester Area Trial for Counseling in Hyperlipidemia (WATCH). *Arch Int Med* 1999; 159:725-31.
33. Peveler R, George C, Kinmonth A, Campbell M, Thompson C. Effect of antidepressant drug counselling and information leaflets on adherence to drug treatment in primary care: randomised controlled trial. *BMJ* 1999; 319:612-5.
34. Scott C, Tacchi MJ, Jones R, Scott J. Acute and one-year outcome of a randomised controlled trial of brief cognitive therapy for major depressive disorder in primary care. *Br J Psychiatry* 1997; 171:131-4.
35. Skinner P. Skills not pills: learning to cope with anxiety symptoms. *J Royal Coll Gen Practitioners* 1984; 34:258-60.
36. Swinson R, Soulios C, Cox B, Kuch K. Brief treatment of emergency room patients with panic attacks. *Am J Psychiatry* 1992; 149:944-6.
37. Wadden T, Berkowitz R, Vogt R, Steen S, Stunkard A, Foster G. Lifestyle modification in the pharmacologic treatment of obesity: a pilot investigation of a potential primary care approach. *Obesity Res* 1997; 5:218-26.
38. White J, Keenan M. Stress control: a pilot study of large group therapy for generalized anxiety disorder. *Behav Psychotherapy* 1990; 18:143-6.
39. Beusterien K, Buesching D, Robison R, et al. Evaluation of an information exchange program for primary care patients with depression. *Disease Management* 2000; 3:1-9.
40. Bowman M, Dignan M, Crandall S, Baier M. Changes in functional status related to health maintenance visits to family physicians. *J Fam Pract* 2000; 49:428-33.
41. Durand M, King M. Specialist treatment versus self-help for bulimia nervosa: a randomised controlled trial in general practice. *Personal communication* 2002.

42. Katzelnick D, Simon G, Pearson S, et al. Randomized trial of a depression management program in high utilizers of medical care. *Arch Fam Med* 2000; 9:345-51.
43. King M, Hindler C, Nazareth I, Farmer R, Gerada C, Cohen J. A controlled evaluation of small group education of general practitioners in the management of drug users. *Br J Gen Pract* 1998; 48:1159-60.
44. Kotke T, Brekke M, Solberg L, Hughes J. A randomized trial to increase smoking intervention by physicians: doctors helping smokers, round 1. *JAMA* 1989; 261:2101-6.
45. Wallace P, Cutler S, Haines A. Randomised controlled trial of general practitioner intervention in patients with excessive alcohol consumption. *BMJ* 1988; 297:663-8.
46. Blomhoff S, Haug TT, Hellström K, et al. Randomised controlled general practice trial of sertraline, exposure therapy and combined treatment in generalised social phobia. *Br J Psychiatry* 2001; 179:23-30.
47. Fleming M, Lawton Barry K, Baier Manwell L, Johnson K, London R. Brief physician advice for problem alcohol drinkers: a randomized controlled trial in community-based primary care practices. *JAMA* 1997; 277:1039-45.
48. Fleming M, Baier Manwell L, Lawton Barry K, Adams W, Stauffacher E. Brief physician advice for alcohol problems in older adults: a randomized community-based trial. *J Fam Pract* 1999; 48:378-84.
49. Goldstein M, Pinto, Marcus B, Lynn H, Jette A, Rakowski W. Physician-based physical activity counseling for middle-aged and older adults: a randomized trial. *Ann Behav Med* 1999; 21:40-7.
50. Patel C, Marmot M. Can general practitioners use training in relaxation and management of stress to reduce mild hypertension. *BMJ* 1988; 296:21-4.
51. Pill R, Stott NC, Rollnick SR, Rees M. A randomized controlled trial of an intervention designed to improve the care given in general practice to Type II diabetic patients: patient outcomes and professional ability to change behaviour. *Fam Pract* 1998; 15:229-35.
52. Simkin-Silverman L, Wing R. Management of obesity in primary care. *Obesity Res* 1997; 5:603-12.
53. Anderson P, Scott E. The effect general practitioners' advice to heavy drinking men. *Br J Addiction* 1992; 87:891-900.
54. Brody DS, Lerman CE, Wolfson HG, Caputo GC. Improvement in physicians' counseling of patients with mental health problems. *Arch Intern Med* 1990; 150:993-8.
55. Calfas K, Long B, Sallis J, Wooten W, Pratt M, Patrick K. A controlled trial of physician counseling to promote the adoption of physical activity. *Prev Med* 1996; 25:225-33.
56. Horst HE van der, Eijk JTMv, Schellevis FG, Devillé WLJM. The effectiveness of patient education and promotion of self-care for patients with irritable bowel syndrome in primary care. Irritable bowel syndrome in general practice: how effective is patient education and counselling? (PhD thesis). Amsterdam: HE van der Horst, 1997.
57. Catalan J, Gath D, Edmonds G, Ennis J. The effects of non-prescribing of anxiolytics in general practice: I. controlled evaluation of psychiatric and social outcome. *Br J Psychiatry* 1984; 144:593-602.
58. Catalan J, Gath D, Bond A, Martin P. The effects of non-prescribing anxiolytics in general practice II. Factors associated with outcome. *Br J Psychiatry* 1984; 144:603-10.
59. King M, Davidson O, Taylor F, Haines A, Sharp D, Turner R. Effectiveness of teaching general practitioners skills in brief cognitive behaviour therapy to treat patients with depression: randomised controlled trial. *BMJ* 2002; 324:947-53.

60. Baillargeon L, Demers M, Ladouceur R. Stimulus-control: nonpharmacologic treatment for insomnia. *Can Fam Physician* 1998; 44:73-9.
61. Gask L. Training general practitioners to detect and manage emotional disorders. *Int Rev Psychiatry* 1992; 4:293-300.
62. Morriss R, Gask L, Ronalds C, et al. Clinical and patient satisfaction outcomes of a new treatment for somatized mental disorders taught to general practitioners. *Br J Gen Pract* 1999; 49:263-7.
63. Malt UF, Robak OH, Madsbu HP, Bakke O, Loeb M. The Norwegian naturalistic treatment study of depression in general practice (NORDEP)-I: randomised double blind study. *BMJ* 1999; 318:1180-4.
64. Wilson DM, Taylor W, Gilbert JR, et al. A randomized trial of a family physician intervention for smoking cessation. *JAMA* 1988; 260:1570-4.
65. Mynors-Wallis LM, Gath DH, Day A, Baker F. Randomised controlled trial of problem solving treatment, antidepressant medication, and combined treatment for major depression in primary care. *BMJ* 2000; 320:26-30.
66. Blankenstein AH, Horst van der HE, Schilte AF, et al. Effectiveness of reattribution for somatisation in general practice, a randomised controlled trial. *Somatizing patients in general practice (PhD thesis)*. Amsterdam: AH Blankenstein, 2001:49-65.
67. Richmond RL, Webster IW. A smoking cessation programme for use in general practice. *Med J Austr* 1985; 142:190-4.
68. Segnan N, Pont A, Battista RN, et al. A randomized trial of smoking cessation interventions in general practice in Italy. *Canc Causes Contr* 1991; 2:239-46.
69. McIntosh MC, Leigh G, Baldwin NJ, Marmulak J. Reducing alcohol consumption: comparing three brief methods in family practice. *Can Fam Physician* 1997; 43:1959-67.
70. Richmond R, Heather N, Wodak A, Kehoe L, Webster I. Controlled evaluation of a general practice-based brief intervention for excessive drinking. *Addiction* 1995; 90:119-32.
71. Silagy C, Stead LF. Physician advice for smoking cessation. *The Cochrane Library*. Oxford: Update Software, 2003.
72. Windt DAWM van der, Koes BW, Van Aarst M, Heemskerk MAMB, Bouter LM. Practical aspects of conducting a pragmatic randomised trial in primary care: patient recruitment and outcome assessment. *Br J Gen Pract* 2000; 50:371-4.
73. Huibers MJH, Bleijenberg G, Beurskens AJHM, et al. An alternative trial design to overcome validity and recruitment problems in primary care intervention studies. *Fam Pract* 2003; in press.
74. Moher D, Schulz KF, Altman DG, for the CONSORT group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001; 357:1191-4.
75. Knottnerus JA, Dinant GJ. Medicine based evidence, a prerequisite for evidence based medicine. *BMJ* 1997; 315:1109-10.

An alternative trial design to overcome validity and recruitment problems in primary care research

Marcus J.H. Huibers ^{1,2,3}

Gijs Bleijenberg ²

Anna J.H.M. Beurskens ¹

Ijmert Kant ¹

J. André Knottnerus ³

Daniëlle A.W.M. van der Windt ⁴

Ellen Bazelmans ²

Constant P. van Schayck ³

FAMILY PRACTICE 2003, IN PRESS

¹ DEPARTMENT OF EPIDEMIOLOGY, MAASTRICHT UNIVERSITY, THE NETHERLANDS

² DEPARTMENT OF MEDICAL PSYCHOLOGY, UMC NIJMEGEN, THE NETHERLANDS

³ DEPARTMENT OF GENERAL PRACTICE, MAASTRICHT UNIVERSITY, THE NETHERLANDS

⁴ DEPARTMENT OF GENERAL PRACTICE, INSTITUTE FOR RESEARCH IN EXTRAMURAL MEDICINE, VUMC, THE NETHERLANDS

Abstract

BACKGROUND

Although the randomised controlled trial is widely accepted as the best design to investigate new interventions, conducting a trial in primary care may face researchers with many methodological problems.

OBJECTIVE

To present an alternative trial design to overcome internal validity and recruitment problems.

METHODS

In a randomised controlled trial, fatigued employees absent from work were selected among the population of a occupational health service in the South of the Netherlands. Patients randomly assigned to the experimental condition received cognitive behaviour therapy by a research GP nearby their home address, whereas patients in the control group received no intervention. We describe our considerations for building an alternative design. Research GPs and patients were recruited separately for the study. The prerandomisation design was applied.

RESULTS

Nine research GPs performed all the interventions. 76 experimental patients and 75 control patients were selected for study participation. Of these, only 6 patients in the experimental group and 7 patients in the control group withdrew from the study at some point during follow-up.

CONCLUSION

Results suggest that recruitment and randomisation procedures in the alternative design served their purpose well. The alternative design proposed here might have several advantages compared to conventional trial procedures. However, our design is not widely applicable and there are ethical aspects involved that should be considered. Researchers should address their creativity when trying to minimize the problems they may encounter in designing a study.

Introduction

The randomised controlled trial is widely accepted as the best design to investigate new interventions. In order to successfully run a randomised trial, it is vital to use effective and practical research strategies.¹ However, researchers may face many methodological problems when a trial is conducted in a general practice setting.^{1, 2} Here, we will focus on a few problems that relate to recruitment and randomisation procedures.

The recruitment of general practitioners and patients is a general concern in primary care research. Obtaining the voluntary participation of GPs in research projects has become a major challenge.²⁻⁷ Lack of available time is the most cited reason for the reluctance among GPs to become involved in research activities. Once a sufficient number of GPs has agreed to become active in the research project, a phenomenon referred to as Lasagna's Law⁸ almost inevitably emerges. It entails that researchers and clinicians invariably overestimate the number of patients available for study participation so that recruitment of eligible patients falls short of the expectations.² Factors that might explain this phenomenon are several: eligibility criteria are often overly strict and do not reflect the prevalence of the disorder on which sample size calculations are based, recruiting physicians are not always motivated or able to enrol eligible patients^{2, 9, 10} and eligible patients are not always willing to participate. Ultimately, studies might be delayed, suspended or even prematurely terminated due to unsuccessful recruitment.¹⁰

Recruitment of patients becomes virtually impossible when interventions are complex and, as a consequence, only a small number of GPs can be trained and supervised in performing the intervention. If patients are then to be recruited by their own general practitioners, as is customary in primary care research, only a small number of patients can be approached for study participation. These recruitment problems get even worse if the disorder under investigation is rare and/or if the information needed to identify eligible patients is not directly available to the GP.

In some instances, the randomisation procedure is likely to cause internal validity problems. In most primary care intervention studies, patients are recruited by their own GPs and randomly allocated to one of the treatment conditions after informed consent. As widely used as it is, this conventional approach has problematic features if 'usual care' or 'no treatment' is the control condition¹¹, as these conditions usually cannot be blinded.

With 'usual care' or 'no treatment' as the control condition, it can become very difficult to find patients who are willing to be randomised, especially if the experimental condition is an appealing option to patients.¹¹ More importantly, disappointment because of control group allocation among those who do agree with randomisation might lead to selective withdrawal from the study, selective non-compliance¹¹ or contamination¹² if control patients, inspired by their research participation, start to look for treatment elsewhere. The internal validity, and thereby the justification of the study, would then be seriously threatened.

Researchers foreseeing recruitment, internal validity or other methodological problems should try and minimize these problems. The question here is how. Obviously, there is no single answer to this question. Instead, we wish to share our particular experiences by presenting an alternative study design that enabled us to overcome some of the difficulties mentioned here. We conducted a trial in primary care in which we assessed the efficacy of cognitive behaviour therapy by general practitioners for persistent fatigue. In this paper, our trial will serve as case illustration. Following the description of the design, the results of the recruitment and randomisation procedures will be presented and discussed.

Methods

STARTING POINT FOR BUILDING THE DESIGN

Several considerations were taken into account when we started designing the study. First, we wished to conduct a randomised trial with an experimental and a control condition and a follow-up period of twelve months. Since there is no standard treatment for persistent fatigue and our intervention is new and unfamiliar to GPs, we decided to assess the efficacy (can it work under ideal circumstances?) of the intervention rather than the effectiveness (does it work in ordinary settings?).¹³ We therefore chose to create optimal conditions for the experimental intervention in order to assess its potential effect on the one hand, and to merely follow up the natural course in the control group on the other hand.

Choosing efficacy over effectiveness presented us with two problems however. The rather complex experimental intervention consists of five to seven 30-minute sessions by a GP. In order to administer the intervention under the best circumstances, GPs should be extensively trained and supervised throughout the trial. Our time and resources enabled us to train and supervise up to ten GPs. If these GPs had to recruit all patients in their own practice, only a small number of patients could be enrolled in the study. Also, given the fact that the compared conditions could not be blinded, we feared that classical randomisation of selected patients after informed consent might lead to selective withdrawal and contamination of the trial conditions, which could pose a serious threat to the internal validity of the study. Consequently, we were challenged to design alternative recruitment and randomisation procedures.

RECRUITMENT OF GENERAL PRACTITIONERS AND PATIENTS

We decided to recruit patients and GPs separately. GPs were recruited among the population of GPs in the Southeast of the Netherlands. We aimed to deploy up to ten research GPs whose practices were equally distributed across the recruitment area. We decided that research GPs would only become active if a sufficient number of patients from their area could be enrolled in the study. However, instead of having these research GPs recruit their own patients for the trial, patients were recruited in collaboration with a local occupational health service (OHS) that monitors a working population of 80.000 employees who live in the recruitment area. We needed 150 patients in total (75 patients per group) to detect a clinically relevant difference on the main outcome measure, the Checklist Individual Strength (CIS).¹⁴⁻¹⁶ Eligibility criteria for the trial were severe fatigue (measured with the CIS) for more than four months as one of the main health problems and complete absenteeism from work for more than five weeks.

On a monthly basis, employees monitored by the OHS who were on sick leave for more than two weeks, were sent study information by the OHS, irrespective of the reason for sick leave, followed

by a 'blind' reminder two weeks later. The recruitment letter included a screening list that could be sent back to the research team. This recruitment route secured the privacy of selected employees and enabled us to identify eligible patients in an early stage of absenteeism. Based on the screening lists sent back to us, we contacted potential candidates who were willing to participate. Eligible patients were invited to visit the university research centre.

PRERANDOMISATION DESIGN

As stated earlier, we foresaw internal validity problems like contamination and selective withdrawal if patients would be randomised after informed consent. Therefore, we used the prerandomisation design¹², also referred to as randomised consent or Zelen design.¹⁷ In the original Zelen design, informed consent is obtained after randomisation and only from patients in the experimental group (see figure 1, a.) in order to prevent selective withdrawal and contamination during the trial. Although the use of this design has raised many objections among researchers¹⁸, others have suggested it as an alternative for classical randomisation if the experimental intervention is attractive to all participants and the control group will receive standard (or no) treatment.¹¹

FIGURE 1
ORIGINAL ZELEN DESIGN AND ADAPTED PRERANDOMISATION DESIGN

R=randomisation

For the purpose of our study, we used the adapted prerandomisation design (see figure 1, b.).¹² Informed consent was obtained from participants in both groups. However, randomisation took place before detailed information about the study was provided, and participants allocated to one group were kept blind to the randomisation procedure and thus, the existence of the other group. In the initial recruitment letter sent by the OHS, we invited fatigued employees absent from work to participate in a 'health-related' study, without mentioning the experimental intervention. At their visit to the research centre, patients who were selected for study participation and randomised, were handed an envelope containing written information about the allocated condition only: patients in both groups learned that they would be followed up over the next twelve months, while patients in the experimental group in addition were offered the experimental intervention. Experimental patients who entered the study were assigned to a research GP nearby their home addresses and a first appointment was made by the researchers. Regular GPs of patients in both groups were contacted, informed about the study participation of their patients and asked to register the medical consumption for the following four months. The medical ethics committee of Maastricht University approved the study protocol.

Results

RECRUITMENT OF GPs

Figure 2 presents the results of the recruitment of GPs that started in March 1999. Of the 1030 GPs addressed by letter, 90 GPs (8.7%) responded. Based on their geographical position, we invited 30 GPs to attend our training sessions. 22 GPs attended the first training session and 15 GPs attended the second training session also, which was a requirement to become active as research GP. Of these 15 trained GPs, we eventually deployed nine research GPs - based on the local intake of patients - who performed all the interventions.

FIGURE 2
FLOW CHART: RECRUITMENT OF GENERAL PRACTITIONERS

FIGURE 3

FLOW CHART: RECRUITMENT OF PATIENTS

RECRUITMENT OF PATIENTS

Figure 3 presents the results of the recruitment of patients that started in November 1999. 8736 employees who were absent from work for more than two weeks, were addressed within the course of twenty months (November 1999-June 2001). 4242 employees (48.5%) responded to the initial recruitment letter. Of the 2454 employees absent from work for more than *two* weeks who were willing to participate in the study, only 195 employees (7.9%) met our strict eligibility criteria (e.g. severe fatigue, complete absenteeism for more than *five* weeks) after screening and were invited to visit the research centre. Eventually, 151 employees entered the study.

REFUSAL OF TREATMENT AFTER RANDOMISATION

As a result of the prerandomisation procedure, a phenomenon we refer to as 'refusal of treatment after randomisation' emerged: since patients were not informed beforehand that an experimental intervention could be offered to them, five patients refused the intervention once they were informed about the experimental condition (see Figure 3). Since they were randomised into the study, these patients were encouraged to take part in the follow-up assessments and were included in intention-to-treat analyses.

WITHDRAWAL RATES

Of the 151 patients who entered the study, thirteen patients (8.6%) withdrew from the study at some point during the twelve-month follow-up period (see figure 3). Data were lost for two experimental patients (1.3%) and three control patients (2%) at four-month assessment, for three experimental patients (2%) and six control patients (4%) at eight-month assessment and for six experimental patients (4%) and seven control patients (4.6%) at twelve-month assessment. Of the five experimental patients who refused the intervention directly after randomisation, two patients did not complete data collection.

Discussion

Results from the recruitment and randomisation procedures applied in our study suggest that the developed research design was successful. By approaching a considerable number of GPs for study participation, we were able to select a small group of highly motivated research GPs who were willing to undergo the extensive training and supervision and perform complex interventions. It also seems likely that we would have selected fewer patients in the same period of time if conventional recruitment procedures had been used, although Lasagna's Law⁸ suggests that no reliable estimate of the number of patients potentially recruited can be made. The non-selective and relatively low withdrawal rates throughout the trial suggest that the use of prerandomisation had the desired effect. However, in the absence of similar studies using classical randomisation, we can merely speculate what would have happened if our randomisation procedures had been different. The patients' own GPs did not voice any complaints regarding the study participation of their patients.

Since we only by-passed recruitment by a physician, the influence of Lasagna's Law was not eliminated in our trial. According to the Maastricht Cohort Study¹⁹, an affiliated large-scale prospective study of fatigue among 12.000 employees, the point prevalence of persons in the working population meeting all our eligibility criteria is 0.8%. In twenty months time however, we were only able to select 151 prevalent *and* incident cases (0.2%) among the 80.000 employees of our source population. The intake of patients varied between four and sixteen patients per month, another illustration of the fact that recruitment remained a challenge in our design also.

The trial design proposed here has several advantages compared to more conventional trial procedures: researchers are encouraged to open up new recruitment markets, e.g. through collaboration with the OHS in our example, through secondary care services, or newspaper advertisements; researchers maintain control over all recruitment procedures, which contributes to creating optimal and efficient study conditions; recruitment is facilitated of patients who are otherwise more difficult to identify due to strict eligibility criteria, low prevalence of the disorder and/or the fact that information needed for detection is not directly available to the GP; and internal validity problems likely to occur due to patient preference (i.e. selective withdrawal, selective non-compliance and contamination) are reduced.

However, in weighing these advantages against the use of more conventional trial procedures, one should keep in mind that we tailored the design to our particular research needs. Not all of the problems that we encountered are universal problems in primary care research, and the solutions we have advocated here are only applicable under certain circumstances. The proposed design is of no use if recruitment in the general practice is essential to the aim of the study or when the deployment of a research GP instead of the patients' own GP is considered to be a threat to the external validity. But more importantly, the decision to use a prerandomisation design should not be made light-heartedly. The prerandomisation design has been the topic of an unresolved debate concerning the ethical

aspects of informed consent procedures.²⁰ As stated in the recently published Research Governance Framework, informed consent of patients is at the heart of ethical research and should be taken as agreed with the ethics committee.²¹ In 1999, the Health Council in The Netherlands reported on the ethical aspects of informed consent.²⁰ It was concluded that application of a prerandomisation design is only ethically justified if all of the following conditions are met: the study has a clinically relevant objective; it is likely that the study will lead to important new insights; an internally valid evaluation of the effects is impossible using classical randomisation and informed consent procedures; and the evaluated interventions bear no potential harm to participants. In an earlier publication, Knottnerus mentioned in addition to these criteria the necessity to offer at least usual care to patients in the control group and the approval of an independent medical ethics committee with participation from patients.¹² An important issue that deserves special attention is the perspective of participants: how do patients feel about participating in a prerandomised trial?

We decided to apply a prerandomisation design because we feared the lack of blinding would pose a serious threat to the internal validity of our study, i.e. would lead to contamination of the research groups and selective withdrawal from the study. We made sure that all of the conditions for using the design were met, including approval from a medical ethics committee. Until now, we were not able to assess how participants felt about our randomisation procedure because we are still awaiting a final follow-up round.

However, it is important to realise that use of the prerandomisation design is by no means justified when the objective is merely to speed up the recruitment of patients. We used a separate strategy to tackle the recruitment problems we foresaw, something that should not be confused with our choice to use a prerandomisation design.

Despite these limitations, our trial design may present an interesting starting point for researchers who are challenged to be creative in dealing with the methodological and practical problems they encounter. In that sense, our design may serve as a basis for some efficacy studies and is less applicable when the effectiveness of an intervention is evaluated.

Acknowledgments

This research project was funded by the Health Research and Development Council (ZorgOnderzoek Nederland), The Netherlands. The recruitment of patients was made possible by the donation of services by 'Arbo Unie Limburg' (Occupational Health Service Limburg). The authors wish to thank Prof. Paul Knipschild for his useful comments on the manuscript.

References

1. Ward E, King M, Lloyd M, Bower P, Friedli K. Conducting randomized trials in general practice: methodological and practical issues. *Br J Gen Pract* 1999; 49:919-22.
2. Van der Windt DAWM, Koes BW, Van Aarst M, Heemskerk MAMB, Bouter LM. Practical aspects of conducting a pragmatic randomised trial in primary care: patient recruitment and outcome assessment. *Br J Gen Pract* 2000; 50:371-74.
3. Huibers MJH, Van der Windt DAWM, Boeke AJP. De deelname van huisartsen aan wetenschappelijk onderzoek: problemen en oplossingen (Participation of general practitioners in primary care research: problems and solutions). *Huisarts en Wetenschap* 2002; 45:454-8.
4. Borgiel AE, Dunn EV, Lamont CT, et al. Recruiting family physicians as participants in research. *Fam Pract* 1989; 6:168-72.
5. Levinson W, Dull VT, Roter DL, Chaumenton N, Frankel RM. Recruiting physicians for office-based research. *Med Care* 1998; 36:934-7.
6. Gray RW, Woodward NJ, Carter YH. Barriers to the development of collaborative research in general practice: a qualitative study. *Br J Gen Pract* 2001; 51:221-2.
7. Heywood A, Mudge P, Ring I, Sanson-Fisher R. Reducing systematic bias in studies of general practitioners: the use of a medical peer in the recruitment of general practitioners in research. *Fam Pract* 1995; 12:227-31.
8. Harris EL, Fitzgerald JD. The principles and practices of clinical trials. In: Harris EL, Fitzgerald JD, eds. *Edinburgh and London: E. & S. Livingstone*, 1970.
9. Peto V, Coulter A, Bond A. Factors affecting general practitioners' recruitment of patients into a prospective study. *Fam Pract* 1993; 10:207-11.
10. Tognoni G, Alli C, Avanzini F, et al. Randomised clinical trials in general practice: lessons from a failure. *BMJ* 1991; 303:969-71.
11. Schellings R, Kessels AGH, Ter Riet G, Sturmans F. The Zelen design may be the best choice for a heroin-provision experiment. *J Clin Epidemiol* 1999; 52:503-7.
12. Knottnerus JA. Gezondheidszorgonderzoek in extramurale settings (Health care research in extramural settings). In: Ten Have HAMJ, Blijham G, Engberts DP, eds. *Ethiek en recht in de gezondheidszorg (Ethics and law in health care)*. Deventer: Kluwer, 1997:151-98.
13. Fletcher RH, Fletcher SW, Wagner EH. *Clinical epidemiology: the essentials*. Baltimore, MD: Williams & Wilkins, 1996.
14. Beurskens AJHM, Bültmann U, Kant IJ, Vercoulen JHMM, Bleijenberg G, Swaen GMH. Fatigue among working people: validity of a questionnaire measure. *Occup Environ Med* 2000; 57:353-7.
15. Vercoulen JHMM, Alberts M, Bleijenberg G. De Checklist Individual Strength (CIS). *Gedragstherapie* 1999; 32:31-6.
16. Vercoulen JHMM, Swanink CM, Fennis JF, Galama JM, Van der Meer JW, Bleijenberg G. Dimensional assessment of chronic fatigue syndrome. *J Psychosom Res* 1994; 38:383-92.
17. Zelen M. Randomised consent designs for clinical trials: an update. *Statistics in medicine* 1990; 9:645-56.
18. Ellenberg SS. Randomization designs in comparative clinical trials. *N Eng J Med* 1984; 310:1404-8.
19. Bültmann U, Kant IJ, Kasl SV, Beurskens AJHM, Van den Brandt PA. Fatigue and psychological distress in the working population: psychometrics, prevalence and correlates. *J Psychosom Res* 2002; 52:445-52.

20. *Gezondheidsraad (Health Council of The Netherlands). Informed consent en prerandomisatie (Informed consent and prerandomisation). The Hague: Health Council of The Netherlands (Gezondheidsraad), 1999.*
21. *Department of Health (UK). Research Governance Framework for Health and Social Care, second edition (autumn) 2003.*

Efficacy of cognitive behaviour therapy by general practitioners for unexplained fatigue among employees: a randomised controlled trial

Marcus J.H. Huibers ^{1,2,3}

Anna J.H.M. Beurskens ¹

Constant P. van Schayck ²

Ellen Bazelmans ³

Job F.M. Metsemakers ²

J. André Knottnerus ²

Gijs Bleijenberg ³

THE BRITISH JOURNAL OF PSYCHIATRY 2003, IN PRESS

¹ DEPARTMENT OF EPIDEMIOLOGY, MAASTRICHT UNIVERSITY, THE NETHERLANDS

² DEPARTMENT OF GENERAL PRACTICE, MAASTRICHT UNIVERSITY, THE NETHERLANDS

³ DEPARTMENT OF MEDICAL PSYCHOLOGY, UMC ST RADBOUD NIJMEGEN, THE NETHERLANDS

Abstract

BACKGROUND

Fatigue is a common complaint that may lead to long-term sick leave and work disability.

AIMS

To assess the efficacy of cognitive behaviour therapy by general practitioners for unexplained, persistent fatigue among employees.

METHOD

A randomised controlled trial using a prerandomisation design in primary care. 151 fatigued employees on sick leave entered the study. Patients in the experimental group were offered five to seven 30-minute sessions of cognitive behaviour therapy by a trained research GP, while participants in the control group were offered no treatment. Main outcome measures: fatigue severity, self-reported absenteeism, registered absenteeism, and clinical recovery assessed at 4, 8 and 12 months.

RESULTS

At baseline, 44% of the patients already met research criteria for chronic fatigue syndrome. There were no significant differences between the experimental group and the control group on primary or secondary outcomes at any point.

CONCLUSIONS

Cognitive behaviour therapy by GPs for unexplained, persistent fatigue did not prove to be an effective intervention. Since research GPs were not able to deliver CBT effectively under ideal circumstances, it is unlikely that GPs in routine practice would be more successful in doing so.

Introduction

Fatigue is a common complaint in the general and working population, with a reported prevalence varying from 7 to 45%.¹⁻³ In primary care, 5 to 10% of patients present with fatigue as their main complaint.⁴ In most of these patients, fatigue lacks a clear somatic cause⁴ and appears to be a functional symptom.^{5,6} Fatigue can best be understood as a continuum, ranging from mild complaints frequently seen in the community to severe, disabling fatigue like the chronic fatigue syndrome (CFS).¹ When fatigue becomes severe and persistent, it may lead to long-term sick leave⁷ and work disability.⁸ Although persistently fatigued patients tend to rely heavily on the care of their general practitioner (GP), evidence-based treatment options for GPs are hardly available while referral to secondary care is often no option. Studies in secondary care have shown that cognitive behaviour therapy (CBT) by skilled therapists is effective in the treatment of chronic fatigue syndrome.^{9, 10} It has been suggested that some GPs might also provide CBT in the treatment of (chronic) fatigue.⁴ However, it is unknown whether CBT is effective in less advanced fatigue cases and whether GPs are able to deliver CBT.

In this study, we aimed to assess the efficacy of CBT delivered by GPs for *unexplained, persistent* fatigue. We hypothesized that delivering CBT to fatigued employees on sick leave would reduce fatigue complaints and stimulate the resumption of work.

Methods

DESIGN AND PROCEDURES

To study our hypothesis, we conducted a randomised controlled trial in primary care using an adapted prerandomisation or randomised consent design¹¹ to prevent contamination and selective withdrawal.¹² The trial was designed to assess the *efficacy*¹³ of the intervention, i.e. to assess the potential benefit of the intervention under ideal circumstances. Randomisation took place before detailed information about the study was provided, and patients allocated to one group were kept blind to the randomisation procedure and thus, the existence of the other group. After randomisation, patients were informed only about the procedures in the condition they were allocated to and informed consent was obtained. Patients were randomised to receive cognitive behaviour therapy by a GP or to be followed up in a control group. Allocation to group was carried out by the principal investigator (MH) using cards in sealed opaque envelopes. An independent person prepared envelopes by coding them according to a computer-generated list of random numbers. Randomisation was performed in blocks of four and pre-stratified according to 'duration of absenteeism'. Data were collected on four occasions: at baseline measurement, at the end of the treatment period (4 months) and twice at follow up (8 and 12 months). The medical ethics committee of Maastricht University approved the study protocol.

PATIENTS

Patients were recruited in collaboration with a local occupational health service (OHS) that monitors a working population of 80,000 employees. On a monthly basis, employees monitored by the OHS who were on sick leave were sent limited study information (including a screening list) by the OHS, irrespective of the reason for sick leave, followed by a reminder two weeks later. Based on the screening lists that were sent back to the research team, we invited potential candidates who were willing to participate to visit the university research centre.

Eligibility was assessed by the principal investigator (MH) in a clinical interview. Inclusion criteria were severe fatigue (a score of 35 or more on the subscale fatigue of the Checklist Individual Strength^{14,15}; for four months or more as one of the main health problems and complete absenteeism from work for 6 to 26 weeks. Patients were excluded from participation if they had medical conditions that explained fatigue (e.g. cancer); received co-interventions for fatigue (e.g. treatment for burnout); had a previously classified psychiatric disorder; or received concurrent psychological treatment (e.g. cognitive behaviour therapy for major depression). In addition, absenteeism should not be caused primarily by problems unrelated to health (e.g. work conflict).

INTERVENTIONS

Patients in both groups were asked if they could be followed up over the next twelve months. In addition, patients allocated to the experimental treatment group were offered five to seven 30-minute sessions of cognitive behaviour therapy (CBT) within the course of four months. Nine GPs delivered all interventions aside their regular practice. These 'research' GPs were recruited among the GP population in the Southeast of the Netherlands based on their geographical position. Patients who agreed to receive CBT were assigned to a research GP nearby their home address. The intervention was partly based on the CBT programme for chronic fatigue syndrome developed by members of our group.¹⁰ Box 1 summarizes the goal and content of the intervention. The intervention was written out in a treatment manual. Research GPs – none of which had previous experience with CBT – were trained in delivering the intervention in two five-hour workshops and supervised in monthly two-hour sessions throughout the trial by two experienced behaviour therapists (GB&EB). No research interventions were offered to patients in the control group. Patients in both groups were free to visit their regular GP for usual care.

BOX 1

GP-DELIVERED CBT PROGRAMME FOR UNEXPLAINED, PERSISTENT FATIGUE AMONG EMPLOYEES

goal of the intervention

- diminish fatigue and other complaints
- establishing work resumption and other personal goals
- establishing self-perceived recovery

steps in the intervention

- assessment of perpetuating factors on the level of:
 - cognitions (e.g. non-acceptance fatigue, lost sense of control over symptoms)
 - overt behaviour (e.g. disturbed sleeping pattern, unbalanced physical activities)
 - social factors (e.g. lack of social support, dysfunctional work environment)
- modification of identified perpetuating factors by:
 - explaining the perpetuating circle
 - setting goals for activities and other problem areas
 - providing helpful cognitions
 - planning systematic and gradual work resumption
 - planning achievement of other personal goals
 - involving the social environment

OUTCOME ASSESSMENT

Researchers and research GPs were not blind to the group allocation. All self-reported outcomes were assessed using computerized questionnaires at the research centre (baseline and post-treatment measurement) and postal questionnaires (8 and 12 month follow up).

Primary outcomes were fatigue severity, absenteeism and 'clinical recovery'. Fatigue was measured with the subscale fatigue severity of the CIS ^{14, 15}, high scores indicating high severity. Absenteeism was assessed in two ways: self-reported work resumption or absenteeism (yes/no) at each measurement and absenteeism registered by the occupational health service (number of partial or complete sick days in 365 days). Clinical recovery was defined as having a CIS fatigue score of 34 or lower in combination with self-reported work resumption.

Secondary outcomes were global perceived effect (GPE), physical functioning and psychological distress. GPE was rated on a 7-point scale ¹⁶ and then dichotomised into 'recovered' and 'not recovered'. Physical functioning was measured with the subscale physical functioning of the Short Form Health Survey (SF-36) ¹⁷, high scores indicating high functioning. Psychological distress was measured with the Symptom Checklist 90 (SCL-90) ¹⁸, high scores indicating high distress.

To assess adherence of experimental patients and research GPs to the treatment manual, GPs used standardized registration forms on which they registered the actions taken in each therapy session and commented on the proceedings. Based on these forms and the information gathered in supervision, the CBT supervisors performed a quality check. For each experimental patient it was assessed whether the CBT received was according to 'protocol'. This was defined as being exposed to all essential steps of the intervention (box 1), something that might have been achieved in less than five sessions. In addition, treatment process indicators from our previous fatigue studies ^{10, 19, 20} were used to assess the cognitive impact of the intervention: self-efficacy (sense of control in relation to complaints), psychological attributions (beliefs regarding psychological cause of complaints) and somatic attributions (beliefs regarding somatic cause of complaints), high scores indicating high self-efficacy or strong attributions.

Additional measures included psychosocial co-interventions, usual care by the regular GP and physical activity. Psychosocial co-interventions (defined as treatment by a psychiatrist, a psychologist and/or psychoactive drug treatment) were measured by self-report and expressed in numbers of patients who received co-interventions during the first four months after baseline. Usual GP care was expressed as number of visits to the regular GP during the first four months after baseline. Physical activity was measured with the actometer, a motion-sensing device attached to the ankle and worn for 12 days continuously.²¹ Scores reflect the average number of accelerations per 5-minute period. High scores indicate high activity, with a mean of 91 (SD 25) for healthy subjects.

SAMPLE SIZE AND DATA ANALYSIS

In calculating the sample size we chose the fatigue scale CIS as our main outcome measure. Based on the baseline data of the affiliated Maastricht Cohort Study ²² we estimated the SD to be 12. To detect a clinically significant difference of 6 or more on the CIS at a two-sided significance level of 5% and a power of 80% we would need 63 patients per group. We extended the sample size to 75 patients per group to take into account potential withdrawal from the trial.

In the main analysis, data were analysed according to the intention-to-treat principle using all available data. The t-test for independent samples (continuous data) and the chi-square test (dichotomous data) were used to compare the two conditions. Difference scores on continuous outcomes (to-score minus tx-score) were used in the analysis. The influence of baseline differences was assessed using linear and logistic regression. An analysis of variance was used to assess the effect of research GP on treatment outcomes. In addition to intention-to-treat analyses, explanatory subgroup analyses were performed. All analyses were performed using SPSS (version 11.0).

Results

Figure 1 presents the flow of patients through the trial. Recruitment started January 2000 and was completed in July 2001. Since all employees on sick leave, most of whom were not fatigued, were addressed by the OHS, only a small fraction of the responders was eligible for study participation. 151 patients were thus enrolled in the study. After 12 months of follow up (July 2002), data were available for 138 patients who completed the study (withdrawal rate 8.6%).

Baseline demographic and clinical characteristics were fairly similar in both groups, except for physical functioning and psychological distress (table 1). Adjustment for these two factors did not change the findings. Patients were identified as potential chronic fatigue syndrome cases (CFS-like cases) if they met all of the following research criteria²³: a CIS score of 40 or higher, a duration of fatigue complaints of six months or more and a SF-36-score on the subscale physical functioning of 60 or lower. At baseline, in total 66 patients (44%) met research criteria for CFS, none of whom reported to be diagnosed with CFS previously. One patient indicated she believed she suffered from CFS.

TABLE 1

BASELINE CHARACTERISTICS OF PATIENTS. DATA ARE MEAN VALUES (SD) UNLESS STATED OTHERWISE

characteristics	CBT n=76	control n=75
sex, number (%) m/f	37/39 (49/51)	31/44 (41/59)
age (years)	43.6 (8.9)	43.3 (7.7)
education (1=low to 7=high)	3.8 (1.7)	3.8 (1.5)
fatigue severity (CIS, scale 35-56)	48.8 (5.3)	48.5 (7.1)
duration of fatigue complaints (months)	27.9 (31.4)	26.9 (27.6)
duration of absenteeism (weeks)	12.6 (5.9)	12 (3.7)
other health complaints, number (%)		
painful joints/limbs	30 (40)	34 (45)
stress	22 (29)	20 (27)
headache	17 (22)	16 (21)
backpain	12 (16)	17 (23)
poor concentration or memory	14 (18)	13 (17)
depressive mood	10 (13)	16 (21)
sleep disturbances	12 (16)	11 (15)
dizziness	12 (16)	10 (13)
lack of energy	8 (11)	11 (15)
bowel, abdominal or stomach complaints	8 (11)	7 (9)
other	17 (22)	17 (23)
physical functioning (SF-36, scale 0-100)	56.7 (25)	67.1 (22)
psychological distress (SCL-90, scale 90-450)	175 (42)	190 (57)
general physical activity (actometer) *	67 (22)	69.5 (22)
CFS-like cases, number (%) †	36 (47)	30 (40)

* Valid actometer data were available for 60 (CBT group) and 60 (C group) patients.

† CFS-like cases=patients who met research criteria for chronic fatigue syndrome

FIGURE 1

FLOWCHART OF PATIENT RECRUITMENT, TREATMENT ALLOCATION AND OUTCOME ASSESSMENTS

TABLE 2

CLINICAL OUTCOMES. DATA ARE MEAN VALUES (SD) UNLESS STATED OTHERWISE

clinical outcome	CBT (n=76)	control (n=75)	difference in difference score (to-tx) or percentage (95% CI)
<i>fatigue severity (CIS)</i>			
to*	48.8 (5.3)	48.5 (7)	
t1	38.2 (12.7)	39.8 (13.7)	2.2 (-1.7 to 6.1)
t3	35.5 (12.9)	33.9 (13.8)	-0.5 (-5 to 3.9)
<i>work resumers, % (nr)</i>			
t1	50 (37)	61 (44)	-11% (-23 to 1)
t3	59 (41)	65 (44)	-6% (-23 to 10)
<i>registered absenteeism (days)</i>			
t3	234 (116)	230 (116)	4 (-35.9 to 44)
<i>clinical recovery, % (nr)</i>			
t1	24 (18)	28 (20)	-4% (-18 to 11)
t3	33 (23)	43 (29)	-10% (-26 to 6)
<i>perceived recovery (GPE), % (nr)</i>			
t1	38 (28)	39 (28)	-1% (-17 to 15)
t3	49 (34)	53 (36)	-4% (-21 to 12)
<i>physical functioning (SF-36)</i>			
to	56.8 (25)	67.1 (22)	
t1	65.2 (25.8)	72.5 (19.1)	-4.4 (-10.9 to 2.1)
t3	70.1 (24.7)	77.4 (20.9)	-3.6 (-10.4 to 3.1)
<i>psychological distress (SCL-90)</i>			
to	175 (42)	190 (57)	
t1	156 (42)	163 (55)	-10.4 (-23.5 to 2.7)
t3	152 (51)	153 (62)	-11.4 (-27.1 to 4.3)

* to=baseline, t1=post-treatment (4 months), t3=follow-up (12 months)

TABLE 3

TREATMENT PROCESS INDICATORS. DATA ARE MEAN VALUES (SD)

indicator scale	CBT (n=76)	control (n=75)	difference in difference score (to-tx) or percentage (95% CI)
<i>self-efficacy (5-24)</i>			
to*	15.4 (3.5)	14.9 (3.5)	
t1	16.6 (3.9)	16.4 (3.6)	0.2 (-0.9 to 1.3)
<i>psychological attributions (7-28)</i>			
to	18.3 (4.5)	19.4 (4.1)	
t1	18.4 (4.8)	19.2 (3.8)	-0.3 (-1.3 to 0.8)
<i>somatic attributions (4-16)</i>			
to	9.3 (2.5)	9.3 (2.6)	
t1	9.5 (2.6)	9.4 (2.6)	-0.1 (-0.8 to 0.6)

* to=baseline, t1=post-treatment (4 months)

OUTCOMES OF INTERVENTION

Table 2 presents the clinical outcomes at baseline, 4 months and 12 months. There were no significant differences between the CBT group and the control group on any clinical outcome at any point (including the 8-month assessment).

TREATMENT RECEIVED

Five patients allocated to the CBT group refused the treatment offered to them instantaneously (figure 1). Of the remaining 71 patients who agreed to receive treatment, 51 patients (71.8%) completed the CBT according to protocol.

Mean number of CBT sessions attended was 5.3 (SD 1.9), with a mean duration of 33.7 minutes (SD 9.4). In addition, 13 patients (18% of 74) in the CBT group underwent psychosocial co-interventions during the first four months (intervention period) compared to 22 patients (31% of 72) in the control group (95% CI for the difference, -1% to 27%). Mean number of visits to the regular GP in the intervention period was 4.2 (SD 4.5) in the CBT group and 3.5 (SD 4.3) in the control group (95% CI for the difference, -0.8 to 2.1). No adverse events attributable to CBT were reported.

Table 3 presents scores on the treatment process indicators assessed at baseline and post-treatment measurement. There were no differences between the CBT and the control group. Moreover, post-treatment scores remained stable in both groups compared to baseline.

OUTCOMES IN CBT COMPLETERS

We compared the baseline characteristics of those who completed CBT according to protocol (CBT completers, $n=51$) to those who refused CBT or withdrew from CBT (CBT non-completers, $n=25$) and found that CBT completers were similar to CBT non-completers at baseline. Clinical outcomes and scores on treatment process indicators of CBT completers ($n=51$) were then compared to scores in the control group ($n=75$). There were no significant or considerable differences between these groups.

SUBGROUP ANALYSES

While some differences appeared between individual research GPs in treatment effect, no overall GP effect on clinical outcomes was found. Patients were stratified according to fatigue severity (CIS 35-49, $n=74$; CIS 50-56, $n=77$), duration of absenteeism (6-12 weeks, $n=90$; ≥ 13 weeks, $n=61$), duration of fatigue complaints (4-12 months, $n=79$; ≥ 13 months, $n=72$), CFS-like caseness (yes, $n=66$; no, $n=85$), sex (men, $n=68$; women, $n=83$), age (≤ 44 , $n=81$; ≥ 45 , $n=77$) and education (low, $n=109$; high, $n=42$) at baseline. There were no significant or considerable effects of CBT compared to control in any of these subgroups.

Discussion

Cognitive behaviour therapy delivered by GPs to fatigued employees who were absent from work had no substantial effect on clinical outcomes or cognitive processes during 12 months of follow-up. This finding could not be explained by a difference in effect among GPs or withdrawal from treatment in the CBT group. Exploratory analyses in relevant subgroups revealed no significant or considerable effects of CBT.

METHODOLOGICAL CONSIDERATIONS

This is the second study of CBT by GPs for fatigue. In a first, preliminary study, it was concluded that GPs were unable to treat CFS effectively²⁴, but the strength of this study should be questioned because of poor recruitment and high drop out rates. We designed our trial to assess whether the intervention works under ideal circumstances. Consequently, we chose to train and deploy a small number of research GPs instead of a large sample of GPs who would have to treat their own patients. Blinding is virtually impossible in this type of research, and therefore we used a prerandomisation design to control for contamination between groups and to prevent selective withdrawal from the study.¹² In our view, the low withdrawal rate (8.6% of 151 patients) outweighs the occasional refusal of treatment as a result of the design.

As could be expected, patients in the control group received more co-interventions, which may have led to an underestimation of the effect of CBT. A limitation of this study is the nature of our quality check. We used registration forms instead of audiovisual recordings to assess the quality of the CBT that was delivered. As a result, we cannot be certain that GPs actually did what they claimed to have done.

CBT FOR FATIGUE AND CBT BY GPs

This is the first study on CBT for fatigue in primary care that adequately controlled for usual care. CBT for chronic fatigue by skilled therapists was no more effective than counselling in a study by Ridsdale et al.²⁵, but their findings were not compared to no treatment or usual care. It is interesting to find that the portion of patients who failed to complete CBT in our study (33%) is comparable to therapy dropout rates in primary²⁵ and secondary¹⁰ care studies on CBT for chronic fatigue.

Few earlier studies have addressed the effectiveness of CBT by GPs. In a recent study by King et al.²⁶, teaching CBT skills to GPs proved ineffective. However, the extent to which GPs actually applied their skills was not assessed in this study. In a review for the Cochrane Collaboration²⁷, we found little evidence for the effectiveness of psychosocial interventions by GPs, except for a favourable effect of problem-solving treatment by a small number of experienced GPs on major depression.^{28, 29} In general, positive effects of psychosocial interventions seemed to be associated with a small number of GPs (10 or less) treating many patients.²⁷

CAN THE LACK OF EFFICACY BE EXPLAINED?

The lack of efficacy in this study cannot be attributed to clear methodological flaws: interventions were delivered under 'ideal circumstances' as opposed to routine care evaluations, study conditions were highly contrasted, and the majority of patients and GPs appeared compliant with the protocol. Rather, the lack of efficacy is likely to result from a disturbance in the interaction between the patient, the doctor and the intervention. Our study was not designed to reveal the source of the disturbance in this triad.

However, a plausible explanation for our findings would be that research GPs did not treat enough patients to gain the necessary experience in delivering the complex intervention. Also, since many patients at baseline were characterised by an advanced degree of impairment in terms of fatigue, psychological distress and physical functioning at baseline, they might have had better chances for improvement in secondary mental health care.¹⁰ This might be particularly true for those who met the research criteria for CFS.

On the other hand, one might claim that the substantial (natural) recovery in both groups is partly accountable for the lack of differences in effect between the conditions. Another explanation would then be that the brief CBT intervention in itself was not adequate, specific or intensive enough for those who did not recover in the course of time. An additional difficulty might have been that the intervention was not only targeted at fatigue but also at the resumption of work, a complex process that is determined by a wide variety of factors.

How do our findings reflect on the evidence for the effectiveness of CBT in patients with CFS? We should keep in mind that our study differs from previous studies^{9,10} in important ways. We did not deploy experienced CBT therapists but relatively inexperienced GPs, our study population consisted of individuals with a different profile (e.g. impairment, attitudes, illness beliefs, prognosis) than 'classical' CFS patients in secondary care, and our CBT intervention consisted of only half the number of sessions administered in other studies. Apparently, these factors have to be matched appropriately for CBT to be successful in CFS patients.

IN CONCLUSION

A recent review concluded that it cannot not be assumed that psychosocial treatments in secondary care will produce the same magnitude of effect in primary care³⁰, due to differences in prognostic spectrum, treatment regimen and treatment provision. Effective CBT for chronic fatigue syndrome in secondary care consisted of approximately 16 sessions by skilled and supervised specialists.¹⁰ In primary care however, more treatment sessions or more training²⁶ seem no feasible option.

Even under ideal circumstances, our research GPs were not able to deliver CBT effectively. In our opinion, it is unlikely that GPs in routine practice would be more successful in delivering a complex psychosocial treatment such as CBT.

Acknowledgments

This research project was funded by the Health Research and Development Council (ZorgOnderzoek Nederland), The Netherlands (grant no. 2830180). The recruitment of patients was made possible by the donation of services by 'Arbo Unie Limburg' (Occupational Health Service Limburg). The authors wish to thank Dr. Ludovic van Amelsvoort for statistical advice.

References

1. Lewis G, Wessely S. The epidemiology of fatigue: more questions than answers. *J Epidemiol Community Health* 1992; 46:92-7.
2. Pawlikowska T, Chalder T, Hirsch S, Wallace P. Population based study of fatigue and psychological distress. *BMJ* 1994; 308:763-6.
3. Bültmann U, Kant IJ, Kasl SV, Beurskens AJHM, Van den Brandt PA. Fatigue and psychological distress in the working population: psychometrics, prevalence and correlates. *J Psychosom Res* 2002; 52:445-52.
4. Sharpe M, Wilks D. ABC of psychological medicine: Fatigue. *BMJ* 2002; 325:480-3.
5. Wessely S, Nimnuan C, Sharpe M. Functional somatic syndromes: one or many? *Lancet* 1999; 354:936-9.
6. Mayou R, Farmer A. ABC of psychological medicine: Functional somatic symptoms and syndromes. *BMJ* 2002; 325:265-8.
7. Janssen N, Kant IJ, Swaen GMH, Janssen PPM, Schröer CAP. Fatigue as a predictor of sickness absence: results from the Maastricht cohort study on fatigue at work. *Occup Environ Med* 2003; 60 (Suppl 1):i71-i6.
8. Amelsvoort van LGPM, Kant IJ, Beurskens AJHM, Schröer CAP, Swaen GMH. Fatigue as a predictor of work disability. *Occup Environ Med* 2002; 59:712-3.
9. Whiting P, Bagnall AM, Sowden AJ, Cornell JE, Mulrow CD, Ramirez G. Interventions for the treatment and management of chronic fatigue syndrome: a systematic review. *JAMA* 2001; 286:1360-8.
10. Prins JB, Bleijenberg G, Bazelmans E, et al. Cognitive behaviour therapy for chronic fatigue syndrome: a multicentre randomised controlled trial. *Lancet* 2001; 357:841-7.
11. Zelen M. Randomised consent designs for clinical trials: an update. *Stat in Med* 1990; 9:645-56.
12. Knottnerus JA. Gezondheidszorgonderzoek in extramurale settings (Health care research in extramural settings). In: Ten Have HAMJ, Blijham G, Engberts DP, eds. *Ethiek en recht in de gezondheidszorg (Ethics and law in health care)*. Deventer: Kluwer, 1997:151-98.
13. Scott J, Sensky T. Methodological aspects of randomized controlled trials of psychotherapy in primary care. *Psychol Med* 2003; 33:191-6.
14. Beurskens AJHM, Bültmann U, Kant IJ, Vercoulen JHMM, Bleijenberg G, Swaen GMH. Fatigue among working people: validity of a questionnaire measure. *Occup Environ Med* 2000; 57:353-7.
15. Vercoulen JHMM, Alberts M, Bleijenberg G. De Checklist Individual Strength (CIS). *Gedragstherapie* 1999; 32:31-6.
16. Feinstein AR. *Clinimetrics*. New Haven, Conn.: Yale University Press, 1987:91-103.
17. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; 30:473-83.
18. Derogatis LR, Lipman RS, Covi L. SCL-90: an outpatient psychiatric rating scale—preliminary report. *Psychopharmacol Bull* 1973; 9:13-28.
19. Huibers MJH, Beurskens AJHM, Prins JB, et al. Fatigue, burnout and chronic fatigue syndrome among employees on sick leave: do attributions make the difference? *Occup Environ Med* 2003; 60 (Suppl 1):i26-i31.
20. Vercoulen JHMM, Swanink CM, Fennis JF, Galama JM, Van der Meer JW, Bleijenberg G. Prognosis in chronic fatigue syndrome: a prospective study on the natural course. *J Neurol Neurosurg Psychiatry* 1996; 60:489-94.
21. Van der Werf SP, Prins JB, Vercoulen JH, Van der Meer JW, Bleijenberg G. Identifying physical activity patterns in chronic

- fatigue syndrome using actigraphic assessment. J Psychosom Res* 2000; 49:373-9.
22. Kant IJ, Bültmann U, Schroer CAP, Beurskens AJHM, Van Amelsvoort LGPM, Swaen GMH. An epidemiological approach to study fatigue in the working population: The Maastricht Cohort Study. *Occup Environ Med* 2003; 60 (Suppl 1):i32-i9.
 23. Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JC, 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.
 24. Whitehead L, Campion P. Can general practitioners manage chronic fatigue syndrome? A controlled trial. *J Chron Fatigue Syn* 2002; 10:55-64.
 25. Ridsdale L, Godfrey E, Chalder T, et al. Chronic fatigue in general practice: is counselling as good as cognitive behaviour therapy? A UK randomised trial. *Br J Gen Pract* 2001; 51:19-24.
 26. King M, Davidson O, Taylor F, Haines A, Sharp D, Turner R. Effectiveness of teaching general practitioners skills in brief cognitive behaviour therapy to treat patients with depression: randomised controlled trial. *BMJ* 2002; 324:947-53.
 27. Huibers MJH, Beurskens AJHM, Bleijenberg G, Schayck van CP. The effectiveness of psychosocial interventions delivered by general practitioners: a Cochrane Review. *The Cochrane Library*. Vol. Issue 2. Oxford: Update Software, 2003.
 28. Mynors Wallis LM, Gath DH, Lloyd Thomas AR, Tomlinson D. Randomised controlled trial comparing problem solving treatment with amitriptyline and placebo for major depression in primary care. *BMJ* 1995; 310:441-5.
 29. Mynors-Wallis LM, Gath DH, Day A, Baker F. Randomised controlled trial of problem solving treatment, antidepressant medication, and combined treatment for major depression in primary care. *BMJ* 2000; 320:26-30.
 30. Raine R, Haines A, Sensky T, Hutchings A, Larkin K, Black N. Systematic review of mental health interventions for patients with common somatic symptoms: can research evidence from secondary care be extrapolated to primary care? *BMJ* 2002; 325:1082-93.

Why did cognitive behaviour therapy by general practitioners for fatigue not work? A qualitative analysis of the delivery of the intervention

Ellen Bazelmans ¹

Marcus J.H. Huibers ²

Gijs Bleijenberg ¹

SUBMITTED

¹ DEPARTMENT OF MEDICAL PSYCHOLOGY, UMC ST RADBOUD NIJMEGEN, THE NETHERLANDS

² DEPARTMENTS OF EPIDEMIOLOGY AND GENERAL PRACTICE, MAASTRICHT UNIVERSITY, THE NETHERLANDS

Abstract

BACKGROUND

Five to ten percent of the patients in primary care present with fatigue as their main complaint. We conducted a randomised controlled trial to assess the efficacy of cognitive behaviour therapy (CBT) delivered by general practitioners (GPs) to severely fatigued employees on long-term sick leave. The CBT intervention had no substantial effect on clinical outcomes. In the present study, we aim to describe the treatment protocol and present a qualitative analysis of the delivery of the intervention.

METHODS

We analysed the execution of the treatment protocol in relation to outcomes in individual patients. To assess protocol execution, GPs used standardized registration forms in each treatment session. A quality check was performed to assess whether CBT was completed according to protocol. Immediately after the intervention, the progress of individual patients was assessed and the intervention was evaluated by patients and GPs.

RESULTS

Of the 71 patients who started CBT, 51 patients completed the intervention according to protocol. There were no striking differences in protocol execution between those who recovered after CBT and those who did not recover after CBT. Although there were differences in the performance and delivery of the CBT intervention, there was no association between protocol execution and treatment results of individual GPs. Despite the lack of efficacy, the intervention received a positive evaluation from both patients and GPs.

CONCLUSION

CBT completers who recovered did not receive a clearly different treatment than those who did not recover. In addition, successful GPs did not deliver a clearly different treatment than less successful GPs. The lack of efficacy can at least partly be attributed to the inadequacy of the intervention, whether it is the intervention itself or its delivery by GPs.

Introduction

In primary care, 5 to 10% of patients present with fatigue as their main complaint.¹ In most of these patients, fatigue lacks a clear somatic cause¹ and appears to be a functional symptom.^{2,3} Fatigue can best be understood as a continuum, ranging from mild complaints frequently seen in the community to severe, disabling fatigue like the chronic fatigue syndrome (CFS).⁴ When fatigue becomes severe and persistent, it may lead to long-term sick leave⁵ and work disability.⁶

Recently, we conducted a randomised controlled trial to assess the efficacy of cognitive behaviour therapy (CBT) delivered by nine general practitioners to severely fatigued employees who were on long-term sick leave.⁷ Studies in secondary care have shown that CBT by skilled therapists is effective in the treatment of chronic fatigue syndrome.^{8,9} Although it was suggested that some GPs might also provide CBT in the treatment of (chronic) fatigue¹, it was unknown whether CBT is effective in less advanced fatigue cases and whether GPs are able to deliver CBT.

CBT had no substantial effect on clinical outcomes or cognitive processes during 12 months of follow-up. Our findings could not be explained by a difference in effect among GPs, withdrawal from the CBT intervention or lack of compliance with treatment. We concluded that the lack of efficacy was likely to result from a disturbance in the interaction between the patient, the doctor and the intervention. Our study was not designed to reveal the source of the disturbance in this triad. However, it is possible to investigate whether there is any association between the recovery of individuals in our trial and characteristics of the treatment received.

In this study, we present a qualitative analysis of the delivery of the intervention. We will do this in four steps. Firstly, a description of the intervention and the treatment protocol will be given. Secondly, the execution of the treatment protocol will be compared in four subgroups of patients. Thirdly, treatment results and protocol execution of the participating GPs will be compared. Finally, we will describe how patients and GPs evaluated the intervention.

Description of the intervention

STRUCTURE OF THE INTERVENTION

The intervention consisted of five to seven 30-minute sessions of cognitive behaviour therapy (CBT) within the course of four months. The intervention was partly based on our CBT protocol for chronic fatigue syndrome.⁹ The intervention was written out in a treatment protocol. Participating GPs - none of which had previous experience with CBT - were trained in delivering the intervention in two five-hour workshops (theory, paper assignments, role playing with simulation patients) and supervised in monthly two-hour sessions throughout the trial by an experienced behaviour therapist (EB). An acquaintance meeting between GP and patient preceded the start of each intervention because patients were not treated by their own GP. In most cases, spouses of patients were invited to attend one or more sessions. Patients were free to visit their regular GP for usual care.

STEPS IN THE INTERVENTION: TREATMENT PROTOCOL

Goal and content of the intervention are briefly summarized in box 1. Principal aim of the intervention was to diminish fatigue (and other health complaints) and to establish work resumption (and other personal goals).

After explanation of the treatment process in the first session, the distinction between causal and perpetuating factors was discussed. Although complaints might have been initiated by somatic, psychological or a combination of factors, psychological factors probably perpetuate the complaints. The intervention was aimed at these perpetuating factors.

GPs were trained to use a diagnostic schedule to assess the perpetuating factors of each patient in a clinical interview. Perpetuating factors could be classified using the diagnostic schedule in three pre-specified categories. Cognitive factors were non-acceptance of fatigue, somatic attributions, lost sense of control over symptoms, having high personal standards, and fear that activity will aggravate symptoms. Behavioural factors were the amount of physical, mental and social activities and sleeping pattern. Social factors were the work environment and the amount of social support.

Once the perpetuating factors were fully assessed, interventions were tailored to fit the diagnostic schedule. Specific interventions included providing helpful cognitions by an 'accepting-fatigue' exercise; explaining the role of perpetuating factors; helping to put things in perspective; instructions to restore a balance in the activity levels; an activity program to gradually increase activities; instructions for a normal sleep pattern. If the work environment was an impeding factor, restoring relations with the work environment was discussed.

GPs encouraged patients to start developing a work resumption plan from the second session on. Patients were asked to write down work activities, the problems to be expected when starting these activities and possible solutions for these problems. In following sessions, GPs encouraged patients

to execute the work resumption plan in steps by gradually building up activities, in time and in difficulty. Once patients had actually resumed work, each step was carefully evaluated and problems encountered were discussed.

An important aim during the intervention was enhancement of self-activity. Patients were encouraged to take responsibility for their own actions. GPs tried to assist in this process by asking questions such as 'What did you learn from what we discussed this session?', 'What can you do differently now to positively influence your complaints?', and 'What exactly are you going to do differently from now on?'.

GPs handed out standard forms to patients for the self-observation of perpetuating factors, and for the registration of cognitions and behaviours to be changed. For all steps in the intervention standard forms were offered to GPs, to be used as homework assignments for patients.

BOX 1

GP-DELIVERED CBT PROGRAMME FOR FATIGUE AMONG EMPLOYEES

goal of the intervention

- diminish fatigue and other complaints
- establishing work resumption and other personal goals

steps in the intervention

- assessment of perpetuating factors on the level of:
 - cognitions (e.g. non-acceptance fatigue, lost sense of control over symptoms)
 - overt behaviour (e.g. disturbed sleeping pattern, unbalanced physical activities)
 - social factors (e.g. lack of social support, dysfunctional work environment)
- modification of identified perpetuating factors by:
 - explaining the perpetuating circle
 - setting goals for activities and other problem areas
 - providing helpful cognitions
 - planning systematic and gradual work resumption
 - planning achievement of other personal goals
 - involving the social environment

Methods of the present analysis

RANDOMISED TRIAL

Details of the randomised trial we conducted are described elsewhere.⁷ In short, we selected 151 employees with severe, medically unexplained fatigue for 4 months or more who were absent from work on sick leave for 6 to 26 weeks. Employees were recruited in collaboration with a local occupational health service. In total 76 patients were assigned to the experimental condition consisting of CBT delivered by a GP nearby their home address. The control group received usual GP care. For practical reasons, we chose to train and deploy a small number of 'research' GPs instead of a large sample of GPs who would have to treat their own patients. GPs delivering the CBT were recruited among the GP population in the Southeast of the Netherlands based on their geographical position. To become a candidate for study participation, GPs had to attend both training sessions. A further requirement was willingness to treat up to 10 patients under supervision of an experienced therapist (EB). Fifteen of the 25 GPs that we invited attended both training sessions. Based on their geographical position we selected nine GPs who delivered all CBT interventions aside their regular practice.

STUDY VARIABLES

Protocol execution

To assess the actual execution of the treatment protocol, GPs used standardized registration forms on which they registered the duration and the steps addressed in each session.

CBT completers and non-completers

Based on these registration forms and the information gathered in supervision, the CBT supervisor (EB) performed a quality check. For each patient it was assessed whether the CBT received was according to 'protocol'. If for the individual patient essential steps could not be addressed before the intervention was terminated, the intervention was considered not completed. It should be noted that these 'essential steps' were not necessarily similar for all patients. Patients who completed CBT according to protocol will be referred to as 'CBT completers'.

Recovered and non-recovered cases

Immediately following the intervention (four months after baseline), we assessed whether patients had recovered or not. Recovery was defined as having a CIS fatigue score below the cut-off of 35 (high CIS scores indicate high fatigue severity)^{10,11} in combination with self-reported work resumption.

Evaluation of the intervention

At the post-treatment follow-up (four months after baseline), we asked patients to evaluate the intervention using a short questionnaire. GPs were asked to evaluate the intervention after the entire intervention period was terminated.

Results

In total 76 fatigued employees were randomly allocated to receive CBT. Five of these patients refused the CBT offered to them instantaneously. Of the remaining 71 patients who agreed to receive treatment, 51 patients (72%) completed the CBT according to protocol. Reasons for non-completion or drop out were: work resumption/too busy ($n=6$); not satisfied with intervention ($n=5$); psychiatric complications ($n=2$); unknown reasons ($n=7$).

PROTOCOL EXECUTION

CBT completers versus non-completers

In table 1, protocol execution in those who completed CBT according to protocol (CBT completers) and in those who did not complete CBT (non-completers) is compared. As can be expected, percentages of patients exposed to steps in the intervention were lowest in the non-completers group.

TABLE 1
CHARACTERISTICS OF THE INTERVENTION RECEIVED BY CBT COMPLETERS AND NON-COMPLETERS

	PATIENTS STARTING CBT (N=71)			
	non-completers (n=18*)		CBT completers (n=51)	
<i>number and duration</i>				
mean (sd) nr of sessions	3.05 (1.23)		6.2 (1.2)	
mean (sd) session duration	35.4 (7.2)		33.8 (6.4)	
<i>steps in the intervention</i>				
<i>percentage of patients exposed and mean number of sessions</i>				
<i>in which step was addressed</i>	%	m (sd)	%	m (sd)
explaining treatment goal	100	1.9 (0.7)	100	2.1 (0.8)
explaining distinction causal and perpetuating factors	94	1.6 (0.8)	98	1.9 (1.0)
discussing treatment goal	83	1.6 (1.1)	100	2.2 (0.9)
assessment using diagnostic schedule	72	1.2 (1.0)	98	2.5 (1.0)
intervening based on diagnostic schedule	50	0.8 (0.9)	90	2.0 (1.1)
exercising to accept fatigue	61	0.8 (0.8)	74	1.4 (1.1)
making a plan for work resumption	44	0.6 (0.8)	88	2.5 (1.4)
discussing problems accompanying work resumption	38	0.6 (0.9)	96	3.2 (1.4)
handing out self-observation assignments	61	0.8 (0.9)	91	1.5 (0.8)
handing out homework assignments	16	0.2 (0.5)	86	2.3 (1.5)

CBT completers = those who completed CBT according to protocol

non-completers = those who did not complete CBT according to protocol

* two patients missing

Recovered cases versus non-recovered cases

In table 2, protocol execution in CBT completers who had recovered directly after treatment (recovered cases) and in CBT completers who had not recovered (non-recovered cases) is compared. Visual inspection of the data revealed no striking differences that might explain the different outcomes in both groups.

TABLE 2

CHARACTERISTICS OF THE INTERVENTION RECEIVED BY RECOVERED CASES AND NON-RECOVERED CASES AMONG CBT COMPLETERS

	CBT COMPLETERS (N=51)			
	recovered cases (n=12)		non-recovered cases (n=39)	
<i>number and duration</i>				
mean (sd) nr of sessions	6.3 (1.3)		6.2 (1.2)	
mean (sd) session duration	34.9 (6.2)		33.5 (6.5)	
<i>steps in the intervention</i>				
<i>percentage of patients exposed and mean number of sessions</i>				
<i>in which step was addressed</i>	%	m (sd)	%	m (sd)
explaining treatment goal	100	1.9 (0.5)	100	2.2 (0.9)
explaining distinction causal and perpetuating factors	100	1.6 (0.7)	97	2.0 (1.1)
discussing treatment goal	100	2.0 (0.4)	100	2.2 (1.0)
assessment using diagnostic schedule	100	2.4 (0.8)	97	2.5 (1.0)
intervening based on diagnostic schedule	83	1.8 (1.1)	92	2.1 (1.1)
exercising to accept fatigue	75	1.0 (0.7)	74	1.5 (1.2)
making a plan for work resumption	92	2.7 (1.2)	87	2.5 (1.5)
discussing problems accompanying work resumption	100	3.7 (1.1)	95	3.0 (1.5)
handing out self-observation assignments	83	1.3 (0.8)	92	1.5 (0.8)
handing out homework assignments	92	2.4 (1.6)	85	2.3 (1.6)

recovered cases = CBT completers who scored < 35 on the CIS fatigue severity scale and resumed work at follow-up

non-recovered cases = CBT completers who scored ≥ 35 on the CIS fatigue severity score and/or did not resume work at follow-up

TABLE 3

PERFORMANCE OF INDIVIDUAL GPs AMONG PATIENTS STARTING CBT (N=71) AND CHARACTERISTICS OF THE INTERVENTIONS DELIVERED BY INDIVIDUAL GPs TO CBT COMPLETERS (N=51)

general practitioner $\ddagger$	1	2	3	4	5	6	7	8	9
performance									
nr patients treated (%)	8 (100)	6 (100)	9 (100)	7 (100)	7 (100)	7 (100)	10 (100)	7 (100)	10 (100)
nr CBT completers (%)	8 (100)	5 (83)	5 (56)	6 (86)	5 (71)	4 (57)	7 (70)	6 (86)	5 (50)
nr recovered CBT completers (%)	3 (38)	2 (33)	3 (33)	1 (14)	1 (14)	1 (14)	1 (10)	0 (0)	0 (0)
number and duration									
mean (sd) nr of sessions	5.9 (0.8)	7.4 (1.3)	6.8 (0.8)	6.2 (0.8)	5.8 (1.3)	5.3 (0.5)	5.9 (1.4)	5.5 (1.1)	7.4 (0.9)
mean (sd) session duration	30.8 (2.2)	36.4 (4.8)	40.6 (2.5)	33.1 (3.3)	30.9 (1.3)	43.9 (4.1)	33.7 (4.2)	22.8 (2.2)	38.1 (2.5)
steps in the intervention									
percentage of patients exposed and mean number of sessions in which step was addressed	%	m	%	m	%	m	%	m	%
	(sd)	(sd)	(sd)	(sd)	(sd)	(sd)	(sd)	(sd)	(sd)
explaining treatment goal	100	2.0	100	1.8	100	2.2	100	1.3	100
	(0.0)	(0.4)	(0.4)	(0.4)	(0.8)	(0.7)	(0.5)	(0.6)	(1.1)
explaining distinction causal and perpetuating factors	100	1.4	100	1.4	100	3.3	100	1.5	100
	(0.5)	(0.4)	(0.5)	(0.5)	(1.0)	(0.0)	(1.3)	(0.8)	(0.5)
discussing treatment goal	100	2.0	100	1.8	100	2.0	100	1.5	100
	(0.8)	(0.0)	(0.8)	(0.8)	(0.6)	(0.4)	(0.6)	(1.2)	(1.0)
assessment using diagnostic schedule	100	2.9	100	1.6	100	2.8	100	1.3	100
	(0.6)	(0.5)	(0.0)	(0.0)	(0.8)	(1.6)	(0.5)	(1.2)	(0.6)
intervening based on diagnostic schedule	100	2.1	100	2.0	100	2.3	60	0.5	100
	(0.8)	(0.7)	(0.7)	(0.6)	(0.8)	(1.0)	(0.6)	(1.1)	(0.7)
exercising to accept fatigue	63	1.0	80	1.2	60	0.6	100	2.0	40
	(0.9)	(0.8)	(0.8)	(0.5)	(0.9)	(0.9)	(0.9)	(0.9)	(0.9)
making a plan for work resumption	100	2.6	100	2.2	100	4.0	100	3.8	71
	(0.7)	(1.1)	(1.1)	(1.0)	(0.8)	(1.9)	(1.0)	(1.4)	(2.4)
discussing problems accompanying work resumption	88	3.3	100	3.4	100	3.6	100	2.8	86
	(1.5)	(1.8)	(1.5)	(1.5)	(1.7)	(2.1)	(1.0)	(1.8)	(1.5)
handing out self-observation assignments	100	1.8	100	1.6	83	1.5	80	0.8	86
	(0.7)	(0.8)	(1.0)	(1.0)	(0.5)	(0.7)	(1.0)	(0.8)	(0.7)
handing out homework assignments	88	2.3	80	3.6	100	3.4	100	1.0	43
	(1.7)	(1.1)	(1.5)	(1.5)	(0.8)	(0.8)	(1.5)	(1.1)	(1.1)

CBT completers = those who completed CBT according to protocol

 $\ddagger$ ordering of GPs according to % recovered CBT completers (first order) and % CBT completers (second order)

Results of individual GPs

In table 3, the individual performance of the nine GPs and the characteristics of the interventions delivered by these GPs to CBT completers are presented. Despite differences in performance and delivery of the CBT intervention, there were no distinct associations between (patterns of) intervention characteristics and treatment results of individual GPs.

EVALUATION BY PATIENTS AND GPs

Table 4 presents the overall evaluation of the intervention by patients who started the CBT and the nine GPs who delivered the intervention. In general, both patients and GPs evaluated the intervention positively. Most patients (98%) were positive about the GP assigned to them and satisfied (90%) and completely or almost completely compliant (78%) with the intervention. GPs rated the intervention to be effective and feasible, training to be sufficient (78%) and supervision to be needed (100%).

TABLE 4
EVALUATION OF THE CBT INTERVENTION BY PATIENTS AND GPs

PATIENTS STARTING CBT (N=71)		
overall impression of GP delivering CBT	positive	98%
	neutral	2%
	negative	0%
satisfaction with intervention	very satisfied	17%
	satisfied	73%
	dissatisfied	10%
	very dissatisfied	0%
self-rated compliance with advice and instructions in intervention	completely compliant	32%
	almost completely compliant	46%
	partly compliant	15%
	hardly complaint	5%
	not compliant at all	2%
GPs DELIVERING CBT (N=9)		
overall rating of the effectiveness of the intervention (0=low, 10=high)	mean (range)	7.1 (5-8)
overall rating of the feasibility of the intervention (0=low, 10=high)	mean (range)	7.2 (6-8)
nr of 'training' patients needed before intervention can be delivered optimally	mean (range)	3.7 (2-5)
CBT training was sufficient	yes	78%
supervision meetings were necessary	yes	100%
nr of sessions was adequate	yes	89%
	no, less would be sufficient	11%
	no, more sessions needed	0%

Discussion

In the present report, we investigated the association between the effects of CBT in fatigued employees on sick leave and the execution of the treatment protocol. Overall, the protocol was reasonably well performed. 72% of the patients starting CBT completed the intervention according to protocol, a compliance rate that is acceptable in psychotherapy research.

However, we found no association between recovery after CBT and protocol execution. Furthermore, despite a broad range in individual performance (recovery rates from 0% to 38%), there was no association between the treatment results of individual GPs and characteristics of the CBT delivered by them. In general, the range of success among CBT completers was rather modest. Of the 51 patients who completed the CBT according to protocol only 12 patients (23%) had recovered. Surprisingly, the intervention received a positive evaluation from both patients and GPs.

An important limitation of the present report is that we were able to assess the quantity of the steps in the intervention but not the quality of the intervention delivered. Additional information on how well the intervention was performed might have illuminated our trial results. Another limitation pertains to the qualitative approach we took. Our analysis lacked a formal validity basis, especially in the analysis of small groups of participants. However, while conducting the analysis we found that statistical parameters did not contribute to our findings.

How can our findings be explained? One explanation would be that our CBT approach was simply too difficult to be delivered by GPs. Although GPs were trained in delivering the intervention, they might have lacked the necessary experience to deliver the intervention effectively. After all, our GPs were no psychotherapists. Furthermore, GPs stated that they would need to treat three or four 'training' patients before the intervention could be delivered optimally. However, a 'training' effect could not be detected in the subsequent treatment results of individual GPs (results not shown). The fact that some GPs performed better than others might relate to differences in 'attitude' or 'treatment style', factors that are very difficult to assess. Perhaps the intervention itself required too much tailoring for it to be effective, a complex process that is not easily mastered. Another explanation would be that we have underestimated the impact of being assigned to an unfamiliar research GP instead of the usual family doctor. Results from our evaluation however indicate that patients did not perceive this arrangement as unpleasant or disturbing.

As mentioned, we reported the lack of efficacy of CBT delivered by GPs compared to usual care in an earlier paper.⁷ A per-protocol analysis yielded no significant differences between CBT completers and the control group. We concluded that the lack of efficacy was likely to result from a disturbance in the interaction between the patient, the doctor and the intervention.

In the present report, we focused on the intervention and found that CBT completers who recovered did not receive a clearly different treatment than those who did not recover. In addition, successful GPs did not deliver a clearly different treatment than less successful GPs. Consequently, we come

to the conclusion that the lack of efficacy can at least partly be attributed to the inadequacy of the intervention, whether it is the intervention itself or the delivery by GPs.

Our findings are in line with previous studies. King et al. reported that teaching CBT skills to GPs proved ineffective.¹² In a review for the Cochrane Collaboration¹³, we found little evidence for the effectiveness of psychosocial interventions by GPs, except for a favourable effect of problem-solving treatment by a small number of experienced GPs on major depression.^{14,15}

The main finding in our efficacy study was that a small group of trained and highly motivated 'research' GPs was not able to deliver CBT effectively. We concluded that it is unlikely that GPs in routine practice would be more successful in delivering a complex psychosocial treatment such as CBT. The findings in the present report seem to underline this conclusion.

Acknowledgments

This research project was funded by the Health Research and Development Council (ZorgOnderzoek Nederland), The Netherlands (grant no. 2830180). The recruitment of patients was made possible by the donation of services by 'Arbo Unie Limburg' (Occupational Health Service Limburg).

References

1. Sharpe M, Wilks D. ABC of psychological medicine: *Fatigue*. *BMJ* 2002; 325:480-3.
2. Wessely S, Nimnuan C, Sharpe M. Functional somatic syndromes: one or many? *Lancet* 1999; 354:936-9.
3. Mayou R, Farmer A. ABC of psychological medicine: *Functional somatic symptoms and syndromes*. *BMJ* 2002; 325:265-8.
4. Lewis G, Wessely S. The epidemiology of fatigue: more questions than answers. *J Epidemiol Community Health* 1992; 46:92-7.
5. Janssen N, Kant IJ, Swaen GMH, Janssen PPM, Schröer CAP. Fatigue as a predictor of sickness absence: results from the Maastricht cohort study on fatigue at work. *Occup Environ Med* 2003; 60 (Suppl 1):i71-i6.
6. Amelsvoort van LGPM, Kant IJ, Beurskens AJHM, Schröer CAP, Swaen GMH. Fatigue as a predictor of work disability. *Occup Environ Med* 2002; 59:712-713.
7. Huibers MJH, Beurskens AJHM, Schayck van CP, et al. Efficacy of cognitive behaviour therapy by general practitioners for unexplained fatigue among employees: a randomised controlled trial. *Br J Psychiatry* 2003; in press.
8. Whiting P, Bagnall AM, Sowden AJ, Cornell JE, Mulrow CD, Ramirez G. Interventions for the treatment and management of chronic fatigue syndrome: a systematic review. *JAMA* 2001; 286:1360-8.
9. Prins JB, Bleijenberg G, Bazelmans E, et al. Cognitive behaviour therapy for chronic fatigue syndrome: a multicentre randomised controlled trial. *Lancet* 2001; 357:841-7.
10. Beurskens AJHM, Bültmann U, Kant IJ, Vercoulen JHMM, Bleijenberg G, Swaen GMH. Fatigue among working people: validity of a questionnaire measure. *Occup Environ Med* 2000; 57:353-7.
11. Vercoulen JHMM, Alberts M, Bleijenberg G. De Checklist Individual Strength (CIS). *Gedragstherapie* 1999; 32:31-36.
12. King M, Davidson O, Taylor F, Haines A, Sharp D, Turner R. Effectiveness of teaching general practitioners skills in brief cognitive behaviour therapy to treat patients with depression: randomised controlled trial. *BMJ* 2002; 324:947-3.
13. Huibers MJH, Beurskens AJHM, Bleijenberg G, Schayck van CP. The effectiveness of psychosocial interventions delivered by general practitioners: a Cochrane Review. *The Cochrane Library*. Vol. Issue 2. Oxford: Update Software, 2003.
14. Mynors Wallis LM, Gath DH, Lloyd Thomas AR, Tomlinson D. Randomised controlled trial comparing problem solving treatment with amitriptyline and placebo for major depression in primary care. *BMJ* 1995; 310:441-5.
15. Mynors-Wallis LM, Gath DH, Day A, Baker F. Randomised controlled trial of problem solving treatment, antidepressant medication, and combined treatment for major depression in primary care. *BMJ* 2000; 320:26-30.

Predictors of outcome in fatigued employees on sick leave: results from a randomised trial

Marcus J.H. Huibers ^{1,2,3}

Gijs Bleijenberg ²

Ludovic G.P.M. van Amelsvoort ¹

Anna J.H.M. Beurskens ¹

Constant P. van Schayck ³

Ellen Bazelmans ²

J. André Knottnerus ³

JOURNAL OF PSYCHOSOMATIC RESEARCH 2003, ACCEPTED PENDING REVISION

¹ DEPARTMENT OF EPIDEMIOLOGY, MAASTRICHT UNIVERSITY, THE NETHERLANDS

² DEPARTMENT OF MEDICAL PSYCHOLOGY, UMC ST RADBOUD NIJMEGEN, THE NETHERLANDS

³ DEPARTMENT OF GENERAL PRACTICE, MAASTRICHT UNIVERSITY, THE NETHERLANDS

Abstract

OBJECTIVE

The main objective of this study was to identify predictors of fatigue caseness, work resumption and chronic fatigue syndrome-like caseness in a sample of fatigued employees on sick leave.

METHODS

151 fatigued employees on sick leave, 44% of whom met research criteria for CFS at baseline, were followed up during 12 months. Measures included fatigue, health aspects, psychological problems, burnout, causal attributions and self-efficacy. Logistic regression analysis was used to determine associations between predictor variables at baseline and outcome at follow-up.

RESULTS

After 12 months, 43% of the patients were no longer fatigue cases and 62% had resumed work. Recovery from fatigue caseness was predicted by stronger psychological attributions and other perception-related factors, whereas work resumption was predicted by lower age, male sex, CFS-like caseness and less cognitive difficulties. Lower physical functioning scores predicted (the development of) CFS-like caseness.

CONCLUSION

Recovering from persistent fatigue and work resumption seem to result from different underlying processes and do not necessarily fall together. As many factors associated with outcome in fatigue reflect illness perception, prevention of persistent fatigue and CFS may partly be achieved by the modification of perception.

Introduction

Fatigue is a common complaint in the general, primary care and working population, with a reported prevalence varying from 7 to 45%.¹⁻³ In most cases, fatigue lacks a clear somatic cause⁴ and appears to be a functional symptom.^{5,6} Fatigue can best be understood as a continuum, ranging from mild complaints frequently seen in the community to severe, disabling fatigue like chronic fatigue (syndrome).¹ When fatigue becomes severe and persistent, it may lead to long-term sick leave⁷ and work disability.

In the present study, we report predictors of outcome in a sample of persistently fatigued employees absent from work on sick leave. To our knowledge, predictors of outcome have never been studied in this particular group of patients. The study sample was originally selected to assess the efficacy of cognitive behaviour therapy (CBT) by general practitioners in fatigued employees on sick leave. CBT had no detectable effect on the course and duration of fatigue and absenteeism compared to usual GP care.⁸ One explanation for this finding might be that impairment in our sample was too advanced for a short treatment in a primary care setting. At baseline, 44% of the patients already met research criteria for chronic fatigue syndrome (CFS).⁹

Previously, we studied the relation between persistent fatigue among employees and CFS and found that the two conditions share many characteristics.¹⁰ Interesting in this context is the potential role of causal attributions. Studies have shown that outcomes in CFS patients are predicted by the extent to which patients attribute their illness to a somatic or psychological cause.¹¹⁻¹⁷ In a recent study, Chalder *et al.*¹⁸ found that fatigue caseness and social disability after CBT or counselling were predicted by stronger somatic causal attributions in a sample of fatigued primary care patients, 28% of whom met research criteria for CFS.⁹

We examined the role of several factors, among which causal attributions, as predictors of outcome after one year. Outcome was defined in terms of (recovery from) fatigue caseness, work resumption and CFS-like caseness (meeting research criteria for CFS) at follow-up. As fatigued employees who meet criteria for CFS have an advanced level of impairment compared to fatigued employees who do not meet CFS criteria¹⁰, we hypothesized that different predictors for outcome would be found in CFS-like cases and non-CFS cases at baseline.

Methods

DESIGN

We conducted a prospective study as part of a randomised controlled trial. Data were collected on four occasions: at baseline assessment and post-treatment assessment (4 months) using computerized questionnaires and twice at follow-up (8 and 12 months) using postal questionnaires.

SUBJECTS

Patients were recruited in collaboration with a local occupational health service (OHS) that monitors a working population of 80,000 employees. On a monthly basis, employees monitored by the OHS who were on sick leave were sent limited study information (including a screening list) by the OHS, irrespective of the reason for sick leave, followed by a reminder two weeks later. Based on the screening lists sent back to the research team, we invited potential candidates who were willing to participate to visit the university research centre.

Inclusion criterion was severe fatigue (a score of 35 or more on the Checklist Individual Strength ^{19, 20}) for four months or more as one of the main health problems in combination with complete absenteeism from work for six weeks or more. Patients were excluded from participation if they: had medical conditions that explained fatigue; received co-interventions for fatigue; had a previously classified psychiatric disorder; or received psychological treatment. In addition, absenteeism should not be caused by problems unrelated to health (e.g. work conflict). In total 151 fatigued employees were thus selected for participation.

STUDY VARIABLES

Checklist Individual Strength (CIS)

Fatigue severity was measured with the subscale fatigue severity of the Checklist Individual Strength (CIS) ^{19, 20}. Higher scores indicate a higher severity of fatigue, a score of 35 or higher is indicative of severe fatigue. ^{19, 21}

Short Form Health Survey (SF-36)

Three subscales of the Short Form Health Survey (SF-36) ²² were used: physical functioning, pain and general health perception. Higher scores on these scales indicate higher levels of physical functioning and pain and a better perception of general health.

Symptom Checklist 90 (SCL-90)

Three subscales of the Symptom Checklist 90 (SCL-90) ²³ were used: depression, cognitive difficulties and somatisation. Higher scores on these scales indicate higher levels of depression, cognitive difficulties and somatisation.

Maslach Burnout Inventory – General Survey (MBI-GS)

Aspects of burnout were measured with two subscales of the Maslach Burnout Scale – General Survey (MBI-GS) ²⁴: exhaustion and professional efficacy. Higher scores on these subscales indicate higher levels of work-related emotional exhaustion and (perceived) professional efficacy.

Self-efficacy Scale and Causal Attributions List

To measure self-efficacy (sense of control in relation to complaints), psychological attributions (beliefs regarding the psychological cause of complaints) and somatic attributions (beliefs regarding the somatic cause of complaints), two scales administered in our previous fatigue studies ^{10,16,25} were used: the self-efficacy scale (SES) and a modified version of the Causal Attributions List (CAL). The CAL consists of two subscales, psychological attributions and somatic attributions. Higher scores on these (sub)scales indicate higher levels of self-efficacy and stronger causal attributions.

Other measures

Other variables included self-reported duration of fatigue complaints and self-reported duration of absenteeism, both assessed at baseline.

DEFINITION OF OUTCOME

We defined two types of *recovery* based on fatigue caseness and work resumption. Recovery A (no fatigue caseness) was defined as having a CIS score *lower* than 35. Recovery B (work resumption) was defined as self-reported work resumption.

In addition, subjects were identified as *CFS-like cases* if they met all of the following research criteria: a CIS score of 40 or higher ²⁵, a duration of fatigue complaints of six months or more and a SF-36 score on the subscale physical functioning of 60 or lower. These research criteria were based on the CDC criteria for CFS. ⁹

It is emphasized that participants who met CFS criteria did not necessarily qualify as CFS patients: a final diagnosis of CFS can only be made by a physician after a sufficient physical examination. None of the fatigued employees had a CFS diagnosis made by a physician at entry in the study. Therefore, those who met research criteria will be referred to as CFS-like cases (the suffix “like” was used to clarify that participants met criteria by self-report only ⁹). A CFS-like status is widely regarded as a good proxy for true CFS. ^{26,27}

STATISTICAL ANALYSIS

Analyses were conducted for the entire sample and for the sample stratified according to CFS-like caseness at baseline. To test whether differences in proportions were significant, the chi-square test was used.

Multiple logistic regression models were used to determine associations between potential predictors at baseline and the two types of recovery or CFS-like caseness at 12-month follow-up. Predictors were chosen based on their univariate association with the dependent variable. Treatment was not associated with outcome.⁸ However, we controlled for it by including 'group allocation' (experimental CBT=1, control=0) in all regression models, although the effects of this control were only marginal. In a first step, predictors were entered in the model and eliminated in a backwards procedure until the model consisted of significant predictors only ($p\text{-value} < 0.05$). In a second step, sex, age, education, and group allocation were entered as covariates and the backwards elimination procedure was continued.

An explanatory regression analysis using Generalized Estimating Equations (GEE) was performed to identify factors that were cross-sectionally associated with CFS-like caseness (yes=1, no=0), which enabled us to use all data available at each measurement (0, 4, 8 and 12 months) most efficiently. Odds ratios were standardized to represent change per *SD* on the scale. Logistic regression analysis was performed using SPSS (version 11.0). Regression analysis using GEE was performed using the GENMOD procedure in SAS, with the correlation matrix defined as 'unstructured'.

Results

CHARACTERISTICS OF THE STUDY SAMPLE

In table 1, demographic and clinical variables that were entered in multivariate models are presented. 66 patients (44%) met the research criteria for CFS at baseline. Data were available for 146 patients at 4 months, 143 patients at 8 months and 138 patients at 12 months.

TABLE 1
CHARACTERISTICS OF FATIGUED EMPLOYEES AT BASELINE

	FATIGUED EMPLOYEES (N=151)
<i>demographic variables</i>	
sex (m/f)*	68/83 (45%)
age in years	43.4 (8.3)
education (1=low to 7=high)	3.8 (1.6)
<i>clinical variables (scale)</i>	
CIS, fatigue severity (8-56)	48.7 (6.2)
duration of fatigue complaints in months	27.4 (29.5)
duration of absenteeism in weeks	12.3 (4.9)
SF-36, physical functioning (0-100)	59.9 (25.3)
SF-36, pain (0-100)	66.4 (30.6)
SF-36, general health perception (0-100)	51.9 (17.9)
SCL-90, depression (16-80)	36.6 (11.7)
SCL-90, cognitive difficulties (9-45)	23.7 (7.8)
SCL-90, somatisation (12-60)	28.2 (7.9)
MBI-GS, exhaustion (0-6)	3.4 (1.5)
MBI-GS, professional efficacy (0-6)	4.0 (1.1)
SES, self-efficacy (5-24)	15.2 (3.5)
CAL, somatic attributions (4-16)	9.3 (2.5)
CAL, psychological attributions (7-28)	18.9 (4.3)
CFS-like cases*	66 (44%)

Data are mean (SD) or *number of patients (%).

CFS-like cases= cases who met research criteria for CFS at baseline.

RECOVERY AND CFS-LIKE CASENESS

Recovered cases and CFS-like caseness among the patients in the course of 12 months are presented in table 2. After 12 months, 43% of the participants were no longer fatigue cases and 62% had resumed work. The percentage of CFS-like cases dropped significantly from 44% at baseline to 20% at 12-month

follow-up ($\chi^2=9.98$, $p=0.002$). In total 60 participants (40%) did not meet CFS criteria at any point in time, while only 11 CFS-like cases at baseline (7%) remained CFS-like cases throughout the 12-month follow-up (not in table). Overall, CFS-like employees at baseline (CFS+) had significantly worse outcomes in the course of 12 months compared to non-CFS employees (CFS-) (p -values for differences in proportion in table 2 ranging from 0.019 to <0.000).

TABLE 2

RECOVERED CASES AND CFS-LIKE CASES AMONG FATIGUED EMPLOYEES (N=151) IN THE COURSE OF 12 MONTHS

follow-up	DEFINITION OF OUTCOME		
	recovery A (no fatigue caseness)	recovery B (work resumption)	CFS-like caseness (CFS research criteria)
<i>4 months</i>			
total group (n=146)	50 (34%)	81 (55%)	46 (31%)
CFS+ (n=63)	13 (21%)	28 (44%)	33 (52%)
CFS- (n=83)	37 (45%)	53 (64%)	13 (16%)
<i>8 months</i>			
total group (n=143)	56 (37%)	87 (61%)	38 (27%)
CFS+ (n=62)	15 (24%)	26 (42%)	24 (39%)
CFS- (n=81)	41 (51%)	61 (75%)	14 (17%)
<i>12 months</i>			
total group (n=138)	63 (43%)	85 (62%)	28 (20%)
CFS+ (n=57)	18 (32%)	27 (47%)	19 (33%)
CFS- (n=81)	45 (56%)	58 (72%)	9 (11%)

All values are number of valid cases (%) available at follow-up.

Recovery A = number of cases who were no longer fatigue cases (CIS score below 35).

Recovery B = number of cases who had resumed work.

CFS-like caseness = number of cases who met research criteria for CFS at follow-up.

CFS+ = fatigued employees who met research criteria for CFS at baseline.

CFS- = fatigued employees who did not meet research criteria for CFS at baseline.

PREDICTORS OF RECOVERY

Table 3 presents factors at baseline that were significantly associated with the two types of recovery after 12 months, for the entire sample (total group) and stratified according to CFS-like caseness at baseline (CFS+ and CFS-). It should be noted that odds ratios of continuous variables were standardized to represent change per SD on the scale.

Recovery A (no fatigue caseness) and recovery B (work resumption) were associated with different predictors at baseline. Whereas recovery A in the total group was predicted by psychological attributions, pain, general health perception, professional efficacy and cognitive difficulties, recovery B was predicted by sex, age, CFS-like caseness at baseline and cognitive difficulties.

TABLE 3
SIGNIFICANT BASELINE PREDICTORS OF RECOVERY AT 12-MONTH FOLLOW-UP

predictors	RECOVERY A (NO FATIGUE CASENESS) ¹		RECOVERY B (WORK RESUMPTION) ²	
	total group	CFS+ CFS-	total group	CFS+ CFS-
sex (male=1)	-	-	3.25 (1.34 to 7.86)	4.31 (1.07 to 17.39)
age	-	-	0.41 (0.25 to 0.65)	0.50 (0.25 to 0.84)
psychological attributions	1.62 (1.04 to 2.43)	1.75 (1.05 to 3.08)	-	0.34 (0.17 to 0.71)
self-efficacy	-	2.50 (1.19 to 5.18)	-	-
pain (SF-36)	0.54 (0.40 to 0.74)	0.40 (0.22 to 0.74)	-	-
general health perception (SF-36)	1.69 (1.19 to 2.39)	-	-	0.39 (0.22 to 0.83)
professional efficacy (MBI)	0.61 (0.39 to 0.94)	-	-	-
cognitive difficulties (SCL)	0.57 (0.37 to 0.92)	-	0.51 (0.33 to 0.78)	0.47 (0.28 to 0.85)
somatisation (SCL)	-	0.33 (0.17 to 0.66)	-	0.36 (0.17 to 0.78)
CFS-like caseness (yes=1)	-	-	0.32 (0.14 to 0.72)	-
Nagelkerke R ² (model)	0.31	0.30	0.29	0.32

Predictor values are significant odds ratios (95% CI) adjusted for all other predictors in the model. Values of continuous predictors are expressed as OR per SD.

CFS+ = fatigued employees who met research criteria for CFS at baseline.

CFS- = fatigued employees who did not meet research criteria for CFS at baseline.

¹ Full model included self-efficacy, somatic attributions, psychological attributions, fatigue severity, duration of fatigue complaints, physical functioning, CFS-like caseness, pain, general health perception, somatisation, cognitive difficulties and professional efficacy (step 1), age, sex, education and allocated condition (step 2).

² Full model included duration of absenteeism, physical functioning, CFS-like caseness, pain, general health perception, depression, somatisation, cognitive difficulties and emotional exhaustion (step 1), age, sex, education and allocated condition (step 2).

As hypothesized, predictors of the two types of recovery were different for CFS-like employees (CFS+) and non-CFS employees (CFS-) at baseline. Recovery A (no fatigue caseness) for example was predicted by higher scores on self-efficacy and lower scores on the SF-36 pain scale in CFS-like employees, while higher scores on psychological attributions and lower scores on somatisation were predictive of recovery in non-CFS employees.

PREDICTORS OF CFS-LIKE CASENESS

In logistic regression analysis, lower baseline scores on physical functioning were found to predict CFS-like caseness 12 months later, in the total group (OR per SD 0.27, 95%CI: 0.15 to 0.46) and in the subgroup of non-CFS employees at baseline (OR per SD 0.15, 95%CI: 0.06 to 0.46). In CFS-like employees at baseline however, CFS-like caseness at follow-up was predicted by higher fatigue scores (OR per SD 2.25, 95%CI: 1.06 to 5.31) and higher pain scores (OR per SD 2.56, 95%CI: 1.37 to 4.76).

DIFFERENCES BETWEEN CFS-LIKE EMPLOYEES AND NON-CFS EMPLOYEES IN TIME

An explanatory GEE regression analysis was used to identify factors that were associated with CFS-like caseness in the course of 12 months. Since CFS-like caseness is based on fatigue (CIS), physical functioning (SF-36) and duration of fatigue complaints, these variables were left out of the analysis. Nine factors were significantly associated with CFS-like caseness (OR per SD): sex (male vs female, OR 0.48), psychological attributions (OR 0.57), somatic attributions (OR 1.61), self-efficacy (OR 0.71), pain (OR 1.88), general health perception (OR 0.57), somatisation (1.58), and work-related exhaustion (OR 1.91). A decreasing trend over time was also found (OR 3.4, 1.94 and 1.40 on respectively 0, 4 and 8 months versus 12 months).

Discussion

PREDICTORS OF RECOVERY

We found that a substantial portion of the fatigued employees in our sample did not recover in the course of 12 months, which indicates the advanced level of impairment in this group of patients. Subsequently, we assessed predictors of recovery. We found that recovery from fatigue caseness (recovery A) was predicted by stronger psychological attributions and other perception-related factors like a better perception of health and lower professional efficacy. Work resumption on the other hand (recovery B) was predicted by male sex, lower age, CFS-like caseness and lower levels of cognitive difficulties. Surprisingly, selection variables like fatigue severity, duration of complaints and duration of absenteeism were no longer associated with recovery after adjustment for other determinants.

CFS-LIKE CASENESS

We examined the role of CFS-like caseness, both as determinant and outcome, and found that the proportion of patients who met research criteria for CFS decreased significantly in time, with 67% of the CFS-like cases at baseline no longer meeting criteria for CFS one year later. This illustrates that CFS-like caseness was not a stable condition for all who met the criteria at some point. CFS-like caseness as baseline predictor was univariately associated with worse outcomes (i.e. no recovery, CFS-like caseness at follow-up) in time. In a GEE regression analysis, causal attributions and other perception-related factors were found to differentiate CFS-like and non-CFS cases in the course of 12 months. This finding suggests that fatigued employees who met CFS criteria at some point and those who did not also differed on important characteristics other than CFS criteria. Our hypothesis that predictors of outcome would be different for CFS-like patients and non-CFS patients at baseline was confirmed. In the CFS-like subgroup at baseline, perpetuation of, or reappearance of, CFS-like caseness one year later was predicted by higher levels of fatigue and pain, while becoming a CFS-like case after an initial non-CFS status at baseline was predicted by lower levels of physical functioning.

METHODOLOGICAL CONSIDERATIONS

Our prospective study offered us the unique opportunity to observe the development of CFS in a sample of fatigued employees. As stated, a CFS-like status is regarded as a good proxy for true CFS.²⁶ Nevertheless, caution should be exercised in the interpretation of our findings relating to differences between CFS-like and non-CFS.

Another limitation pertains to the generalization of findings to other populations. We investigated a sample of fatigued employees who chose to participate in a clinical trial with rather stringent entry

criteria. Accordingly, key characteristics of the sample might not be representative for the entire spectrum of persons suffering from fatigue due to selection bias, although our findings are consistent with findings from other studies of fatigue.

THE ROLE OF CAUSAL ATTRIBUTIONS

Previous studies have provided ample evidence that most CFS patients do not recover spontaneously^{13, 16, 17, 28-30}, with stronger somatic attributions or weaker psychological attributions predicting worse outcomes.¹¹⁻¹⁷ In the present study, recovery in the subgroup of CFS-like cases at baseline was significantly worse compared to the subgroup of non-CFS cases while stronger somatic and weaker psychological attributions differentiated CFS-like cases from non-CFS cases. In addition, stronger psychological attributions were found to predict recovery from fatigue caseness in non-CFS cases. This raises questions on the role of causal attributions: are they modified by the nature, persistence and duration of complaints or do they determine illness outcome? ¹⁰ Butler *et al.*³¹ have suggested that the tendency to make somatic attributions might be a vulnerability factor for the development of CFS. Chalder *et al.*¹⁸ have suggested that a more pessimistic view of the illness may encourage symptom focusing, which in turn may lead to the perpetuation of fatigue. Future cohort studies free from selection bias may clarify causal relationships of this kind.

CONCLUSIONS

Several conclusions can be drawn from the findings in this study. First, recovering from persistent fatigue and going back to work seem to entangle different underlying processes. Work absenteeism or resumption is a complex phenomenon that is determined by a wide variety of demographic, social, work-related, organisational and individual factors.³²⁻³⁴ Consequently, recovering from health complaints and resuming work after a period of sickness absence do not necessarily fall together. Furthermore, our findings suggest that CFS and persistent fatigue, although different, are no distinct categories but rather stages of a fatigue severity continuum¹ on which people can actually shift back and forth.

Thirdly, our findings fit the notion that many key factors associated with persistent fatigue and CFS reflect the perception of illness or condition (e.g. causal attributions, health perception, self-efficacy, professional efficacy). Possibly, the prevention of persistent fatigue and CFS can partly be achieved by the modification of cognitions and perception, an approach that has shown to be effective in the treatment of CFS.^{25, 35}

Acknowledgments

This research project was funded by the Health Research and Development Council (ZorgOnderzoek Nederland), The Netherlands (grant nr. 2830180). The recruitment of patients was made possible by the donation of services by 'Arbo Unie Limburg' (Occupational Health Service Limburg).

References

- Lewis G, Wessely S. The epidemiology of fatigue: more questions than answers. *J Epidemiol Community Health* 1992; 46:92-7.
- Pawlikowska T, Chalder T, Hirsch S, Wallace P. Population based study of fatigue and psychological distress. *BMJ* 1994; 308:763-6.
- Bültmann U, Kant IJ, Kasl SV, Beurskens AJHM, Van den Brandt PA. Fatigue and psychological distress in the working population: psychometrics, prevalence and correlates. *J Psychosom Res* 2002; 52:445-452.
- Sharpe M, Wilks D. ABC of psychological medicine: Fatigue. *BMJ* 2002; 325:480-3.
- Wessely S, Nimnuan C, Sharpe M. Functional somatic syndromes: one or many? *Lancet* 1999; 354:936-9.
- Mayou R, Farmer A. ABC of psychological medicine: Functional somatic symptoms and syndromes. *BMJ* 2002; 325:265-8.
- Janssen N, Kant IJ, Swaen GMH, Janssen PPM, Schröer CAP. Fatigue as a predictor of sickness absence: results from the Maastricht cohort study on fatigue at work. *Occup Environ Med* 2003; 60 (Suppl 1):i71-i6.
- Huibers MJH, Beurskens AJHM, Schayck van CP, et al. Efficacy of cognitive behaviour therapy by general practitioners for unexplained fatigue among employees: a randomised controlled trial. *Br J Psychiatry* 2003; in press.
- 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.
- Huibers MJH, Beurskens AJHM, Prins JB, et al. Fatigue, burnout and chronic fatigue syndrome among employees on sick leave: do attributions make the difference? *Occup Environ Med* 2003; 60 (Suppl 1):i26-i31.
- Sharpe M, Hawton K, Seagroatt V, Pasvol G. Follow up of patients presenting with fatigue to an infectious diseases clinic. *BMJ* 1992; 305:147-52.
- Cope H, David A, Pelosi A, Mann A. Predictors of chronic "postviral" fatigue. *Lancet* 1994; 344:864-8.
- Wilson A, Hickie I, Lloyd A. Longitudinal study of outcome of chronic fatigue syndrome. *BMJ* 1994; 308:756-9.
- Clark MR, Katon W, Russo J, Kith P, Sintay M, Buchwald D. Chronic fatigue: risk factors for symptom persistence in a 2 1/2-year follow-up study. *Am J Med* 1995; 98:187-95.
- Chalder T, Power MJ, Wessely S. Chronic fatigue in the community: 'a question of attribution'. *Psychol Med* 1996; 26:791-800.
- Vercoulen JHMM, Swanink CM, Fennis JF, Galama JM, Van der Meer JW, Bleijenberg G. Prognosis in chronic fatigue syndrome: a prospective study on the natural course. *J Neurol Neurosurg Psychiatry* 1996; 60:489-94.
- Van der Werf SP, De Vree B, Alberts M, Van der Meer JWM, Bleijenberg G. Natural course and predicting self-reported improvement in patients with chronic fatigue syndrome with a relatively short illness duration. *J Psychosom Res* 2002; 53:749-53.
- Chalder T, Godfrey E, Ridsdale L, King M, Wessely S. Predictors of outcome in a fatigued population in primary care following a randomized controlled trial. *Psychol Med* 2003; 33:283-7.
- Vercoulen JHMM, Alberts M, Bleijenberg G. De Checklist Individual Strength (CIS). *Gedragstherapie* 1999; 32:31-6.
- Beurskens AJHM, Bültmann U, Kant I, Vercoulen JHMM, Bleijenberg G, Swaen GMH. Fatigue among working people: validity of a questionnaire measure. *Occup Environ Med* 2000; 57:353-7.
- Vercoulen JHMM, Swanink CM, Fennis JF, Galama JM, Van der Meer JW, Bleijenberg G. Dimensional assessment of chronic fatigue syndrome. *J Psychosom Res* 1994; 38:383-92.

22. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; 30:473-83.
23. Derogatis LR, Lipman RS, Covi L. SCL-90: an outpatient psychiatric rating scale—preliminary report. *Psychopharmacol Bull* 1973; 9:13-28.
24. Schaufeli WB, Leiter MP, Maslach C, Jackson SE. The Maslach Burnout Inventory-General Survey. In: Maslach C, Jackson SE, eds. *Maslach Burnout Inventory Manual*. Palo Alto, CA: Consulting Psychologists Press, 1996.
25. Prins JB, Bleijenberg G, Bazelmans E, et al. Cognitive behaviour therapy for chronic fatigue syndrome: a multicentre randomised controlled trial. *Lancet* 2001; 357:841-7.
26. Taylor RR, Jason LA, Torres A. Fatigue rating scales: an empirical comparison. *Psychol Med* 2000; 30:849-56.
27. Jason LA, Taylor RR. Applying cluster analysis to define a typology of chronic fatigue syndrome in a medically-evaluated, random community sample. *Psychol Health* 2002; 17:323-37.
28. Peterson PK, Schenck CH, Sherman R. Chronic fatigue syndrome in Minnesota. *Minn Med* 1991; 74:21-6.
29. Tirelli U, Marotta G, Improta S, Pinto A. Immunological abnormalities in patients with chronic fatigue syndrome. *Scand J Immunol* 1994; 40:601-8.
30. Bombardier CH, Buchwald D. Outcome and prognosis of patients with chronic fatigue vs chronic fatigue syndrome. *Arch Intern Med* 1995; 155:2105-10.
31. Butler JA, Chalder T, Wessely S. Causal attributions for somatic sensations in patients with chronic fatigue syndrome and their partners. *Psychol Med* 2001; 31:97-105.
32. Smulders PGW. Comments on employee absence/attendance as a dependent variable in organizational research. *J Appl Psychology* 1980; 65:368-71.
33. Kristensen TS. Sickness absence and work strain among danish slaughterhouse workers: an analysis of absence from work regarded as coping behaviour. *Soc Sci Med* 1991; 32:15-27.
34. Niedhammer I, Bugel I, Goldberg M, Leclerc A, Gueguen A. Psychosocial factors at work and sickness absence in the Gazel cohort: a prospective study. *Occup Environ Med* 1998; 55:735-41.
35. Whiting P, Bagnall AM, Sowden AJ, Cornell JE, Mulrow CD, Ramirez G. Interventions for the treatment and management of chronic fatigue syndrome: a systematic review. *JAMA* 2001; 286:1360-8.

Predicting the two-year course of unexplained fatigue and the onset of long-term sickness absence in fatigued employees: results from the Maastricht Cohort Study

Marcus J.H. Huibers ^{1, 2}

Ute Bültmann ^{1, 2}

Stanislav V. Kasl ³

Ijmert Kant ^{1, 2}

Ludovic G.P.M. van Amelsvoort ^{1, 2}

Constant P. van Schayck ²

Gerard M.H. Swaen ^{1, 2}

SUBMITTED

¹ DEPARTMENT OF EPIDEMIOLOGY, MAASTRICHT UNIVERSITY, THE NETHERLANDS

² CARE AND PUBLIC HEALTH RESEARCH INSTITUTE (CAPHRI), MAASTRICHT UNIVERSITY, THE NETHERLANDS

³ DEPARTMENT OF EPIDEMIOLOGY AND PUBLIC HEALTH, YALE UNIVERSITY SCHOOL OF MEDICINE, NEW HAVEN CT, USA

Abstract

AIM

To examine baseline predictors of the two-year course of unexplained fatigue and the onset of long-term sickness absence in a sample of severely fatigued employees.

METHODS

Analyses were based on the Maastricht Cohort Study (MCS), a prospective population-based cohort study among more than 12.000 employees. 2108 severely fatigued employees not on sick leave were selected at baseline and followed up at six time points over two years.

RESULTS

Point prevalences of severe fatigue (59 to 63%) and long-term sickness absence (1.8 to 3.1%) among participants were fairly stable at all consecutive time points. Low fatigue severity, low work-related exhaustion, low anxious mood, absence of conflicts with colleagues and good self-rated health at baseline were predictors of the onset of recovery from fatigue in survival analyses. High age, low decision authority, female sex, working in nightshift, a physical attribution of fatigue and a history of absenteeism were predictors of the onset of long-term absenteeism.

CONCLUSION

We conclude that predictors of both unexplained fatigue and long-term absenteeism displayed a broad range of factors. In accordance with previous studies, our findings underscore the notion that prevention and treatment of fatigue should be aimed at health perception and emotional well-being.

Introduction

Fatigue is a common complaint in the general, primary care and working population, with reported prevalence rates varying from 7 to 45%.¹⁻³ In most cases, fatigue lacks a clear somatic cause⁴ and appears to be a functional symptom or complaint.^{5,6} Fatigue can best be understood as a continuum, ranging from mild complaints frequently seen in the community to severe, disabling fatigue like the chronic fatigue syndrome.¹

Fatigue in the working population has been the subject of growing interest because of its high reported prevalence of 22%³ and the many work-related consequences of severe fatigue such as reduced performance, sickness absence⁷ and work disability.⁸ Because of these serious adverse consequences of severe fatigue, it is important to identify factors that determine the prognosis of fatigue.

Several prospective studies have investigated the course and outcome in fatigue.⁹⁻²⁰ However, these studies were all conducted in clinical settings among individuals seeking help and most fatigued patients were diagnosed with the chronic fatigue syndrome, a condition from which most patients do not recover spontaneously.²¹

In a recently submitted paper, we studied the prognosis of fatigued employees on long-term sick leave following a randomised controlled trial of cognitive behaviour therapy. After one year of follow-up, the majority of the study sample had not recovered from fatigue. We found that recovery from fatigue was predicted by factors related to illness perception, such as the tendency to attribute fatigue complaints to a psychological cause and good self-rated health. Interestingly, these findings are in keeping with previous studies on the outcome in chronic fatigue (syndrome) patients.⁹⁻²⁰

In the present study, we examine predictors of outcome in a large sample of severely fatigued employees who were not on sick leave. Our analysis is based on data from the Maastricht Cohort Study of Fatigue at Work, a population-based sample frame in which more than 12,000 employees were followed prospectively over three years using a broad range of measures.²² We define two research questions:

1. What are the predictors of the course of unexplained fatigue (recovery, persistence or reappearance) in employees who were fatigued at baseline?
2. What are the predictors of the onset of long-term sickness absence in employees who were fatigued at baseline?

Methods

STUDY DESIGN OF THE MAASTRICHT COHORT STUDY

We used prospective data from the Maastricht Cohort Study on Fatigue at Work (MCS). In May 1998, 26,978 employees from 45 Dutch companies and organizations were invited to participate in the MCS. 12,161 employees (45%) responded by providing informed consent and completing the postal self-report questionnaire at baseline. There were no significant differences between baseline respondents and non-respondents in non-response analyses, except that non-respondents were less likely to be fatigued or absent from work than respondents. Respondents in the MCS were followed up four-monthly for three years. Further details on the MCS are described elsewhere.²²

In addition to the self-report questionnaire data, participating companies were asked to provide data on individual sickness absence days during the follow-up period. Data were obtained by record linkages to company sickness absence and work disability registry systems. At the time of the present report, the sickness absence registry contained complete data for 548 days following baseline.

STUDY POPULATION

Participants were selected at baseline of the MCS. Inclusion criterion for our analysis was a score of 35 or higher on the subscale fatigue severity of the Checklist Individual Strength^{23, 24}, a cut-off that is indicative of severe fatigue.^{23, 25}

Fatigued employees at baseline (n=3129) were excluded if they: reported to be on (partial) sick leave (n=502); had more than one job (n=121); had changed their working hours in the past year (n=365); were pregnant (n=33); or reported a somatic condition that could explain fatigue complaints (i.e. cerebrovascular accidents, liver dysfunctions, diabetes mellitus, hyperthyroidism or cancer, n=136). In total 2108 employees were thus included in the analysis of fatigue.

Employees were selected for the analysis of company-registered absenteeism using identical in- and exclusion criteria, with the exception that the exclusion of employees on sick leave was not based on self-report but on data from the sickness absence registry. However, not all companies were able to provide sickness absence data in the appropriate electronic format, and sickness absence data were only available for 1681 employees.

STUDY VARIABLES

Demographic factors

Demographic variables included sex, age and educational level (high, medium, low) and were collected at baseline.

Fatigue-related factors

Fatigue was assessed with the four subscales of the Checklist Individual Strength (CIS) ^{23, 24}: fatigue severity; concentration; motivation; and physical activity. Higher scores on these subscales indicate a higher level of fatigue severity and lower levels of concentration, motivation and physical activity respectively. The CIS-fatigue severity scale served as an outcome measure.

Burnout was assessed with the three subscales of the Maslach Burnout Inventory-General Survey (MBI) ²⁶: exhaustion; cynicism; and professional efficacy, with higher scores indicating higher levels on these scales.

Need for recovery was assessed with the Need for Recovery scale ^{27, 28} that indicates the short-term effects of a day of work, with a higher score indicating a higher need for recovery.

Participants were asked to indicate whether they felt they had suffered from *fatigue complaints* in the previous four months (yes or no) and, if so, whether they felt the cause of their fatigue complaints was psychological (*psychological attribution*, yes or no), somatic (*physical attribution*, yes or no) or of an unknown nature (*unknown attribution*, yes or no).

Mental health factors

Psychological distress was assessed with the 12-item version of the General Health Questionnaire (GHQ) ²⁹, with a higher score indicating a higher level of psychological distress.

Depressed mood was assessed with a single item stating "did you feel down almost every day during the past two weeks, yes or no?"

Anxious mood was assessed by adding the scores of three items on anxiety (excessive worrying, compulsive behaviour, compulsive thinking), with a higher score indicating a higher level of anxious mood. The internal consistency was modest but satisfactory, with a Cronbach's alpha of 0.66.

Participants were also asked if they had experienced shocking *life events* in the past 12 months (yes or no).

Health factors

Self-rated health was assessed with one item from the Short Form Health Survey (SF-36) ³⁰, which is a widely used measure of general health status. ³¹ The original 5-point scale (1=excellent health; 2=very good health; 3=good health; 4=moderate health; 5=bad health) was dichotomised into 'good health' (1-3) and 'bad health' (4-5).

In addition, participants were asked if they: had visited their general practitioner (*visits to GP*) and/or their occupational physician (*visits to OP*) in the past 4 months because of problems at work (yes or no); suffered from *health complaints* (yes or no); suffered from one or more *chronic diseases* (yes or no); or suffered from *sleep disturbances* (yes or no). It should be noted that participants with fatigue-related chronic diseases were already excluded from the analysis.

Sickness absence

Sickness absence was an outcome measure and was assessed in two ways: as self-reported sickness absence from work and as number of sick leave days according to the sick leave registry provided by the companies. We chose 42 consecutive days of sickness absence or more as a cut-off for company-registered long-term absenteeism in line with Dutch standards. The highest category of the duration of self-reported sickness absence available in the MCS questionnaire stated 'five weeks or more'. Self-reported *history of absenteeism* in the past four months was assessed in three categories: no history of absenteeism; a recent sickness absence episode of two weeks or less; a recent sickness absence episode of more than two weeks.

Work-related factors

Psychosocial work characteristics were measured with five subscales of the Job Content Questionnaire (JCQ) ³²: psychological job demands; decision authority; skill discretion; co-worker social support; and supervisor social support, with higher scores indicating higher levels on these scales.

Participants were also asked to indicate whether they: worked in *day or nightshift*; experienced *conflicts with colleagues* (yes or no) or *conflicts with supervisor* (yes or no); and how many hours they worked each week (*working hours*, assessed in five categories: <16, 16-25, 26-35, 36-40, >40).

STATISTICAL ANALYSES

We used data collected in the course of two years (at baseline, 4, 8, 12, 16, 20 and 24 months). Number of fatigue cases (those scoring 35 or higher on the subscale fatigue severity of the CIS) and cases on *self-reported* long-term sickness absence (those being absent from work on sick leave for five weeks or more) at follow-up were calculated.

Cox proportional hazards regression models were used to determine associations between predictors at baseline and the first onset of recovery from fatigue (scoring 34 or lower on the subscale fatigue severity of the CIS) in the course of two years ($n=2108$). Predictors that were bivariately associated (p -value <0.05) with recovery were entered in a multiple survival model, together with age, sex and education, and eliminated in a backwards procedure until the model consisted of significant predictors only (p -value <0.05). Relative risks and 95% confidence intervals were calculated and standardized to represent change per SD on the scale. The proportional hazards assumption was tested using Cox regression with time-dependent covariates and was confirmed for all models.

An explanatory regression analysis using Generalized Estimating Equations (GEE) was performed to identify baseline predictors that were associated with the persistence or reappearance of fatigue (scoring 35 or higher on the subscale fatigue severity of the CIS) at follow-up (4-24 months). We used GEE to facilitate the analysis of repeated measurements of fatigue over time. Predictors that were bivariately associated with fatigue were eliminated until the model consisted of significant predic-

tors only (p -value < 0.05) and interactions with time were investigated. Odds ratios and 95% confidence intervals were calculated and standardized to represent change per SD on the scale.

Cox proportional hazards regression models were also used to determine associations between predictors at baseline and the first onset of long-term sickness absence (42 consecutive days of company-registered sickness absence or more) in the course of 548 days ($n=1681$).

Survival analyses were performed using SPSS. The GEE analysis was performed using the GENMOD procedure in SAS, with the correlation matrix defined as 'unstructured'.

Results

Table 1 presents the demographic characteristics of the study sample. The majority of the employees was male. Educational level was distributed fairly equally in the three categories which we used.

TABLE 1
DEMOGRAPHIC CHARACTERISTICS OF FATIGUED EMPLOYEES AT BASELINE

<i>characteristic</i>	FATIGUED EMPLOYEES (N=2108)
male sex, nr (%)	1574 (75%)
age in years, mean (sd)	40.6 (8.7)
educational level	
• high, nr (%)	744 (35%)
• medium, nr (%)	886 (42%)
• low, nr (%)	478 (23%)

COURSE OF FATIGUE AND LONG-TERM SICKNESS ABSENCE

Fatigue cases and cases on long-term sickness absence in the course of two years are presented in table 2. A stable percentage of fatigue cases of approximately 60% appeared at each follow-up. The percentage of persistent fatigue cases since baseline however decreased in time, from 61% after 4 months to 23% after 24 months. Number of sickness absence cases increased slightly over time, from 1.8% at 4-month follow-up to 3.1% at 24-month follow-up. As with fatigue caseness, the percentage of employees who were sickness absence cases at consecutive time points decreased rapidly, until after one year there were no cases left.

PREDICTORS OF FATIGUE

Table 3 presents predictors of the onset of recovery from fatigue and the persistence or reappearance of fatigue. Five baseline predictors were associated with the onset of recovery from fatigue in survival analyses: low fatigue severity, low work-related exhaustion, low anxious mood, absence of conflicts with colleagues and good self-rated health. These five predictors were also associated in the opposite direction with persistence or reappearance of fatigue over time, together with need for recovery, supervisor social support, self-rated suffering from fatigue complaints and the absence of a psychological attribution of fatigue complaints in an explanatory GEE regression analysis. In addition, we found an interaction between age and time: during the two-year follow-up, younger employees were less likely to be fatigue cases while older employees were more likely to be fatigue cases (figure 1).

TABLE 2

NUMBER OF FATIGUE AND LONG-TERM SICKNESS ABSENCE CASES IN THE COURSE OF TWO YEARS AMONG 2108 EMPLOYEES WHO WERE FATIGUED AT BASELINE

cases, nr (%)	FOLLOW-UP MEASUREMENTS					
	T1 (4 mths)	T2 (8 mths)	T3 (12 mths)	T4 (16 mths)	T5 (20 mths)	T6 (24 mths)
valid cases available at follow-up Tx	1841 (100%)	1769 (100%)	1647 (100%)	1516 (100%)	1438 (100%)	1332 (100%)
fatigue cases at follow-up Tx	1124 (61%)	1084 (61%)	984 (60%)	951 (63%)	849 (59%)	799 (60%)
persistent fatigue cases since baseline	1124 (61%)	779 (44%)	567 (34%)	465 (31%)	376 (26%)	300 (23%)
sickness absence cases at follow-up Tx	33 (1.8%)	34 (1.9%)	38 (2.3%)	34 (2.2%)	38 (2.6%)	41 (3.1%)
consecutive sickness absence cases since T1	33 (1.8%)	9 (0.5%)	1 (0.06%)	0	0	0

fatigue cases = employees who scored ≥ 35 on CIS fatigue severity scale

sickness absence cases = employees who reported to be absent from work on sick leave for 5 weeks or more

TABLE 3

SIGNIFICANT BASELINE PREDICTORS OF THE FIRST ONSET OF RECOVERY FROM FATIGUE (COX REGRESSION) AND THE PERSISTENCE OR REAPPEARANCE OF FATIGUE (GEE REGRESSION ANALYSIS) DURING TWO YEARS OF FOLLOW-UP IN EMPLOYEES FATIGUED AT BASELINE (N=2108)

predictors†	OUTCOME	
	recovery from fatigue RR (per SD) (95% CI)‡	persistence or reappearance of fatigue OR (per SD) (95%)‡
<i>continuous</i>		
age	-	0.96 (0.85 to 1.09)
age*time interaction	-	1.03 (1.01 to 1.06)
fatigue severity (CIS)	0.79 (0.74 to 0.84)	1.39 (1.28 to 1.51)
need for recovery	-	1.12 (1.04 to 1.22)
exhaustion (MBI)	0.85 (0.80 to 0.91)	1.24 (1.13 to 1.36)
supervisor social support (JCQ)	-	1.07 (1.01 to 1.15)
anxious mood	0.93 (0.88 to 0.98)	1.09 (1.01 to 1.17)
<i>dichotomous</i>		
self-rated fatigue complaints (yes=1)	-	1.81 (1.35 to 2.41)
conflicts with colleagues (yes=1)	0.77 (0.64 to 0.93)	1.36 (1.09 to 1.71)
psychological attribution of fatigue complaints (yes=1)	-	0.84 (0.73 to 0.97)
self-rated health (SF36) (good=1)	1.31 (1.16 to 1.48)	0.58 (0.51 to 0.67)

† Full model included age, sex, education, all four CIS-subcales, all three MBI-subcales, all five JCQ-subcales, all three attribution of fatigue variables, fatigue complaints, need for recovery, conflicts with colleagues, conflicts with supervisor, psychological distress, depressed mood, anxious mood, life events, visits to GP, visits to OP, self-rated health, health complaints, chronic diseases, sleep disturbances, working hours, dayshift, history of absenteeism.

‡ Predictor values are significant relative risks or odds ratios (95% CI) adjusted for all other predictors in the model. Values of continuous predictors are expressed as RR or OR per SD.

FIGURE 1

INTERACTION OF AGE AND TIME AS PREDICTOR OF FATIGUE CASENESS IN GEE REGRESSION ANALYSIS DURING TWO YEARS OF FOLLOW-UP

TABLE 4

SIGNIFICANT BASELINE PREDICTORS OF THE FIRST ONSET OF LONG-TERM SICKNESS ABSENCE (COX REGRESSION) DURING 548 DAYS FOLLOW-UP IN EMPLOYEES FATIGUED AT BASELINE (N=1681)

predictors [†]	OUTCOME
	long-term sickness absence (≥ 42 days) RR (per SD) (95% CI) [‡]
<i>continuous</i>	
age	1.41 (1.18 to 1.52)
decision authority (JCQ)	0.78 (0.67 to 0.92)
<i>dichotomous/categorical</i>	
sex (male=1)	0.52 (0.38 to 0.72)
dayshift (yes=1)	0.56 (0.41 to 0.77)
physical attribution of fatigue complaints (yes=1)	1.41 (1.05 to 1.89)
history of absenteeism • ≤ 2 weeks vs no previous absenteeism	2.19 (1.59 to 3.04)
• > 2 weeks vs no previous absenteeism	1.70 (1.06 to 2.74)

[†] Full model included age, sex, education, CIS-fatigue, CIS-motivation, all three MBI-subscales, JCQ-decision authority, JCQ-skill discretion, JCQ-supervisor social support, physical attribution of fatigue, fatigue complaints, psychological distress, depressed mood, anxious mood, life events, self-rated health, health complaints, chronic diseases, sleep disturbances, dayshift, history of absenteeism.

[‡] Predictor values are significant relative risks or odds ratios (95% CI) adjusted for all other predictors in the model. Values of continuous predictors are expressed as RR per SD.

PREDICTORS OF LONG-TERM SICKNESS ABSENCE

Table 4 presents predictors of the onset of long-term sickness absence. Six factors were associated with the onset of long-term sickness absence in survival analyses: high age, low decision authority, female sex, working in nightshift, a physical attribution of fatigue and a history of absenteeism in the previous four months.

Discussion

We found that unexplained fatigue was a stable and persistent condition in many fatigued employees at baseline, with 23% of the sample scoring above the cut-off for fatigue at all consecutive time points in the course of two years. The point prevalence of self-reported long-term sickness absence in this group ranged from approximately 2 to 3% at consecutive time points.

A good self-rated health was found to be an important predictor of recovery from fatigue, together with low fatigue severity, low work-related exhaustion, low anxious mood and the absence of conflicts with colleagues. Since self-rated health was such a strong predictor, we conducted an additional analysis and found that the bivariate association between self-rated health and fatigue could not be affected by controlling for other health variables (results not shown), suggesting that self-rated health is more of a subjective experience rather than a proxy for objective health status.

An explanatory analysis of the persistence or reappearance of fatigue yielded the same predictors as for recovery from fatigue, with associations in the opposite direction as expected. The additional predictors of the persistence or reappearance of fatigue we found and the differences in magnitude between predictors of recovery and persistence/reappearance indicate that recovering from fatigue entails a somewhat different process than remaining fatigued throughout time. However, these differences may very well result from the increase of power of the explanatory GEE analysis that was based on several observations per person, compared to the survival analysis that was based on a single occurrence per person.

The onset of long-term sickness absence in our sample was found to be predicted by demographic factors (i.e. high age and female sex) and work-related factors (i.e. working in nightshift, a history of absenteeism and low decision authority). Although other fatigue-related factors were *not* predictive of sickness absence, a physical attribution of fatigue complaints emerged as a strong predictor. The association of physical attribution and sickness absence was not affected by controlling for other factors in an additional analysis (results not shown), indicating that physical attribution stands on its own as an independent predictor of absenteeism.

The design of the Maastricht Cohort Study²² offered us a unique opportunity to monitor and investigate the course of fatigue and sickness absence in a large sample of employees. Dropout and loss to follow-up on the other hand are inevitable consequences of our longitudinal approach. Baseline respondents of the MCS who were no longer available for data collection at follow-up were more fatigued than those who continued to participate.²² This may have led to an underestimation of fatigue caseness at follow-up. An obvious limitation and potential source of bias is the use of self-report measures. However, a postal data collection is the only feasible option when conducting longitudinal research on this scale. In contrast, the analysis of 'objective' company-registered sick leave data free from employees' perception bias is a major advantage of the present study.

Study samples in previous longitudinal studies of fatigue are not easily compared with the sample of fatigued employees in the present study. Most studies were carried out in hospital settings and subjects were mostly patients diagnosed with the chronic fatigue syndrome. The most important difference therefore relates to selection bias: the participants in previous studies were help-seekers²⁰ and for this reason findings can not be transferred to the population in our community-based study.

However, our findings are in accordance with some previously reported findings. Studies have shown that most chronic fatigue syndrome patients do not recover spontaneously^{9, 12-14, 17, 18}, with stronger somatic attributions or weaker psychological attributions predicting worse outcome.^{10, 11, 13, 15-18, 20} A substantial portion of the present study sample was still fatigued after two years and outcome was partly predicted by self-rated health and psychological attributions of fatigue. Similarly, in a recently submitted paper we reported that most fatigued employees on long-term sickness absence had not recovered after one year, with illness perception-related factors (attributions, self-rated health) predicting the course of fatigue.

Although predictors of unexplained fatigue displayed a broad range of fatigue-, health- and work-related factors, our findings underscore the notion that prevention and treatment of fatigue in employees should be aimed at the perception of illness (or health) and emotional well-being in general, an approach that is already successful in the treatment of the chronic fatigue syndrome.^{33, 34} An interesting question for future research that follows from this is whether persistent, unexplained fatigue in employees can be a precursor of the development of the chronic fatigue syndrome.

Acknowledgments

This study was funded by the Health Research and Development Council (ZorgOnderzoek Nederland), The Netherlands (grant nr. 2830180). The Maastricht Cohort Study is part of The Netherlands concerted action on 'Fatigue at Work' granted by The Netherlands Organization for Scientific Research (NWO). The authors wish to thank Prof. J.A. Knottnerus for his useful comments on the manuscript.

References

1. Lewis G, Wessely S. *The epidemiology of fatigue: more questions than answers.* *J Epidemiol Community Health* 1992; 46:92-7.
2. Pawlikowska T, Chalder T, Hirsch S, Wallace P. *Population based study of fatigue and psychological distress.* *BMJ* 1994; 308:763-6.
3. Bültmann U, Kant IJ, Kasl SV, Beurskens AJHM, Van den Brandt PA. *Fatigue and psychological distress in the working population: psychometrics, prevalence and correlates.* *J Psychosom Res* 2002; 52:445-52.
4. Sharpe M, Wilks D. *ABC of psychological medicine: Fatigue.* *BMJ* 2002; 325:480-3.
5. Wessely S, Nimmuan C, Sharpe M. *Functional somatic syndromes: one or many?* *Lancet* 1999; 354:936-9.
6. Mayou R, Farmer A. *ABC of psychological medicine: Functional somatic symptoms and syndromes.* *BMJ* 2002; 325:265-8.
7. Janssen N, Kant IJ, Swaen GMH, Janssen PPM, Schröer CAP. *Fatigue as a predictor of sickness absence: results from the Maastricht cohort study on fatigue at work.* *Occup Environ Med* 2003; 60 (Suppl 1):i71-i6.
8. Amelsvoort van LGPM, Kant IJ, Beurskens AJHM, Schröer CAP, Swaen GMH. *Fatigue as a predictor of work disability.* *Occup Environ Med* 2002; 59:712-3.
9. Peterson PK, Schenck CH, Sherman R. *Chronic fatigue syndrome in Minnesota.* *Minn Med* 1991; 74:21-6.
10. Sharpe M, Hawton K, Seagroatt V, Pasvol G. *Follow up of patients presenting with fatigue to an infectious diseases clinic.* *BMJ* 1992; 305:147-52.
11. Cope H, David A, Pelosi A, Mann A. *Predictors of chronic "postviral" fatigue.* *Lancet* 1994; 344:864-8.
12. Tirelli U, Marotta G, Improta S, Pinto A. *Immunological abnormalities in patients with chronic fatigue syndrome.* *Scand J Immunol* 1994; 40:601-8.
13. Wilson A, Hicki I, Lloyd A. *Longitudinal study of outcome of chronic fatigue syndrome.* *BMJ* 1994; 308:756-9.
14. Bombardier CH, Buchwald D. *Outcome and prognosis of patients with chronic fatigue vs chronic fatigue syndrome.* *Arch Intern Med* 1995; 155:2105-10.
15. Clark MR, Katon W, Russo J, Kith P, Sintay M, Buchwald D. *Chronic fatigue: risk factors for symptom persistence in a 2 1/2-year follow-up study.* *Am J Med* 1995; 98:187-95.
16. Chalder T, Power MJ, Wessely S. *Chronic fatigue in the community: 'a question of attribution'.* *Psychol Med* 1996; 26:791-800.
17. Vercoulen JHMM, Swanink CM, Fennis JF, Galama JM, Van der Meer JW, Bleijenberg G. *Prognosis in chronic fatigue syndrome: a prospective study on the natural course.* *J Neurol Neurosurg Psychiatry* 1996; 60:489-94.
18. Van der Werf SP, De Vree B, Alberts M, Van der Meer JWM, Bleijenberg G. *Natural course and predicting self-reported improvement in patients with chronic fatigue syndrome with a relatively short illness duration.* *J Psychosom Res* 2002; 53:749-53.
19. Bentall RP, Powell P, Nye FJ, Edwards HT. *Predictors of response to treatment for chronic fatigue syndrome.* *Br J Psychiatry* 2002; 181:248-52.
20. Chalder T, Godfrey E, Ridsdale L, King M, Wessely S. *Predictors of outcome in a fatigued population in primary care following a randomized controlled trial.* *Psychol Med* 2003; 33:283-7.
21. Afari N, Buchwald D. *Chronic Fatigue Syndrome: a Review.* *Am J Psychiatry* 2003; 160:221-36.
22. Kant IJ, Bültmann U, Schroer CAP, Beurskens AJHM, Van Amelsvoort LGPM, Swaen GMH. *An epidemiological approach to study fatigue in the working population: The Maastricht Cohort Study.* *Occup Environ Med* 2003; 60 (Suppl 1):i32-i9.

23. Vercoulen JHMM, Swanink CM, Fennis JF, Galama JM, Van der Meer JW, Bleijenberg G. Dimensional assessment of chronic fatigue syndrome. *J Psychosom Res* 1994; 38:383-92.
24. Beurskens AJHM, Bültmann U, Kant IJ, Vercoulen JHMM, Bleijenberg G, Swaen GMH. Fatigue among working people: validity of a questionnaire measure. *Occup Environ Med* 2000; 57:353-7.
25. Vercoulen JHMM, Alberts M, Bleijenberg G. De Checklist Individual Strength (CIS). *Gedragstherapie* 1999; 32:31-36.
26. Schaufeli WB, Leiter MP, Maslach C, Jackson SE. The Maslach Burnout Inventory-General Survey. In: Maslach C, Jackson SE, eds. *Maslach Burnout Inventory Manual*. Palo Alto, CA: Consulting Psychologists Press, 1996.
27. Sluiter JK, Van der Beek AJ, Frings-Dresen MHW. The influence of work characteristics on the need for recovery and experienced health: a study on coach drivers. *Ergonomics* 1999; 42:573-83.
28. Jansen NWH, Kant IJ, Van den Brandt PA. Need for Recovery in the working population: descriptions and associations with fatigue and psychological distress. *Int J Behav Med* 2002; 9:322-40.
29. Goldberg DP, Williams P. *The user's guide to the General Health Questionnaire*. Windsor: NFER-Nelson, 1991.
30. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; 30:473-83.
31. Aaronson NK, Muller M, Cohen PD, et al. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. *J Clin Epidemiology* 1998; 51:1055-68.
32. Karasek RA. *The Job Content Questionnaire and user's guide*. Los Angeles: Department of Industrial and Systems Engineering, USC, 1985.
33. Prins JB, Bleijenberg G, Bazelmans E, et al. Cognitive behaviour therapy for chronic fatigue syndrome: a multicentre randomised controlled trial. *Lancet* 2001; 357:841-7.
34. Whiting P, Bagnall AM, Sowden AJ, Cornell JE, Mulrow CD, Ramirez G. Interventions for the treatment and management of chronic fatigue syndrome: a systematic review. *JAMA* 2001; 286:1360-8.

Fatigue, burnout and chronic fatigue syndrome among employees on sick leave: do attributions make the difference?

Marcus J.H. Huibers ^{1,2,3}

Anna J.H.M. Beurskens ¹

Judith B. Prins ²

Ijmert Kant ¹

Ellen Bazelmans ²

Constant P. van Schayck ³

J. André Knottnerus ³

Gijs Bleijenberg ²

OCCUPATIONAL & ENVIRONMENTAL MEDICINE 2003, 60(SUPPL 1);i26-i31

¹ DEPARTMENT OF EPIDEMIOLOGY, MAASTRICHT UNIVERSITY, THE NETHERLANDS

² DEPARTMENT OF MEDICAL PSYCHOLOGY, UMC NIJMEGEN, THE NETHERLANDS

³ DEPARTMENT OF GENERAL PRACTICE, MAASTRICHT UNIVERSITY, THE NETHERLANDS

Abstract

BACKGROUND

Persistent fatigue among employees, burnout and chronic fatigue syndrome (CFS) are three fatigue conditions that share some characteristics in theory. However, these conditions have not been compared in empirical research, despite conceptual similarities.

METHODS

This cross-sectional study aimed to investigate relations between persistent fatigue, burnout and CFS by describing the clinical features of a sample of 151 fatigued employees on sick leave. Using validated instruments, subgroups based on research criteria for CFS and burnout within the sample of fatigued employees and a reference group of 97 diagnosed CFS patients were compared. Analyses of covariance were performed.

RESULTS

66 (43.7%) fatigued employees met research criteria for CFS (except symptom criteria) and 76 (50.3%) met research criteria for burnout. 'CFS-like employees' (fatigued employees who met CFS criteria) reported stronger somatic attributions than 'non-CFS-like employees'. Burned-out CFS-like employees were more depressed and distressed than CFS-like employees who were not burned-out. Burnout cases among the non-CFS-like employees had stronger psychological attributions than fatigued employees who were not burned-out. Compared to diagnosed CFS patients, CFS-like employees merely had a shorter duration of fatigue complaints. Burned-out CFS-like employees had stronger psychological attributions and were more distressed than CFS patients.

CONCLUSIONS

Fatigued employees shared many important characteristics with CFS patients, regardless of burnout status, and many fatigued employees met CFS criteria and/or burnout criteria. Differences however concerned the causal attributions that were made. This raises questions about the role of causal attributions: are they modified by fatigue complaints or do they determine illness outcome?

Introduction

Fatigue is a common complaint in the general and working population, with a reported prevalence varying from 7 to 45%.^{1,2,3} Fatigue can best be understood as a continuum, ranging from mild complaints frequently seen in the community on the one hand to severe, disabling fatigue, like chronic fatigue syndrome on the other.¹ When fatigue among employees becomes severe and persistent, it may lead to long-term sick leave and work disability.⁴ Conceptually linked with fatigue and absenteeism is the phenomenon of burnout. In general, burnout can be described as a persistent, negative, work-related state of mind characterised by work-related emotional exhaustion and accompanied by distress, (perceived) reduced effectiveness, decreased motivation and dysfunctional attitudes and behaviours at work.⁵ Burnout symptoms are mostly psychological and burned-out workers often causally attribute their complaints to problems at work, blaming their jobs for their condition.⁵ A conservative estimate of the prevalence of 'clinical' burnout is 4.2% in the working population.⁶ Like persistent fatigue, burnout can lead to long-term sick leave. However, it is important to realise that persistently fatigued workers are not burned-out by definition, and that burned-out workers might not experience fatigue as a major complaint.

Both persistent fatigue and burnout are reported to be serious conditions⁶, but little research is available on their clinical features. In that respect, chronic fatigue syndrome (CFS) has received a wider range of attention.⁷ CFS is characterised by persistent medically unexplained fatigue for at least six months, several unexplained symptoms and severe functional impairment⁸, with a reported prevalence ranging from 0.2 to 0.5%.^{9,10} The cause of CFS is unknown and most patients do not recover spontaneously.¹¹⁻¹⁵ Of the various somatic and psychological treatments, only cognitive behaviour therapy and graded exercise seem to be promising.¹⁶ Interesting in that respect is that CFS patients tend to make causal attributions by assuming their illness results from a physical process.^{17,18} Several studies have shown that stronger physical causal attributions predict a longer duration of fatigue complaints in CFS.^{13,15,19} It has been hypothesized that, as fatigue complaints persist in time and no medical answers can be given, the need for a single, conclusive physical causal explanation strengthens in CFS patients.¹⁹ Whether this applies to fatigued or burned-out employees as well has yet to be investigated.

To our knowledge, there is no literature available in which CFS, burnout and persistent fatigue among employees are compared, despite their similarities (e.g. fatigue and impairment, potentially leading to work disability). A first assumption is that there is some (conceptual) overlap between these fatigue conditions, as illustrated in figure 1. Furthermore, as the duration of fatigue complaints in CFS patients increases, the chances of recovery diminish, until after 15 months complete recovery is rare.¹⁹ This leads us to hypothesize that fatigued employees with a long duration of complaints might already fulfil criteria for CFS. Also, since causal attributions seem to play an important role in CFS, we might

FIGURE 1

PROPOSED RELATIONS BETWEEN PERSISTENT FATIGUE AMONG EMPLOYEES, BURNOUT AND CFS

hypothesize that differences between the three fatigue conditions are reflected in the causal attributions that are made.

In this study, we aim to investigate relations between persistent fatigue, burnout and CFS by describing the clinical features of a sample of fatigued employees on sick leave. Subgroups based on research criteria for CFS and burnout within the sample of fatigued employees will be compared. In addition, comparisons will be made between subgroups of fatigued employees and a sample of diagnosed CFS patients that serve as a reference group.

Methods

SUBJECTS

Fatigued employees on sick leave were recruited among 80,000 employees under the registration of a major occupational health service (OHS) in the South of the Netherlands. On a monthly basis, persons on sick leave for more than two weeks, irrespective of the reason, were sent a screening list. Employees were eligible for participation if they met the following inclusion criteria: severe fatigue (a score of 35 or higher on the subscale fatigue severity of the Checklist Individual Strength (CIS)^{20, 21}) for more than four months; fatigue is reported to be one of the major health complaints; complete absenteeism from work for 6 to 26 weeks. Relevant exclusion criteria were a somatic explanation for fatigue and a history of or current psychiatric co-morbidity. In total 151 fatigued employees were thus included.

In addition, 97 patients diagnosed with CFS who were on sick leave, either temporary or due to work disability, served as a reference group. These patients were part of a larger sample of diagnosed CFS patients described elsewhere.²²

INSTRUMENTS

Fatigue

Fatigue severity was measured with the subscale fatigue severity of the Checklist Individual Strength (CIS).^{20, 21} This subscale contains eight items rated on a seven-point scale, with scores ranging from 8 to 56, higher scores indicating a higher severity of fatigue.

Functional impairment

Functional impairment was measured with the subscale physical functioning of the Short Form Health Survey (SF-36).²³ Scores on the SF-36 can range from 0 to 100, higher scores indicating higher levels of physical functioning. For specificity reasons, we chose a rather conservative cut-off point of 60 or lower for functional impairment, although others have suggested higher values (less impairment) as cut-off point.²⁴

Burnout

Burnout was measured with the Maslach Burnout Scale – General Survey (MBI-GS).^{25 26} Scores on the three subscales of the MBI-GS (exhaustion, professional efficacy and cynicism) can range from 0 to 6. Higher scores on these subscales indicate respectively higher levels of emotional exhaustion, (perceived) professional efficacy and cynicism towards work in general.

BOX 1

SAMPLE ITEMS OF THE SELF-EFFICACY SCALE AND THE CAL

self-efficacy*"Whatever I do, I cannot change my complaints"*

1= yes, I am convinced of that

4=no, I am convinced that is not the case

psychological attributions*"Do you think your complaints are related to stress?"*

1= no, I am convinced that is not the case

4=yes, I am convinced of that

somatic attributions*"Do you think your complaints are caused by something physical?"*

1= no, I am convinced that is not the case

4=yes, I am convinced of that

Physical activity

Physical activity was measured with the actometer, a motion-sensing device attached to the ankle and worn for 12 days continuously.²⁷ Scores reflect the average physical activity over 12 days, expressed in the average number of accelerations per 5-minute period. Higher scores indicate higher levels of physical activity, with a mean of 91 (SD 25) for healthy subjects.²⁷

Psychological problems

Psychological problems were assessed by two related constructs, depression and psychological distress in general. Depression was measured with the Beck Depression Inventory (BDI).²⁸ Scores on the BDI can range from 0 to 63, higher scores indicating higher levels of depression with a score of 16 as cut-off point for clinical depression. Psychological distress was measured with the Symptom Checklist 90 (SCL-90).²⁹ Scores on the SCL-90 can range from 90 to 450, higher scores indicating higher levels of psychological distress.

Self-efficacy and attributions

To measure self-efficacy, the self-efficacy scale^{22 30} administered in our previous CFS studies was used. The scale consists of four questions on a five-point scale and one question on a four-point scale that measure sense of control in relation to complaints, with scores ranging from 5 to 24, high scores indicating high levels of self-efficacy. To measure beliefs regarding the somatic and psychological causes of fatigue complaints, a *modified* version of the Causal Attributions List (CAL)^{22 30} administered in our previous CFS studies was used. The CAL consists of two subscales, psychological attributions and somatic attributions. In this study, the subscale psychological attributions contained five questions on a four-point scale, with scores ranging from 5 to 20. The subscale somatic attributions contained four questions on a four-point scale, with scores ranging from 4 to 16. Higher scores on these

two subscales indicate stronger psychological or somatic attributions. Sample items of the self-efficacy scale and the CAL are presented in box 1.

FORMATION OF SUBGROUPS

Fatigued employees were identified as potential CFS cases if they met all of the following operational research criteria: a CIS score of 40 or higher, duration of fatigue complaints six months or more and a SF-36-score on physical functioning of 60 or lower. These operational research criteria were based on the CDC criteria⁸ although CDC symptom criteria were not applied. A somatic explanation for fatigue and current psychiatric co-morbidity, contraindications for the diagnosis of CFS, already served as exclusion criteria for the selection of fatigued employees.

It is emphasized here that those participants who met CFS criteria did not necessarily qualify as CFS patients: a final diagnosis of CFS can only be made by a physician after a sufficient physical examination. None of the fatigued employees had a CFS diagnosis at entry in the study. Therefore, those who met research criteria will be referred to as CFS-like employees (the suffix "like" was used to clarify that participants met criteria by self-report only³¹)

Fatigued employees were identified as potential burnout cases if they met operational research criteria for burnout based on the Maslach Burnout Inventory – General Survey (MBI-GS). The MBI-GS contains three subscales that cover relevant aspects of burnout; exhaustion (ex), professional efficacy (pe) and cynicism (cy).²⁶ A clinically derived cut-off point for burnout - determined on data from the Maastricht Cohort Study - was used, based on the following criteria: $ex > 2.78$ and $(pe < 3.55$ or $cy > 1.75)$ indicating potential burnout caseness.³²

STATISTICAL ANALYSIS

To test differences between groups, a series of Analysis of Covariance (ANCOVA) was performed. ANCOVA was used as a statistical matching procedure to adjust group means to what they would be if all subjects scored equally on the covariates of choice.³³ Since it is essential to select a small number of covariates to prevent reduction of power, a limited number of demographic variables were considered as potential covariates. Thus, 'age' and 'sex' were included as covariates in the analysis. The assumptions of the ANCOVA model (normality of sampling distributions, homogeneity of variance, reliability of covariates, absence of multicollinearity between covariates, linearity, absence of outliers, homogeneity of regression) were evaluated before each analysis. If Levene's test for equality of error variances was significant, the alpha was adjusted to 0.01. After evaluation of the main effect, pair wise comparisons between adjusted means were performed. Due to the large number of comparisons, the alpha was adjusted using the Bonferonni method.

Absence of statistically significant differences does not imply clinical comparability or equivalence.

Therefore, equivalence of means was assessed by evaluation of the confidence interval (CI) for the true difference between adjusted group means.³⁴ First, for every measure a clinically relevant difference (Δ) was determined based on clinical reasoning of experts.³⁴ A range of equivalence was then predefined by setting up an interval that corresponds to a difference of no clinical importance, ranging from $-\Delta$ to $+\Delta$. Clinically relevant differences and ranges of equivalence were defined for the following instruments: CIS (Δ 5), SF-36 (Δ 10), actometer (Δ 10), self-efficacy (Δ 3), psychological attributions (Δ 2.5), somatic attributions (Δ 2), BDI (Δ 5) and SCL-90 (Δ 30). If the CI for the true difference lied entirely between $-\Delta$ to $+\Delta$, equivalence was assumed.

Results

In table 1, demographic and study variables of the 151 fatigued employees and the reference group of CFS patients are displayed. In both groups, the average duration of fatigue complaints was relatively long (fatigued employees 27.4 months, CFS patients 62 months). Among fatigued employees, the average score for physical functioning was below the cut-off point (60 or lower) for functional impairment. Fatigued employees had high scores on exhaustion and cynicism, but also on professional efficacy.

TABLE 1
CHARACTERISTICS OF PARTICIPANTS IN THE STUDY

	fatigued employees on sick leave N=151	CFS patients on sick leave N=97
<i>demographic variables</i>		
sex (m/f)*	68/83 (45/55)	24/73 (25/75)
age in years	43.4 (8.3)	38.8 (8.9)
educational attainment (1=low to 7=high)	3.8 (1.6)	4.4 (1.5)
<i>study variables</i>		
fatigue severity (CIS)	48.7 (6.2)	53.0 (3.5)
duration of fatigue complaints in months	27.4 (29.5)	62.0 (64.5)
duration of absenteeism in weeks	12.3 (4.9)	**
physical functioning (SF-36)	59.9 (25.3)	**
burnout (MBI-GS) • exhaustion	3.4 (1.5)	**
• professional efficacy	4.0 (1.1)	**
• cynicism	2.2 (1.3)	**
psychological distress (SCL-90)	182.7 (50.5)	171.3 (41.8)
depression (BDI)	14.8 (7.4)	14.2 (6.9)
general physical activity (actometer)	63.7 (25.7)	61.7 (20.1)
self-efficacy (self-efficacy scale)	15.1 (3.5)	14.3 (3.4)
psychological attributions (CAL)	12.3 (2.6)	9.6 (2.4)
somatic attributions (CAL)	9.3 (2.5)	11.3 (2.3)

Data are mean (SD) or *number of participants (%). ** Data not available for CFS patients

Table 2 presents adjusted group means of fatigued employees who did not meet CFS criteria, stratified according to burnout (non-CFS-like employees, groups 1 and 2), fatigued employees who met CFS criteria, stratified according to burnout (CFS-like employees, groups 3 and 4) and diagnosed CFS patients who were absent from work on (long-term) sick leave (group 5).

TABLE 2
ADJUSTED MEANS* (95% CI) FOR FATIGUED EMPLOYEES ACCORDING TO CFS CASESNESS AND BURNOUT CASESNESS, CFS PATIENTS ON SICK LEAVE AND GROUP COMPARISONS

variable [scale]	non-CFS-like employees		CFS-like employees		CFS patients on sick leave (5)		group comparisons ‡
	burnout (1) n=42	no burnout (2) n=43	burnout (3) n=34	no burnout (4) n=32	n=97		
self-efficacy [5-24]	15.5 (14.5-16.6)	16.2 (15.1-17.2)	14.1 (12.9-15.2)	14.8 (13.6-15.9)	14.2 (13.5-14.9)		D E 2-5 1-2 3-5 3-5 4-5 4-5
psychological attributions [5-20]	13.4 (12.6-14.1)	11.7 (10.9-12.5)	12.5 (11.7-13.3)	11.0 (10.2-11.9)	9.7 (9.2-10.3)		1-2 1-3 1-4 2-3 1-5 2-4 2-5 2-5 3-5 3-5
somatic attributions [4-16]	7.9 (7.2-8.6)	8.8 (8.1-9.5)	10.6 (9.8-11.3)	10.6 (9.8-11.4)	11.3 (10.8-11.8)		1-3 3-4 1-4 3-5 1-5 4-5 2-3 2-3 2-4 2-4 2-5 2-5
depression† [0-63]	15.5 (12.6-18.4)	14.4 (11.6-17.2)	18.3 (15.1-21.4)	11.4 (8.2-14.7)	14.0 (12.1-15.9)		3-4 2-5
psychological distress† [90-450]	185.7 (166.9-204.5)	176.7 (158.3-195.1)	206.9 (186.4-227.5)	160.0 (138.7-181.2)	171.7 (159.2-184.2)		3-4 3-5
physical activity [no scale]	63.1 (55.8-70.4)	71.8 (64.8-78.9)	73.9 (66.3-81.5)	64.1 (56.1-72.1)	61.7 (57.3-66.2)		3-5 ** **

* means are adjusted for age and sex

‡ group comparisons are shown if means are significantly different at $p < 0.05$ = D and/or equivalent = E (***) = no significance or equivalence

† confidence interval (CI) is adjusted to 99% after interpretation of Levene's test of equality of error variances

non-CFS-like employees = fatigued employees who do not meet research criteria for CFS; CFS-like employees = fatigued employees who meet research criteria for CFS

How many fatigued employees on sick leave met operational research criteria for CFS and burnout?

In total 66 (43.7%) fatigued employees fulfilled criteria for CFS and 76 (50.3%) fatigued employees fulfilled criteria for burnout. CFS-like employees were as likely to be a burnout case as non-CFS-like employees.

Did fatigued employees on sick leave who met CFS criteria differ from those who did not meet CFS criteria?

Since CFS-like employees and non-CFS-like employees were divided based on fatigue, duration of fatigue complaints and physical functioning, differences between these groups on these variables were not tested. Both in CFS-like employees and non-CFS-like employees, differences in fatigue (group means: (1)=47.2; (2)=47.0; (3)=51.8; (4)=50.2), duration of complaints (group means: (1)=23.8; (2)=18.8; (3)=38.5; (4)=25.9) and physical functioning (group means: (1)=78.3; (2)=71.9; (3)=42.6; (4)=37.9) between burnout cases and non-burnout cases (group 1 vs 2, group 3 vs 4) were not significant (not in table). In fact, some group means were assumed to be equivalent for burnout cases and non-burnout cases (fatigue, groups 1 and 2; physical functioning, groups 1 and 2, 3 and 4).

Differences emerged in causal attributions: burned-out non-CFS-like employees (group 1) had stronger psychological attributions than the non-burnout cases among CFS-like and non-CFS-like employees (groups 2 and 4), while CFS-like employees had stronger somatic attributions than non-CFS-like employees, regardless of burnout status. In addition, equivalence of psychological attributions scores was assumed between the two burnout groups (1 and 3) and between the two non-burnout groups (2 and 4), although this finding is counteracted by the equivalence between groups 2 and 3. Finally, burned-out CFS-like employees (group 3) were significantly more depressed and psychologically distressed than non-burned-out CFS-like employees (group 4).

There were no significant differences in self-efficacy and physical activity, between CFS-like employees and non-CFS-like employees nor between burnout cases and non-burnout cases.

Did fatigued employees on sick leave who met CFS criteria differ from diagnosed CFS patients on sick leave?

Fatigue did not differ in CFS-like employees and CFS patients on sick leave (group means: (3)=51.8; (4)=50.2; (5)=52.8); burned-out CFS-like employees (group 3) even had fatigue scores equivalent to those of CFS patients (not in table). CFS patients did have a longer duration of fatigue complaints than CFS-like employees (group means in months: (3)=38.5; (4)=25.9; (5)=64.1), but the difference only reached significance between non-burned-out CFS-like employees (group 4) and CFS patients (not in table).

Burned-out CFS-like employees (group 3) had stronger psychological attributions and more psychological distress than CFS patients. Observed differences between CFS-like employees and CFS patients in somatic attributions, self-efficacy and depression were not significant and scores on somatic attributions and self-efficacy between CFS patients and CFS-like employees were assumed to be equivalent.

Discussion

In this study, we investigated associations with CFS and burnout in a sample of fatigued employees on sick leave.

First, we detected potential CFS cases among our sample of fatigued employees. More than 40 per cent met the research criteria for CFS. Besides differences in fatigue, duration of complaints and physical functioning due to the selection procedure, these 'CFS-like employees' had stronger somatic attributions than those who did not meet CFS criteria. Compared to diagnosed CFS patients, CFS-like employees had comparable scores on fatigue, self-efficacy and somatic attributions but a shorter duration of fatigue complaints, indicating that CFS-like employees generally were in an earlier stage of illness than CFS patients.

To illuminate relations further, we stratified participants according to burnout status and found that half of the fatigued employees fulfilled criteria for burnout, almost equally distributed among CFS-like employees and non-CFS-like employees. Burnout among fatigued employees was only associated with psychological attributions and psychological problems. This was to be expected, since we defined burnout primarily as a negative state of mind accompanied by psychological problems. However, these associations were different for CFS-like employees and non-CFS-like employees: within the group of CFS-like employees, those who met burnout criteria were more depressed and distressed than those who did not meet burnout criteria, while burnout cases in the non-CFS-like group had stronger psychological attributions than non-burnout cases in either groups. In addition, burned-out CFS-like employees had stronger psychological attributions and were more distressed than CFS patients.

As to the methodological restrictions of our study, it might have been insightful if we would have been able to examine larger samples of independently recruited fatigued employees, diagnosed burned-out employees and CFS patients using the same instruments. However, our approach provided us with sufficient information to answer several important questions. Since SF-36 data were not available for CFS patients, we were not able to determine the extent to which fatigued employees and CFS patients might differ in physical functioning.

In review of the results, we conclude that our sample of fatigued employees shared many important characteristics with CFS patients: severe fatigue, functional impairment and a relatively long duration of complaints. Differences however concern the attribution to causal factors. CFS-like employees and CFS patients were more likely to attribute their condition to somatic factors than fatigued employees who did not meet CFS criteria, while burned-out employees had stronger psychological attributions than non-burned-out employees and CFS patients. This raises questions about the role of causal attributions: are they modified by the nature, persistence and duration of complaints or are they factors that determine illness outcome?

Causal attributions are predictors of the course of CFS.^{13, 15, 19} The manner in which a person labels his fatigue complaints, for instance as a temporary state due to temporary pressure at work (like

burnout) versus a chronic, persistent state that may last endlessly due to unknown somatic factors (like CFS), might influence the course and duration of complaints. It is important to bear in mind that fatigued employees in this study who applied to CFS criteria were not diagnosed as such, and may never be aware of this 'status': CFS-like employees (none had received a prior CFS diagnosis) were identified using operational research criteria for CFS, CDC symptom criteria⁸ were not applied and fatigued employees did not undergo physical examination. We are inclined to believe that many of these fatigued employees might have a better prognosis than 'regular' CFS patients, not only because of the shorter duration of their fatigue complaints but also due to a different labelling of these complaints. Generally, not many CFS patients recover spontaneously from their complaints without adequate treatment.¹¹⁻¹⁵ If most fatigued employees do recover spontaneously, causal attributions might have something to do with it. Butler *et al.*¹⁸ have suggested that the tendency to make somatic attributions might be a vulnerability factor for the development of CFS. We do not know if and how this applies to fatigued employees. In that sense, it is of great importance to identify the determinants of recovery in fatigued employees: if causal attributions can determine the course and outcome of fatigue complaints in employees, it might be an indication that early prevention of chronic fatigue lies partly in alterations of the labelling of fatigue complaints, for example with the use of cognitive behavioural techniques.

Acknowledgments

The authors would like to thank Pieter Leffers for methodological advice. This research project was funded by the Health Research and Development Council (ZorgOnderzoek Nederland), The Netherlands (grant no. 2830180). The recruitment of patients was made possible by the donation of services by 'Arbo Unie Limburg' (Occupational Health Service Limburg).

References

1. Lewis G, Wessely S. The epidemiology of fatigue: more questions than answers. *J Epidemiol Community Health* 1992; 46:92-7.
2. Pawlikowska T, Chalder T, Hirsch S, Wallace P. Population based study of fatigue and psychological distress. *BMJ* 1994; 308:763-6.
3. Bültmann U, Kant IJ, Kasl SV, Beurskens AHJM, Van den Brandt PA. Fatigue and psychological distress in the working population: psychometrics, prevalence and correlates. *J Psychosom Res*; 52:445-52.
4. Schröer CAP. De toename van arbeidsongeschiktheid wegens psychische aandoeningen [The increase of work disability due to mental disorders]. *Tijdschrift voor Bedrijfs- en Verzekeringsgeneeskunde* 1997; 5:16-23.
5. Schaufeli WB, Enzmann D. *The burnout companion to study and practice: a critical analysis*. London: Taylor & Francis, 1998.
6. Houtman ILD, Schaufeli WB, Taris T. *Psychische vermoeidheid en werk [Mental fatigue and work]*. Alphen a/d Rijn: NWO-prioriteiten programma PVA/Samsom, 1998.
7. Wessely S, Hotopf M, Sharpe M. *Treatment of CFS: the evidence. Chronic Fatigue and its Syndromes*. Oxford New York Tokyo: Oxford University Press, 1998:357-98.
8. 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.
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. Bleijenberg G, Bazelmans E, Prins JB. Chronisch vermoeidheidssyndroom [Chronic fatigue syndrome]. *Praktijkreeks gedragstherapie [Practice series behavioural therapy]*. Houten/Diegem: Bohn Stafleu Van Loghum, 2001.
11. Peterson PK, Schenck CH, Sherman R. Chronic fatigue syndrome in Minnesota. *Minn Med* 1991; 74:21-6.
12. Tirelli U, Marotta G, Improta S, Pinto A. Immunological abnormalities in patients with chronic fatigue syndrome. *Scand J Immunol* 1994; 40:601-8.
13. Wilson A, Hickie I, Lloyd A. Longitudinal study of outcome of chronic fatigue syndrome. *BMJ* 1994; 308:756-9.
14. Bombardier CH, Buchwald D. Outcome and prognosis of patients with chronic fatigue vs chronic fatigue syndrome. *Arch Intern Med* 1995; 155:2105-10.
15. Vercoulen JHMM, Swanink CM, Fennis JF, Galama JM, Van der Meer JW, Bleijenberg G. Prognosis in chronic fatigue syndrome: a prospective study on the natural course. *J Neurol Neurosurg Psychiatry* 1996; 60:489-94.
16. Whiting P, Bagnall AM, Sowden AJ, Cornell JE, Mulrow CD, Ramirez G. Interventions for the treatment and management of chronic fatigue syndrome: a systematic review. *JAMA* 2001; 286:1360-8.
17. Wessely S, Powell R. Fatigue syndromes: a comparison of chronic "postviral" fatigue with neuromuscular and affective disorders. *J Neurol Neurosurg Psychiatry* 1989; 52:940-8.
18. Butler JA, Chalder T, Wessely S. Causal attributions for somatic sensations in patients with chronic fatigue syndrome and their partners. *Psychol Med* 2001; 31:97-105.
19. Van der Werf SP, De Vree B, Alberts M, Van der Meer JWM, Bleijenberg G. Natural course and predicting self-reported improvement in patients with chronic fatigue syndrome with a relatively short illness duration. *J Psychosom Res* 2002; 53:749-53.
20. Vercoulen JHMM, Alberts M, Bleijenberg G. De Checklist Individual Strength (CIS). *Gedragstherapie* 1999; 32:31-6.

21. Beurskens AJHM, Bultmann U, Kant IJ, Vercoulen JH, Bleijenberg G, Swaen GM. *Fatigue among working people: validity of a questionnaire measure.* *Occup Environ Med* 2000; 57:353-7.
22. Prins JB, Bleijenberg G, Bazelmans E, et al. *Cognitive behaviour therapy for chronic fatigue syndrome: a multicentre randomised controlled trial.* *Lancet* 2001; 357:841-7.
23. Ware JE, Jr., Sherbourne CD. *The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection.* *Med Care* 1992; 30:473-83.
24. Powell P, Bentall RP, Nye FJ, Edwards RH. *Randomised controlled trial of patient education to encourage graded exercise in chronic fatigue syndrome.* *BMJ* 2001; 322:387-90.
25. Schaufeli WB, Leiter MP, Maslach C, Jackson SE. *The Maslach Burnout Inventory-General Survey.* In: Maslach C, Jackson SE, eds. *Maslach Burnout Inventory Manual.* Palo Alto, CA: Consulting Psychologists Press, 1996.
26. Schaufeli WB, Van Dierendonk D. *UBOS, Utrechtse Burnout Schaal, Handleiding (UBOS, Dutch translation of the MBI, Manual).* Lisse: Swets Test Publishers, 2000.
27. Van der Werf SP, Prins JB, Vercoulen JH, Van der Meer JW, Bleijenberg G. *Identifying physical activity patterns in chronic fatigue syndrome using actigraphic assessment.* *J Psychosom Res* 2000; 49:373-9.
28. Beck AT, Steer RA, Garbin MC. *Psychometric properties of the beck depression inventory: twenty-five years of evaluation.* *Clin Psychol Rev* 1988; 8:77-100.
29. Derogatis LR, Lipman RS, Covi L. *SCL-90: an outpatient psychiatric rating scale—preliminary report.* *Psychopharmacol Bull* 1973; 9:13-28.
30. Vercoulen JHMM, Hommes OR, Swanink CM, et al. *The measurement of fatigue in patients with multiple sclerosis. A multidimensional comparison with patients with chronic fatigue syndrome and healthy subjects.* *Arch Neurol* 1996; 53:642-9.
31. Jason LA, Taylor RR. *Applying cluster analysis to define a typology of chronic fatigue syndrome in a medically-evaluated, random community sample.* *Psychol Health* 2002; 17:323-37.
32. Kant IJ, Bultmann U, Schroer CAP, Beurskens AJHM, Van Amelsvoort LGPM, Swaen GMH. *An epidemiological approach to study fatigue in the working population: The Maastricht Cohort Study.* *Occup Environ Med* 2003; 60 (Suppl 1):i32-i9.
33. Tabachnick BG, Fidell LS. *Using multivariate statistics.* Needham Heights, MA: Allyn & Bacon, 2001.
34. Jones B, Jarvis P, Lewis JA, Ebbutt AF. *Trials to assess equivalence: the importance of rigorous methods.* *BMJ* 1996; 313:36-9.

Prevalence of chronic fatigue syndrome-like caseness in the working population: results from the Maastricht Cohort Study

Marcus J.H. Huibers ¹

Ijmert Kant ¹

Gerard M.H. Swaen ¹

Stanislav V. Kasl ²

OCCUPATIONAL & ENVIRONMENTAL MEDICINE 2003, IN PRESS

¹ DEPARTMENT OF EPIDEMIOLOGY, MAASTRICHT UNIVERSITY, THE NETHERLANDS

² DEPARTMENT OF EPIDEMIOLOGY AND PUBLIC HEALTH, YALE UNIVERSITY SCHOOL OF MEDICINE, NEW HAVEN CT, USA

Abstract

BACKGROUND

Fatigue is a common complaint in the working population. Until now, the relation between fatigue in the working population and the chronic fatigue syndrome (CFS) has seldom been the focus of study. In this study, we aim to determine the prevalence of CFS-like caseness in the working population.

METHODS

We used data from the prospective Maastricht Cohort Study on Fatigue at Work. The prevalence and the incidence of CFS-like cases (employees meeting research criteria for CFS) was determined among 5499 employees who responded to the follow-up assessment 3 years and 8 months after baseline.

RESULTS

Of the 5499 employees, 199 (3.6%) were identified as CFS-like cases. By deleting possible CFS-like cases at baseline, the annual incidence of CFS-like caseness was estimated to be 85 per 10,000. 20 employees (0.36%) reported having been diagnosed with CFS by a physician.

CONCLUSIONS

The prevalence of CFS-like cases in our study (3.6%) was considerably higher than the prevalence of chronic fatigue syndrome reported in previous studies (0.006% to 3%). These findings suggest that the chronic fatigue syndrome-like caseness may be underdetected in the working population and perhaps in other populations as well.

Introduction

Fatigue is a common complaint in the working population, with a reported prevalence of 22%.¹ Fatigue can best be understood as a continuum, ranging from mild complaints frequently seen in the community to severe, disabling fatigue like chronic fatigue syndrome.² When fatigue becomes severe and persistent, it may lead to long-term sick leave³ and work disability.⁴

Although the body of literature on fatigue in the working population is growing substantially, the relation to the chronic fatigue syndrome (CFS) has barely received any attention in occupational health. CFS is characterized by persistent medically unexplained fatigue for at least six months, several unexplained symptoms (e.g. sore throat, headache, painful joints) and severe functional impairment, with a reported prevalence ranging from 0.006% to 3.0% depending on setting and definitions used.⁵ The cause of CFS is unknown and most patients do not recover spontaneously.⁵

The aim of the present report is to determine the prevalence of CFS-like caseness in the working population. We believe we are the first to do so.

Methods

DESIGN AND STUDY POPULATION

We report data from the prospective Maastricht Cohort Study on Fatigue at Work (MCS). In May 1998, 26,978 employees from 45 companies and organizations representing the main work sectors in The Netherlands were invited to participate in the MCS. 12,161 employees (45%) responded by providing informed consent and completing the postal self-report questionnaire at baseline. There were no demographic differences between respondents and non-respondents at baseline. Further details on the MCS are described elsewhere.⁶

The present report is based on follow-up data that were obtained 3 years and 8 months after baseline measurement. Variables necessary to determine whether employees met research criteria for CFS were only available at this point. Consequently, 5499 employees who completed all CFS-related questions were included in the analysis. Respondents at follow-up ($n=5499$) on average had a higher educational level compared to baseline respondents who did not complete follow-up ($n=6662$). Other demographic factors did not differ considerably.

STUDY VARIABLES

Fatigue was assessed with the subscale fatigue severity of the Checklist Individual Strength (CIS).⁷ Scores can range from 8 to 56, with higher scores indicating a higher severity of fatigue. Functional impairment was assessed with the subscale physical functioning of the Short Form Health Survey (SF-36).⁸ Scores can range from 0 to 100, with higher scores indicating higher levels of physical functioning. Burnout was assessed with the subscale exhaustion of the Maslach Burnout Inventory – General Survey (MBI-GS).⁹ Scores can range from 0 to 6, with higher scores indicating higher levels of work-related emotional exhaustion. Psychological distress was assessed with the Hospital Anxiety and Depression Scale (HAD) that consists of an anxiety and a depression scale.¹⁰ Scores can range from 0 to 21 on each scale, with higher scores indicating higher levels of anxiety and/or depression. A score of 11 or higher on either of the two scales is indicative of being a psychological distress case.¹¹ Other variables included self-reported absence from work and demographic variables.

CFS-LIKE CASENESS AND CFS DIAGNOSIS

Employees were identified as *CFS-like cases* if they met all of the following research criteria that approximate the CDC criteria for CFS¹²: a CIS score of 40 or higher, a self-reported duration of fatigue complaints of six months or more, a SF-36 score on physical functioning of 60 or lower and the self-reported presence of four or more CFS symptoms during the past six months, in line with previous studies of CFS.^{13, 14} Somatic or psychiatric conditions that might explain fatigue complaints

could not be excluded. Since CFS criteria could only be fully assessed at follow-up, *potential CFS-like cases at baseline* were identified based on provisional criteria: a CIS score of 40 or higher at baseline, scoring "yes" on at least one of two questions on health-related impairment at baseline and a reported duration of fatigue complaints of four years or more at follow-up. It is emphasized here that *CFS-like cases* did not necessarily qualify as *CFS patients*: a CFS diagnosis can only be made by a physician after a sufficient physical examination. However, CFS-like caseness is regarded as a good proxy for true CFS.¹⁵

In addition, the follow-up questionnaire included the following question: "Did you at any time receive a diagnosis of chronic fatigue syndrome or ME from a qualified physician?" However, this information was not used to determine CFS-like caseness.

STATISTICAL ANALYSIS

Differences between CFS-like employees and non-CFS employees were tested using multiple logistic regression models, adjusting for demographic factors.

Results

PREVALENCE AND INCIDENCE

Of the 5499 employees at follow-up, 199 employees (3.6%) met the CFS research criteria (see table). Of these 199 CFS-like cases, 28 employees might have been CFS-like cases at baseline. By deleting these 28 employees from the population at risk, we estimated the yearly incidence of CFS-like caseness to be 85 cases per 10,000. Prevalence and incidence indicated a mean duration of CFS-like caseness of approximately four years ($P = I \times D$).

Of the 20 employees (0.36%) who reported a lifetime diagnosis of CFS, only two employees were identified as CFS-like cases.

CHARACTERISTICS

There were no gender differences between CFS-like employees and non-CFS employees (see table). Compared to employees who did not meet CFS research criteria, CFS-like cases on average were older, had a lower education and were more likely to suffer from burnout, depression or anxiety, be a psychological distress case (HAD) or be absent from work at follow-up. In the group of employees who reported having received a CFS diagnosis, the mean age was 43.6 (SD 8.1); 11 employees (55%) were female; and the majority had a high (45%) or medium educational level (45%; not in table).

TABLE 2

PREVALENCE, INCIDENCE AND CHARACTERISTICS OF CHRONIC FATIGUE SYNDROME-LIKE CASES IN THE MAASTRICHT COHORT STUDY AT FOLLOW-UP

POPULATION	NUMBER OF CASES		
working population at follow-up	5499 (10,000/10,000)		
<i>Employees meeting CFS research criteria</i>			
prevalence at follow-up	199 (362/10,000)		
estimated incidence since baseline	171 (311/10,000)		
estimated incidence per year	47 (85/10,000)		
<i>Employees reporting CFS diagnosis</i>			
lifetime prevalence at follow-up	20 (36/10,000)		
number of employees meeting CFS criteria at follow-up who reported lifetime CFS diagnosis	2 (4/10,000)		

<i>characteristic †</i>	<i>CFS-like employees</i> n=199	<i>non-CFS employees</i> n=5300	<i>adjusted odds ratio ‡</i> (95% CI)
female sex, nr (%)	55 (28%)	1426 (27%)	1.28 (0.92 to 1.79)
age in years	48.4 (7.6)	45.8 (8.3)	1.04 (1.02 to 1.05)
educational level • high, nr (%)	31 (16%)	2054 (40%)	0.21 (0.13 to 0.33)
• Medium, nr (%)	100 (51%)	2290 (44%)	0.6 (0.43 to 0.84)
• low, nr (%)	64 (33%)	833 (16%)	1.00
fatigue severity (CIS)	47.2 (4.6)	23.1 (11.7)	*
physical functioning (SF-36)	44.2 (14.4)	91.2 (13.2)	*
Exhaustion (MBI-GS)	3.6 (1.2)	1.6 (1.2)	2.78 (2.48 to 3.12)
anxiety (HAD)	9.3 (4.4)	4.8 (3.5)	1.28 (1.24 to 1.33)
depression (HAD)	9.1 (4.2)	3.7 (3.5)	1.33 (1.28 to 1.37)
HAD case, nr (%)	100 (50%)	517 (10%)	8.59 (6.37 to 11.6)
absent from work, nr (%)	49 (25%)	341 (7%)	3.74 (2.57 to 5.41)
duration absenteeism, nr (%)			
• 0-4 months	11 (22% of 49)	109 (32% of 341)	1.00
• 5-6 months	4 (8% of 49)	39 (11% of 341)	1.07 (0.32 to 3.62)
• 7-12 months	11 (22% of 49)	66 (19% of 341)	1.68 (0.68 to 4.17)
• 13 months or longer	23 (48% of 49)	127 (38% of 341)	1.82 (0.81 to 4.07)

† data are mean (sd) unless stated otherwise.

‡ adjusted for sex, age and educational level

* odds ratios not presented because groups were defined based on these variables

CFS-like employees: employees who met research criteria for chronic fatigue syndrome.

non-CFS employees: employees who did not meet research criteria for chronic fatigue syndrome.

Discussion

We found a prevalence of CFS-like caseness in the working population of 3.6%. Although prevalence estimates vary greatly due to differences in CFS definitions and study settings, the criteria-based prevalence we found is still considerably higher than CFS prevalences reported across previous studies.⁵ Also interesting in this context is our finding that only 0.36% of the employees claimed to have received a CFS diagnosis, while ten times as many employees (3.6%) actually met CFS criteria, with a poor overlap between both groups. It was surprising to find that there were no gender differences between CFS-like and non-CFS employees since most diagnosed CFS patients are women.⁵

Most studies of the prevalence of CFS are conducted in clinical settings among diagnosed and/or help-seeking CFS patients, which inevitably introduces self-selection bias. The major advantage of the present report is that we detected (undiagnosed) CFS-like cases using research criteria free from bias regarding health-care seeking or illness recognition by doctors and patients. This screening procedure may explain the high prevalence of CFS-like caseness we found in contrast to previous studies and the number of employees reporting to have received a CFS diagnosis. The prevalence reported here may even be an underestimation. Prospective studies in the working population are usually biased by the healthy-worker effect. In the Maastricht Cohort Study, fatigued employees at baseline were more likely to drop out in time than non-fatigued employees.⁶ On the other hand, the use of self-report questionnaires clearly has limitations in the assessment of chronic fatigue syndrome. A major disadvantage of the present report is that we could not exclude somatic or psychiatric conditions that might explain fatigue, which in turn may have led to an overestimation of CFS-like caseness. The considerable percentage of psychological distress cases among CFS-like employees seems to confirm this notion, although HAD caseness does not qualify as indicator of a psychiatric disorder that might explain fatigue. Furthermore, we are not sure if the selection of potential CFS-like cases at baseline was accurate since we used a provisional set of necessary but insufficient criteria.

Our findings lead us to suggest that chronic fatigue syndrome-like caseness may be underdetected in the working population and possibly in other populations as well. A possible explanation would be that the road to receiving a diagnosis of CFS usually takes a lengthy process of vigorous health care-seeking. Whether the male gender, educational level and active work status of most undiagnosed CFS-like cases in this study play a role in this is unclear.

Acknowledgments

This study was funded by the Health Research and Development Council (ZorgOnderzoek Nederland), The Netherlands (grant nr. 2830180). The Maastricht Cohort Study is part of The Netherlands concerted action on 'Fatigue at Work' granted by The Netherlands Organization for Scientific Research (NWO).

References

1. Bültmann U, Kant IJ, Kasl SV, Beurskens AJHM, Van den Brandt PA. *Fatigue and psychological distress in the working population: psychometrics, prevalence and correlates.* *J Psychosom Res* 2002; 52:445-52.
2. Lewis G, Wessely S. *The epidemiology of fatigue: more questions than answers.* *J Epidemiol Community Health* 1992; 46:92-7.
3. Janssen N, Kant IJ, Swaen GMH, Janssen PPM, Schröer CAP. *Fatigue as a predictor of sickness absence: results from the Maastricht cohort study on fatigue at work.* *Occup Environ Med* 2003; 60 (Suppl 1):i71-i6.
4. Amelsvoort van LGPM, Kant IJ, Beurskens AJHM, Schröer CAP, Swaen GMH. *Fatigue as a predictor of work disability.* *Occup Environ Med* 2002; 59:712-13.
5. Afari N, Buchwald D. *Chronic Fatigue Syndrome: a Review.* *Am J Psychiatry* 2003; 160:221-36.
6. Kant IJ, Bültmann U, Schroer CAP, Beurskens AJHM, Van Amelsvoort LGPM, Swaen GMH. *An epidemiological approach to study fatigue in the working population: The Maastricht Cohort Study.* *Occup Environ Med* 2003; 60 (suppl 1):i32-i9.
7. Beurskens AJHM, Bültmann U, Kant IJ, Vercoulen JHHM, Bleijenberg G, Swaen GMH. *Fatigue among working people: validity of a questionnaire measure.* *Occup Environ Med* 2000; 57:353-7.
8. Ware JE, Jr., Sherbourne CD. *The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection.* *Med Care* 1992; 30:473-83.
9. Schaufeli WB, Leiter MP, Maslach C, Jackson SE. *The Maslach Burnout Inventory-General Survey.* In: Maslach C, Jackson SE, eds. *Maslach Burnout Inventory Manual.* Palo Alto, CA: Consulting Psychologists Press, 1996.
10. Zigmond AS, Snaith RP. *The Hospital Anxiety and Depression Scale.* *Acta Psychiatr Scand* 1983; 67:361-70.
11. Wilkinson MJB, Barczak P. *Psychiatric screening in general practice: comparison of the general health questionnaire and the hospital anxiety and depression scale.* *J R Coll Gen Pract* 1988; 38:311-13.
12. Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JC, 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.
13. Powell P, Bental RP, Nye FJ, Edwards RH. *Randomised controlled trial of patient education to encourage graded exercise in chronic fatigue syndrome.* *BMJ* 2001; 322:387-90.
14. Prins JB, Bleijenberg G, Bazelmans E, et al. *Cognitive behaviour therapy for chronic fatigue syndrome: a multicentre randomised controlled trial.* *Lancet* 2001; 357:841-7.
15. Jason LA, Taylor RR. *Applying cluster analysis to define a typology of chronic fatigue syndrome in a medically-evaluated, random community sample.* *Psychol Health* 2002; 17:323-37.

The development of the chronic fatigue syndrome in fatigued employees: predictors of outcome in the Maastricht Cohort Study

Marcus J.H. Huibers ^{1,2}

Ijmert Kant ^{1,2}

J. André Knottnerus ²

Gijs Bleijenberg ³

Gerard M.H. Swaen ^{1,2}

Stanislav V. Kasl ⁴

SUBMITTED

¹ DEPARTMENT OF EPIDEMIOLOGY, MAASTRICHT UNIVERSITY, THE NETHERLANDS

² CARE AND PUBLIC HEALTH RESEARCH INSTITUTE (CAPHRI), MAASTRICHT UNIVERSITY, THE NETHERLANDS

³ DEPARTMENT OF MEDICAL PSYCHOLOGY, UMC ST RADBOUD NIJMEGEN, THE NETHERLANDS

⁴ DEPARTMENT OF EPIDEMIOLOGY AND PUBLIC HEALTH, YALE UNIVERSITY SCHOOL OF MEDICINE, NEW HAVEN CT, USA

Abstract

OBJECTIVE

To identify predictors of the development of the chronic fatigue syndrome (CFS), the persistence or recurrence of fatigue, or recovery from fatigue in a large sample of fatigued employees.

METHOD

1143 employees with unexplained fatigue were followed prospectively for 44 months. Multiple regression models were used to identify predictors of three potential outcomes at follow-up: CFS-like caseness (meeting research criteria for CFS), non-CFS fatigue caseness or no fatigue caseness.

RESULTS

At 44-month follow-up, 8% of the employees were CFS-like cases (none of who reported to have received a CFS diagnosis), 40% were non-CFS fatigue cases and 52% were no longer fatigue cases. Factors that predicted CFS-like caseness *vs.* non-CFS fatigue caseness were high age, exhaustion, female sex, low education and visits to the general practitioner (GP). Factors that predicted CFS-like caseness *vs.* no fatigue caseness were fatigue, exhaustion, low education, visits to the GP and occupational physician and bad self-rated health. Factors that predicted non-CFS fatigue caseness *vs.* no fatigue caseness were fatigue, low self-perceived activity, exhaustion, anxious mood and bad self-rated health.

CONCLUSIONS

Severe, persistent, unexplained fatigue among employees can be a precursor of the development of self-reported CFS. The prognostic role of self-rated health suggests that prevention and treatment of chronic fatigue should be aimed at changing the perception of health or illness. Less clear is the role of health care seeking and/or receiving a CFS diagnosis.

Introduction

The chronic fatigue syndrome (CFS) is a condition characterized by persistent medically unexplained fatigue lasting for at least six months, several other unexplained symptoms (e.g. sore throat, headache, painful joints) and severe functional impairment.¹ The reported prevalence ranges from 0.006% to 3.0%.² The cause of CFS is unknown² and most patients do not recover spontaneously.³⁻⁸

Fatigue can best be understood as a continuum, ranging from mild complaints frequently seen in the community to severe, disabling fatigue like the chronic fatigue syndrome.⁹ From this perspective, the question comes up what the precursors of CFS might be.

In recent years, fatigue in the working population has been the subject of growing interest, not in the least because of its high prevalence (22%)¹⁰ and the adverse consequences of severe fatigue, like reduced performance, sickness absence¹¹ and work disability¹² that have been reported.

Recently, we studied the relation between unexplained persistent fatigue among employees and CFS and found that the two conditions shared many characteristics¹³, with 44% of the employees with unexplained fatigue meeting the research criteria for CFS.¹ In a second study that was recently submitted, we concluded that CFS and unexplained persistent fatigue among employees seem different but related stages on the fatigue continuum.¹⁴ In this second study, we examined the predictors of outcome in fatigued employees on long-term sick leave and found that recovery from fatigue was associated with factors reflecting the perception of illness, such as self-rated health or the tendency to attribute fatigue complaints to a psychological cause. Interestingly, studies have shown that the extent to which CFS patients attribute their illness to a somatic or psychological cause is predictive of the course of CFS.^{5, 7, 8, 15-18}

In the present study, we seek to identify predictors of long-term outcomes in a large sample of fatigued employees, with 'outcome' defined as either the development of self-reported CFS, the persistence or recurrence of fatigue, or recovery from fatigue. Our study is based on data from the Maastricht Cohort Study of Fatigue at Work, a population-based sample frame in which more than 12.000 employees were followed prospectively for three years and eight months.¹⁹ In line with previous studies²⁰⁻²², we hypothesized that severity of fatigue at baseline would be a predictor of fatigue at follow-up. Considering the prognostic role of causal attributions in CFS, we also hypothesized that factors reflecting the perception of illness (self-rated health, causal attributions) would be predictive of the course of fatigue.

Methods

STUDY DESIGN OF THE MAASTRICHT COHORT STUDY

We used prospective data from the Maastricht Cohort Study on Fatigue at Work (MCS). In May 1998, 26,978 employees from 45 Dutch companies and organizations were invited to participate in the MCS. 12,161 employees (45%) responded by providing informed consent and completing the postal self-report questionnaire at baseline. There were no significant differences between respondents and non-respondents in non-response analyses, except that non-respondents were less likely to be fatigued or absent from work than respondents. Respondents in the MCS were followed every four months for three years, with an additional follow-up after 3 years and 8 months. Further details on the MCS are described elsewhere.¹⁹

STUDY POPULATION

Participants were selected at baseline of the MCS if they scored 35 or higher on the subscale fatigue severity of the Checklist Individual Strength^{23, 24}, a cut-off that is indicative of severe fatigue.^{23, 25} Fatigued employees at baseline were excluded if they reported a somatic condition that could explain fatigue complaints (e.g. cerebrovascular accidents, liver dysfunctions, diabetes mellitus, thyroid gland dysfunctions, cancer) or if they were potential CFS-like cases at baseline. Variables necessary to determine whether employees met research criteria for CFS were only available at the additional follow-up 3 years and 8 months (44 months) after baseline measurement. Consequently, we chose to select fatigued employees at baseline who were also available at the 44-month follow-up. Since CFS criteria could only be fully assessed at the 44-month follow-up, we identified potential CFS-like cases *at baseline* using provisional criteria (see definitions of caseness). 1143 employees were thus included in the analyses.

PREDICTOR VARIABLES

Demographic factors

Demographic variables included sex, age and educational level (high, medium, low) and were assessed at baseline.

Fatigue-related factors

Fatigue was assessed with the four subscales of the Checklist Individual Strength (CIS)^{23, 24}: fatigue severity; concentration; motivation; and self-perceived activity. Higher scores on these subscales indicate a higher level of fatigue severity and lower levels of concentration, motivation and self-perceived activity respectively.

Burnout was assessed with the three subscales of the Maslach Burnout Inventory-General Survey (MBI) ²⁶: exhaustion; cynicism; and professional efficacy, with higher scores indicating higher levels on these scales.

Need for recovery was assessed with the Need for Recovery scale ^{27, 28} that indicates the short-term effects of a day of work, with a higher score indicating a higher need for recovery.

Participants were asked to indicate whether they felt they had suffered from *fatigue complaints* in the previous four months (yes or no) and, if so, whether they felt the cause of their fatigue complaints was physical (*physical attribution*, yes or no) or psychological in nature (*psychological attribution*, yes or no).

Mental health factors

Psychological distress was measured with the 12-item version of the General Health Questionnaire (GHQ) ²⁹, with a higher score indicating a higher level of psychological distress.

Depressed mood was assessed with a single item stating "did you feel down almost every day during the past two weeks, yes or no?"

Anxious mood was measured by adding the scores of three items on anxiety (excessive worrying, compulsive behavior, compulsive thinking), with a higher score indicating a higher level of anxious mood. The internal consistency was modest but satisfactory, with a Cronbach's alpha of 0.66.

Participants were also asked if they had experienced any shocking *life events* in the past 12 months (yes or no).

Health factors

Self-rated health was assessed with one item from the Short Form Health Survey (SF-36) ³⁰, which is a widely used measure of general health status. ³¹ The original 5-point scale (1=excellent health; 2=very good health; 3=good health; 4=moderate health; 5=bad health) was dichotomized into 'good health' (1-3) and 'bad health' (4-5).

In addition, participants were asked if they: were pregnant (*pregnancy*, yes or no); had visited their general practitioner (*self-reported visits to GP*) and/or their occupational physician (*self-reported visits to OP*) in the past 4 months because of problems at work (yes or no); suffered from *health complaints* (yes or no); or suffered from *sleep disturbances* (yes or no).

Miscellaneous factors

Work status was assessed as self-reported sickness absence from work (yes or no). Participants were asked whether their health had impaired them in performing their work (*impairment in work*, yes or no) and/or their non-work-related activities (*impairment in activities*, yes or no).

DEFINITIONS OF CASENESS

We assessed the outcome in our sample of fatigued employees ($n=1143$) at the 44-month follow-up and classified participants according to three mutually exclusive types of caseness: CFS-like caseness ($n=94$), non-CFS fatigue caseness ($n=457$) and no fatigue caseness ($n=592$).

Employees were identified as *CFS-like cases* if they met all of the following research criteria that approximate the CDC criteria for CFS¹: a CIS score of 40 or higher, a reported duration of fatigue complaints of six months or more, a SF-36 score on physical functioning of 60 or lower³⁰ and the presence of four or more CFS symptoms. The suffix "like" was used to clarify that participants met criteria by self-report only but were not diagnosed as such. It is emphasized here that *CFS-like cases* did not necessarily qualify as *CFS patients*: a CFS diagnosis can only be made by a physician after a sufficient physical examination. However, CFS-like caseness is regarded as a good proxy for CFS diagnosed by a physician.^{32, 33}

Employees were identified as *non-CFS fatigue cases* if they scored 35 or higher on the CIS. Fatigue cases who were CFS-like cases as well were excluded from this group. Employees were identified as *no fatigue cases* if they scored 34 or lower on the CIS.

Potential CFS-like cases at baseline were identified and excluded from analysis based on provisional criteria: a CIS score of 40 or higher at baseline, scoring "yes" on at least one of the two questions on health-related impairment at baseline and a reported duration of fatigue complaints of four years or more at 44-month follow-up.

In addition, the follow-up questionnaire included the following question: "*Did you at any time receive a diagnosis of chronic fatigue syndrome or ME from a qualified physician?*" However, this information was not used to determine CFS-like caseness.

STATISTICAL ANALYSES

Group differences in baseline characteristics were tested using ANOVA (continuous variables), the chi-square test (dichotomous variables) and the Kruskal-Wallis test (categorical variables).

Multiple logistic regression models were used to determine separately the associations between predictors and three dichotomies of caseness: CFS-like cases vs. non-CFS fatigue cases ($n=551$); CFS-like cases vs. no fatigue cases ($n=686$); and non-CFS fatigue cases vs. no fatigue cases ($n=1049$). Analyses were performed in two steps. In a first step, baseline predictors that were bivariately associated ($p\text{-value} < 0.05$) with caseness were entered in a multiple logistic model, together with the covariates age, sex, educational level, work status and pregnancy, and eliminated in a backwards procedure until the model consisted of significant predictors only ($p\text{-value} < 0.05$).

In a second step, prospective change scores of significant predictors were added to the model and the backwards elimination procedure was repeated. Change scores of continuous variables were calculated by subtracting baseline scores from scores at one year (T3) and two years (T6). Change scores

of dichotomous variables were also constructed, with the direction of the transition since the previous measurement retained.

Odds ratios and 95% confidence intervals were calculated and standardized to represent change per SD on the scale. Logistic regression analyses were performed using SPSS.

FIGURE 1

STUDY DESIGN: SELECTION OF PARTICIPANTS AND ASSESSMENTS

TABLE 1
BASELINE CHARACTERISTICS OF THE SAMPLE OF FATIGUED EMPLOYEES (N=1143) CLASSIFIED ACCORDING TO OUTCOME (CASENESS) AT 44-MONTH FOLLOW-UP

variable	group 1 CFS-like cases N=94	group 2 non-CFS fatigue cases N=457	group 3 no fatigue cases N=592	<i>p-values for differences between the groups</i> [†]		
				1 VS 2	1 VS 3	2 VS 3
<i>demographic</i>						
age, mean (sd)	43.8 (7.5)	41.2 (7.9)	41.3 (7.9)	0.009*	0.010*	1.000
male sex, nr (%)	61 (65%)	336 (73%)	397 (67%)	0.090	0.679	0.024*
educational level • high, nr (%)	18 (20%)	187 (42%)	231 (40%)			
• medium, nr (%)	45 (48%)	184 (41%)	250 (43%)		0.000*	
• low, nr (%)	30 (32%)	77 (17%)	100 (17%)			
<i>continuous, mean (sd)</i>						
fatigue severity (scale 35-56)	44.7 (5.9)	42.7 (5.6)	40.6 (5.0)	0.005*	0.000*	0.000*
emotional exhaustion (scale 0-6)	3.5 (1.2)	3.0 (1.2)	2.6 (1.1)	0.005*	0.000*	0.000*
psychological distress (scale 0-12)	5.5 (4.2)	4.4 (3.8)	3.6 (3.6)	0.022*	0.000*	0.002*
anxious mood (scale 3-12)	6.1 (1.8)	6.0 (1.7)	5.5 (1.5)	1.000	0.002*	0.000*
<i>dichotomous, nr of cases (%)</i>						
depressed mood	34 (37%)	106 (24%)	120 (21%)	0.009*	0.001*	0.266
physical attribution of fatigue	44 (47%)	157 (34%)	200 (34%)	0.022*	0.014*	0.847
psychological attribution of fatigue	45 (48%)	240 (53%)	282 (48%)	0.412	0.966	0.117
no specific attribution of fatigue	21 (23%)	89 (19%)	111 (19%)	0.527	0.412	0.767
self-rated health (good)	42 (45%)	256 (57%)	431 (73%)	0.045*	0.000*	0.000*
health complaints	71 (76%)	252 (55%)	283 (48%)	0.000*	0.000*	0.022*
absent from work	9 (10%)	33 (7%)	37 (6%)	0.434	0.231	0.532

[†] pairwise comparisons for continuous and dichotomous variables, overall comparison for categorical variable educational level

* *p*-value is statistically significant at *p*<0.05

CFS-like cases = fatigue employees at baseline who met CFS research criteria at follow-up; Non-CFS fatigue cases=fatigued employees at baseline who were (still) fatigued but did not meet CFS research criteria at follow-up; No fatigue cases = fatigued employees at baseline who were no longer fatigued at follow-up.

Results

In table 1, baseline characteristics of 1143 fatigued employees classified according to caseness at follow-up are presented. Three years and eight months after baseline, 94 employees (8%) were CFS-like cases, 457 employees (40%) were non-CFS fatigue cases and 592 employees (52%) were no longer fatigue cases. None of the CFS-like cases reported that they had received a CFS or ME diagnosis (not in table). Compared to non-CFS cases (groups 2 and 3), CFS-like cases at follow-up had a higher age, were more fatigued, exhausted and psychologically distressed and were more likely to report health complaints, a depressed mood, a physical attribution of fatigue and bad self-rated health at baseline.

Five baseline factors were found to be predictive of CFS-like caseness at follow-up as opposed to non-CFS fatigue caseness: high age, high exhaustion, female sex, a low educational level and a self-reported visit to the general practitioner (GP) because of problems at work (table 2). When change scores of significant predictors were added to the logistic regression model in a second step, an increase in exhaustion over time (i.e. a higher score at two-year assessment compared to the baseline score) also emerged as a significant predictor.

TABLE 2
SIGNIFICANT PREDICTORS OF CFS-LIKE CASENESS VERSUS NON-CFS FATIGUE CASENESS AT 44-MONTH FOLLOW-UP IN EMPLOYEES FATIGUED AT BASELINE (N=551)

predictors	MULTIPLE LOGISTIC REGRESSION MODELS	
	baseline predictors only† OR (per SD) (95% CI)	baseline and follow-up predictors‡ OR (per SD) (95% CI)
<i>continuous</i>		
age	1.36 (1.08 to 1.84)	1.36 (1.08 to 1.71)
exhaustion (MBI)	1.33 (1.04 to 1.67)	1.52 (1.16 to 2.02)
• change score t6-to	-	1.36 (1.02 to 1.81)
<i>dichotomous/categorical</i>		
sex (male=1)	0.41 (0.24 to 0.70)	0.36 (0.21 to 0.64)
educational level • low versus high	3.82 (1.92 to 7.61)	4.02 (1.99 to 8.11)
• middle versus high	2.48 (1.35 to 4.51)	2.48 (1.34 to 4.59)
self-reported GP visit (yes=1)	1.98 (1.19 to 3.29)	1.87 (1.11 to 3.14)

Predictor values are significant odds ratios (95% CI) adjusted for all other predictors in the model. Values of continuous predictors are expressed as OR per SD.

t3 = one-year assessment; t6 = two-year assessment; GP = general practitioner.

† Full model included age, sex, educational level, work status, pregnancy, CIS-fatigue, CIS-motivation, need for recovery, MBI-exhaustion, impairment in work, impairment in activities, physical attribution of fatigue, psychological distress, depressed mood, life events, visits to GP, visits to occupational physician, self-rated health, health complaints, all assessed at baseline.

‡ Full model included age, sex, educational level, MBI-exhaustion (baseline), MBI-exhaustion change score (t3-to), MBI-exhaustion change score (t6-to), visits to GP (baseline), visits to GP (t3), visits to GP (t6).

When CFS-like caseness at follow-up was compared to no fatigue caseness, six baseline predictors of CFS-like caseness were found: high fatigue, high exhaustion, a low educational level, a self-reported visit to the GP because of problems at work, the *absence* of a self-reported visit to the occupational physician (OP) and bad self-rated health (table 3). When change scores were added to the model, increases of fatigue and exhaustion in time were found to be significant predictors. A change of self-rated health from bad to good between baseline and one-year assessment was protective of CFS-like caseness, while a change from good to bad between one and two years of follow-up was predictive of CFS-like caseness. When self-reported visits to GP and OP at follow-up were added to the model, visits to GP and OP at baseline lost significance and a self-reported visit to the GP at one-year assessment emerged as a predictor.

TABLE 3

SIGNIFICANT PREDICTORS OF CFS-LIKE CASENESS VERSUS NO FATIGUE CASENESS AT 44-MONTH FOLLOW-UP IN EMPLOYEES FATIGUED AT BASELINE (N=686)

predictors	MULTIPLE LOGISTIC REGRESSION MODELS	
	baseline predictors only† OR (per SD) (95% CI)	baseline and follow-up predictors‡ OR (per SD) (95%)
<i>continuous</i>		
fatigue severity (CIS)	1.37 (1.11 to 1.86)	1.86 (1.31 to 2.61)
• change score t6-to	-	1.70 (1.11 to 2.31)
exhaustion (MBI)	1.66 (1.28 to 2.18)	1.90 (1.33 to 2.74)
• change score t6-to	-	1.55 (1.09 to 2.24)
<i>dichotomous/categorical</i>		
educational level • low versus high	2.61 (1.29 to 5.24)	2.29 (1.08 to 4.86)
• middle versus high	2.38 (1.27 to 4.45)	2.75 (1.38 to 5.46)
self-reported GP visit (yes=1)	3.06 (1.71 to 5.50)	-
• recent visit to GP (t3)	-	2.01 (1.07 to 3.76)
self-reported OP visit (yes=1)	0.46 (0.22 to 0.95)	-
self-rated health (good=1)	0.56 (0.33 to 0.95)	0.27 (0.14 to 0.51)
• change to-t3 'bad to good'	-	0.11 (0.03 to 0.35)
• change t3-t6 'good to bad'	-	2.54 (1.14 to 5.66)

Predictor values are significant odds ratios (95% CI) adjusted for all other predictors in the model. Values of continuous predictors are expressed as OR per SD.

t3 = one-year assessment; t6=two-year assessment; GP=general practitioner; OP=occupational physician.

† Full model included age, sex, educational level, work status, pregnancy, CIS-fatigue, CIS-motivation, CIS-concentration, need for recovery, all three MBI-subscales, impairment in work, impairment in activities, physical attribution of fatigue, psychological distress, depressed mood, anxious mood, life events, visits to GP, visits to occupational physician, self-rated health, health complaints, sleep disturbances, all assessed at baseline.

‡ Full model included educational level, CIS-fatigue (baseline), CIS fatigue change score (t3-to), CIS fatigue change score (t6-to), MBI-exhaustion (baseline), MBI-exhaustion change score (t3-to), MBI-exhaustion change score (t6-to), visits to GP (baseline), visits to GP (t3), visits to GP (t6), visits to OP (baseline), visits to OP (t3), visits to OP (t6), self-rated health (baseline), self-rated health change bad to good (t3-t6), self-rated health change good to bad (t3-t6), self-rated health change bad to good (t3-t6), self-rated health change good to bad (t3-t6).

We found five baseline predictors of non-CFS fatigue caseness as outcome at follow-up as opposed to no fatigue caseness: high fatigue, low self-perceived activity, high exhaustion, high anxious mood and bad self-rated health (table 4). When change scores were added to the model, increases of fatigue and self-perceived activity in time and a change of self-rated health from bad to good between baseline and one-year assessment were found to be significant predictors. In the second step however, exhaustion and self-perceived activity at baseline lost their significance as predictors.

TABLE 4

SIGNIFICANT PREDICTORS OF NON-CFS FATIGUE CASENESS VERSUS NO FATIGUE CASENESS AT 44-MONTH FOLLOW-UP IN EMPLOYEES FATIGUED AT BASELINE (N=1049)

predictors	MULTIPLE LOGISTIC REGRESSION MODELS	
	baseline predictors only† OR (per SD) (95% CI)	baseline and follow-up predictors‡ OR (per SD) (95%)
<i>continuous</i>		
fatigue severity (CIS)	1.24 (1.05 to 1.37)	1.96 (1.61 to 2.26)
• change score t3-to	-	1.35 (1.23 to 1.65)
• change score t6-to	-	2.31 (1.88 to 2.82)
self-perceived activity (CIS)	1.18 (1.04 to 1.35)	-
• change score t6-to	-	0.78 (0.67 to 0.91)
exhaustion (MBI)	1.31 (1.13 to 1.53)	-
anxious mood	1.22 (1.06 to 1.40)	1.29 (1.10 to 1.49)
<i>dichotomous</i>		
self-rated health (good=1)	0.64 (0.48 to 0.85)	0.58 (0.40 to 0.84)
• change to-t3 'bad to good'		0.57 (0.34 to 0.97)

Predictor values are significant odds ratios (95% CI) adjusted for all other predictors in the model. Values of continuous predictors are expressed as OR per SD.

t3 = one-year assessment; t6 = two-year assessment.

† Full model included age, sex, educational level, work status, pregnancy, all four CIS-subscales, all three MBI-subscales, need for recovery, self-rated fatigue complaints, impairment in work, impairment in activities, psychological distress, anxious mood, visits to GP, self-rated health, health complaints, sleep disturbances, all assessed at baseline.

‡ Full model included CIS-fatigue (baseline), CIS-fatigue change score (t3-to), CIS-fatigue change score (t6-to), CIS-activity (baseline), CIS-activity change score (t3-to), CIS-activity change score (t6-to), MBI-exhaustion (baseline), MBI-exhaustion change score (t3-to), MBI-exhaustion change score (t6-to), anxiety (baseline), anxiety change score (t3-to), anxiety change score (t6-to), self-rated health (baseline), self-rated health change bad to good (t0-t3), self-rated health change good to bad (t0-t3), self-rated health change bad to good (t3-t6), self-rated health change good to bad (t3-t6).

Discussion

PREDICTORS OF OUTCOME

If fatigue can best be understood as a continuum of severity ⁹, it is likely that fatigue among employees and the chronic fatigue syndrome are different but related stages on the continuum. ¹⁴ In line with this notion, we found that (persistent) unexplained fatigue can indeed be a precursor of a chronic syndrome fatigue-like status. 48% of the 1143 fatigued employees at baseline were (still) fatigue cases 44 months later, and 94 of them (8% of the total group) met the criteria for CFS by that time, none of who had previously received a CFS diagnosis.

In table 5, we have summarized the most important findings from our prediction analyses. A few of these findings are notable; the magnitude of the effects of fatigue-related factors, educational level and self-rated health, the role of self-reported visits to health care providers and the role of self-perceived activity.

TABLE 5
SUMMARY OF FINDINGS: SIGNIFICANT PREDICTORS ACROSS GROUP COMPARISONS

GROUP COMPARISONS IN LOGISTIC REGRESSION ANALYSIS		
<i>CFS-like cases versus non-CFS fatigue cases</i>	<i>CFS-like cases versus no fatigue cases</i>	<i>non-CFS fatigue cases versus no fatigue cases</i>
<i>demographic variables</i>	<i>demographic variables</i>	
sex		
age		
educational level	education level	
<i>fatigue-related variables</i>	<i>fatigue-related variables</i>	<i>fatigue-related variables</i>
exhaustion	exhaustion	exhaustion
	fatigue severity	fatigue severity
		self-perceived activity
<i>health care seeking</i>	<i>health care seeking</i>	
visiting GP	self-reported GP visits	
	self-reported OP visits	
	<i>self-rated health</i>	<i>self-rated health</i>
		anxiety

GP = general practitioner; OP = occupational physician.

As hypothesized, we found that fatigue severity was a strong predictor of fatigue caseness (both CFS-like and non-CFS) four years later almost, despite a limited range in scores at baseline. Work-related exhaustion emerged as an important predictor as well, underlining the important contribution of fatigue-related factors to the prediction model.

Having a CFS-like status at follow-up (compared to both non-CFS fatigue and no fatigue) was strongly associated with a low educational level, with odds ratios ranging from 2.3 to 4.0. Similarly, having a good self-rated health was a strong predictor of the absence of fatigue at follow-up as opposed to being fatigued (both CFS-like and non-CFS), and the effect increased when changes in self-rated health over time were included as predictors.

We were surprised to find that self-reported visits to general practitioners (GPs) and occupational physicians (OPs) emerged as predictors of outcome 44 months later. CFS-like cases were almost twice as likely to report visiting their GP at baseline compared to non-CFS cases (both fatigued and non-fatigued). Self-reported visits to the OP on the other hand appeared protective of CFS-like caseness at follow-up. However, since visiting an OP was predictive of CFS-like caseness in bivariate analyses (e.g., CFS-like versus no fatigue, $OR=1.91$, results not shown) but turned protective when adjusted for visits to the GP ($OR=0.75$, results not shown), the direction of the association after adjustment should be regarded as a possible artefact resulting from collinearity of predictors in multiple regression. A similar explanation may hold for the direction of self-perceived activity (table 4; change score t6-to, $OR=0.78$). Adding the change score to the prediction model seemed to 'flip' the direction of the initial association between self-perceived activity and fatigue caseness. In fact, additional analyses revealed that *lower* levels of self-perceived activity at two-year follow-up compared to baseline were *negatively* associated with fatigue caseness only after adjusting for change scores in fatigue severity (results not shown).

METHODOLOGICAL CONSIDERATIONS

The design of the Maastricht Cohort Study ¹⁹ offered us the unique opportunity to investigate the development of self-reported chronic fatigue syndrome in a large sample of employees. An obvious limitation and potential source of bias is the use of self-report measures. Self-report questionnaires clearly have limitations in the assessment of the chronic fatigue syndrome, especially when it comes to somatic conditions that might explain fatigue complaints. However, the CFS criteria we applied are regarded as a good proxy for clinically validated CFS. ^{32, 33} Moreover, our approach enables us to identify CFS-like cases using research criteria free from bias regarding health care seeking or illness recognition by doctors and patients, which is a major advantage of the present study. Unfortunately, the full set of CFS criteria was lacking at baseline assessment, and we are not sure whether the selection of potential CFS-like cases at baseline was accurate since we used a provisional set of necessary but insufficient criteria.

PREVIOUS FINDINGS: THE ROLE OF ILLNESS PERCEPTION

Causal attributions have been found to be associated with the course of CFS in CFS patients, with stronger physical attributions or weaker psychological attributions predicting worse outcomes.^{5, 7, 8, 15-18} Chalder *et al.* studied the course of fatigue in a sample of fatigued primary care patients.³⁴ They found that poor outcome was predicted by stronger physical attributions, a long perceived future illness duration and, notably, *not* visiting the GP for an emotional reason. In a submitted paper, we recently reported that recovery from fatigue in employees on long-term sick leave was associated with good self-rated health and a psychological attribution of fatigue.¹⁴ In the present study, self-rated health was predictive of fatigue caseness (CFS-like or non-CFS) but causal attributions were not. Our hypothesis regarding the role of illness perception-related factors was therefore only partially confirmed.

In review of our findings, we conclude that severe, persistent, unexplained fatigue can be a precursor of the development of self-reported chronic fatigue syndrome. In understanding the prominent role of self-rated health as a predictor of CFS-like caseness and the persistence or recurrence of fatigue, Chalder *et al.* have provided an interesting rationale. They have suggested that a more pessimistic view of the illness may encourage symptom focusing, which in turn may lead to the perpetuation of fatigue.³⁴ How other factors play a role in this is yet unknown, but common sense leads us to suggest that the prevention of the chronic fatigue syndrome and the perpetuation of fatigue should in part be aimed at changing the perception of health complaints. Interestingly, such an approach has already shown to be successful in the treatment of CFS.^{35, 36}

Less clear is the role of self-reported visits to the GP or OP as predictor of CFS-like caseness, especially because of the large time delay between exposure and outcome. One explanation may be that visiting a health care provider is an indicator of an underlying concept, e.g. proneness to focus on health complaints. Hamilton *et al.* concluded that CFS patients consulted their GP more frequently in the 15 years before the development of their condition, possibly because symptoms were perceived more readily as illness or due to an increase in the tendency to consult a doctor.³⁷ Afari *et al.* on the other hand have suggested that the probability of receiving a diagnosis of CFS is related to the degree of health care access and utilization.² However, one should keep in mind that none of the CFS-like cases in this study had received a CFS diagnosis. Interesting in that respect is the fact that most CFS-like cases in this study were males with low education, as opposed to most diagnosed CFS patients who are usually professionally successful women.²

The considerations described here suggest several questions. Is health care seeking a differential factor in the persistence of fatigue or the development of a CFS-like status? What are the factors contributing to receiving a diagnosis of CFS? Do sociodemographic factors, access and utilization of health care or illness perception have a role in this? And if so, what does it tell us about potential approaches to prevention and treatment? These are important questions for further research.

Acknowledgments

This study was funded by the Health Research and Development Council (ZorgOnderzoek Nederland), The Netherlands (grant nr. 2830180). The Maastricht Cohort Study is part of The Netherlands concerted action on 'Fatigue at Work' granted by The Netherlands Organization for Scientific Research (NWO).

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. Afari N, Buchwald D. Chronic Fatigue Syndrome: a Review. *Am J Psychiatry* 2003; 160:221-36.
3. Peterson PK, Schenck CH, Sherman R. Chronic fatigue syndrome in Minnesota. *Minn Med* 1991; 74:21-6.
4. Tirelli U, Marotta G, Improtà S, Pinto A. Immunological abnormalities in patients with chronic fatigue syndrome. *Scand J Immunol* 1994; 40:601-8.
5. Wilson A, Hickie I, Lloyd A. Longitudinal study of outcome of chronic fatigue syndrome. *BMJ* 1994; 308:756-759.
6. Bombardier CH, Buchwald D. Outcome and prognosis of patients with chronic fatigue vs chronic fatigue syndrome. *Arch Intern Med* 1995; 155:2105-10.
7. Vercoulen JHMM, Swanink CM, Fennis JF, Galama JM, Van der Meer JW, Bleijenberg G. Prognosis in chronic fatigue syndrome: a prospective study on the natural course. *J Neurol Neurosurg Psychiatry* 1996; 60:489-94.
8. Van der Werf SP, De Vree B, Alberts M, Van der Meer JWM, Bleijenberg G. Natural course and predicting self-reported improvement in patients with chronic fatigue syndrome with a relatively short illness duration. *J Psychosom Res* 2002; 53:749-53.
9. Lewis G, Wessely S. The epidemiology of fatigue: more questions than answers. *J Epidemiol Community Health* 1992; 46:92-7.
10. Bültmann U, Kant IJ, Kasl SV, Beurskens AJHM, Van den Brandt PA. Fatigue and psychological distress in the working population: psychometrics, prevalence and correlates. *J Psychosom Res* 2002; 52:445-52.
11. Janssen N, Kant IJ, Swaen GMH, Janssen PPM, Schröer CAP. Fatigue as a predictor of sickness absence: results from the Maastricht cohort study on fatigue at work. *Occup Environ Med* 2003; 60 (Suppl 1):i71-i6.
12. Amelsvoort van LGPM, Kant IJ, Beurskens AJHM, Schröer CAP, Swaen GMH. Fatigue as a predictor of work disability. *Occup Environ Med* 2002; 59:712-3.
13. Huibers MJH, Beurskens AJHM, Prins JB, et al. Fatigue, burnout and chronic fatigue syndrome among employees on sick leave: do attributions make the difference? *Occup Environ Med* 2003; 60 (Suppl 1):i26-i31.
14. Huibers MJH, Bleijenberg G, Amelsvoort van L, et al. Predictors of outcome in fatigued employees on sick leave: results from a randomised trial. Submitted 2003.
15. Sharpe M, Hawton K, Seagroatt V, Pasvol G. Follow up of patients presenting with fatigue to an infectious diseases clinic. *BMJ* 1992; 305:147-52.
16. Cope H, David A, Pelosi A, Mann A. Predictors of chronic "postviral" fatigue. *Lancet* 1994; 344:864-8.
17. Clark MR, Katon W, Russo J, Kith P, Sintay M, Buchwald D. Chronic fatigue: risk factors for symptom persistence in a 2 1/2-year follow-up study. *Am J Med* 1995; 98:187-95.
18. Chalder T, Power MJ, Wessely S. Chronic fatigue in the community: 'a question of attribution'. *Psychol Med* 1996; 26:791-800.
19. Kant IJ, Bültmann U, Schroer CAP, Beurskens AJHM, Van Amelsvoort LGPM, Swaen GMH. An epidemiological approach to study fatigue in the working population: The Maastricht Cohort Study. *Occup Environ Med* 2003; 60 (Suppl 1):i32-i9.
20. Wessely S, Chalder T, Hirsch S, Pawlikowska T, Wallace P, Wright DJ. Postinfectious fatigue: prospective cohort study in primary care. *Lancet* 1995; 345:1333-8.
21. Taylor RR, Jason LA, Curie CJ. Prognosis of chronic fatigue in a community-based sample. *Psychosom Med* 2002; 64:319-27.

22. Skapinakis P, Lewis G, Mavreas V. One-year outcome of unexplained fatigue syndromes in primary care: results from an international study. *Psychol Med* 2003; 33:857-66.
23. Vercoulen JHMM, Swanink CM, Fennis JF, Galama JM, Van der Meer JW, Bleijenberg G. Dimensional assessment of chronic fatigue syndrome. *J Psychosom Res* 1994; 38:383-92.
24. Beurskens AJHM, Bültmann U, Kant IJ, Vercoulen JHMM, Bleijenberg G, Swaen GMH. Fatigue among working people: validity of a questionnaire measure. *Occup Environ Med* 2000; 57:353-7.
25. Vercoulen JHMM, Alberts M, Bleijenberg G. De Checklist Individual Strength (CIS). *Gedragstherapie* 1999; 32:31-36.
26. Schaufeli WB, Leiter MP, Maslach C, Jackson SE. The Maslach Burnout Inventory-General Survey. In: Maslach C, Jackson SE, eds. *Maslach Burnout Inventory Manual*. Palo Alto, CA: Consulting Psychologists Press, 1996.
27. Sluiter JK, Van der Beek AJ, Frings-Dresen MHW. The influence of work characteristics on the need for recovery and experienced health: a study on coach drivers. *Ergonomics* 1999; 42:573-83.
28. Jansen NWH, Kant IJ, Van den Brandt PA. Need for Recovery in the working population: descriptions and associations with fatigue and psychological distress. *Int J Behav Med* 2002; 9:322-40.
29. Goldberg DP, Williams P. *The user's guide to the General Health Questionnaire*. Windsor: NFER-Nelson, 1991.
30. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; 30:473-83.
31. Aaronson NK, Muller M, Cohen PD, et al. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. *J Clin Epidemiol* 1998; 51:1055-68.
32. Taylor RR, Jason LA, Torres A. Fatigue rating scales: an empirical comparison. *Psychol Med* 2000; 30:849-56.
33. Jason LA, Taylor RR. Applying cluster analysis to define a typology of chronic fatigue syndrome in a medically-evaluated, random community sample. *Psychol Health* 2002; 17:323-37.
34. Chalder T, Godfrey E, Ridsdale L, King M, Wessely S. Predictors of outcome in a fatigued population in primary care following a randomized controlled trial. *Psychol Med* 2003; 33:283-7.
35. Prins JB, Bleijenberg G, Bazelmans E, et al. Cognitive behaviour therapy for chronic fatigue syndrome: a multicentre randomised controlled trial. *Lancet* 2001; 357:841-7.
36. Whiting P, Bagnall AM, Sowden AJ, Cornell JE, Mulrow CD, Ramirez G. Interventions for the treatment and management of chronic fatigue syndrome: a systematic review. *JAMA* 2001; 286:1360-8.
37. Hamilton WT, Hall GH, Round AP. Frequency of attendance in general practice and symptoms before development of chronic fatigue syndrome: a case-control study. *Br J Gen Pract* 2001; 51:553-8.

General discussion

In this final chapter, I will reflect on some of the major findings described in previous chapters. One approach would be to address and discuss all of our findings in great detail. Instead, I choose to highlight several issues that seem important and essential to me personally and search for explanations, interpretations and integration of findings. For the sake of transparency, I will do so by reviewing the research themes and questions of this thesis that were introduced in the first chapter.

The randomised controlled trial

MAIN FINDINGS

Perhaps the most striking finding of this thesis is the lack of efficacy in our randomised controlled trial. Cognitive behaviour therapy (CBT) delivered by general practitioners (GPs) had no substantial effect on any of the outcomes, as was described in *chapter 4*, despite our efficacy design (does it work under ideal research circumstances?) and the fact that the two research conditions were highly contrasted. In addition, analyses in subgroups of patients stratified according to prognostic factors (e.g., fatigue severity, duration absenteeism, CFS-like status) yielded no effects of CBT compared to the control condition.

The three main research questions related to the efficacy of the intervention (*chapter 1*) can now be answered. Firstly, our research GPs were not able to deliver the CBT intervention appropriately. Secondly, the short CBT programme that we modified for use in primary care was not effective in the treatment of fatigued employees. And thirdly, the resumption of work of patients in the experimental group could not be attributed to our CBT intervention. These findings should be interpreted in relation to our choice of design. An additional benefit of efficacy studies is that they are very informative when the results are negative: if it does not work under ideal circumstances, it is unlikely the intervention will be of value in daily practice.¹

THE TRIAD OF EFFICACY

In *chapter 4*, we introduced a 'triad' of efficacy as a conceptual framework for the interpretation of trial findings. In short, we argued that findings in efficacy (or effectiveness) studies always result from the interaction between the intervention-giver (i.e., the doctor), the intervention-receiver (i.e., the patient) and the intervention itself. The lack of efficacy in our trial should be considered as a disturbance in this triad. We addressed the issue briefly in the discussion section of *chapter 4*, and here I take the opportunity to elaborate a little further on potential sources of the disturbance. Before I do, some food for thought: how come we only start to look for explanations when the results of a trial are negative?

In *chapter 5*, we examined the delivery of the CBT intervention by our research GPs. We found no qualitative association between recovery among patients in our trial on the one hand and the way in which the intervention was delivered on the other. This finding is a stern indication that the lack of efficacy can partly be attributed to the inadequacy of the intervention, whether it is the intervention itself or its delivery by our research GPs. At the same time, we concluded it is very likely that the intervention was simply too difficult or sophisticated for our relatively inexperienced GPs to deliver, a notion that we could not confirm based on data.

A third explanation pertains to the selection of patients in our trial. In designing the study, we were faced with the difficulty of defining the 'stage of disease' the intervention should be aimed at for optimal benefits. In *chapter 1*, we introduced the concept of fatigue as a continuum. If we were to select employees on the far left side of the continuum, with mild fatigue complaints for a short duration of time, spontaneous recovery in the control condition would be high and we would have a hard time upholding the necessity of our intervention towards fellow researchers, doctors and patients. However, if we were to select employees on the far right side of the continuum, with severe, disabling fatigue that resembles the chronic fatigue syndrome (CFS) in intensity and duration, a short treatment in primary care would be inappropriate. In the absence of empirical information to guide our decision, we chose to select fatigued employees with severe fatigue, but not too severe, who were on sick leave for some time, but not too long. Much to our surprise, the general impairment of selected patients turned out more advanced than we initially expected, with 44% of the patients at baseline already meeting criteria for CFS. Perhaps, the severity of fatigue and the duration of sick leave in our trial population were too advanced after all for a potential response to our intervention in primary care.

IMPLICATIONS OF OUR TRIAL FINDINGS

Do our findings mean that cognitive behaviour therapy delivered by general practitioners is ineffective in general? A saying states that one swallow does not make a summer. We were the first to study this particular intervention delivered by general practitioners, and in that sense, the evidence-base on the subject has to grow substantially before definite conclusions can be drawn. However, I feel that we should not go down this road any further.

GPs are the gatekeepers of health care. Their medical responsibilities are extensive, the many demands that are placed on them are intensive and their available time is limited and costly. Proceeding insight has led me to believe that complex interventions such as CBT should not belong to the task of the already overloaded general practitioner.

This view is partly based on the findings of the systematic review that we have described in *chapter 2*. Evidence on the effectiveness, or ineffectiveness, of psychosocial interventions such as CBT delivered by GPs was scarce and unconvincing. No reason to discourage initiatives to investigate the effects

of psychosocial interventions by GPs, one could say. However, the great difficulty we and others before us experienced in trying to establish a positive effect of psychosocial interventions by GPs in a highly controlled research setting should raise some doubts about the potential benefits in everyday practice. In a study that was not included in our review because it lacked clinical outcomes, King et al.² found that it was ineffective to teach GPs cognitive behavioural skills. This finding corresponds with the notion that it is difficult to master complex interventions skills such as CBT, even when GPs are motivated, trained and experienced to some extent.

Our negative trial findings do not reflect on the necessity to offer psychosocial interventions to patients in primary care. Many patients visit their GP because of problems that are psychological in origin, and many of these patients do not receive the help they need in primary care. There are indications that psychosocial interventions in primary care can be helpful. In a review for the Cochrane Collaboration, Bower and colleagues for example concluded that counselling in primary care is effective in the short-term.³

A relevant question in Dutch primary care is, who should deliver these interventions? In some local settings, GPs actively cooperate with primary care psychologists and social workers, but unfortunately, this has not led to a structural form of cooperation in the Dutch primary care system. Another possibility we have described in *chapter 2* is the deployment of nurse practitioners, a professional discipline still in its development in The Netherlands but whose place in primary care is greatly advocated by health care policy makers. In a recent review, Horrocks and colleagues concluded that nurse practitioners in primary care provide equivalent care and excite more satisfaction among patients compared to GPs.⁴ Consequently, a challenge for future research might be to generate more evidence on the effectiveness of psychosocial interventions delivered by nurse practitioners, primary care psychologists and social workers in close collaboration with the gate keeping GP.

Considerations for conducting (psychotherapy) trials in primary care

Conducting the randomised trial described in this thesis has been a major experience-base for all of us who were involved. Here, I would like to share a few lessons that we have learned along the way. Some of these issues are already addressed in *chapter 3*. I will focus on the *blinding of conditions*, *recruitment procedures*, *deployment of research GPs* and *parameters of study quality*.

A major methodological limitation of psychotherapy research is the fact that *blinding of research conditions* is virtually impossible. By definition, treatment givers and treatment receivers in psychotherapy cannot be oblivious to the nature of the intervention. In fact, one of the mechanisms of psychotherapy is becoming aware of the intervention process. Unfortunately, the blinding of research

conditions is essential for internal validity. It serves two purposes: standardisation of treatment circumstances and measurement of effects unaffected by treatment allocation bias.⁵ Although true blinding is not an option, we can try and think of ways to compensate for the lack of it. In our trial for example, we came up with a prerandomisation design. By informing patients only about the condition they were allocated to, thus blinding them for the existence of another condition, we tried to prevent selective withdrawal from the study and contamination of the research conditions. Similar solutions to minimize the problem of blinding should be encouraged in psychotherapy research. In *chapter 3*, we described our considerations for applying an alternative trial design. One of the obstacles we were faced with was the *recruitment* of patients and GPs, a common challenge in primary care research.^{6,7} We, and many others before us, experienced that things always went wrong in situations where we had to depend on the efforts and goodwill of others. An explanation for this universal phenomenon is that no GP, practice assistant or psychotherapist is as convinced of the importance to undertake the necessary research actions (e.g., screening, recruitment, selection, intervention, measurements) as the researchers whose careers depend on output. A general strategy for researchers to overcome recruitment problems and other obstacles is to maintain control over as many research procedures as possible and limit the involvement of clinicians, something that might not always be feasible.⁸

An important element in our alternative trial design was the *deployment of a small number of research GPs* who delivered the intervention, instead of the regular GPs of patients, which is customary in primary care research. Our available time and resources enabled us to train and supervise only a limited number of GPs. If these GPs had to recruit patients from their own practice, our study sample would inevitably be too small. As a result, we decided to recruit patients and research GPs separately. When we first introduced our plans, we received a lot of critical remarks from colleagues who questioned the transferability of findings of such a trial. It was argued that the personal relationship between the patient and his regular GP is an essential element in any kind of intervention, and that our trial would violate this aspect. Our persistence to deploy research GPs despite these comments relates to the objective of our trial. We wanted to assess whether our intervention could be delivered by GPs in the first place, a question of efficacy; transferability to daily practice was not our main concern. In our case, practical reasons outweighed the criticism from colleagues. Other researchers may face a similar dilemma. Their objections to having many regular GPs recruiting and treating a few patients can be several: recruitment by regular GPs usually falls short of expectation^{7,9} and may even lead to a premature ending of the trial¹⁰; the costs and time of training and supervising may mount to unacceptable levels; it may be necessary to have GPs gain substantial experience in delivering a (complex) intervention by treating many patients. If these researchers aim to conduct efficacy research, the deployment of research GPs may solve some of their problems.

A final issue pertains to psychotherapy research in specific. In the systematic review on the effectiveness of psychosocial interventions by general practitioners (*chapter 2*), we concluded that select-

ed studies reported poorly on the *parameters of study quality*, such as the supervision and training of GPs, the use of treatment protocols, the application of integrity checks (i.e., assessing the quantity but also the quality of the intervention delivered), the compliance of patients, the content of usual care, the co-interventions received by patients, and the subgroup or prognostic analyses that were performed. This lack of vital information made it difficult to interpret the results, and we have tried to do better in the report of our own trial (*chapter 4*). Especially in the field of psychotherapy research, in which the lack of blinding is a threat to internal validity, high-quality descriptions of all factors that might introduce bias are of great importance.

Fatigue among employees and the chronic fatigue syndrome

The second research theme in this thesis concerns the relation between our study subject, fatigue among employees, and the chronic fatigue syndrome. The rationale for this research theme and the underlying main research questions are described in *chapter 1*. A general overview of the findings presented in this thesis confirms the notion that fatigue among employees and the chronic fatigue syndrome (CFS) are linked in many different ways, in line with the concept of fatigue as a continuum of severity.¹¹

MAIN FINDINGS

Are fatigue among employees and CFS different but related stages on the continuum of fatigue severity? In *chapter 6* and *8*, we presented evidence in favour of this hypothesis. We found that fatigued employees and diagnosed CFS patients shared many important characteristics. Interestingly, 44% of the fatigued employees already met the criteria for CFS (i.e., were CFS-like fatigue cases) at baseline of the trial (*chapter 8*). This finding suggests that the conditions of fatigue among employees and CFS can overlap. Also, in the course of 12 months, some of the employees in our trial shifted back and forth between a CFS-like status and non-CFS fatigue caseness (*chapter 6*).

Can fatigue among employees eventually lead to the development of CFS-like caseness? Again, our hypothesis was confirmed. In *chapter 6*, we observed fatigued employees in our trial who developed a CFS-like status in the course of time, some of who returned to be non-CFS fatigue cases or had recovered after 12 months. In *chapter 10*, we followed fatigued employees for almost four years in the Maastricht Cohort Study and found that a substantial portion of the participants had developed a CFS-like status at follow-up. Based on these findings, we concluded that fatigue among employees can be a precursor of chronic fatigue syndrome.

A third important finding emerged from our studies, a finding that was indirectly associated with

the two main research questions. In previous studies of CFS, it has been reported that most CFS patients do not recover spontaneously.¹²⁻¹⁷ In three separate prospective analyses, we found that severe fatigue at baseline persisted in many participants over an extended period of time. In *chapter 6*, we reported that 57% of the patients in our trial had not recovered from fatigue after 12 months. In *chapter 7*, we found that 60% of the fatigued employees that were followed in the Maastricht Cohort Study were fatigue cases after 24 months of follow-up. Similarly, in *chapter 10*, we found that 48% of the fatigued employees we selected at baseline of the Maastricht Cohort Study were fatigue cases after 44 months of follow-up. Overall, we found that high severity of fatigue at baseline, in the present example defined as scoring above the cut-off of 35 on the CIS^{18,19}, is a strong predictor of the persistence of fatigue in employees, in line with prospective studies of CFS.²⁰⁻²² These findings are particularly convincing because they emerged in two different study populations.

CFS PATIENTS VERSUS CFS-LIKE FATIGUED EMPLOYEES: A MATTER OF SELECTION

There is a vast body of literature devoted to the chronic fatigue syndrome.²³ In many instances, it is tempting to compare the findings from our studies of fatigued employees to findings reported in previous CFS studies. However, these comparisons lack validity to a certain degree because of a fundamental methodological problem, the difficulty of transference. Most studies of CFS are conducted in clinical settings among diagnosed CFS patients who have actively sought help in primary or secondary care. This approach inevitably introduces some form of (self-)selection bias.

In the studies presented in this thesis, study participants were selected in the working population. We selected fatigued employees for our trial by screening the population of an occupational health service, a procedure that was initially based on duration of absenteeism, not on health complaints. In the Maastricht Cohort Study, employees from varying companies were invited to participate in a prospective study regardless of their fatigue status.

This lack of bias regarding health care seeking or illness recognition as a result of our recruitment strategies is reflected by the fact that the majority of the CFS-like fatigue cases we identified using objective criteria for CFS had not received a prior diagnosis of CFS, i.e., were not recognized as CFS patients. Hence, our CFS-like fatigue cases may differ significantly from the diagnosed CFS patients of other studies. This is not to say that the CFS-like fatigue cases in our studies are no true CFS cases. As was stated in previous chapters, a CFS-like caseness is regarded as a good proxy for CFS diagnosed by a physician^{24,25} and CFS-like fatigue cases are likely to meet the CFS criteria even after a physical examination. Rather, one of the major differences between undiagnosed CFS-like fatigue cases and diagnosed CFS patients in the present thesis relates to health care seeking, an important factor as I will discuss later on. Consequently, whenever we draw comparisons between our study population and CFS patients, some caution is in place.

ROLE OF ILLNESS PERCEPTION

In studying the relation between fatigue among employees and CFS from different angles, we came to focus on the role of illness perception. Earlier studies of the prognosis of CFS already reported on the role of illness perception in the chronic fatigue syndrome. Overall, Afari and Buchwald stated, studies of perception suggest that the ways in which CFS patients perceive themselves, label their symptoms, and appraise stressors may perpetuate or exacerbate their physical and psychosocial dysfunction.²³ Most of these studies focused mainly on the role of causal attributions, and found that stronger physical attributions or weaker psychological attributions predicted worse outcomes.^{14, 16, 17, 26-30} Following from this, we hypothesized that factors reflecting the perception of illness, such as self-rated health and causal attributions, would be play an important role in fatigued employees as well.

In *chapter 8*, we found that diagnosed CFS patients and CFS-like fatigued employees were more likely to attribute their fatigue complaints to a somatic cause compared to fatigued employees who did not meet research criteria for CFS. Similarly, in *chapter 6*, psychological attributions, somatic attributions and other perception-related factors such as self-rated health were found to differentiate CFS-like fatigue cases and non-CFS fatigued cases in the course of time. In line with the aforementioned studies of CFS, illness perception also emerged as an important predictor of outcome in fatigued employees. A psychological attribution of the fatigue complaints (in *chapters 6 & 7*) and a good self-rated health (in *chapters 6, 7 & 10*) were found to be predictors of recovery from fatigue in time. Our finding that illness perception plays a role in fatigued employees provides additional evidence that fatigue among employees and CFS are different but related stages on the fatigue continuum.

How can the role of illness perception be explained? Explanations offered by other authors are restricted to causal attributions. Van der Werf and colleagues have suggested that, as fatigue persists in time and no medical answers become available, the need for a somatic causal explanation will strengthen.¹⁷ Others view illness attribution as predisposing factor. Butler and colleagues have suggested that the tendency to make somatic attributions might be a vulnerability factor for the development of CFS.³¹ Chalder and colleagues consecutively have suggested that a more pessimistic view of the illness may encourage symptom focusing, which in turn may lead to the perpetuation of fatigue.³⁰

ILLNESS PERCEPTION, HEALTH CARE SEEKING AND THE EFFECT OF DIAGNOSIS

In this chapter, we came to distinguish three fatigue states: fatigue that does not meet the criteria for CFS, fatigue that meets the criteria for CFS but is not diagnosed as such, and diagnosed chronic fatigue syndrome. We described the prevalence of a CFS-like status in the working population in *chapter 9*. Interestingly, the vast majority of fatigued employees who met criteria for CFS had not received a diagnosis of chronic fatigue syndrome. In addition, we found that the majority of CFS-like fatigued employees were low educated males, a surprising finding considering that most diag-

nosed CFS patients are professionally successful women.²³ Afari and Buchwald have suggested that the probability of receiving a diagnosis of CFS is related to the degree of health care access or utilization.²³ Embroidering on the latter suggestion, one could suggest that proneness to focus on health complaints and readiness to engage in a lengthy process of vigorous health care seeking are important determinants of the probability of receiving a CFS diagnosis. Conceptually, these factors can both be considered as reflections of illness perception.

These considerations lead me to propose a hypothesis on illness perception and diagnosis. Starting point is the assumption that focusing on health complaints and readiness to engage in vigorous health care seeking are related reflections of a pessimistic illness perception that are potentially harmful, and will stimulate the determination to receive a diagnosis that matches one's fatigue complaints. If the diagnosis thus acquired is chronic fatigue syndrome, receiving the CFS diagnosis will trigger and/or validate perceptions of ill health, which in turn leads to the exacerbation and perpetuation of fatigue complaints (see illustration below). This hypothesis implies that those who receive a diagnosis of CFS are more likely to have a worse prognosis compared to CFS-like fatigue cases who do not receive a diagnosis, partly due to differences in illness perception.

THE PROPOSED HYPOTHESIS ON ILLNESS PERCEPTION AND DIAGNOSIS IN CFS-LIKE FATIGUE CASES

The rationale behind the proposed hypothesis is that receiving the diagnosis of CFS may contain a 'sickening' message to those who are receiving it. In account of the enormous amount of information on CFS that can be found on the Internet for instance, some CFS patients tend to view their complaints as an awful, mysterious, unexplainable illness from which they most likely will never recover.

If this message takes root in a person suffering from fatigue, whether it is prompted by personal beliefs, comments by others or the negative attitude of a physician who is inexperienced with the diagnosis of CFS, the consequences can be comprehensive. Interestingly, Hazemeijer and Rasker recently have argued that physicians should stop diagnosing the syndrome because the diagnosis itself generally leads to an aggravation of complaints.³² They were not discussing the diagnosis of CFS, notably, but the diagnosis of fibromyalgia, a medically unexplained condition that is documented to be a functional syndrome as well.^{33, 34}

The hypothesis stated here provides an additional, more controversial explanation for the lack of efficacy in our randomised controlled trial. We selected our patients by screening the population of an occupational health service and inviting eligible patients for research participation. Consequently, our trial patients were no help seekers by definition. As was described in *chapter 3*, we applied a randomised consent design in which patients were only informed about the research condition they were allocated to. Patients in the control group were merely asked to be followed for 12 months. Patients in the experimental group however were offered an experimental intervention for their fatigue complaints, something that came as a surprise to many of them. Is it possible that being offered a treatment for a thus far unlabelled complaint has an impact similar to receiving a diagnosis? And is it possible then that we unknowingly have encouraged or confirmed perceptions of ill health in our experimental group, creating an inequality between the two research conditions in favour of the control group? Alas, we did not design our trial to expose the source of disturbance accountable for the lack of efficacy, and it seems somewhat unlikely that the inequality suggested here could have had a major part in it. However, from a methodological standpoint, we should take into account that participation in a research project may have an effect in itself, a phenomenon that has been referred to as the Hawthorne-effect. If these effects are distributed unequally among the research conditions, a serious form of bias is introduced.

Future research

In general, the findings presented in this thesis are often just a first step and many of the important issues discussed in this chapter remain unclear. Therefore, I will present a few loosely defined suggestions for future research.

In review of the findings discussed here, I conclude that individuals with severe, persistent fatigue need professional help, whether they are employees or not, meet CFS criteria or not, or can be classified differently. As we have seen, severe, persistent fatigue generally has a poor prognosis. But what are the treatment options that are eligible for evaluation?

One possibility would be to offer fatigued employees the cognitive behaviour therapy (CBT) by professional therapists that has proven to be effective in the treatment of chronic fatigue syndrome.³⁵ ³⁶ We found that the fatigued employees in our studies already occupied an advanced stage on the fatigue continuum, approximating the severity of fatigue that can be observed in CFS patients. We also found that illness perception plays an important role in fatigued employees as well. Markedly, an important element in CBT for CFS is changing the illness beliefs that perpetuate the fatigue complaints, an approach that is likely to be effective when performed by experienced, well-trained professionals. With this suggestion for future research, we now come full circle in our reasoning: initially, we chose to evaluate a modified CBT programme for fatigued employees *because* CBT by professional therapists had proven to be effective in CFS patients.

Another challenge for future research that builds on our findings relates to the alleged differences between CFS patients and CFS-like fatigue cases. Research questions that follow from the hypothesis proposed in the previous paragraph are several: which factors determine whether fatigued individuals who meet the criteria for CFS receive a CFS diagnosis? Is this process mediated by illness perception? What is the impact of receiving a CFS diagnosis on illness perception? Do diagnosed CFS patients have worse outcomes in time compared to CFS-like fatigue cases? An appropriate manner to address these research questions is to compare diagnosed CFS patients and undiagnosed CFS-like fatigue cases directly in a prospective cohort study.

Little is known about potential prevention strategies for the development of the chronic fatigue syndrome. One approach to this complex research issue would be to aim (secondary) prevention strategies at fatigue conditions that were found to be preliminary stages of CFS. We found that fatigue among employees can be a precursor of CFS-like caseness, with a negative illness perception as a possible vulnerability for the development of CFS. Accordingly, early detection and prevention of CFS could be aimed at fatigued employees, with a special focus on illness perception.

References

1. Huibers MJH, Knipschild P. Methodological aspects of randomized controlled trials of psychotherapy research in primary care: correspondence. *Psychol Med* 2003; in press.
2. King M, Davidson O, Taylor F, Haines A, Sharp D, Turner R. Effectiveness of teaching general practitioners skills in brief cognitive behaviour therapy to treat patients with depression: randomised controlled trial. *BMJ* 2002; 324:947-53.
3. Bower P, Rowland N, Hardy R. The clinical effectiveness of counselling in primary care: a systematic review and meta-analysis. *Psychol Med* 2003; 33:203-15.
4. Horrocks S, Anderson E, Salisbury C. Systematic review of whether nurse practitioners working in primary care can provide equivalent care to doctors. *BMJ* 2002; 324:819-23.
5. Knipschild P. The false positive therapeutic trial. *J Clin Epidemiology* 2002; 55:1191-5.
6. Ward E, King M, Lloyd M, Bower P, Friedli K. Conducting randomized trials in general practice: methodological and practical issues. *Br J Gen Pract* 1999; 49:919-22.
7. Van der Windt DAWM, Koes BW, Van Aarst M, Heemskerk MAMB, Bouter LM. Practical aspects of conducting a pragmatic randomised trial in primary care: patient recruitment and outcome assessment. *Br J Gen Pract* 2000; 50:371-4.
8. Huibers MJH, Van der Windt DAWM, Boeke AJP. De deelname van huisartsen aan wetenschappelijk onderzoek: problemen en oplossingen (Participation of general practitioners in primary care research: problems and solutions). *Huisarts Wet* 2002; 45:454-8.
9. Peto V, Coulter A, Bond A. Factors affecting general practitioners' recruitment of patients into a prospective study. *Fam Pract* 1993; 10:207-11.
10. Tognoni G, Alli C, Avanzini F, et al. Randomised clinical trials in general practice: lessons from a failure. *BMJ* 1991; 303:969-71.
11. Lewis G, Wessely S. The epidemiology of fatigue: more questions than answers. *J Epidemiol Community Health* 1992; 46:92-7.
12. Peterson PK, Schenck CH, Sherman R. Chronic fatigue syndrome in Minnesota. *Minn Med* 1991; 74:21-6.
13. Tirelli U, Marotta G, Improta S, Pinto A. Immunological abnormalities in patients with chronic fatigue syndrome. *Scand J Immunol* 1994; 40:601-8.
14. Wilson A, Hicki I, Lloyd A. Longitudinal study of outcome of chronic fatigue syndrome. *BMJ* 1994; 308:756-9.
15. Bombardier CH, Buchwald D. Outcome and prognosis of patients with chronic fatigue vs chronic fatigue syndrome. *Arch Intern Med* 1995; 155:2105-10.
16. Vercoulen JHMM, Swanink CM, Fennis JF, Galama JM, Van der Meer JW, Bleijenberg G. Prognosis in chronic fatigue syndrome: a prospective study on the natural course. *J Neurol Neurosurg Psychiatry* 1996; 60:489-94.
17. Van der Werf SP, De Vree B, Alberts M, Van der Meer JWM, Bleijenberg G. Natural course and predicting self-reported improvement in patients with chronic fatigue syndrome with a relatively short illness duration. *J Psychosom Res* 2002; 53:749-53.
18. Vercoulen JHMM, Swanink CM, Fennis JF, Galama JM, Van der Meer JW, Bleijenberg G. Dimensional assessment of chronic fatigue syndrome. *J Psychosom Res* 1994; 38:383-92.
19. Beurskens AJHM, Bültmann U, Kant IJ, Vercoulen JHMM, Bleijenberg G, Swaen GMH. Fatigue among working people: validity of a questionnaire measure. *Occup Environ Med* 2000; 57:353-7.

20. Wessely S, Chalder T, Hirsch S, Pawlikowska T, Wallace P, Wright DJ. Postinfectious fatigue: prospective cohort study in primary care. *Lancet* 1995; 345:1333-8.
21. Taylor RR, Jason LA, Curie CJ. Prognosis of chronic fatigue in a community-based sample. *Psychosom Med* 2002; 64:319-27.
22. Skapinakis P, Lewis G, Mavreas V. One-year outcome of unexplained fatigue syndromes in primary care: results from an international study. *Psychol Med* 2003; 33:857-66.
23. Afari N, Buchwald D. Chronic Fatigue Syndrome: a Review. *Am J Psychiatry* 2003; 160:221-36.
24. Taylor RR, Jason LA, Torres A. Fatigue rating scales: an empirical comparison. *Psychol Med* 2000; 30:849-56.
25. Jason LA, Taylor RR. Applying cluster analysis to define a typology of chronic fatigue syndrome in a medically-evaluated, random community sample. *Psychol Health* 2002; 17:323-37.
26. Sharpe M, Hawton K, Seagroatt V, Pasvol G. Follow up of patients presenting with fatigue to an infectious diseases clinic. *BMJ* 1992; 305:147-52.
27. Cope H, David A, Pelosi A, Mann A. Predictors of chronic "postviral" fatigue. *Lancet* 1994; 344:864-8.
28. Clark MR, Katon W, Russo J, Kith P, Sintay M, Buchwald D. Chronic fatigue: risk factors for symptom persistence in a 2 1/2-year follow-up study. *Am J Med* 1995; 98:187-95.
29. Chalder T, Power MJ, Wessely S. Chronic fatigue in the community: 'a question of attribution'. *Psychol Med* 1996; 26:791-800.
30. Chalder T, Godfrey E, Ridsdale L, King M, Wessely S. Predictors of outcome in a fatigued population in primary care following a randomized controlled trial. *Psychol Med* 2003; 33:283-287.
31. Butler JA, Chalder T, Wessely S. Causal attributions for somatic sensations in patients with chronic fatigue syndrome and their partners. *Psychol Med* 2001; 31:97-105.
32. Hazemeijer I, Rasker JJ. Fibromyalgia and the therapeutic domain. A philosophical study on the origins of fibromyalgia in a specific social setting. *Rheumatology* 2003; 42:507-15.
33. Wessely S, Nimnuan C, Sharpe M. Functional somatic syndromes: one or many? *Lancet* 1999; 354:936-9.
34. Mayou R, Farmer A. ABC of psychological medicine: Functional somatic symptoms and syndromes. *BMJ* 2002; 325:265-8.
35. Prins JB, Bleijenberg G, Bazelmans E, et al. Cognitive behaviour therapy for chronic fatigue syndrome: a multicentre randomised controlled trial. *Lancet* 2001; 357:841-7.
36. Whiting P, Bagnall AM, Sowden AJ, Cornell JE, Mulrow CD, Ramirez G. Interventions for the treatment and management of chronic fatigue syndrome: a systematic review. *JAMA* 2001; 286:1360-8.

Samenvatting (Dutch summary)

Vermoeidheid onder werknemers wordt alom erkend als een belangrijk gezondheidsprobleem. In dit proefschrift staat de behandeling van ernstig vermoeide werknemers in de eerste lijn centraal. Aan de hand van een aantal wetenschappelijke studies hebben wij geprobeerd om een beter zicht te krijgen op de effectiviteit van cognitieve gedragstherapie door huisartsen bij vermoeide werknemers die verzuimen en op de relatie tussen vermoeidheid onder werknemers en het chronisch vermoeidheidssyndroom. In deze appendix geef ik een beknopte samenvatting van de afzonderlijke hoofdstukken zoals die in dit proefschrift beschreven zijn. Voor een volledig beeld van de materie verwijs ik de lezer echter met nadruk naar de hoofdstukken zelf. Het discussie hoofdstuk aan het einde (hoofdstuk 11) is daarbij een goede leidraad om zicht te krijgen op de belangrijkste bevindingen van ons onderzoek.

Hoofdstuk 1 vormt de inleiding van het proefschrift. Allereerst ga ik in op het concept vermoeidheid. Vermoeidheid kan het beste beschouwd worden als continuüm, variërend van milde klachten die iedereen kent tot zeer ernstige klachten zoals we die zien bij het chronisch vermoeidheidssyndroom. In dit proefschrift richten wij ons op ernstige, aanhoudende, medisch onverklaarde vermoeidheid onder werknemers, afgekort als vermoeidheid onder werknemers.

Een eerste thema in het proefschrift betreft de behandeling van deze vorm van vermoeidheid in de eerstelijns gezondheidszorg. Hierbij staat een gerandomiseerd, gecontroleerd experiment centraal waarbij huisartsen vermoeide werknemers die verzuimen behandelen met een kortdurende vorm van cognitieve gedragstherapie (CGT). Eerder werd al aangetoond dat CGT door ervaren psychotherapeuten effectief is bij patiënten met het chronisch vermoeidheidssyndroom. Onze belangrijkste onderzoeksvragen zijn: zijn huisartsen in staat op CGT op een effectieve manier toe te passen?; en is onze kortdurende vorm van CGT effectief in de behandeling van vermoeide werknemers die verzuimen, in termen van een vermindering van vermoeidheid en het hervatten van werk? In de hoofdstukken 2 tot en met 7 komen we terug op deze en aanverwante vragen.

Een tweede thema betreft de relatie tussen onze onderzoekspopulatie, vermoeide werknemers, en het chronisch vermoeidheidssyndroom (CVS). De aandoeningen CVS en 'vermoeidheid onder werknemers' hebben op het eerste gezicht een aantal overeenkomsten. In concept zijn het twee verschillende maar gerelateerde stadia op het continuüm van vermoeidheid. De eerste onderzoeksvraag is dan ook of een dergelijke relatie kan worden aangetoond. Een tweede onderzoeksvraag is: kan vermoeidheid onder werknemers uiteindelijk leiden tot CVS? Deze en andere vragen worden aan de orde gesteld in de hoofdstukken 6 en 8 tot en met 10:

Is er bewijs voorhanden dat psychosociale interventies, zoals cognitieve gedragstherapie, effectief door huisartsen kunnen worden toegepast? Om deze vraag, die in feite voorafgaat aan ons experiment met CGT door huisartsen, te beantwoorden hebben wij een systematisch literatuuronderzoek verricht dat in **hoofdstuk 2** wordt beschreven. Het onderzoek werd uitgevoerd onder auspiciën van

de Cochrane Collaboration, een wereldwijd netwerk ter bevordering van systematische literatuurstudies van een hoge kwaliteit.

Eerder onderzoek heeft uitgewezen dat een psychosociale interventie een effectieve behandeling kan zijn bij een reeks van klachten in de eerste lijn. In de meeste van deze studies werd de behandeling toegepast door psychologen. We hebben echter gezocht naar studies waarin huisartsen de interventie uitvoerden. Door op een systematische wijze verschillende databanken te doorzoeken zijn uiteindelijk acht studies gevonden die aan onze selectiecriteria voldoen, met uiteenlopende onderwerpen als depressie, somatisatie, stoppen met roken en alcoholisme.

In het algemeen kan gesteld worden dat bewijs voor de effectiviteit van psychosociale interventies door huisartsen beperkt en niet eenduidig is. 'Problem-solving treatment' (een vorm van CGT) door huisartsen bij depressie lijkt een veelbelovende interventie, maar er is meer onderzoek nodig om zicht te krijgen op de effectiviteit van psychosociale interventies door huisartsen.

In **hoofdstuk 3** beschrijven we de opzet van ons gerandomiseerd, gecontroleerd experiment. In de aanloop naar ons onderzoek stuiten we op een aantal methodologische problemen.

Een van onze doelen was het uitvoeren van het experiment onder 'ideale' omstandigheden om zo het maximale effect van de behandeling te kunnen onderzoeken. Door onze beperkte tijd en middelen waren we slechts in staat een klein aantal huisartsen te trainen en te superviseren in kortdurende cognitieve gedragstherapie bij vermoeidheid. In onderzoek binnen de eerstelijns gezondheidszorg is het gebruikelijk dat huisartsen hun eigen patiënten voor het onderzoek rekruteren. Echter, als het kleine aantal huisartsen in ons onderzoek de patiënten uit hun eigen praktijk had moeten rekruteren zouden er te weinig patiënten geselecteerd kunnen worden.

Een tweede probleem betreft de aantrekkelijkheid van onze experimentele interventie (CGT) ten opzichte van de controle conditie (geen interventie vanuit het onderzoek, dus in feite niets doen). De meeste patiënten met klachten willen graag een behandeling ontvangen. Wij voorzagen dat er mogelijk een vertekening van de resultaten zou ontstaan wanneer patiënten aan de controle groep toegewezen worden in de wetenschap dat zij ook een behandeling hadden kunnen loten: het is waarschijnlijk dat deze patiënten gaan 'shoppen' bij andere hulpverleners of uit ontevredenheid uit het onderzoek stappen, in tegenstelling tot patiënten in de experimentele groep.

Om op deze voorziene problemen te anticiperen hebben wij twee aanpassingen aan de gangbare onderzoeksopzet van een behandelstudie in de eerste lijn bedacht: de rekrutering van patiënten afzonderlijk van de huisartsen in ons onderzoek en de toepassing van het zogeheten 'prerandomisatie-design'.

De patiënten in ons onderzoek zijn gerekruteerd uit de populatie van een grote Arbodienst en de helft is toegewezen aan een door ons getrainde onderzoekshuisarts. In de meeste experimenten worden patiënten geïnformeerd over de onderzoekscondities en de willekeurige toewijzing aan een van die condities door middel van loting. Echter, wij hebben de patiënten in ons onderzoek alleen geïn-

formeerd over de conditie waar wij hen door loting aan hebben toegewezen (prerandomisatie), waardoor zij niet op de hoogte zijn van het bestaan van een andere conditie.

In hoofdstuk 3 beschrijven wij de resultaten van deze aanpassingen en de ethische aspecten die aan onze onderzoeksopzet verbonden zijn. In totaal zijn 151 patiënten gerekruteerd voor het onderzoek, waarvan 9% zich tijdens het onderzoek op een willekeurig moment terugtrok, gelijk verdeeld over de beide groepen. Deze cijfers wijzen erop dat onze strategie tot een beperking van de uitval heeft geleid. Onze onderzoeksopzet is wellicht ook interessant voor andere onderzoekers. Daarbij mag echter niet uit het oog worden verloren dat elk onderzoek zijn eigen methodologische uitdagingen kent, en dat met name het prerandomisatie-design niet lichtvaardig mag worden toegepast.

De resultaten van ons onderzoek naar de effectiviteit van cognitieve gedragstherapie (CGT) door huisartsen bij vermoeide werknemers zijn beschreven in hoofdstuk 4, de 'hoofdstudie' van dit proefschrift. Zoals gezegd, eerder onderzoek heeft uitgewezen dat CGT door ervaren therapeuten een effectieve behandeling voor het chronisch vermoeidheidssyndroom is. Wij op onze beurt hebben onderzocht of een dergelijke behandeling ook geschikt is om door huisartsen te worden toegepast bij mildere vermoeidheidsklachten.

In totaal deden 151 vermoeide, verzuimende werknemers en 9 onderzoekshuisartsen mee aan de studie. De patiënten zijn door loting (randomisatie) toegewezen aan een van de twee condities: de experimentele groep, waarin de deelnemers 5 tot 7 sessies CGT door een onderzoekshuisarts werd aangeboden, of de controle groep, waarin alleen het natuurlijk beloop van de klachten werd gevolgd. Alle patiënten zijn gedurende een jaar 4 keer gemeten. De belangrijkste uitkomstmaten zijn de ernst van vermoeidheid, werkverzuim of werkhervatting, en 'klinisch herstel' (gedefinieerd als niet meer moe zijn in combinatie met werkhervatting).

Aan het begin van het onderzoek blijkt 44% van de patiënten te voldoen aan de internationale criteria voor het chronisch vermoeidheidssyndroom (CVS). Hoewel veel deelnemers verbeteren zijn er na 4, 8 en 12 maanden geen significante of relevante verschillen tussen de groepen op een van de uitkomstmaten. Ook aanvullende analyses in verschillende subgroepen laten geen verschillen zien. Dit brengt ons tot de conclusie dat kortdurende CGT door huisartsen in ons onderzoek geen effectieve behandeling is voor vermoeidheid onder verzuimende werknemers. Aangezien onze onderzoekshuisartsen niet in staat zijn geweest de interventie onder 'ideale' omstandigheden met succes toe te passen, moet het sterk betwijfeld worden of huisartsen in de dagelijkse praktijk meer succes zullen hebben met deze interventie.

In hoofdstuk 5 beschrijven wij het behandelprotocol dat aan onze kortdurende CGT ten grondslag ligt en proberen wij een kwalitatieve analyse te maken van de wijze waarop de interventie is uitgevoerd. Dit hoofdstuk ligt daarmee in het verlengde van hoofdstuk 4.

Het is moeilijk om achteraf te onderzoeken waarom de interventie in ons geval niet heeft gewerkt. Echter, het is wel mogelijk om te onderzoeken of er een relatie is tussen het herstel van de deelnemers in de experimentele groep en de wijze waarop de interventie is uitgevoerd. Daartoe gebruiken we de sessieregistratieformulieren waarop de onderzoekshuisartsen de voortgang in elke sessie hebben genoteerd.

Van de 71 patiënten die aan de CGT begonnen zijn hebben er 51 de interventie 'volgens protocol' doorlopen. Vergelijken wij in deze groep van 51 patiënten de 'herstelden' en de 'niet-herstelden', dan zijn er tussen beide groepen geen opmerkelijke verschillen in de kenmerken van de ontvangen interventie. Hoewel er verschillen zijn in de uitvoering van de interventie en de prestaties tussen huisartsen onderling, zien we geen samenhang tussen de manier waarop de interventie is uitgevoerd en de behandelresultaten per huisarts. Opmerkelijk is dat zowel de patiënten als de huisartsen de interventie als positief evalueren, ondanks het gebrek aan werkzaamheid.

Aangezien 'herstelden' en 'niet-herstelden' een vergelijkbare interventie ontvingen, en succesvolle huisartsen en minder succesvolle huisartsen een vergelijkbare interventie toepasten, zijn er aanwijzingen dat de interventie zelf niet adequaat genoeg is om effectief te zijn. Of dit aan de interventie zelf of aan de toepassing door huisartsen heeft gelegen is niet duidelijk.

De interventie bleek dus niet effectief te zijn, maar zijn er andere factoren te ontdekken die het herstel van de patiënten in onze onderzoekspopulatie kunnen voorspellen? Die vraag komt aan de orde in **hoofdstuk 6** in een zogeheten predictieve of voorspellende analyse.

Na 12 maanden blijkt 43% van de vermoeide werknemers in ons onderzoek herstelt van vermoeidheid en heeft 62% inmiddels het werk hervat. 20% van de deelnemers voldoet na 12 maanden aan de criteria voor het chronisch vermoeidheidssyndroom, afgekort als een CVS-achtige status (voor het label CVS is een diagnose door een arts nodig), in vergelijking met 44% aan het begin van het onderzoek. Door middel van multiële logistische regressieanalyse hebben wij vervolgens gekeken welke factoren aan het begin van het onderzoek de uitkomsten na 12 maanden voorspellen.

Hersteld zijn van vermoeidheid wordt voorspeld door het hebben van psychologische attributies (denken dat de klachten een psychische oorzaak hebben) en andere aan ziekteperceptie gerelateerde factoren zoals ervaren gezondheid. Werknemers die na 12 maanden het werk hebben hervat daarentegen zijn onder andere vaker man, zijn jonger en hebben minder vaak een CVS-achtige status aan het begin van het onderzoek. Het hebben van een CVS-achtige status na 12 maanden tenslotte wordt voorspeld door fysieke dysfuncties. Vergelijkbare analyses zijn uitgevoerd met de onderzoekspopulatie maar nu gestratificeerd naar het hebben van een CVS-achtige status aan het begin van de studie.

De resultaten laten zien dat herstel van vermoeidheid en werkhervatting door andere (onderliggende) processen worden bepaald. Fluctuaties in het aantal deelnemers dat voldoet aan de criteria voor CVS onderstrepen het vermoeden dat CVS en vermoeidheid onder werknemers verschillende maar gerelateerde stadia op een vermoeidheidscontinuüm zijn waarop individuen kunnen 'schuiven'. Daarnaast

vormen de factoren die het herstel van vermoeidheid voorspellen een aanwijzing dat de preventie en behandeling van vermoeidheid ten dele in het veranderen van perceptie moeten worden gezocht.

In **hoofdstuk 7** wisselen we van onderzoekspopulatie en maken we voor het eerst gebruik van de Maastrichtse Cohort Studie (MCS), een grootscheeps onderzoek onder 12.000 werknemers (vermoeiden en niet-vermoeiden) die bijna vier jaar lang gevolgd zijn. In onze experimentele behandelstudie hebben we vermoeide werknemers bestudeerd die al enige tijd op ziekteverzuim waren. In hoofdstuk 7 kijken we naar vermoeide werknemers uit de MCS die (nog) aan het werk zijn. In aanvulling op de resultaten uit hoofdstuk 6 beschrijven we een prospectieve studie naar de factoren die het beloop van vermoeidheid en de aanvang van langdurig verzuim in deze groep voorspellen.

In totaal zijn 2108 vermoeide, niet-verzuimende deelnemers aan de MCS geselecteerd aan het begin van de studie en zes maal opgevolgd over een periode van twee jaar. De punt prevalenties van vermoeidheid (59 tot 63%) en langdurig verzuim (2 tot 3%) blijken op al deze vervolgmetingen redelijk stabiel te zijn.

Met behulp van survival analyse (de analyse van 'overlevingsduur') worden vervolgens een aantal predictiemodellen opgesteld. De factoren bij de beginmeting die herstel van vermoeidheid in de loop van twee jaar voorspellen zijn beperkte ernst van vermoeidheid en uitputting, weinig angst, de afwezigheid van conflicten met collega's en een goede ervaren gezondheid. Een exploratieve GEE analyse van het aanhouden of weer opduiken van vermoeidheid levert dezelfde predictoren op, maar dan met associaties in de omgekeerde richting. Kijken we naar de voorspellers bij de beginmeting van de aanvang van langdurig ziekteverzuim in deze groep, dan vinden we dat verzuimers ouder zijn, vaker van het vrouwelijk geslacht zijn, minder beslissingsautoriteit hebben, vaker in nachtdienst werken, hun vermoeidheid vaker aan een somatische aandoening toeschrijven en vaker een geschiedenis hebben van ziekteverzuim in vergelijking met deelnemers die niet langdurig verzuimen in de loop van twee jaar.

Evenals in hoofdstuk 6 komen we tot de conclusie dat vermoeidheid en in dit geval ziekteverzuim door andere factoren wordt bepaald. In lijn met vorige studies die zijn uitgevoerd onder patiënten met het chronisch vermoeidheidssyndroom onderschrijven onze resultaten dat de preventie en behandeling van vermoeidheid gericht moet worden op de perceptie van gezondheid en op emotioneel welbevinden.

Vermoeidheid kent vele verschijningsvormen. In **hoofdstuk 8** onderzoeken we de relatie tussen vermoeidheid, burnout en het chronisch vermoeidheidssyndroom (CVS) door de profielen van de vermoeide deelnemers uit onze behandelstudie te vergelijken met CVS patiënten die eerder werden onderzocht bij het Kenniscentrum Chronische Vermoeidheid in Nijmegen. Op basis van de criteria voor CVS en burnout onderscheiden we een aantal subgroepen in onze oorspronkelijke onderzoekspopulatie. 44% van de deelnemers blijkt bij de voormeting te voldoen aan de criteria voor CVS (heeft een CVS-achtige status) en 50% voldoet aan de criteria voor burnout.

Vervolgens vergelijken we deze subgroepen op hun klinische kenmerken. CVS-achtige werknemers hebben sterkere somatische attributies dan werknemers die niet voldoen aan de criteria voor CVS. Binnen de groep CVS-achtigen zijn werknemers met een burnout depressiever dan werknemers zonder burnout. Werknemers met een burnout die niet voldeden aan de criteria voor CVS hebben sterkere psychologische attributies dan vermoeide werknemers zonder burnout. In vergelijking met CVS patiënten hebben CVS-achtige werknemers gemiddeld alleen maar een kortere klachtenduur. CVS-achtige werknemers met een burnout daarentegen hebben sterkere psychologische attributies en meer psychische klachten dan CVS patiënten.

De resultaten beschouwend concluderen we dat vermoeide werknemers veel overeenkomsten vertonen met CVS patiënten, ongeacht de burnout status, en dat veel vermoeide werknemers in de behandelstudie voldoen aan de criteria voor CVS en burnout. De belangrijkste verschillen tussen de onderzochte subgroepen betreffen echter de causale attributies (het toeschrijven van de vermoeidheidsklachten aan een psychische en/of somatische oorzaak) die gemaakt worden. Dit roept de vraag op welke rol causale attributies spelen: worden ze beïnvloed door vermoeidheidsklachten of zijn het factoren die het klachtenbeloop bepalen?

In hoofdstuk 9 keren we terug naar de populatie van de Maastrichtse Cohort Studie (MCS) om de prevalentie van een CVS-achtige status in de werkende populatie te bepalen. Daarvoor maken we gebruik van de laatste meting van de MCS die 3 jaar en 8 maanden na het begin van de studie heeft plaatsgevonden, het enige moment waarop we informatie over de criteria voor CVS hebben verzameld.

Van de 5499 werknemers die op dat moment nog deelnemen aan de MCS voldoen 199 werknemers (3.6%) aan de criteria voor CVS. Na het excluseren van werknemers die bij de beginmeting van de MCS wellicht hebben voldaan aan de criteria voor CVS kunnen we de jaarlijkse incidentie van een CVS-achtige status schatten op 0.85%. In totaal 20 werknemers (0.36%) geven aan dat ze ooit de diagnose CVS of ME van een arts hebben ontvangen, waarvan maar 2 werknemers ook daadwerkelijk voldoen aan de criteria voor CVS.

We vinden dus tien keer zoveel werknemers die aan de criteria voor CVS voldoen dan werknemers die verklaren gediagnosticeerd te zijn met CVS. Ook is onze prevalentie veel hoger dan wordt gevonden in andere prevalentiestudies naar patiënten met de diagnose CVS. Deze resultaten suggereren dat CVS ondergedetecteerd wordt in de werkende bevolking, en wellicht ook in andere populaties. Verklaringen voor deze constatering moeten wellicht gezocht worden in het hulpzoekgedrag van individuen met vermoeidheidsklachten.

In de hoofdstukken 6 en 8 zagen we reeds dat vermoeidheid onder werknemers en CVS gerelateerde verschijnselen zijn, in lijn met het concept van vermoeidheid als continuüm. Maar is het ook zo dat vermoeidheid onder werknemers een voorstadium van CVS kan zijn? Om deze vraag te beantwoor-

den maken we in **hoofdstuk 10** wederom gebruik van data uit de Maastrichtse Cohort Studie (MCS). Bij de beginmeting van de MCS selecteren we 1143 werknemers met lichamelijk onverklaarde vermoeidheid die niet voldoen aan de criteria voor CVS en die bij de laatste meting nog deelnemen aan de studie. Na 44 maanden gevolgd te zijn in de MCS bekijken we de ontwikkeling van deze groep deelnemers, uitgesplitst naar drie mogelijke uitkomsten: het hebben ontwikkeld van een CVS-achtige status (8%, waarvan niemand aangeeft met CVS gediagnosticeerd te zijn), (nog steeds) moe zijn maar niet voldoen aan de criteria voor CVS (40%), en hersteld zijn van vermoeidheid (52%). Met behulp van multiële logistische regressieanalyse onderzoeken we vervolgens welke factoren bij de beginmeting de uitkomst 44 maanden later voorspellen door de drie subgroepen paarsgewijs te vergelijken. Het hebben van een CVS-achtige status versus moe zijn zonder aan de criteria voor CVS te voldoen wordt voorspeld door hoge leeftijd, uitputting, van het vrouwelijk geslacht zijn, een lage opleiding en eerdere bezoeken aan de huisarts voor problemen op het werk. Het hebben van een CVS-achtige status versus hersteld zijn van vermoeidheid wordt voorspeld door vermoeidheid, uitputting, een lage opleiding, eerdere bezoeken aan de huisarts en aan de bedrijfsarts en een slechte ervaren gezondheid. Factoren bij de beginmeting die moe zijn zonder aan de CVS criteria te voldoen versus hersteld zijn van moeheid voorspellen zijn: vermoeidheid, uitputting, lage ervaren activiteit, angstig zijn en een slechte ervaren gezondheid.

Uit deze resultaten blijkt dat vermoeidheid inderdaad een voorstadium kan zijn van een CVS-achtige status onder werknemers. De prognostische rol die ervaren gezondheid speelt suggereert wederom dat de preventie en behandeling van (chronische) vermoeidheid gezocht moet worden in het veranderen van de ziekteperceptie. Minder duidelijk is de rol van hulpzoekgedrag of het krijgen van een CVS diagnose in deze.

Hoofdstuk 11 vormt de slotbeschouwing waarin ik de belangrijkste bevindingen in dit proefschrift nog eens uitlicht en deze in een breder kader probeer te plaatsen.

Een belangrijke conclusie die uit dit proefschrift kan worden getrokken is dat wij de effectiviteit van cognitieve gedragstherapie door huisartsen bij vermoeidheid niet kunnen aantonen. Om dit gegeven te interpreteren hebben we in hoofdstuk 4 de 'triade van effectiviteit' geïntroduceerd: de bevindingen in behandelstudies zijn altijd het resultaat van de interactie tussen de behandelde, de behandelaar en de behandeling. Het gebrek aan effectiviteit in onze studie is dus op te vatten als een verstoring van deze interactie. De bron van deze verstoring kunnen we helaas niet vaststellen, maar we kunnen wel een aantal mogelijke verklaringen opperen aan de hand van onze resultaten. Maar wat betekenen deze bevindingen nu voor de dagelijkse praktijk? De resultaten van de behandelstudie maar ook van de systematische review zijn aanwijzingen dat complexe interventies zoals cognitieve gedragstherapie niet tot de taak van de huisarts moeten behoren. Dat neemt niet weg dat patiënten met lichamelijk onverklaarde of psychische klachten hulp nodig hebben, ook in de eerstelijns gezondheidszorg. Meer onderzoek naar de effectiviteit van psychosociale behandelingen in de eer-

ste lijn door andere hulpverleners (eerstelijns psychologen, praktijkverpleegkundigen) is dan ook aan te bevelen.

De behandelstudie die in dit proefschrift is beschreven heeft ons onderzoekers ook een aantal methodologische inzichten verschaft. Aan de orde komen het probleem van blinding in psychotherapeutisch onderzoek en de mogelijkheden ter compensatie, de uitdagingen met betrekking tot de rekrutering van huisartsen en patiënten en de controle van onderzoeksactiviteiten door onderzoekers, de participatie van onderzoekshuisartsen en de overdraagbaarheid naar de praktijk, en de informatie over de kwaliteitskenmerken van de behandeling die noodzakelijk is voor een adequate beoordeling van psychotherapeutisch onderzoek.

Een tweede onderzoeksthema in dit proefschrift betreft de relatie tussen vermoeidheid onder werknemers en het chronisch vermoeidheidssyndroom (CVS). De resultaten in ogenschouw nemend kunnen we concluderen dat vermoeidheid onder werknemers in veel gevallen een langdurig aanhoudende conditie is, dat vermoeidheid onder werknemers en CVS veel overeenkomsten vertonen en dat vermoeidheid onder werknemers kan leiden tot de ontwikkeling van een CVS-achtige status, in lijn met het concept van vermoeidheid als continuüm. We hebben ons daarbij met name gericht op de prognostische rol van ziekteperceptie als persisterende, verergerende factor, iets dat al eerder werd gedaan in studies over CVS.

De mogelijke verklaringen voor onze bevindingen vormen de aanleiding tot een nieuwe hypothese over de rol van ziekteperceptie en de diagnose van CVS. Kort samengevat luidt deze hypothese dat een ongezonde perceptie van de vermoeidheidsklachten in combinatie met de bereidheid tot aanhoudend hulpzoekgedrag bij sommige vermoeiden zal leiden tot het vinden van een diagnose die past bij de klachten, en dat de diagnose CVS op haar beurt bij sommigen de ongezonde perceptie van vermoeidheid zal bekrachtigen, wat weer kan leiden tot een verergering en voortzetting van de klachten. Het krijgen van de diagnose CVS kan dus een ziekermakende betekenis hebben voor de ontvanger die lijdt onder zijn klachten.

Tot slot van hoofdstuk 11 volgt een aantal aanbevelingen voor verder onderzoek, waarbij de behandeling van vermoeide werknemers, de veronderstelde verschillen tussen CVS-achtigen en CVS patiënten en de preventie van CVS besproken worden.

Dankwoord

Onderzoek doe je niet in je eentje, in tegenstelling tot wat sommigen beweren. Daarom wil ik op deze plaats vooral diegenen noemen die aan de totstandkoming van dit proefschrift hebben bijgedragen.

Allereerst de personen die mij geholpen hebben bij de dataverzameling en -verwerking. Ik noem Karin Aretz, Marliese Saya, Marieke Verhulsdonk, Esther Collaris, Ria te Winkel, Babette Doorn en Jos Slangen. Ieder van jullie heeft het nodige werk voor het project verzet, sommigen vrijwel belangenloos, en ik ben jullie daar bijzonder dankbaar voor.

Daarnaast heb ik niet alles zelf hoeven te bedenken. Ik bedank de leden van mijn projectgroep voor hun richtingbiedende bijdragen, tijdens onze vaak verhitte bijeenkomsten en daarbuiten: Onno van Schayck, Gijs Bleijenberg, André Knottnerus, Sandra Beurskens, Ellen Bazelmans, Gerard Swaen en Job Metsemakers. Een speciaal woord van dank voor IJmert Kant en Raymond Ostelo: een officiële rol in het project hadden jullie niet, maar zonder de vele uren waarin we de wereld, de wetenschap en onze projecten in het bijzonder hebben besproken had dit boekje er heel anders uitgezien. Daniëlle van der Windt, Judith Prins, Ute Bültmann en Ludovic van Amelsvoort waren zo vriendelijk om mee te schrijven aan een aantal artikelen uit dit proefschrift. Stan Kasl, thank you for valuable contribution to the papers we worked on during my visit to New Haven.

Elk behandelonderzoek gaat gepaard met logistieke uitdagingen, in ons geval de recrutering van patiënten en het onderbrengen van die patiënten bij een behandelaar. Jos Valks, Clement Colen, Jos Jacobs en Arthur Ramaekers van Arbo Unie Limburg hebben ervoor gezorgd dat ik de patiënten voor dit onderzoek kon verzamelen, waarvoor mijn dank. Eenmaal in het onderzoek werden de patiënten in de experimentele groep toegewezen aan een onderzoekshuisarts. Mijn dank gaat uit naar Fons Alberts, Mieke Bruggeman, Jean Dolhain, Dick Kuijpers, Jan Nijhof, Cees Ruissen, Godfried Saes, Rob Schonck en Toine Schuivens voor het behandelen van vermoede werknemers die niet tot hun eigen huisartsenpraktijk behoorden. Jullie inspanningen over een periode van bijna twee jaar zijn de ruggengraat van dit onderzoek geweest.

De jongens en meisjes bij Huisartsgeneeskunde en Epidemiologie in Maastricht en Medische Psychologie in Nijmegen dank ik voor de tijd die we samen hebben doorgebracht op de vakgroep, in kroegen (hoe zal het jou vergaan, Thembi?) en restaurants, op feesten en congressen. Het ga jullie goed allemaal.

Tot slot, nog een woord aan enkele dierbaren. Henning Tiemeier, Casper Schoemaker en Lex van Naerssen, dank voor lessen in het verleden. Voor mijn vrienden in Utrecht en Amsterdam: kijk nou, heb ik tòch nog een boek geschreven. Suzanne en Mikon, jullie hebben samen de vormgeving van dit boekje verzorgd, het mooiste promotiegeschenk dat ik kan bedenken. Voor mijn familie, ik noem in het bijzonder Myrte, Mathijs, Monique en Jean: dank je wel dat jullie er zijn. Liefje, jou fluister ik aanstonds allerlei zoetigheden in het oor. Maar niet nu. Pappa, jij hebt mij geleerd om verder te kijken dan mijn neus lang is. Mijn liefde voor de wetenschap en het debat draag ik op aan jou. Lieve mamma, duizend dingen die hier niet op zijn plek zijn. Ik laat het bij, ik mis je.

Maastricht, september 2003

M.H.

P.S.

Voor alle mensen die zich de moeite getroosten er op 17 december bij te zijn: dank, dank, dank.

About the author

The academic life of Marcus Huibers (1972) began in 1991 when he started his college education at Utrecht University. Initially, he wanted to study Philosophy, but his father talked him into taking up what he felt was a more promising profession, so he switched to Psychology. In that same year, he joined student union Collegium Studiosorum Veritas.

He took several courses in Philosophy at the University of Amsterdam in 1994 and 1995. To make ends meet, he worked as a cook in restaurants and as a research assistant at the Faculty of Social Sciences. In November 1995, he started his research internship at the Program in Human Sexuality of the University of Minnesota Medical School in Minneapolis for which he received a Free Mover Scholarship from Utrecht University.

In February 1997, he graduated from Utrecht University with a master's degree in Clinical Psychology, specializing in Sexology. Six months later, he took a position as research associate at the Netherlands Institute of Mental Health and Addiction (Trimbos-instituut) in Utrecht. In September 1998, he transferred to Maastricht University to work on the present dissertation as a researcher at the Departments of Epidemiology and General Practice and the Department of Medical Psychology of the UMC St Radboud in Nijmegen. In March and April 2003, he visited with Professor Stanislav Kasl at the Yale University School of Medicine in New Haven to work on three of the studies presented in this thesis.

In July 2003, he was appointed assistant professor at the Department of Medical, Clinical and Experimental Psychology of Maastricht University where he is involved in research, teaching and patient care, the latter at the Community Mental Health Centre (RIAGG) Maastricht. He lives happily apart-together with Ivette van Ooijen in Utrecht and Maastricht, with no children or pets in the foreseeable future.

List of publications

- Huibers MJH, Beurskens AJHM, Schayck CP van, Metsemakers JFM, Bazelmans E, Knottnerus JA, Bleijenberg G. Efficacy of cognitive behaviour therapy by general practitioners for unexplained fatigue among employees: a randomised controlled trial. *Br J Psychiatry* 2003; in press.
- Huibers MJH, Bleijenberg G, Amelsvoort LGPM van, Beurskens AJHM, Schayck CP van, Bazelmans E, Knottnerus JA. Predictors of outcome in fatigued employees on sick leave: results from a randomized trial. *J Psychosom Res* 2003; accepted pending revision.
- Huibers MJH, Bleijenberg G, Beurskens AJHM, Kant IJ, Knottnerus JA, Windt DAWM van der, Bazelmans E, Schayck CP van. An alternative trial design to overcome validity and recruitment problems in primary care research. *Fam Pract* 2003; in press.
- Huibers MJH, Kant IJ, Swaen GMH, Kasl SV. Prevalence of chronic fatigue syndrome-like caseness in the working population: results from the Maastricht Cohort Study. *Occ Environ Med* 2003; in press.
- Huibers MJH, Knipschild P. Methodological aspects of randomized controlled trials of psychotherapy research in primary care: correspondence. *Psychol Med* 2003; in press.
- Huibers MJH, Beurskens AJHM, Bleijenberg G, Schayck CP van. The effectiveness of psychosocial interventions delivered by general practitioners: a Cochrane Review. *The Cochrane Library* 2003, issue 2. Oxford: Update Software.
- Huibers MJH, Beurskens AJHM, Prins JB, Kant IJ, Bazelmans E, Schayck CP van, Knottnerus JA, Bleijenberg G. Fatigue, burnout and chronic fatigue syndrome among employees on sick leave: do attributions make the difference? *Occ Environ Med* 2003; 60(Suppl 1):i26-i31.
- Huibers MJH, van der Windt DAWM, Boeke APJ. De deelname van huisartsen aan wetenschappelijk onderzoek: problemen en oplossingen. *Huisarts Wet* 2002; 45:454-8.
- Huibers MJH, Tiemeier H, Franx G, Hutschemaekers GJM. Common sense, consensus en wetenschap aan de voordeur van de GGZ. *Maandblad Geestelijke volksgezondheid (MGv)* 2000; 55:1023-32.
- Huibers MJH, Chavannes NH, Wagena EJ, Schayck, CP van. Antidepressants for smoking cessation: a promising new approach? *Eur Resp J* 2000; 16:379-80.

- Prins JB, Bos E, Huibers MJH, Servaes P, Van der Werf SP, Van der Meer JW, Bleijenberg G. Social support and the persistence of complaints in chronic fatigue syndrome. *Psychother Psychosom* 2003; in press.
- Wagena EJ, Kant IJ, Huibers MJH, Swaen GMH, Wouters EFM, Schayck CP van. Psychological distress and depressed mood in employees with asthma, chronic bronchitis or emphysema: a population-based observational study on prevalence and the relationship with smoking cigarettes. *Eur J Epid* 2003; in press.
- Wagena EJ, Huibers MJH, Schayck CP van. Possible associations between COPD, smoking and depression. *Thorax* 2001; 56:587-8.
- Wagena EJ, Zeegers MPA, Huibers MJH, Chavannes NH, Schayck CP van. Reactie op: Bupropion: een effectief nieuw hulpmiddel bij het stoppen met roken. *Ned Tijdschr Geneesk* 2001; 145:103-4.
- Wagena EJ, Huibers MJH, Schayck CP van. Therapiën voor stoppen met roken (antidepressiva, nicotinevervanging en counseling) en implicaties voor de behandeling van patiënten met chronische obstructieve longziekte. *Ned Tijdschr Geneesk* 2001; 145:1492-6.
- Franx G, Huibers M, Hutschemaekers G. Landelijke Richtlijn Voordeurprotocol Geestelijke Gezondheidszorg. Utrecht 1999: Trimbo-instituut.
- Tiemeier H, Huibers M. Effectief is niet noodzakelijk: een reactie op Hoogduin en Schaap. *Directieve Therapie (Dth)* 1998; 18:314-8.

SUBMITTED MANUSCRIPTS

- Bazelmans E, Huibers MJH, Bleijenberg G. Why did cognitive behaviour therapy by general practitioners for unexplained fatigue among employees not work? A qualitative analysis of the delivery of the intervention. Submitted.
- Huibers MJH, Bültmann U, Kasl SV, Kant IJ, Amelsvoort LGPM van, Schayck CP van, Swaen GMH. Predicting the two-year course of unexplained fatigue and the onset of long-term sickness absence in fatigued employees: results from the Maastricht Cohort Study. Submitted.
- Huibers MJH, Kant IJ, Knottnerus JA, Bleijenberg G, Swaen GMH, Kasl SV. The development of the chronic fatigue syndrome in fatigued employees: predictors of outcome in the Maastricht Cohort Study. Submitted.