

Economic evaluation of cognitive behavioral rehabilitation for chronic musculoskeletal pain

Citation for published version (APA):

Goossens, M. E. J. B. (1999). *Economic evaluation of cognitive behavioral rehabilitation for chronic musculoskeletal pain*. [Doctoral Thesis, Maastricht University]. Universiteit Maastricht. <https://doi.org/10.26481/dis.19990617mg>

Document status and date:

Published: 01/01/1999

DOI:

[10.26481/dis.19990617mg](https://doi.org/10.26481/dis.19990617mg)

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.

Colophon

Economic evaluation of cognitive
behavioral rehabilitation for chronic
musculoskeletal pain./

M.E.J.B. Goossens. - Thesis Maastricht
University. - With References. -

With summary in Dutch.

ISBN 90-9012767-4

Subject headings: cognitive behavioral
rehabilitation/ economic evaluation/
chronic musculoskeletal pain.

Cover design: Dorit Ben-Ari, Maastricht

Lay-out: Jan van Emmerik, Maastricht

Printed by: Datawyse, Maastricht

© 1999 Mariëlle E.J.B. Goossens, Maastricht

**Economic evaluation of cognitive behavioral rehabilitation
for chronic musculoskeletal pain**

Copyright © 2004 by Lippincott Williams & Wilkins
 All rights reserved. This publication is protected by copyright.
 Any unauthorized reproduction, distribution, or transmission
 in any form or by any means, electronic, mechanical, or
 otherwise, without the prior written permission of the
 publisher, is prohibited. This publication may not be
 reproduced, stored in a retrieval system, or transmitted,
 in any form or by any means, electronic, mechanical,
 photocopying, recording, or otherwise, without the prior
 written permission of the publisher.

Economic evaluation of cognitive behavioral rehabilitation for chronic musculoskeletal pain

Author: [Name], [Institution], [Address], [City], [Country]
 Email: [Email Address]

Abstract: [Summary of the study's findings and objectives]

Introduction: [Background information and study rationale]

Methods: [Description of the study design and data collection]

Results: [Key findings and statistical analysis]

Conclusion: [Summary of the study's implications]

Economic evaluation of cognitive behavioral rehabilitation for chronic musculoskeletal pain

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 donderdag 17 juni 1999 om 14.00 uur**

door

Maria Elisabeth Johannes Barbara Goossens
Geboren te Heerlen op 22 maart 1969

Promotor:

Prof.dr. J.M.J.P. van der Linden

Co-promotores:

Dr. J.W.S. Vlaeyen

Dr. M.P.M.H. Rutten- van Mólken (Erasmus Universiteit Rotterdam)

Beoordelingscommissie:

Prof.dr. C.P. van Schayck (voorzitter)

Dr. S.J. Linton (Örebro Medical Center, Örebro, Sweden)

Prof.dr. J.A.M. Maarse

Prof.dr. F. F.H. Rutten (Erasmus Universiteit Rotterdam)

Dr. Ir. H.C.W. de Vet

PROEFSCHRIFT

naar verrijking van de grond van doctorandus
de Universiteit Maastricht
op grond van de in de Maastrichtse
statuten van de Universiteit Maastricht
verzamelde bepalingen betreffende de
in het bezit van de Universiteit Maastricht
op verzoek van de doctorandus
afgegeven

door

The studies in this thesis are supported by grants from the investigative medicine program of the Health Insurance Executive Board (nr OG 91/088) and from the Dutch Praevention Fund (nr 28-2055).

Financial support for the printing of the thesis has been kindly provided by the 'Anna Fonds, Leiden', the 'Department of Pain Rehabilitation and Rheumatology, Hoensbroeck Rehabilitation Center', the 'Department of Internal Medicine/Division of Rheumatology, University Hospital Maastricht', the 'Nationaal Reumafonds', the 'Dutch Chapter for the Study of Pain' (Nederlandse Vereniging ter Bestudering van Pijn: NVBP), the Institute for Rehabilitation Research, Hoensbroek, 'Rug AdviesCentra Nederland', 'Searle Divisie van Monsanto Nederland B.V.' and 'The Research Institute for Extramural and Transmural Health Care (EXTRA), Maastricht University'.

Contents

1	General Introduction: Medical Technology Assessment of cognitive behavioral therapy in chronic musculoskeletal pain research.....	1
1.1	Introduction	2
1.2	General observation.....	2
1.3	Etiology and diagnosis	3
1.4	Burden of illness: Incidence and prevalence.....	4
1.5	Economic burden of illness	5
1.6	Treatment of chronic musculoskeletal pain.....	7
1.7	Cognitive behavioral therapy: Effectiveness and cost-Effectiveness.....	8
1.8	Aims and outline of the thesis	10
	References	12
2	Principles of economic evaluation for interventions of chronic musculoskeletal pain	15
	Summary.....	16
2.1	Introduction	17
2.2	Economic evaluation designs	18
2.3	Comparison of alternatives.....	19
2.4	Analytic perspective	20
2.5	Relevant costs and outcomes.....	20
2.5.1	Costs	20
2.5.2	Outcomes.....	22
2.6	Sensitivity analysis and discounting.....	24
2.7	Incremental analysis and ratios	24
2.8	Collecting the data.....	25
2.9	Conclusion and discussion	26
	References	27
3	Economic evaluation of back pain interventions.....	29
	Summary.....	30
3.1	Introduction	31
3.2	Principles of economic evaluation	32
3.2.1	Economic evaluation designs	32
3.2.2	Costs and outcomes	33
3.3	Methods	33
3.4	Results	34
3.4.1	Study selection.....	34
3.4.2	Economic evaluation designs	43
3.4.3	Alternatives compared.....	43
3.4.4	Methods used in determining costs	45

3.4.5	Methods used in determining outcomes	46
3.4.6	Discounting and sensitivity analysis.....	47
3.4.7	Some methodological aspects of the studies	47
3.5	Conclusion and discussion	48
	References	50
4	Cost-effectiveness of cognitive-educational treatment of fibromyalgia.....	55
	Summary.....	56
4.1	Introduction	57
4.2	Patients and methods	58
4.2.1	Study design	58
4.2.2	Costs	59
4.2.3	Utilities	61
4.2.4	Quality adjusted life years	62
4.2.5	Statistical analysis	62
4.3	Results	62
4.3.1	Patient characteristics	62
4.3.2	Costs	63
4.3.3	Sensitivity analysis of the cost calculations	66
4.3.4	Utilities	67
4.3.5	Quality Adjusted Life Years.....	68
4.4	Conclusion and discussion	68
	References	71
5	Health economic assessment of behavioral rehabilitation in chronic low back pain	73
	Summary.....	74
5.1	Introduction	75
5.2	Patients and methods	76
5.2.1	Subjects	76
5.2.2	Study design	76
5.2.3	Costs	78
5.2.4	Utilities	79
5.2.5	Global assessment of change.....	80
5.2.6	Statistical analysis	80
5.3	Results	81
5.3.1	Baseline characteristics	81
5.3.2	Costs	81
5.3.3	Utilities	84
5.3.4	Global assessment of change.....	85
5.4	Conclusion and discussion	86
	References	88

6	The cost diary: a method to measure direct and indirect costs in cost effectiveness research	91
	Summary.....	92
6.1	Introduction	93
6.2	Patients and methods	94
6.2.1	Subjects	94
6.2.2	The cost diary	94
6.2.3	Feasibility	98
6.2.4	Influence of different periods of data collection	98
6.2.5	Validity	99
6.3	Results	99
6.3.1	Feasibility	99
6.3.2	Time perspective	100
6.3.3	Validity	100
6.4	Conclusion and discussion	101
	References	104
7	Patient utilities in chronic musculoskeletal pain: how useful is the standard gamble method	107
	Summary.....	108
7.1	Introduction	109
7.2	Patients and methods	111
7.2.1	Subjects	111
7.2.2	Measures.....	111
7.2.3	Analysis	114
7.3	Results	115
7.3.1	Reproducibility	115
7.3.2	Cross-sectional construct validity.....	115
7.3.3	Longitudinal construct validity.....	117
7.3.4	Responsiveness.....	118
7.4	Conclusion and discussion	120
	References	123
8	Conclusions, discussion and recommendations	127
8.1	Introduction	128
8.2	Methodological considerations.....	128
8.3	Consequences for different parties.....	134
8.4	Recommendations for health policy.....	136
8.5	Recommendations for future research.....	137
	References	138

1

GENERAL INTRODUCTION: MEDICAL TECHNOLOGY ASSESSMENT OF COGNITIVE BEHAVIORAL THERAPY IN CHRONIC MUSCULOSKELETAL PAIN RESEARCH

In the field of chronic musculoskeletal pain, the medical technology assessment of two well-established strategies is required to provide optimal treatment for chronic low back pain and fibromyalgia. In the field of chronic low back pain, the medical technology assessment of two well-established strategies is required to provide optimal treatment for chronic low back pain and fibromyalgia. In the field of chronic low back pain, the medical technology assessment of two well-established strategies is required to provide optimal treatment for chronic low back pain and fibromyalgia.

Chronic low back pain is a complex condition that significantly impacts the quality of life of patients. The pathogenesis of chronic low back pain is multifactorial, involving genetic, environmental, and psychosocial factors. The medical technology assessment of cognitive behavioral therapy (CBT) in chronic low back pain research is essential to evaluate its effectiveness and cost-benefit ratio. CBT has been shown to be a promising intervention for chronic low back pain, with studies demonstrating improvements in pain intensity, functional disability, and quality of life. However, the medical technology assessment of CBT in chronic low back pain research is still in progress, and further research is needed to establish its long-term effectiveness and cost-benefit ratio.

The medical technology assessment of CBT in chronic low back pain research is a complex task that requires a multidisciplinary approach. It involves the evaluation of the clinical effectiveness, cost-effectiveness, and acceptability of CBT in the management of chronic low back pain. The medical technology assessment of CBT in chronic low back pain research is essential to provide decision-makers with the information they need to make informed choices about the allocation of healthcare resources. The medical technology assessment of CBT in chronic low back pain research is a complex task that requires a multidisciplinary approach. It involves the evaluation of the clinical effectiveness, cost-effectiveness, and acceptability of CBT in the management of chronic low back pain.

The medical technology assessment of CBT in chronic low back pain research is a complex task that requires a multidisciplinary approach. It involves the evaluation of the clinical effectiveness, cost-effectiveness, and acceptability of CBT in the management of chronic low back pain. The medical technology assessment of CBT in chronic low back pain research is essential to provide decision-makers with the information they need to make informed choices about the allocation of healthcare resources. The medical technology assessment of CBT in chronic low back pain research is a complex task that requires a multidisciplinary approach. It involves the evaluation of the clinical effectiveness, cost-effectiveness, and acceptability of CBT in the management of chronic low back pain.

The medical technology assessment of CBT in chronic low back pain research is a complex task that requires a multidisciplinary approach. It involves the evaluation of the clinical effectiveness, cost-effectiveness, and acceptability of CBT in the management of chronic low back pain. The medical technology assessment of CBT in chronic low back pain research is essential to provide decision-makers with the information they need to make informed choices about the allocation of healthcare resources. The medical technology assessment of CBT in chronic low back pain research is a complex task that requires a multidisciplinary approach. It involves the evaluation of the clinical effectiveness, cost-effectiveness, and acceptability of CBT in the management of chronic low back pain.

1.1 Introduction

Economic evaluation is attracting increasing attention as a tool to inform policy makers, insurers and other payers of the value of existing and new treatment modalities. This thesis focuses on the economic evaluation of treatment interventions in the field of chronic musculoskeletal pain. The content is based mainly on the results of two cost-effectiveness studies of cognitive behavioral treatments for chronic low back pain and fibromyalgia. Part of the thesis deals with methodological aspects of measuring costs at a patient level and health-related quality of life by means of utility measurement.

1.2 General observation

Musculoskeletal pain has a major impact on society in terms of both payment for health care and indirect costs for loss of earnings. The health care costs are considerable, but contribute less than 15% of the total costs, which primarily arise from indirect factors associated with sick leave, compensation and early retirement.^{1,2} It is therefore surprising that most of health care costs arise from hospital visits, physical therapy and diagnostic examinations^{2,3}, rather than from early rehabilitation and pain management in order to stop the further escalation of the indirect costs.

A wide variety of treatments has been developed for chronic musculoskeletal pain, apparently without much (clinical) success. It seems that the majority of the treatments are not getting to the root of the problem and are not addressing the correct issues. The current focus is shifting more and more towards incorporating cognitive behavioral techniques into pain management. However, with decision-makers in health care nowadays striving at efficiency, before these innovative principles of treatment can be recommended for implementation in broader settings, their effectiveness and cost-effectiveness have to be established.

These issues highlight the increasingly important role of Medical Technology Assessment (MTA)^{*} in determining the value of promising interventions to treat chronic musculoskeletal pain. In MTA, careful and systematic evaluation and linkage of the numerous facets of a health care issue are important to determine its full impact, assess the state of the art and rational use of health care services, in order to finally support and guide medical decision making.⁴ In relation to the promotion of efficacy in health care, in the mid-1980s the Dutch Health Insurance Council started to advocate the use of MTA for the review of new therapeutic procedures.⁵ In this process, the Dutch Health Council especially underlines the identification of interventions which warrant assessment, the assessment of new knowledge to be gained through these

^{*} MTA can be described as a multi-disciplinary field of study investigating the consequences of medical technology (e.g. drugs, devices, medical- and surgical procedures, and organization of care) for different groups of the population.

interventions, the implementation of the interventions in the field and finally field evaluation. A conceptual framework to structure the different aspects was introduced by Tugwell *et al.*⁶, including etiology, quantification of the burden of illness, assessment of therapeutic effectiveness and economic evaluation of therapies. In this chapter several chains of the process have been applied to the complex disorder of chronic musculoskeletal pain (figure 1). The purpose of this chapter is to briefly examine various facets of chronic musculoskeletal pain as an introduction to the assessment of economic evaluation of the cognitive behavioral interventions for chronic musculoskeletal pain presented in the following chapters of this thesis.

Medical Technology Assessment and Chronic Musculoskeletal Pain

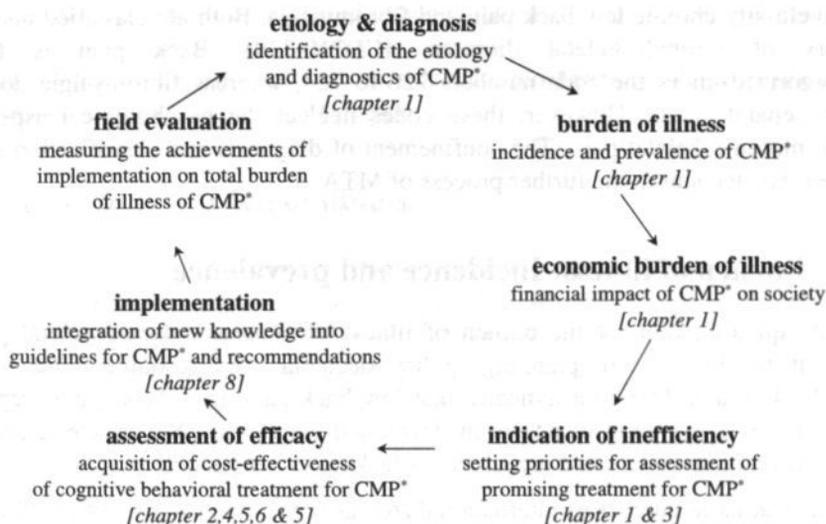


Figure 1. Application of the iterative loop of Medical Technology Assessment in the field of chronic musculoskeletal pain.

* CMP = chronic musculoskeletal pain

1.3 Etiology and diagnosis

In MTA, the disorder on which the research is focused should be defined clearly. In chronic low back pain and fibromyalgia this is not an easy task. Both disorders are regarded as musculoskeletal pain 'syndromes', for which there is no etiology or generally accepted pathophysiology. A major problem in many cases of chronic low back pain and fibromyalgia is the inability to make a precise and reproducible diagnosis.

Chronic low back pain refers to pain of more than three months duration, which is experienced in the lumbosacral region of the spinal column. In about 90 to 95% of the cases the nature of the complaint is nonspecific. In essence, fibromyalgia is defined as

widespread muscle pain (in neck, shoulder, low back and hip) lasting 3 or more months, together with at least 11 of 18 tender points on digital palpation.⁷ A second important overriding symptom in fibromyalgia is fatigue. Both syndromes are associated with many other symptoms, such as sleep disturbance, disturbed mood, unhelpful beliefs, psychosocial problems caused by progressive inactivity at home, socially and at work, secondary physical deconditioning, fears and depression.^{8,9,10} Furthermore, fibromyalgia and chronic low back pain can be secondary to other disorders. The multi-symptomatology has often led to the notion that the underlying problem is partly psychological in nature.¹¹

This multi-symptomatology renders precise classification of both syndromes difficult. In scientific research, the International Classification of Diseases (ICD-9) is most often used to classify chronic low back pain and fibromyalgia. Both are classified under the category of musculoskeletal diseases (ICD710-739). Back pain is further subcategorized under the code numbers 720 to 724, whereas fibromyalgia does not have a separate code. However, these codes neglect the psychological aspects of chronic musculoskeletal pain. The confinement of diagnosis and classification should be taken into account in the further process of MTA.

1.4 Burden of illness: Incidence and prevalence

Accurate quantification of the burden of illness of chronic musculoskeletal pain is important for health care planning, policy decisions and resource allocation for research. Available literature indicates that low back pain and fibromyalgia represent two of the most common disorders in developed countries.¹² Nevertheless, accuracy and precision of incidence and prevalence is lacking in both disorders.

Back pain. A large number of international studies report that 60-80% of people report back pain at some time in their life.¹³ The annual incidence of low back pain was reported to be 5%, with yearly prevalence rates ranging from 12% to 40%.¹⁴ In the Netherlands, the lifetime prevalence of low back pain is 51% in men and 58% in women, while the one-year prevalence is about 30%.¹⁵ In general, low back pain has a short duration. About 90% of the cases improve spontaneously within a month, regardless of treatment. Nevertheless, the relapse rate is very high (70-80%), whereas chronic low back pain, with persistent symptoms after 12 weeks is present in about 6% of patients. Primarily, these chronic cases, where people suffer permanent from back pain, are responsible for the high social cost of back pain, in terms of compensation, absence from work and medical expenses.

Fibromyalgia. The incidence of fibromyalgia is unknown, in part due to the difficulty in classifying and diagnosing the fibromyalgia syndrome. Only in 1990 did formal criteria become available to classify fibromyalgia. However, its association with concurrent medical or psychiatric illness make the classification even more complicated. Wolfe *et al.*^{8,16} revealed that approximately 2% of the total population suffered from fibromyalgia, with a prevalence of 3.4% in women and 0.5% in men. However, the prevalence of chronic widespread pain, the most important symptom in

fibromyalgia, is several times higher than the prevalence of fibromyalgia as defined by the 1990 American College of Rheumatology guidelines.¹¹ To date, no studies at population level have been performed to indicate the exact prevalence of fibromyalgia.

Although some data on prevalence and incidence has been reported, we cannot yet define the true frequency of back pain and fibromyalgia. Discrepancies in figures among studies are largely due to methodological differences, especially in survey techniques, patients groups studied and the difficulty of obtaining a specific diagnosis in most cases. Using the reported incidence and prevalence figures of these disorders as a reliable measure for the burden of illness cannot therefore be recommended. If, for example, prevalence figures were used to indicate the burden of illness, it could easily be concluded that the burden of back pain has not changed over the past 40 years. This in fact is not true. What has increased exponentially is its functional impairment and disability rather than the prevalence of chronic musculoskeletal pain.¹⁰ Further estimates of the economic impact of chronic musculoskeletal pain are necessary for accurately quantifying the burden of illness.

1.5 Economic burden of illness

The economic burden of musculoskeletal pain on society is measurable in terms of direct costs, indicating costs of health care utilization and patient and family costs, and indirect costs, for example due to absenteeism from work and lost productivity.

Back pain. Based on 1991 figures, van Tulder *et al.*² estimated the total costs spent on back pain in the Netherlands to be approximately US \$5 billion, equivalent to 1.7% of the gross national product of the Netherlands. Total direct medical costs of back pain were estimated to be US \$367.6 million, of which 56.5% was related to hospital care, 1% to medical specialist care, 6% to general practitioner care and 36% to paramedical care. Primary care was the most commonly reported source of health care for back pain. General practitioners see about 59% of back pain patients.¹⁵ In the Netherlands, back pain (as defined by the ICD-9 codes 710-739 and subcodes 720-724) accounted for 2% of total GP costs. Furthermore, back pain accounted for 7% of total hospitalization costs and 19% of total paramedical costs. Total indirect costs of back pain due to back pain for the whole labor force were estimated at US \$4.6 billion, of which US \$3.1 billion was due to absenteeism and US \$1.5 billion due to disability. However, these figures do not include the unpaid labor which may have resulted in an underestimation of the total indirect costs.

The economic impact of back pain has also been studied in several other countries, for example the USA, UK, Sweden, Canada and Australia. However, comparisons between these studies and across countries are difficult for several reasons, such as differences in the perception of back pain, differences in health care and insurance systems and social welfare systems, variations in study populations, differences in diagnosis and classification of back pain, socioeconomic variables and differences in wealth. One major factor is the perspective chosen for the study, which can lead to different cost estimates. Most US estimates have not been calculated from a societal

perspective, and workers' compensation rather than wages has been used as the indicator of productivity losses. US studies have mainly been updates from two original reports estimating the total costs of back pain disorders between 1972-78 at US \$16 billion a year¹⁷ and in 1986 at US \$11.1 billion a year.¹⁸ In other European countries, such as the UK and Sweden, economic costs have been calculated from a societal perspective - for the UK between 1991-92 and Sweden for 1995.²⁰

Table 1. Cost of back pain in US dollars in the United Kingdom, Sweden and the Netherlands.

Cost Categories	United Kingdom		Sweden		The Netherlands	
	Cost in million US \$ (% of total)	Cost per inhabitant	Cost in million US \$ (% of total)	Cost per inhabitant	Cost in million US \$ (% of total)	Costs per inhabitant
Direct costs	385 (11.5)	7	213 (8)	24	368 (7.4)	24
Indirect costs	2,948 (88.5)	113	2,262 (92)	266	4,600 (92.6)	299
Total costs	3,333 (100)	120	2,475 (100)	290	4,968 (100)	323

The costs estimated by each study were converted to US dollars for the same base year, i.e. 1991, by using OECD currency and divided per inhabitant according to demographic statistics from the OECD.⁴⁹

Table 1 indicates that costs of back pain per inhabitant are within approximately the same range in the Netherlands and Sweden. For the UK, costs were lower, but using the same method for calculating losses of productivity as was used for Sweden and for the Netherlands, i.e. average wage plus costs for social security, the cost per inhabitant for the UK could be calculated as about 80% of that for the Netherlands. Table 1 also shows that the indirect costs far exceed the direct costs of back pain. Furthermore, the total costs are not normally distributed due to a minority of back pain patients (10 to 25%) who account for more than 75% of costs.^{21,22}

Fibromyalgia. The economic impact of fibromyalgia has not yet been studied in detail. Diagnostic classification problems hamper estimates of the cost of illness. Some studies, however, do report on the frequency of visits to the health care sector. In the US, prevalence of fibromyalgia in categorical hospitals for rheumatic diseases was estimated as high as 15-20% and 11.4% in general hospitals.⁷ In the US fibromyalgia is among the commonest conditions treated by rheumatologists.⁸ In the Netherlands, the standard diagnosis registration system for rheumatic diseases (Standaard Diagnose Registratie van Reumatische Ziekten) estimated that in 1988 about 6% of all new visits and 6% of all follow-up visits to rheumatologists were related to fibromyalgia.²³ The prevalence of fibromyalgia in general practice in the Netherlands has been estimated at approximately 2.96 per 1000 inhabitants.²³ Estimates of indirect costs are lacking. The prevalence of work disability and compensation for fibromyalgia is related to the social insurance and legislation system in each country. Where work disability related to fibromyalgia is limited or not warranted, the number of insurance claims is limited. In the US, as many as 25% of patients with fibromyalgia have received some form of disability or injury compensation.⁸ A recent multicenter study in the US of work and disability in fibromyalgia reported that 27% of fibromyalgia patients received at least some form of disability payments and 17% were receiving social security payments (compared with 2.2% of the US population).⁸

The studies cited above do not provide a complete overview of the economic impact of low back pain and fibromyalgia, they nevertheless indicate that the economic burden of fibromyalgia and low back pain is significant. In the Netherlands, musculoskeletal disorders (ICD 710-739) are ranked as the fourth most expensive disease category regarding health care costs. Together with mental disorders they constitute the most expensive disease category in terms of absenteeism and disablement.²⁴ Although these data demonstrate a substantial economic impact, they may however, underestimate the total costs of musculoskeletal disorders. Although chronic pain is generally believed to have an enormous social and economic impact on the individual and family, none of the studies calculating the economic burden took into consideration direct non-medical costs, such as personal and out-of-pocket expenses. The impact is even more striking considering that expenditures due to pain increase, whereas income due to disability decreases.²⁵ Although difficult to measure, estimates relating to the economic impact of chronic musculoskeletal pain should, whenever possible, assess the financial burdens imposed on the patient and the family. Chapters 4 and 5 report our two studies evaluating cognitive behavioral rehabilitation for fibromyalgia and chronic low back pain. In these studies the economic impact on the individual patient has been assessed, using a cost diary. One year after treatment, direct non-health care costs (such as patient and family costs) accounted for 66% and 71% of the fibromyalgia and chronic low back pain patient's total health care expenditures respectively, as compared to 34% and 29% for direct health care costs (such as doctor visits and hospitalizations).

1.6 Treatment of chronic musculoskeletal pain

Effectiveness. In the process of MTA, identification of treatments which warrant assessment in order to reduce the burden of illness is the next link in the iterative loop. For chronic low back pain and fibromyalgia there is a wide array of treatments but little evidence based upon clinical practice. This raises the question as to whether the proportion of resources provided for the health care of chronic pain patients is being spent on the right treatment.

The majority of health care costs for back pain arise from hospital visits and physical therapy, indicating advanced diagnostic procedures and treatments aimed mainly at symptom reduction.^{2,3} An up-to-date systematic literature review of 80 randomized clinical trials published between 1966 and 1997 on all treatments for chronic low back pain found, in general, sufficient short-term evidence on effectiveness for manipulation, back schools, exercise therapy, multidisciplinary pain treatment and spa therapy and some evidence for behavioral therapy.² However, due to study heterogeneity and low methodological quality of many of the included studies results could not be compared.

In fibromyalgia, the use of a large variety of drugs, physical therapy and tender and trigger point injections also indicates that therapy is often mainly directed at symptom reduction. Some (short-term) efficacy was only demonstrated for low-dose tricyclic therapy, educational and exercise treatment programs.^{26,27}

Cost-effectiveness. In order to inform policy makers about the consequences of adopting an intervention, data on clinical effectiveness needs to be supplemented with data demonstrating cost-effectiveness, particularly because of the pressures on health care costs. In this, cost-effectiveness analyses compare both the costs and outcomes of alternative health care interventions. Three recent reviews of economic evaluation of interventions for musculoskeletal pain syndromes illustrate the growing research interest in incorporating cost issues into study analyses.^{28,30} Chapter 3 discusses the results of a literature search for economic evaluations of back pain interventions, showing that between 1966 and 1997 only 28 such studies were published. The studies all dealt with back-related problems and included injury prevention and back schools, post-incidence management programs, clinical treatments, such as manipulation and physical therapy, exercise therapy and surgery. Only three dealt with chronic low back pain, of which two evaluated multidisciplinary pain treatment (resulting in decrease of health care costs). The two reviews of Ferraz *et al.*²⁸ and Maetzel *et al.*²⁹ limited their search to full economic evaluations published between 1966-1995²⁸ and 1996-1997.²⁹ The studies they identified on back pain were also included in the review described in chapter 3. Our fibromyalgia study, described in chapter 4, appeared to be the only economic evaluation study of fibromyalgia treatment thus far.

Although the reviews were not fully comparable, all concluded that the majority of economic evaluations were of low quality. Some of the studies, however, may represent important intermediate stages in the understanding of the cost and consequences of interventions, but their limited approach in costing and their poor methodological quality does not allow us to answer efficiency questions regarding the different treatments. For this purpose, evidence on the cost-effectiveness of the most promising interventions for chronic musculoskeletal pain is still needed.

1.7 Cognitive behavioral therapy: Effectiveness and cost-Effectiveness

Effectiveness; Given the current evidence summarized above, the majority of existing treatments cannot cure chronic musculoskeletal pain or reduce the pain in the long term. However, this does not necessarily indicate that health care professionals are unable to help chronic musculoskeletal pain patients to deal with their current situation in terms of helping the patients manage their problem, reduce the impact of pain on their life, control the pain, be more active and improve overall quality of life. A wide array of studies has established that psychological factors can influence and worsen pain behavior and disability. In this, two main psychological approaches are important in the reduction of pain disability in chronic musculoskeletal pain: operant-behavioral and cognitive approaches. The operant-behavioral approach is based on the operant learning principle. This principle refers to Thorndikes's 'Law of effect'.³¹ Behavior which is followed by a positive reinforcer has a greater probability to increase in frequency, while behaviors that are not followed by this reinforcement will tend to decrease in frequency. Fordyce³² was the first to apply operant learning principles to chronic pain behavior, and developed a treatment program during which the social

consequences of both pain and healthy behaviors are systematically controlled by specially trained rehabilitation teams. The operant-behavioral treatment program aims at removing positive reinforcers of pain and encouraging wellness behavior, through for example increasing mobility and social activity and reducing medication consumption and health care use.³² Since the social responsiveness of the spouse to pain and healthy behaviors is at least as important than those of the team, spouses are invited to join the treatment as well. They learn to be more aware of their own responses towards the patients pain and healthy behaviors, and to positively reinforce healthy behaviors rather than pain behaviors.³³

Cognitive approaches are initially based on the gate control theory of pain, which established the importance of thoughts in modulating pain.³⁴ Cognitive processes, such as the interpretation of external and internal events, expectations, beliefs, problem solving, interpretation of events and self statements have been shown to affect the experience of pain substantially and may hinder recovery. Generally, two main cognitive treatment approaches can be distinguished: pain coping skills training³⁵ and the modification of maladaptive beliefs.³⁶ In a pain coping skills training, patients are taught to apply imagery, attention diversion, goal setting, stress management and relaxation to gain more feelings of control over pain.^{10,37,38} The cognitive treatment á la Beck is designed to learn patients to be aware of their basic and possibly maladaptive beliefs about themselves and their social environment, to challenge the validity of these beliefs and to finally transform these beliefs into more adaptive ones.

Cognitive behavioral interventions often combine both approaches and emphasize the interdependence of cognitions and behaviors.³⁵ They help patients restructure the way they think about their pain (cognitive) and at the same time change their pattern of behavior (behavioral). The most common cognitive behavioral interventions include relaxation training, behavioral management programs to reduce 'pain behavior', goal setting, role modeling with patient practice and feedback, relabeling pain sensations, problem solving, assertiveness training, stress management, distraction and coping skills and strategies.^{10,37-39} Cognitive behavioral therapy has been incorporated in some multi-dimensional treatment and interdisciplinary programs. To enhance the consistency of the behavioral approach, in this programs patients' conferences attended by the interdisciplinary staff are an essential part. The content of these programs is usually not standardized as there is a variety of cognitive behavioral methods. Difference in the duration of therapy, type of treatment, therapist-related factors, combinations that incorporate other methods and the context and orientation of the program make it difficult to compare the effects of the specific cognitive behavioral aspects within any one approach.⁴⁰ Furthermore, Turk⁴¹ highlighted the need to take account of the recalcitrant nature of chronic pain when considering the success rate of multidisciplinary pain centers. The majority of those who come to try these approaches have not been helped by other therapies: "As such, the population of pain patients referred to and treated at multidisciplinary pain centers are at higher risk to fail any treatment".

Despite this, several overviews and systematic reviews have attempted to draw some general conclusions regarding the effectiveness of multidisciplinary pain management.

Flor *et al.*⁴² reviewed 65 reports of multidisciplinary treatment for chronic low back pain. Van Tulder *et al.*⁴³ reviewed 11 RCTs of behavioral therapy for chronic low back pain. Morley *et al.*⁴⁴ performed a systematic review of psychological treatments for chronic pain, oriented towards the effects of cognitive behavioral treatments. Flor⁴⁵ also reviewed 80 studies on the additional effect of cognitive behavioral techniques to pain management. Although the majority of studies included in these reviews were of poor methodological quality, and despite the limitations mentioned earlier for comparison due to the diversity of programs, several conclusions can be based upon the results of these reviews. Overall, patients who received cognitive behavioral interventions seem to have benefited significantly compared with waiting-list controls (this was the most common type of comparison group), or traditional treatment. The largest effects of these therapies have been found on relevant variables, such as pain perception, distress, daily functioning, impairment and coping skills.

Cost-effectiveness. Despite considerable evidence that cognitive behavioral therapy is effective, the aforementioned reviews on cost-effectiveness in section 1.6 rarely report economic evaluations undertaken for therapies of chronic musculoskeletal pain and none on cognitive behavioral interventions.^{28,29,30} Our cost-effectiveness studies described in chapters 4 and 5 are the first studies performed in this area and might therefore be of great value. Cost factors associated with pain treatment have been minimally explored in well-designed, controlled studies as has been done for many other medical treatments.⁴⁶ Evidence on cost-effectiveness is needed to convince purchasers and payers to include cognitive behavioral treatment in health care benefit packages. The paucity of outcome studies demonstrating cost-effectiveness of cognitive behavioral interventions for chronic pain has also been confirmed by Turk⁴¹, Hagglund and Fillingim³⁹, Linton⁴⁰ and Caudill.⁴⁷ Turk⁴¹ made an attempt to examine the cost-benefit ratio of multidisciplinary pain centers, but due to the methodology used, his conclusions turned out to be far too optimistic. Turk's calculations were based upon assumptions and on extrapolation from treatment outcomes reported in several studies. The studies included were not selected on a systematic basis and only studies with conclusions in favor of multidisciplinary pain centers were included.

Given the current knowledge of effectiveness of CBT, this evidence needs stronger support through the conduct of economic evaluation studies. Information can then be used to make recommendations for implementations of cognitive behavioral principles into the routine chronic musculoskeletal pain management.

1.8 Aims and outline of the thesis

This thesis focuses on the economic evaluation of treatment interventions in the field of chronic musculoskeletal pain. The content of this thesis focuses on the results of two cost-effectiveness studies of cognitive behavioral rehabilitation for chronic low back pain and fibromyalgia. Part of the thesis deals with methodological aspects of measuring costs at a patient level and health-related quality of life.

As a first step, chapter 2 gives a brief overview of the state of the art on how to apply the principles of economic evaluation for the evaluation of interventions in chronic musculoskeletal pain. This will simplify the interpretation of the results of other published studies and make it easier for physicians and other researchers to decide whether the published studies can be applied to their own setting.

In chapter 3, a methodological checklist is used to review the literature on available economic evaluation studies in the field of back pain published between 1984 and 1997. The chapter gives a detailed summary of the studies found and puts further emphasis on the methodological quality of the economic evaluations performed.

Chapter 4 describes the costs and effects of a three-year randomized controlled clinical trial performed by our group, comparing a cognitive-educational treatment (including a pain coping skills training) for fibromyalgia with an educational program plus attention-control and a waiting-list control group. A total of 131 patients were randomly allocated to one of the interventions and followed for 12 months (except for the waiting-list control group).

Chapter 5 presents another cost-effectiveness study in chronic musculoskeletal pain. A combined operant-behavioral program plus pain coping skills training program as provided by our group is compared to operant-behavioral treatment with attention control. Both programs are compared with a waiting-list control group and operant-behavioral rehabilitation as usually provided. Altogether 148 patients with chronic low back pain were randomly assigned to the different conditions and followed over one year. In both studies (as described in chapters 4 and 5) the supplemental effectiveness of the cognitive coping skills training is examined and related to the additional costs. The economic endpoints in both studies were the costs of the program and other health care resource utilization, but also costs for the patient and indirect costs associated with production losses due to the disorder.

Chapter 6 further addresses the measurement of direct and indirect costs. The studies described in chapters 4 and 5 have a societal perspective, which implies that costing data should be gathered as completely as possible. Chronically ill patients consume a large amount of health care and non-health care resources, which to a large extent are not recorded in routine databases. The need to include these costs in the analyses has led to the development of a patient cost diary. Chapter 6 focuses on the development, content and some of the methodological aspects of this patient cost diary.

To satisfy the concept of cost-effectiveness in both studies, the effects are measured in global assessment of change and health state valuations, using rating scale and standard gamble methods. These measures have rarely been used before in chronic musculoskeletal pain. The psychometric properties of both methods are compared in chapter 7. This chapter compares in more detail preference-based instruments with some domain-specific instruments often used in a chronic musculoskeletal pain population and discusses their usefulness in populations with chronic musculoskeletal pain.

Finally, chapter 8 provides a conclusive overview of the different studies presented in this thesis. Furthermore this chapter includes a general discussion and provides recommendations for implementation and further research.

References

1. Coyle, D., Richardson, G. The cost of back pain. In: *Clinical Standards Advisory Group report on back pain*. London: HSMO, 1994: 65-71.
2. van Tulder, M.W., Koes, B.W., Bouter, L.M., A cost of illness study of back pain in the Netherlands. *Pain* 1995; 62: 233-240.
3. Linton, S.J., The Socioeconomic impact of chronic pain: is anyone benefiting? *Pain* 1998; 75: 163-168.
4. Lawrence, V.A., Tugwell, P., Gafni, A., Kosuwon, W., Spitzer O., Acute low back pain and Economics of therapy: the iterative loop approach. *Journal of Clinical Epidemiology* 1992; 3: 301-311.
5. Minister of Health. *Medical Technology Assessment and efficacy in Health care. A policy document from the minister of health*. The Netherlands: Rijswijk, 1996.
6. Tugwell, P., Bennett, K.J., Sackett D.L., Haynes RB., The measurement iterative loop: a framework for the critical appraisal of need, benefits and costs of health care interventions. *Journal of Chronic Diseases* 1985; 38: 339-351.
7. Wolfe, F., Smythe H., Yunus M., et al. The american College of Rheumatology 1990: Criteria for the classification of fibromyalgia. Report of the multicenter criteria committee. *Arthr Rheum* 1990; 33: 160-172.
8. Wolfe, F. The fibromyalgia problem (editorial). *J Rheumatol* 1997; 24:1247-1249.
9. Parziale, J.R., Chen J.J. Fibromyalgia. *Medicine and Health* 1996; 79 (5): 188-192.
10. Waddell, G. *The back pain revolution*. Churchill Livingstone. Edinburgh, 1998.
11. Bennett, R. Fibromyalgia, chronic fatigue syndrome and myofascial pain. *Current Opinion in Rheumatology* 1998; 10: 95-103.
12. Girolamo, de G. Epidemiology and social costs of low back pain and fibromyalgia. *Clin J Pain* 1991; 7 (Suppl. 1): S1-S7.
13. Borenstein, D. Epidemiology, etiology, diagnostic evaluation and treatment of low back pain. *Current Opinion in Rheumatology* 1998; 10: 104-108.
14. Andersson, G.B.J. The Epidemiology of Spinal Disorders. In: Frymoyer J.W. *The Adult Spine: principles and Practice*. New York: Raven Press, 1991: 107-146.
15. Valkenburg, H.A., Haanen, H.C.M. The epidemiology of low back pain. In: A.A. White and S.L. Gordon (Eds.). *Symposium on idiopathic low back pain*. The C.V. Mosby Company: St. Louis, 1982: 9-22.
16. Wolfe, F. The epidemiology of fibromyalgia. *J Musculo Med* 1993; 1: 137-148.
17. Grazier, K.L., Holbrook, T.L., Kelsey, J.L. et al. *The Frequency of Occurrence, Impact, and Cost of Musculoskeletal Conditions in the United States*. Chicago: American Academy of Orthopaedic Surgeons, 1984.
18. Webster, B.S., Snook, S.H. The cost of compensable low back pain. *J Occup Med*. 1990; 32: 13-15.

19. Coyle, D., Richardson, G. The cost of back pain. In: *Clinical Standards Advisory Group report on back pain*. London. HSMO, 1994: 65-71.
20. Norlund, A. Cost of back pain in Sweden. In: SBU. *SBU-report on Back Pain*. Stockholm: SBU (in press).
21. Cats-Baril, W.L., Frymoyer, J.W. The economics of spinal disorders. In: *The adult spine: principles and practice*. First Edition. New York. Raven Press, 1991.
22. Moffet, J.K., Richardson, G., Sheldon, T.A. Maynard A. *Back pain: its management and costs to society*. University of York, Centre for Health Economics. Discussion paper 129. York, 1995.
23. Jacobs, J.W.G., Rasker, J.J., Bijlsma, J.W.J., Classificatie van fibromyalgie: de criteria van het American College of Rheumatology. Ingezonden brief. [Classification of fibromyalgia: the criteria of the American College of Rheumatology. Letter.] *Nederlands Tijdschrift voor Geneeskunde* 1992; (136) 7: 337.
24. Polder, J.J., Meerding, W.J., Koopmanschap, M.A., Bonneux, L., van der Maas, P. J. *Kosten van ziekten in Nederland 1994*. [Costs of illness in the Netherlands, 1994]. Institute of Medical Technology Assessment: Erasmus University Rotterdam, The Netherlands, 1997.
25. Ferrell, B.R., How patients and families pay the price of pain. In: M.J.M. Cohen and J.N. Campbell, editors. *Pain treatment centers at a crossroads: A practical and conceptual reappraisal*. Vol. 7. Seattle, WA: IASP Press, 1996: 229-237.
26. Goldenberg, D.L., Fibromyalgia, chronic fatigue syndrome, and myofascial pain. *Current opinion in Rheumatology* 1996; 8: 113-123.
27. Wolfe, F. and the Vancouver Fibromyalgia Consensus Group. The fibromyalgia syndrome. *J Rheumatology* 1996; 23: 534-539.
28. Ferraz, M.B., Maetzel, A., Bombardier, C. A summary of economic evaluations published in the field of rheumatology and related disciplines. *Arthritis & Rheumatism* 1997; 40 (9): 1587-1593.
29. Maetzel, A., Ferraz, M.B., Bombardier, C. A review of cost-effectiveness analyses in Rheumatology and related disciplines. *Current opinion in Rheumatology* 1998; 10: 136-140
30. Goossens, M.E.J.B., Evers, S.M.A.A. Economic valuation of back pain interventions. *J Occupational Rehabilitation* 1997; 7 (1): 15-32.
31. Thondike, E.L. *The fundamentals of learning*. New York: Teachers College, 1932.
32. Fordyce W.E. *Behavioral methods for chronic pain and illness*. St Louis: The Mosby Company, 1976.
33. Vlaeyen, J.W.S. *Chronic low back pain. Assessment and treatment from a behavioral rehabilitation perspective*. Rockland/Lisse: Swets & Zeitlinger, 1991.
34. Melzack, R., Wall, P.D. *The challenge of pain. Exciting discoveries in the new science of pain control*. New York: Basic Books Inc. Publishers, 1982.
35. Turk, D.C., Meichenbaum, D., Genest, M. *Pain and Behavioral Medicine, a cognitive behavioral perspective*. New York: Guilford Press, 1983.
36. Beck, A.T. *Cognitive therapy and the emotional disorders*. Boston: International Universities Press, 1976.
37. Holzman, A.D., Turk, D.C., Kerns, R.D. The cognitive-behavioral approach to the management of chronic pain. In: A.D. Holzman & D.C. Turk, editors. *Pain management: A handbook of psychological treatment approaches*. Elmsford, NY: Pergamon Press, 1986: 31-50.
38. Bradley L.A. Cognitive-behavioral therapy for chronic pain. In: R.J. Gatchel & D.C. Turk, editors. *Psychological approaches to pain management. A practitioner's handbook*. New York, London: The Guilford Press, 1996: 131-147.

39. Hagglund, K.J., Fillingim, R.B., Cost offset from cognitive-behavioral interventions for chronic pain. *Arch. Phys. Med. Rehabil* 1998; 79: S83-S88.
40. Linton, S. Psychological factors in back and neck pain and its treatment (Submitted).
41. Turk, D.C. Efficacy of multidisciplinary pain centers in the treatment of chronic pain. In: M.J.M. Cohen and J.N. Campell, editors. *Pain treatment centers at a crossroads: A practical and conceptual Reappraisal*, Vol. 7, Seattle, WA: IASP Press, 1996: 257-273.
42. Flor, H., Fydrich, T., Turk, D.C., Efficacy of multidisciplinary pain treatment centers: a meta-analytic review. *Pain* 1992; 49: 221-230.
43. Van Tulder, M.W., Koes, B.W., Bouter, L.M. Conservative treatment of acute and chronic nonspecific low back pain. A systematic review of randomized controlled trials of the most common interventions. *Spine* 1997; 22 (18): 2128 - 2156.
44. Morley, S., Eccleston, C., Williams, A. Systematic review and meta-analysis of randomised controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. *Pain* (in press) 1999;80:1-13.
45. Flor, H. *A way out of the vicious circle: Learning to manage chronic pain through self-control*. Berlin: Humboldt University, 1997.
46. Federico, J.V., The cost of pain centers: Where is the Return? In: In: M.J.M. Cohen and J.N. Campell, editors. *Pain treatment centers at a crossroads: A practical and conceptual reappraisal*, Vol. 7. Seattle, WA: IASP Press, 1996: 249-256.
47. Caudill, M.A. Clinical Implications of the NIH Technology and Assessment Conference Addressing Behavioral Treatment of Chronic Pain. *IASP Newsletter* March/April; 1998: 3-7.
48. Deyo R.A., Cherkin D., Conrad D., Volinn E. Cost, controversy, crisis: low back pain and the health of the public. *Ann Rev Publ Health* 1991; 12: 141-156.
49. Organisation for Economic Co-operation and Development (OECD). *Historical Statistics 1960-1994*. Paris: OECD, 1996.

2 PRINCIPLES OF ECONOMIC EVALUATION FOR INTERVENTIONS OF CHRONIC MUSCULOSKELETAL PAIN

This chapter is accepted for publication as:

Mariëlle E.J.B. Goossens¹, Silvia M.A.A. Evers², Johan W.S. Vlaeyen^{1,2}, Maureen P.M.H. Rutten- van Mólken⁴, Sjeff M.J.P. van der Linden⁵. Principles of economic evaluation for interventions of chronic musculoskeletal pain. *European Journal of Pain*, 1999

¹ Institute for Rehabilitation Research, Hoensbroek

² Department of Neurosurgery, University Hospital of Maastricht

³ Department of Medical, Clinical and Experimental Psychology, Maastricht University

⁴ Institute for Medical Technology Assessment, Erasmus University Rotterdam

⁵ Department of Internal Medicine/Division of Rheumatology, University Hospital of Maastricht

Summary

Economic evaluation is attracting increasing attention to inform policy makers, insurers and other payers about the value of existing and new treatment modalities. Hence it is desirable to assess not only the medical but also the economic consequences the new treatments produce. The available literature on economic evaluation revealed an urgent need for sound economic evaluation studies in the field of chronic musculoskeletal pain. Due to the generally weak methodology, the intended purpose of economic evaluation to help set funding priorities, has often been bypassed. A clear understanding of the main principles of economic evaluation studies might therefore be advantageous. This chapter addresses important methodological issues in economic evaluation research, such as the techniques for economic evaluation studies and the analytic perspective. In addition, the chapter pays attention to the inclusion of costs and outcomes in economic evaluation research, sensitivity analysis, discounting, incremental analysis and ratios, and collecting of data. Further emphasis is placed on the transparent reporting of methods and study results. A clear reporting may help therapists and (other) researchers interpret the results of published studies and apply them to new studies, and it may help decision makers generalize results from one setting to another.

2.1 Introduction

Cost-effectiveness analysis has taken on an increasingly larger role in health care policy debates about interventions for chronic musculoskeletal pain. Growing health care costs and productivity losses, disappointing treatment results and changing beliefs in health and pain have led to this increase in concern about the amount of money spent on chronic musculoskeletal pain. Although therapists have no direct responsibility for allocating scarce resources in the field of chronic musculoskeletal pain, they are confronted with the results of these decisions in their everyday work. It might therefore be advantageous to clearly understand economic evaluation studies and its principles.

More and more studies in the field of chronic musculoskeletal pain incorporate cost issues in their analysis. A MEDLINE-search combining Thesaurus 'musculoskeletal disorders/back pain/spine' and 'cost-and-cost-analysis' shows both an absolute and a relative increase of the number of chronic musculoskeletal pain studies including economic aspects (see figure 1). This does not indicate whether the results of all these studies can ultimately be used for decision making. Government, insurers and other payers will be more willing to implement a certain intervention when its cost-effectiveness is clearly established and compared to other treatments. The field of study in which both the costs and outcomes of alternative interventions are weighed against each other is called 'economic evaluation'.

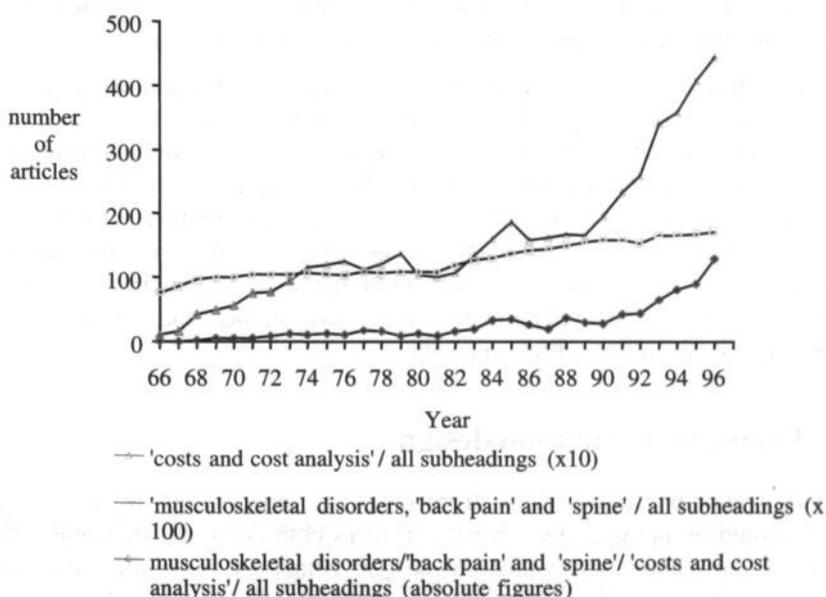


Figure 1. Number of studies in the field of musculoskeletal pain incorporating cost issues.

Three recently performed reviews of economic evaluation studies in the field of musculoskeletal pain showed that besides the already small number of studies included in the reviews, the number of studies which actually demonstrated a sound economic evaluation was even smaller.¹⁻³ Moreover, mainly due to the large differences, lack of clarity and low quality of the costing methodology, it was not possible to reach conclusions on the most cost-effective therapies in this area.

In this context, the low priority given to economic evaluation can partly be explained by the complexity of assessing the cost of pain. The subjective nature of pain (and discomfort), the complexity of assessing quality of life and the little historical weight given to costs for carers and families, make cost evaluation in pain treatment a complex undertaking.⁴ Furthermore, an important explanation of the paucity of good cost-effectiveness studies (which is not typical for chronic musculoskeletal pain) and different interpretations of economic evaluation is related to the evolution of this type of research. Over recent decades, the term 'cost-effectiveness' has come into practice as an instrument to inform policy makers of the value of a certain treatment. Developers and analysts come from a number of academic disciplines, including economics, medicine, operations research, medical sociology, public health, and ethics, bringing their own perspectives, applications and sets of methods.⁵ As a consequence this has led to the current imprecision attached to the term 'cost-effectiveness' and difficulties in the conceptualization of the tasks necessary in undertaking economic evaluations. As noted in the mentioned reviews¹⁻³, the intended purpose of economic evaluations, to help setting funding priorities, has often been bypassed. Especially in chronic musculoskeletal pain, because of increasing economic burden of illness, knowing what treatments actually give value for money is necessary.

This chapter provides an introduction to economic evaluation and discusses the most important issues in the design of such studies, concentrating on the field of chronic musculoskeletal pain. These issues include: the design of the economic evaluation, the analytic perspective, identification, measurement and valuation of costs and outcomes, sensitivity analysis, discounting and collecting of the data. Further emphasis is placed on the transparent reporting of methods and study results. A clear reporting of methods, analysis and results may help both therapists to interpret the results of published studies and to apply them to their own studies and decision makers to generalize results from one setting to another.

2.2 Economic evaluation designs

Economic evaluation studies differ from other outcome studies in that the research question should be phrased such that it considers both costs and outcomes. The gold standard in an economic evaluation design compares both the costs and the consequences of two or more alternative interventions.⁶ When these characteristics are fulfilled, we speak of *full* economic evaluation.

The identification of consequences determines which type of full economic evaluation study is under consideration: cost-minimization analysis (CMA), cost-benefit analysis (CBA), cost-effectiveness analysis (CEA) or cost-utility analysis (CUA). In a CMA the effects of the alternatives are equivalent and are therefore not valued. The evaluation in a CMA is essentially a search for the least costly alternative. In CBA, CEA and CUA, the consequences of the alternatives differ, and these consequences are measured in respectively monetary units, natural units, or utilities (e.g. Quality-Adjusted-Life-Years: QALYs). A CBA tries to express the health improvement in monetary terms. Costs and benefits are then measured in the same unit, and one can see immediately whether the benefits outweigh the costs. To date, the Willingness to Pay (WTP) method is often seen as the best method to value the effects.⁶ In the Willingness to Pay method, individuals are explicitly asked how much money they would be willing to spend to reduce the probability of illness or death. However, the WTP is still experimental and under development and the use of the CBA is therefore not recommended yet.

In chronic musculoskeletal pain, CEA and CUA would be the most appropriate methods to use, since in these studies the effects are measured in natural units and quality of life respectively. The outcome measures used in CEA studies in chronic pain research mainly include outcomes, such as disability days saved, pain-free days or improved quality of life (this will be explained in more detail in the section 'outcomes'). In CUA the consequences are measured in patients' preferences for specific health states or specific health outcomes. In general, in CUA the effects are expressed in QALYs, by which preferences expressed in valuation of quality of life are combined with the number of life-years gained or (more relevant in chronic musculoskeletal pain) the remaining life years. A cost-utility analysis has the advantage (compared to a cost-effectiveness analysis) that not only is a comparison of the results of the two experimental treatments and other programs in the same field possible, but so is a comparison with the results of interventions in other health care fields.

Apart from the full economic evaluation designs mentioned above, several types of *partial* evaluations can be distinguished: cost analysis, cost-outcome description and cost description. Partial means that these studies either do not examine both the costs and the outcomes of the alternatives (cost analysis), or do not compare two alternative interventions (cost-outcome description) or do only examine the costs of one alternative (cost-description). These evaluations may represent important intermediate stages in the understanding of the costs and consequences of health services and programs. However, the label 'partial evaluation' does indicate their inability to yield answers to efficacy questions.⁶

2.3 Comparison of alternatives

As a minimum, in economic evaluation studies, the experimental intervention should be compared to another intervention, preferably to 'standard practice' or the 'best-available alternative'.⁵ However, in chronic musculoskeletal pain, existing practice is not apparent and varies a lot and also the 'best available alternative' is debatable.

Observational studies would take a view of the current practice for chronic musculoskeletal pain. Precaution should be taken as well when comparing an experimental treatment with an 'attention-control' or a waiting-list control. These comparison groups have been subject of debate in psychotherapy and behavioral research since several years.⁷⁻⁹

2.4 Analytic perspective

What costs and outcomes will be examined and what conclusion can be drawn from an economic evaluation study is further determined by the perspective of the study.¹⁰ The perspective is the viewpoint from which the study is conducted, mostly the interesting party that conducts the study. Ideally, the results of economic evaluation studies are to be used for resource allocation on the societal level and therefore economic evaluation studies are preferably performed from the perspective of the society.^{6,11} Nevertheless, clinical managers and third-party payers have also become increasingly interested in the evaluation of new and existing alternatives in chronic musculoskeletal pain and are demanding evidence of costs and effects in monetary and non-monetary terms.¹² The advantage of adopting the broader perspective is that data can be disaggregated and the analysis conducted from a number of viewpoints. Nevertheless, one must keep in mind that the societal perspective does not necessarily represent the complete situation from the viewpoint of particular parties. Furthermore, it is also possible to 'add up' different perspectives to a societal perspective, if attention is made to avoid double counting. For instance the patient perspective may be of a particular interest in order to identify the out-of-pocket charges that a patient has to pay for pharmaceutical drugs, although for the societal perspective these out-of-pocket charges are not included (since they are already included in the cost of the drug). Therefore, it is recommended at study onset to precisely describe the research questions of the different actors involved and take this into account in the design of the study, such as the society, the clinical managers, the Ministry of Health, third party payers and the patients.

2.5 Relevant costs and outcomes

The type of economic evaluation, the perspective and the study frame will determine the structure of the analysis and the cost and outcome data to be used. Figure 2 highlights most of the relevant costs and consequences that may be taken into account when evaluating health care interventions for chronic musculoskeletal pain.

2.5.1 Costs

Costs are the sum of resource use in volumes multiplied by their unit costs. To get some insight into the performance of the costing of a study, the identification, measurement and valuation of the costs should be reported clearly. Figure 2 covers the most relevant concepts in these three steps of evaluating costs.

Identification. In the economic evaluation of chronic musculoskeletal pain interventions costs analysis will focus on costs that are expected to differ between the alternative treatments. These are direct health care costs, such as costs of the intervention and actual changes in health care utilization due to the program, direct non-health care costs, such as costs borne by the patient and the family, and finally indirect costs. Patient and family costs, such as out of pocket costs, paid and unpaid help, travel costs and time costs, are of major importance to include, because of the enormous impact of chronic pain on the patients and the family. As demonstrated in chapters 4 and 5, direct non-health care costs can easily exceed the costs of health care utilization. Indirect costs, which refer to the value of production lost to society due to illness-related absence from (paid and unpaid) work and days of inactivity, are important to take into consideration. For diseases in general, 70% of the total costs are direct and 30% are indirect. For back pain, however, this is the other way around. In several countries more than 70% are indirect, whereas less than 30% are direct.²²⁻²⁴

Measurement. It is dependent on the context of the economic evaluation what data source is the most appropriate, and often several sources are used. There are various sources to obtain data, such as patient records, databases from insurance companies, hospital or provider databases, interviews with patients or providers, questionnaires, diaries, data from other studies or modeling costs into the future.⁵ Guidelines on economic evaluation rarely report on the most adequate method to obtain these data. Routine databases compiled by insurance companies, hospitals and other health care institutions (such as pharmacies and general practitioners) are frequently used. However, these sources may not record either total medical consumption or costs borne by patients and their families.²⁵ In studies evaluating chronic musculoskeletal pain programs one wants to get a complete picture of the costs, which makes these direct costs as well as the indirect cost important to include. Because there are no institutional records tracking information on the type, frequency and magnitude of these expenditures, data must at least partly be obtained through self-reporting channels such as questionnaires and diaries.^{5,6} In this, diaries are prospective, tend to have a better data quality, and are more useful in long-term clinical trials.²⁵ Recently in two CEAs evaluating chronic musculoskeletal pain treatment, as described in chapters 4 and 5, a patient cost diary was developed to collect the most relevant direct (health care and non-health care) and indirect costs as completely as possible.^{26,27}

Valuation. Valuation means the assignment of resource values to the quantities of cost components measured. Theoretically, the proper price of a cost component or resource is its opportunity cost (which means that the costs of choosing the best alternative equal the benefits lost of the next best alternative which is not chosen). Because of the difficulty to estimate opportunity costs, the pragmatic approach to costing is to take existing market or unit prices, for example the charges for a visit to the physiotherapist. The valuation of patient and family costs can be based on real costs reported by the patients or by shadow prices, for example for unpaid help by family or friends. Indirect costs such as production losses often form a major component of a chronic musculoskeletal pain study. Unfortunately, consensus is lacking on the best approach to value indirect costs. Two approaches are used to value the production losses in

chronic musculoskeletal pain research: the Human Capital Approach and the Friction Cost Method. The Human Capital Approach, which values indirect costs based on the expected economic productivity of the individual, is most frequently used.⁵ In the Human Capital Approach the extent of production losses is valued by (market) earnings. According to the newer Friction Cost Method, production losses will be confined to the period needed to replace a sick worker: the so-called friction period.²⁸ In general, estimates based on the Friction Cost Method are about 10% to 20% of those based on the Human Capital Approach.

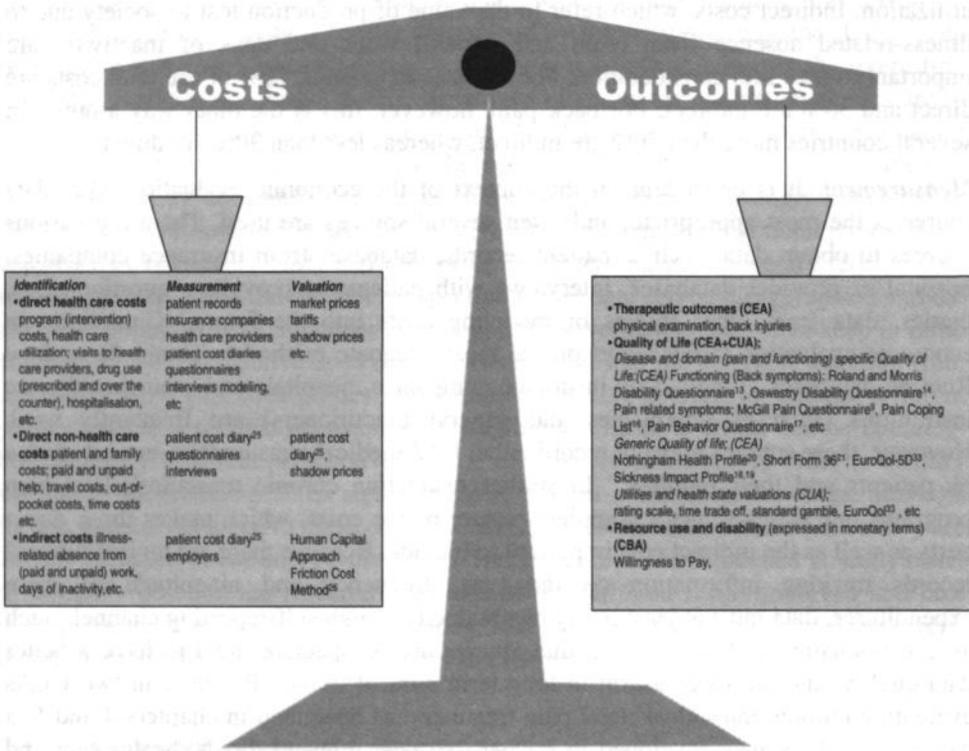


Figure 2. Components of economic evaluation in health care

2.5.2 Outcomes

Apart from the costs, consequences or outcomes of the alternatives should be examined in an economic evaluation study. Comparisons between the outcomes found in economic evaluations of chronic musculoskeletal pain treatments are difficult, because of the large variation in outcome measures used. Stieg *et al.*²⁹ suggested that better agreement about the taxonomy of pain patients and standards of outcome measurement would facilitate comparisons between studies. They classified the most important outcomes for pain patients in therapeutic outcomes (such as physical outcome measures), quality of life of patients and their families and resource use or

benefits (such as reductions in lost work time). Since chronic musculoskeletal pain is not an objectively demonstrable disorder and therapeutic outcomes refer only to an individual's ability to function, it is recommendable to combine therapeutic outcomes with other outcomes such as health related quality of life measures. These quality of life measures should consist of a subset of dimensions which are the most appropriate for chronic musculoskeletal pain patients, such as pain symptoms, functional status, psychological status, social interaction and overall well-being.³⁰ For patients with chronic musculoskeletal pain, loss of quality of life is likely to be perceived as a major consequence of their life. A wide array of measurement instruments have been used within chronic musculoskeletal pain to assess overall quality of life. To improve the comparability between studies, it is recommended to use those measures which are most frequently used in the field, such as those mentioned in figure 2. For a more extensive outline we refer to Deyo *et al.*³¹

Quality of life instruments can be classified in specific and generic measures. In chronic musculoskeletal pain, domain or disease-specific instruments are most frequently used. These measures provide more information in a concise way and are in general highly sensitive to show any change between two interventions. Among these, measures focussed on functional status and psychological aspects related to pain are most important to include. For clinicians who want to decide between two treatments, domain or disease specific instruments can be of major importance. For a more broad comparison, generic measures have gained importance because they reflect the impact of ill health on the lives of people in a variety of populations and can therefore be used to make comparisons within medical domains and across medical domains. Comparison can be made not only between different interventions for chronic musculoskeletal pain, but also between interventions for chronic musculoskeletal pain and for example, asthma. This makes these measures useful in health care decision making and in economic evaluation studies. Subcategories of the generic measures include health profiles and utility measures. Health profiles are single instruments that measure different aspects of quality of life in a variety of conditions. A scoring system permits aggregation of the information collected into a score or index. Among the health profiles, the SF-36 is internationally the most frequently used.²¹ Utility measures are derived from economic theories and go one step further than the health profiles, in that they not only describe a health state but also value this health state. Among utility measures, the standard gamble, rating scale and the time trade-off method have been most often used. These three measures provide us with a quantitative measure of the value or preference the chronic musculoskeletal pain patients attach to their overall health status compared with perfect health (score of 1) and death (score of 0). In cost-utility analysis the utilities are used as correction weights for the remaining years of life or the life-years gained by the program. This calculation results in QALYs or the remaining life-years due to the intervention. To date, there is no agreement on which utility measure is best to be used. Account should be taken that due to differences in scaling and contextual factors the preference scores vary between the methods.³² In general within economic evaluation studies, the EuroQol instrument is recently frequently assessed.³³ Utility measurement is not yet frequently used in chronic musculoskeletal pain. But at this moment, it seems that for chronic musculoskeletal

pain the rating scale is most closely related to other domain (pain) specific measures.³² Nevertheless, it is advisable not to rely just on one type of measure, but to use both specific and generic instruments which best suit the study objectives. Furthermore, to facilitate comparability among studies, it is recommended to use those instruments which are most frequently used in the field. However, which type of quality of life instruments are selected will depend on the perspective taken in the study in question (e.g. physician's, patient's or health care policy perspective).

2.6 Sensitivity analysis and discounting

The information needed for an economic evaluation is never readily available. This means that often estimates are made, based on the best available information. However, estimations of costs and effects contain some degree of imprecision or uncertainty. Sensitivity analysis should be performed, using different assumptions or estimates in order to test the impact of changing these estimates on the results and conclusions. While including such an analysis, it is important to give the reasons for specifying the plausible ranges for the estimates. The results of the sensitivity analysis may show the dependence or otherwise of a conclusion on a particular assumption.

Another important factor is that different studies may have different time profiles of costs and consequences. In chronic musculoskeletal pain in general the benefits of the intervention might occur or extend into the future while the costs will be incurred immediately. In this case one has to discount the benefits and costs at some rate to find their present value. Discounting is not relevant if the study period is less than one year. Several published studies have used a discount rate of 5%.⁵ Given currently available data on real economic growth and corresponding estimates of real consumption rates of interest, a discount rate of 3% was regarded as the most appropriate.⁵ Drummond *et al.*⁶ recommend providing a sensitivity analysis for the discount rate, including 0 percent, 3 percent and 5 percent.

2.7 Incremental analysis and ratios

The core of any economic evaluation is an incremental comparison of the costs and consequences between alternative interventions. The interpretation of this comparison needs to be guided by formal methods of statistical inference of the results of both the costs and outcomes. Conclusions regarding the evidence about cost-effectiveness can only be reliably made with statistical analysis.

The results of comparison are further expressed into one ratio that is calculated by dividing the difference in costs by the difference in effect. Depending on the kind of analysis, this can be either a cost-effectiveness, a cost-benefit or a cost-utility ratio. Examples of cost-effectiveness ratios in chronic musculoskeletal pain may be 'the cost per pain-free day gained', 'the cost per patient reaching the minimal clinically important improvement in quality of life', and 'the cost per work day gained'. An example of the calculation of a ratio is shown in table 1.

Table 1. Example of the calculation of a cost-effectiveness ratio.

Suppose an economic evaluation study compares the cost-effectiveness of two alternative interventions for the treatment of chronic musculoskeletal pain: an experimental treatment (EXP) and a standard (control) treatment (CNTRL):

Costs	
Total direct costs + total indirect costs resulting from the programs	
C_{exp} = total direct costs + total indirect costs resulting from EXP	\$5 000
C_{cntrl} = total direct costs + total indirect costs resulting from CNTRL	\$3 000
$C_{exp} - C_{cntrl}$ = incremental costs for EXP as compared to CNTRL	\$2 000
Effects	
Total improvement in preference scores resulting from the programs	
E_{exp} = total improvement in preference scores resulting from EXP	0.1
E_{cntrl} = total improvement in preference scores resulting from CNTRL	0.03
$E_{exp} - E_{cntrl}$ = incremental improvement in costs for EXP as compared to CNTRL	0.07
Remaining life-years of the patient	
U_{exp} = QALYs following EXP = $0.1 * 34$	3.4
U_{cntrl} = QALYs following CNTRL = $0.03 * 34$	1.02
$U_{exp} - U_{cntrl}$ = incremental improvement in QALYs for EXP as compared to CNTRL	2.38
Cost-effectiveness ratio	
$(C_{exp} - C_{cntrl}) / (E_{exp} - E_{cntrl}) = (\$5000 - \$3000) / (0.1 - 0.03) = \2000 per 0.07 improvement in health state preferences	
Cost-utility ratio	
$(C_{exp} - C_{cntrl}) / (U_{exp} - U_{cntrl}) = (\$5000 - \$3000) / (3.4 - 1.02) = \840.34 per QALY	

2.8 Collecting the data

Besides the chosen perspective and the costs and outcome data needed, the structure of the data collection has to be decided. In general, primary data on costs and effects can be collected from the original study or secondary data from other sources. Often economic evaluation studies are referred to as 'piggyback' studies as these studies are commonly 'piggy-backed' -added on to- randomized controlled trials (RCTs). As a consequence, economic and additional health outcomes are often of secondary concern. Statistical significance for the economic results are often difficult to achieve, because the study power is calculated using clinical endpoints. Only in a few instances RCTs are especially designed to determine the economic significance. RCTs however, may not be representative and generalisable for the real world situation and do not address issues at a more aggregate level.³⁴ Furthermore, these studies are typically large and lengthy and therefore more expensive.⁵

In some cases it may not be possible to gather all the required cost and effectiveness primary data. For example, in chronic musculoskeletal pain it is often not possible to define usual care or existing practice, in which data from other sources, such as observational studies may be required. When data are gathered from separate sources, for the final cost-effectiveness, the information can be combined into a conceptual model.⁵ Especially in interventions with long-term consequences, such as interventions for chronic musculoskeletal pain, existing information from other studies and modeling of outcomes into the future can be useful to conduct cost-effectiveness studies.

2.9 Conclusion and discussion

Over the years economic evaluation has become a substantial part in the evaluation of alternative services. As a result the field has expanded very quickly. This chapter has given a basic insight into the use of the principles of economic evaluation in the field of chronic musculoskeletal pain. An extensive overview of each of these principles was beyond the scope of this chapter. However, by briefly reviewing each of these properties, we highlighted those issues that are important to consider when designing or evaluating an economic evaluation in chronic musculoskeletal pain and to interpret the results of published studies in the field. For further details we refer to the guidebooks of Drummond *et al.*⁶ and Gold *et al.*⁵ in which the methodological criteria for assessing economic evaluation are described in detail. We also refer to the Reference Case proposed by Gold *et al.*⁵ which is a standard set of methodological practices that an analyst would seek to follow in a cost-effectiveness study. Nevertheless, we have to keep in mind that there is no superior strategy for the measurement of both cost and outcomes. We have shown that in chronic musculoskeletal pain some outcomes are less accessible than they may be in other disorders, such as therapeutic outcomes, and some outcomes and costs need to be given higher priority, such as quality of life and patient and family costs.

We have therefore stressed the importance of transparent reporting of methods and study results, especially when the study results have to be informative to other researchers in the field or have to be useful for comparison. Researchers have to be careful to not blindly apply someone's methodology and compare their own results to these other studies. Therefore, when comparing the literature on economic evaluation, an important question to keep in mind is whether all economic evaluations used the same techniques. As we have seen, studies may be performed from different viewpoints and use different research questions; they compare different alternatives, different costs and outcomes; and finally data may be obtained in different ways. Recent reviews in the field of economic evaluation of musculoskeletal pain not only showed the poor methodological quality of the studies, but also confirmed the lack of comparison between the studies because of the differences in methodology.¹⁻³ To resolve this dilemma there is thus a need to work *and* report as closely as possible in accordance with the standards of economic evaluation proposed in the guidebooks. This will ultimately facilitate comparison between studies, because one can be more confident that differences in results reflect the characteristics of the interventions being evaluated, rather than differences in study methodology.

Furthermore, in the phase of developing an economic evaluation the researcher must be aware of the key questions posed by the different parties and to incorporate these in the design of the study. It is recommended that the different parties having an interest in the study, work closely together during the whole process. For clinicians and therapists it is important to be aware that cost-effectiveness analysis demonstrates the consequences of allocative decisions which may have future implications for their own treatment policy. Therefore their involvement in policy making and the process of economic evaluation is needed.

References

1. Ferraz, M.B., Maetzel, A., Bombardier, C. A summary of economic evaluations published in the field of rheumatology and related disciplines. *Arthritis & Rheumatism* 1997; 40 (9): 1587-1593.
2. Maetzel, A., Ferraz, M.B., Bombardier, C. A review of cost-effectiveness analyses in rheumatology and related disciplines. *Current Opinion in Rheumatology* 1998; 10: 136-140
3. Goossens, M.E.J.B., Evers, S.M.A.A. Economic evaluation of back pain interventions. *Journal of Occupational Rehabilitation* 1997; 7 (1): 15-32.
4. Ferrell, B.R. How patients and families pay the price of pain. In: M.J.M. Cohen and J.N. Campbell, editors. *Pain treatment centers at a crossroads: A practical and conceptual reappraisal*, Vol. 7, IASP Press, Seattle, WA, 1996: 229-237.
5. Gold M.R., J.E. Siegel, L.B. Russel, M.C. Weinstein, editors. *Cost-effectiveness in health and medicine*. New York Oxford, Oxford University Press, 1996: 82-134.
6. Drummond, M.F., O'Brien, B., Stoddart, G.L., Torrance G.W., *Methods for the economic evaluation of health care programmes*. Second edition. Oxford, Oxford University Press, 1997.
7. Parloff, M.B. Placebo controls in psychotherapy research: A sine qua non or a placebo for research problems? *Journal of Consulting and Clinical Psychology* 1986; 54: 79-87.
8. Strayhorn, J.M. Control groups for psychosocial intervention outcome studies. *American Journal of Psychiatry* 1987; 144 (3): 275-282.
9. Schwartz, C.E., Chesney, M.A., Irvine M.J., Keefe F.J. The control group dilemma in clinical research: Applications for psychosocial and behavioral medicine trials. *Psychosomatic Medicine* 1997; 59:362-371.
10. Luce, B., Elixhauser, A. Estimating costs in the economic evaluation of medical technologies. *International Journal of Technology Assessment in Health Care* 1990; 6: 57-75.
11. Weinstein, M.C. Principles of cost-effective resource allocation in health care organization. *International Journal of Technology Assessment in Health Care* 1990; 6: 93-103.
12. Spilker B. *Quality of life assessment in clinical trials*. Raven Press. New York, 1990.
13. Roland, M., Morris, R. A study of the natural history of back pain. I: Development of a reliable and sensitive measure of disability in low back pain. *Spine* 1983; 8:141-4
14. Fairbank, J.C.T., Davies, J.B., Mbaot, J.C. O'Brien, J.T. The Oswestry Low Back Pain Disability Questionnaire. *Physiotherapy* 1980; 66:271-273.
15. Melzack, R. *Pain measurement and Assessment*. Raven Press, New York, 1983; 33-37.
16. Vlaeyen, J.W.S., Geurts, S.M., van Eek, H., Kole-Snijders, A.M.J., Schuerman, J.A., Groenman, N.H., What do chronic pain patients think of their pain? Towards a pain cognition questionnaire. *British J of Clinical Psychology* 1990; 29: 95-118.
17. Richards, J.S., Nepomuceno, C., Riles, M. & Suer, Z., Assessing pain behavior: the UAB pain behavior scale. *Pain* 1982; 12: 393-398.
18. Bergner, M., Bobbit, R.A., Carter, W.B., Gilson B.S. The Sickness Impact Profile: Development and final revision of a health status measure. *Medical Care* 1981; 8:787-805.
19. Bruin, de A.F., Buys, M., Witte, de L.P., Diederiks, J.P.M. The Sickness Impact Profile: SIP68; a short generic version, first evaluation of the reliability and reproducibility. *Journal of Clinical Epidemiology* 1994; 44; 863-874.
20. Kind, P., Carr-Hill, R., The Nottingham health profile: A useful tool for epidemiologists? *Social Science of Medicine* 1987; 25: 905-910.

21. Ware, J.E., Sherbourne, C.D., The MOS 36-item short-form health status survey: 1. Conceptual framework and item selection. *Medical Care* 1992; 30 (6): 473-481
22. van Tulder, M.W., Koes, B.W., Bouter, L.M., A cost of illness study of back pain in the Netherlands. *Pain* 1995; 62:233-240.
23. Coyle, D., Richardson, G. The cost of back pain. In: *Clinical standards Advisory Group report on back pain*. London. HSMO, 1994: 65-71.
24. Norlund, A. Cost of back pain in Sweden. In: *SBU-report on back pain*, Stockholm. SBU (accepted).
25. Goossens, M.E.J.B., Rutten-van Mólken, M.P.M.H., Vlaeyen, J.W.S., Linden van der, S.M.J.P. The cost diary: a method to measure direct and indirect costs in cost-effectiveness research. (Provisionally accepted in *Journal of Clinical Epidemiology*).
26. Goossens, M.E.J.B., Rutten-van Mólken, M.P.M.H., Leidl, R.M., Bos, S.G.P.M., Vlaeyen, J.W.S., Teeken-Gruben, N.J.G., Cognitive-educational treatment of fibromyalgia: a randomised clinical trial II. Economic Evaluation. *Journal of Rheumatol* 1996; 23(6): 1246-1254.
27. Goossens, M.E.J.B., Rutten-van Mólken, M.P.M.H., Kole-Snijders, A.M.J., Vlaeyen, J.W.S., van Breukelen, G., Leidl, L., Health economic assessment of behavioural rehabilitation in chronic low back pain; A randomised clinical trial. *Health Economics*, 1998; 7: 39-51
28. Koopmanschap, M.A., Rutten F.F.H. Indirect costs in economic studies; confronting the confusion. *Pharmacoeconomics* 1993; 4(6): 446-454.
29. Stieg, R.L., Williams R.C., Timmermans-Williams G., Tafuro, F., Fallagher L.A., Cost benefits of interdisciplinary chronic pain treatment. *Clinical Journal of Pain* 1986; 1: 189-193.
30. Patrick, D.L., Erickson, P., *Health status and health policy. Quality of life in health care evaluation and resource allocation*. New York Oxford; Oxford University Press 1993.
31. Deyo, RA, Andersson, G, Bombardier, C, Cherkin, DC, Keller, RB, Lee, CK, Liang, MH, Lipscomb, B, Shekelle, P, Spratt, KF, *et al*. Outcome measures for studying patients with low back pain. *Spine* 1994, 15;19 (18 Suppl): 2032S-2036S.
32. Goossens, M.E.J.B., Vlaeyen, J.W.S., Rutten-van Mólken, M.P.M.H., Linden van der, S.M.J. Patient utilities in chronic musculoskeletal pain. How useful is the standard gamble method? *Pain* 1999; 80:365-375.
33. EuroQol Group. EuroQol: A new facility for the measurement of health related quality of life. *Health Policy* 1990; 16:199-208.
34. Baltussen, R.M.P.M., Ament A.J.H.A., Leidl, R.M. Making cost assessments based on RCTs more useful to decision makers. *Health policy* 1996; 37:163-183.

3 ECONOMIC EVALUATION OF BACK PAIN INTERVENTIONS

Parts of this chapter are published as:

Goossens, M.E.J.B.^{1,2}, Evers, S.M.A.A.^{1,3} Economic evaluation in back pain interventions. *Journal of Occupational Rehabilitation* 1997; 7(1); 15-32.

Parts of this chapter are accepted for publication as:

Goossens, M.E.J.B.^{1,2}, Evers, S.M.A.A.^{1,3} Cost-effectiveness of treatment of back pain. In: A. Nachemson, E. Jonsson, Editors. *Evidence based treatment for back pain*. Stockholm: SBU and Philadelphia: Lippincott.

¹ Institute for Rehabilitation Research, Hoensbroek

² Department of Health Organisation, Policy and Economics, Maastricht University

³ Department of Neurosurgery, University Hospital Maastricht

Summary

Despite the rising health care costs of back pain, evaluation of the economic aspects of back pain interventions has received little attention. This chapter reviews 30 economic evaluation studies of back pain interventions looking at costs, consequences and methodological aspects. This paper gives a detailed summary of the economic evaluation data given by these studies and provides insight into the quality of economic evaluation in the field of back pain. Four of the studies dealt with injury prevention programs, 10 with post-incidence management and secondary prevention programs and 16 studies with back injury and pain remedies (clinical treatment). The injury prevention programs (except for back school programs) and post-incidence management and secondary prevention programs appear to produce cost savings due to reduced absenteeism generally for (acute) back pain in the workplace. This paper shows that the small number of comparable treatments and the large differences in costing methodologies and their low quality render it impossible to draw conclusions in favor of any one type of program.

3.1 Introduction

Low back pain is a growing health problem. It is one of the most common and expensive disorders in Western countries. In the literature, the life-time incidence is reported to range from 60-80% and the prevalence rate from 12-35%.¹⁻³ Low back pain is a condition with an early onset (about 25 years) which has a substantial effect on the quality of life of people in their working years. A variety of factors seem to play a role in the development of back pain, such as occupational problems, poor health habits and psychosocial problems.²⁻⁶ The difficulty of making a good diagnosis and the inability to detect the exact cause make it very difficult to find an effective therapy. This is one important reason why, in recent decades, the range of treatments for low back pain has expanded significantly.

Many economic impact studies, in various countries, have reported substantial costs due to back pain.^{3,7-15} In addition to the total costs of back pain, several authors have studied the costs per case.^{7,16-18} They found that the costs of back pain are not normally distributed because a minority of back pain patients (about 10 to 25%) account for most of the expenses.^{14,19} There are enormous costs to society arising from these cases, which are categorized as chronic back pain patients. These costs are primarily due to disability pensions and the patients' 'medical shopping' behavior. Costs can be classified as either direct costs, those arising from medical and non-medical care, or indirect costs, due to reduced productivity due to back pain.²⁰ For diseases in general, 70% of total costs are direct and 30% are indirect¹¹, but for back pain the indirect costs exceed the direct costs.^{9,11,15,19}

Since the incidence and costs of back pain are large and the socioeconomic implications due to disability are growing rapidly, it is becoming important to examine how much of the burden and costs can be avoided by effective therapies, and at what cost. This question can be answered by means of economic evaluation studies. Economic evaluation studies, of which cost-effectiveness analyses are a specific example, compare both the costs and the outcomes of alternative health care interventions. This information can be used to support decision-making about the allocation of scarce resources and to obtain the maximum gain in health. Many studies in the field of low back pain incorporate some cost issues in their analyses. However, previous research has shown that sound economic evaluation in this area has received little attention, in contrast to some other chronic medical problems.²¹

The purpose of this chapter is to provide insight into the status and the quality of economic evaluation studies performed through 1997 in the field of back pain. To start with, a brief outline of the principles of economic evaluation will be given. After that, the methods of this review will be explained. The first result section discusses the alternative treatments and the nature of the economic appraisals found in the literature. The subsequent sections outline the methods used in the measurement of costs and outcomes in these studies and finally discusses the quality of the effectiveness evidence on which the economic evaluations are based will be discussed.

3.2 Principles of economic evaluation

Before discussing the articles in detail it is useful to explain the principles of economic evaluation.

3.2.1 Economic evaluation designs

A full economic evaluation study must compare two or more program alternatives, examining both the costs (inputs) and the consequences (outputs).²² As table 1 shows there are four types of full economic evaluation, depending on the nature of the consequences resulting from the alternative interventions. These four types are cost-minimization analysis, cost-effectiveness analysis, cost-benefit analysis and cost-utility analysis.²²

Table 1. Main characteristics of economic evaluation studies

Method	Comparison of two or more alternatives	Measuring costs	Measuring effects
cost-minimization analysis (CMA)	Yes	monetary value	equal effectiveness; measurement not necessary
cost-effectiveness analysis (CEA)	Yes	monetary value	natural units (e.g., life years gained, pain free days saved, disability days saved, etc.)
cost-utility analysis (CUA)	Yes	monetary value	Quality Adjusted Life Years (QALYs)
cost-benefit analysis (CBA)	Yes	monetary value	monetary value
cost analysis (CA)	Yes	monetary value	effects are not measured
cost outcome description (CO)	No	monetary value	natural units (e.g., life years gained, pain free days saved, disability days saved, etc.)
cost description (CD)	No	monetary value	effects are not measured

A *cost-minimization analysis* (CMA) examines alternative health care programs which are equally effective, so that only the costs need to be compared further. The evaluation is then essentially a search for the 'least-cost' alternative. In a *cost-effectiveness analysis* (CEA) the outcome can be assessed in a variety of ways, such as life-years gained, cases prevented, decreased length of hospital stay or pain-free days. One of the major limitations of the cost-effectiveness analysis is that it can only compare programs that are measured in the same health-related units. Studies cannot be compared if the outcome measures used are not the same. A *cost-utility analysis* (CUA), however, does express health improvement in one uniform measure, the quality-adjusted life year (QALY). A quality of life correction factor called utility is used to adjust life-years for their quality. The utility reflects the preference for a particular health state. Finally, a *cost-benefit analysis* (CBA) expresses the outcomes of the study in monetary terms. Costs and benefits are then measured in the same unit and one can see immediately whether the benefits outweigh the costs.

Studies that do not meet all the conditions for a full economic evaluation are classified as partial evaluations. Such studies may well be worthwhile to perform (and for us to

include in the review), as they may represent important intermediate stages in understanding the costs and consequences of health services or programs.²² But these studies cannot provide a definitive answer to questions of allocation. When drawing conclusions we will take this distinction into account. The major partial evaluations which we included in our review are *costs descriptions* (CD), *costs analysis* (CA) and *cost-outcome* studies (CO). CD and CA are solely aimed at the measurement of costs. In a CD the costs of one program or service are examined. In a CA only the costs of two or more alternatives are compared. In a CO the costs and consequences of only one intervention are examined.

3.2.2 Costs and outcomes

In general, in economic evaluations, the cost factors and consequences should be described, valued and measured accurately. When describing the costs, one has to determine which costs are taken into account. Overall economic evaluation studies are preferably performed from a societal perspective, which means that all costs which are relevant for all participants should be measured, e.g., costs for patients, health care suppliers, insurers and for society as a whole.

Having determined which costs are relevant to the study, one has to determine how these cost can be measured. Cost data can be collected in various ways, such as through suppliers (health organizations, pharmacists and general practitioners), through health insurance companies and through the patients themselves. After identifying which costs are to be included in an economic evaluation and how they are measured, costs have to be valued. This means putting monetary labels on the volumes found. In economic evaluations, charges are often used as a proxy for the costs²³, although one is really interested in the actual economic costs. Charges are not the same as costs: they could better be described as 'negotiated prices'.

The consequences of a study can be determined by great variety of outcome measures. Drummond *et al.*²³ classified the consequences of health programs in three categories. Category one consists of the therapeutic outcomes or effects of the alternatives, such as physical outcome measures. Category two comprises changes in the quality of life of patients and their families, as a result of health care services and programs. Category three covers changes in resource use or benefits. The benefits may accrue to the health sector (e.g., reduced costs, increased flexibility of operations) or to the patients and their families (e.g., reductions in leisure time input or lost work time). In this study, changes in the amount of lost work time will be considered as part of costs, rather than benefits.

After this brief outline in the principles of economic evaluation, the methods will be further explained by using the studies reviewed in this chapter.

3.3 Methods

Several strategies were combined to find literature in the field of the economic evaluation of back pain. Studies published between 1984 and 1997 were identified

with a Medline and EMBASE search, using the keywords 'costs', 'cost-analysis', 'cost-effectiveness analysis', 'back pain', 'spine' and 'economic evaluation'. We also used the references of the NHS economic evaluation database and the Health Economic Evaluation Database (HEED) of the Office of Health Economics and the International Federation of Pharmaceutical Manufacturers Association.^{24,25} Furthermore, additional articles were found by citation tracking and through an economic evaluation bibliography.²¹ Articles were included in the review if the field of study was back pain, the study used an economic evaluation, and the article was published in English.

Table 2. Main items of the checklist

<i>General</i>	<i>Economic</i>
Country	Economic evaluation design used
Perspective of analysis	Costs identification (direct, direct non-medical, indirect costs)
Nature of the study	Costs measurement (sources direct, direct non-medical, indirect costs)
Disease category	Costs valuation (valuation direct, indirect costs)
<i>Epidemiologic</i>	Identification consequences
Epidemiological design used	Measurement consequences
Design characteristics (prestratification/matching, randomization, population size, number of groups)	Valuation consequences
Adequate measurement of effect (blinding)	Discounting, differential timing
Results	Sensitivity analysis

A methodological checklist²⁶ was used to select only those articles describing an economic evaluation study in the field of back pain treatment. Using the checklist the studies were screened on general, epidemiological and economic aspects (see table 2 for the main criteria used) by two independent reviewers (MG and SE). Disagreements between the two reviewers were resolved by consensus.

3.4 Results

3.4.1 Study selection

After excluding editorials, letters, non-English articles and reviews, we included 30 articles referring to economic evaluations of back pain treatments.^{18,24,27-45,47-55,59} In the review we included 3 more studies which referred not to the economic evaluation of a treatment but to either the evaluation of different types of care^{46,56} or to the evaluation of diagnosis.⁵⁷ A large number of other studies were excluded, because they did not appear to be economic evaluation studies. Some studies only mentioned the term costs or cost-effectiveness in their abstract or the conclusion without further report. Most of the studies included in the review were performed in North America, with two studies performed in Australia and ten in Europe (Finland, Germany, the Netherlands, Norway and Belgium). Tables 3 and 4 give an overview of the studies reviewed for this chapter. Table 3 contains the information about the intervention and the conclusions and table 4 summarizes the most important findings on the economic evaluation.

Table 3. Conclusions on the cost-effectiveness (C-E) of various types of interventions in back pain

Author	Intervention	Reference treatment	Diagnosis	Results	Author's conclusion on C-E	Statistics performed on C-E
INJURY PREVENTION AND BACK SCHOOLS						
Brown ²⁹	Back school	No treatment control	Back injury/ Municipal employees (work related)	Significant fewer injuries. Sign. Difference on lost work time, lost time cost and medical costs for back school participants. Dif. Between groups but not statistical.	Negative	No
Versloot ³⁵	Back school	No treatment control	Back pain	Decrease in absenteeism of 5 days, but no sign. Difference between groups	Positive	No
Hochanadel ³²	On site industrial physical therapy program	Situation before	Musculoskeletal injury (50% LBP)	Significant reduction in the absence rate. For the life of the program it results in net savings (program costs - savings)	Positive	Assumptions for the life of the program
Coleman ³⁰	Educational program	No treatment control	Work-related back injury (bus company employees)	Decrease in length of absenteeism by 5 days per year per employee. Significant decrease in cost Per back injury claim during 2nd year of the program. Total cost savings in 3 year; program costs - costs of back injuries (decreased).	Positive	Not clear
POST-INCIDENCE MANAGEMENT/SECONDARY PREVENTION						
Aaras ²⁷	Ergonomic interventions (at the workplace)	No comparison	Musculoskeletal illness	Reduction in sick leave	Positive	No
Linton ³³	Secondary prevention program	Situation before	Sub-acute lbp (female nurses)	Significant decrease in pain and use of medication, increase in activity level, substantial reduction in absenteeism and employer costs.	Positive	Assumptions
Linton ³⁴	Early active intervention	Treatment as usual	Acute musculoskeletal pain	Significant less sick-listing and 8 times lower risk of developing chronic pain; less chronic cases would result in lower costs.	Positive	Assumptions
Mitchell ⁴⁰	Back belts	No treatment control	Low back injuries (warehouse workers)	Back belts are minimally effective in preventing injury, costs of injury are higher while wearing a belt.	Not CE	Yes
Mitchell ³⁹	Functional restoration program	Treatment as usual	Soft tissue and back injuries (injured workers, 71% male)	Difference (not significant) in absence from work, compensation costs and total costs	Positive, but not significant	Yes
Ryan ⁴⁴	Educational and early rehabilitation	No treatment control	Early back pain	Easy to institute and inexpensive program, succeeding in preventing back pain chronicity	Positive	

Author	Intervention	Reference treatment	Diagnosis	Results	Author's conclusion on C-E	Statistics performed on C-E
Ryden ⁴⁵	Health promotion program; I back care program II light duty program	Situation before	Back injuries (hospital employees)	Decrease in back injuries, and decrease in costs of back injuries (savings in medical bills, workers' compensation cost and lost time)	Positive	
Simmons ⁴⁷	Multi-disciplinary therapy	Situation before	Chronic low back pain (with and without radiculopathy)	Reduction in medical costs	Positive	Yes
Stieg ⁵¹	Interdisciplinary pain treatment	Situation before	Chronic pain	Reduction in health care costs	Positive	No
MANIPULATION AND PHYSICAL THERAPY						
Meade ³⁹	Chiropractic treatment	Conventional hospital outpatient management	Low back pain of mechanical origin	Clinical improvement in pain intensity and disability, resulting in savings in social security payments	Positive	Assumptions
Nyiendo ⁶¹	Chiropractic treatment	Other medical care	Low back pain	Increase (not sign. Different) in return to work	Negative	No
Skargen ⁴⁹	Chiropractic treatment	Physiotherapy	Low back and neck pain	Same effectiveness and costs for chiropractic and physiotherapy	Negative	Yes
Tuchin ⁵³	Chiropractic treatment	Conventional medical management	Back pain, spinal injuries	Fewer days off work due to chiropractic treatment	Possibly positive	Pilot study
EXERCISE THERAPY						
Malmivaara ³⁶	Ordinary activity	Bed rest vs exercises	Acute low back pain	Light exercise resulted in slower recovery after 3 weeks, avoiding bed rest leads to rapid recovery. Difference but not significant in costs between groups	Positive	No
Mitchell ⁴¹	Intensive exercises	Treatment as usual	Back injuries, acute soft tissue	Treatment program resulted in earlier recovery, enabling a greater percentage to return to work sooner. The increase in health care costs were offset by savings in wage loss cost	Positive	No
Timm ⁵²	Active treatment	Passive treatment	Chronic low back pain	Active approaches of low-tech and high-tech exercise significantly improved objective functional measures, low tech provides the longest interval of pain relief and may therefore be cost-effective	Positive	Assumptions
SURGERY						
Bell ²⁸	Spinal cord stimulation	'chronic maintenance'	Failed back surgery syndrome	Improvements in scs clinical efficacy would result in less demand of medical care	Positive (long term)	Assumptions

Author	Intervention	Reference treatment	Diagnosis	Results	Author's conclusion on C-E	Statistics performed on C-E
Devulder ³¹	Spinal cord stimulation		Failed back surgery syndrome	In a period of 13 years, SCS resulted in pain relief and return to work. Although SCS is expensive, it may be c-e in the long term	Positive (long term)	Assumptions
Malter ³⁷	Lumbar discectomy	No surgical treatment	Herniated intervertebral disc	For carefully selected patients unresponsive to conservative treatment, it provides substantial benefit in quality of life. Total costs were greater for the surgical treatment	Positive	No
Manucher ³⁸	Chemonucleolysis	Laminectomy	Herniated lumbar disc	Rates of return to work and patient improvement are greater for CNL. No difference in results of clinical examination. CNL resulted in charge savings	Positive	Partly
Ramizer ⁴²	Chemonucleolysis	Laminectomy	Herniated nucleus pulposus, LBP and neck pain	5-day difference in length of hospitalization resulting in cost savings for chemonucleolysis (less reoperation)	Positive	Assumptions
Ray ⁴³	Ray threaded Fusion Cage	Anterior composite interbody grafting with posterior pedicle screw fixation (360° fusion)	Back pain	Improvement in overall surgical and hospitalization costs, surgical time and blood losses provided by threaded fusion cage technique (assumable similar clinical outcomes)	Positive	No
Stevenson ⁵⁰	Automated percutaneous lumbar discectomy	Microrisectomy	Contained lumbar disc herniation	Worse outcomes on a four-point scale for the apld, costs per successful outcome for microdescectomy was 60% of the average cost per apld successful outcome	Negative (for apld)	No (poor analysis)
Tunturi ⁵⁴	Posterior fusion of the lumbar spine	-	Patients subjected to fusion of lumbosacral spine	Profitable at a cost-benefit ratio of 1:29 (costs compared work output)	Positive	No
OTHER THERAPIES/PROCEDURES						
Liang ⁵⁷	Lumbar roentgenograms	Doing nothing	Acute low back pain	To avert one day of physical suffering, the population would have to be subjected to the additional risk of 3,188 mrad of radiation and an additional cost of \$2,072; benefit is too small to justify the costs and risk	Negative	Partly
Lissovoy ³⁵	Intrathecal morphine therapy	Medical management	Failed back surgery syndrome	In a simulation model the average patient achieved 43.8 months of pain relief. IMT may save money relative to conventional therapy	Positive	Model

Author	Intervention	Reference treatment	Diagnosis	Results	Author's conclusion on C-E	Statistics performed on C-E
Maclean ⁵⁶	Family physicians care	Hospital care	Medical back pain	Family physicians treat inpatients with back pain probably cheaper than other physicians without any compromise of effectiveness	Positive	No
Shekelle ⁴⁶	General practitioner	Chiropractor, internist, orthopedist	Back pain	Lowest cost for care provided by gps in comparison with care provided by chiropractors, internists and orthopedists	-- (comparison)	

Table 4. Main characteristics of the studies

First Author & year	Intervention and reference treatment	Diagnosis/subjects	Study design	Economic design	N exp.	N cont.	Costs *	Valuation costs	Main outcome measures	Dis-coun-ting	Sensi-tivity analyses
Injury prevention and back schools											
Brown, 1992 ²⁹	Back school vs no treatment control	Back injury, municipal employees	Cohort	CEA	70	75	1,2,4	Charges HCA ⁷	Sick days, lost time cost, medical expenses, number of back injuries	No	No
Coleman, 1994 ³⁰	Educational program vs no treatment control	Back injury, bus company employees	Cohort	CD	30	-	1,2,4	Unclear	-	No	No
Hochanadel, 1993 ³²	Physical therapy and back school vs situation before	Musculoskeletal illness (50% lbp)	Cohort	CDd	5301		1,4	Estimates	-	No	No
Versloot, 1992 ⁵⁵	Back school vs no treatment control	Back pain	RCT	CEA	200	300	1,4	Unclear	Sick days	No	No
Post-incident management / secondary prevention											
Aarås, 1994 ²⁷	Ergonomic interventions	Musculoskeletal illness	Cohort	CO ⁵	420	-	1,4	Estimates	Reduction in recruitment costs, training costs, instructors' salary costs, sick leave	Yes	No
Linton, 1992 ³³	Secondary prevention program vs situation before	Sub-acute LBP, female nurses	Cohort	CEA ⁵	36	-	1,4	Estimates	Pain intensity, fatigue, anxiety, sleep, ADL, depression, helplessness, medication, sick-absenteeism	No	No
Linton, 1993 (40)	Early active intervention vs treatment as usual	Acute musculoskeletal pain	RCT	CEA	134	106	1	Estimates	Treatment satisfaction, pain discomfort, pain-free days, sleep quality, stress, depression, well-being, pain control, sickness absenteeism	No	No
Mitchell, L., 1994 ⁴⁰	Back belts vs no treatment control	LB injury, warehouse workers	Cohort	CEA	1000	1000	1,2,4	Unclear	Back injuries, lost time due to injuries, limited activity days	No	No
Mitchell R., 1994 ⁵⁹	Functional restoration vs treatment as usual	Back injury (injured workers, 71% male)	RCT	CA	271	271	1,2,4	Unclear	Return to full-time work, number of patients granted income support	No	No

First Author & year	Intervention and reference treatment	Diagnosis/subjects	Study design	Economic design	N exp.	N cont.	Costs	Valuation costs	Main outcome measures	Discounting	Sensitivity analyses
Ryan, 1995 ⁴⁴	Early rehabilitation and education vs no treatment control	Early back pain	Cohort	CA	Un-clear	Un-clear	1,2,4	Charges		No	No
Ryden, 1988 ⁴⁵	2 health promotion programs; 1 educational, 2 light duty program vs situation before	Back injury, hospital employees	Cohort	CEA ⁵	Un-clear	-	1,2,4	Charges	Hospital costs, workers compensation costs and lost time	No	No
Sinclair, 1997 ⁴⁸	Early active intervention vs usual care	Acute back injury	Cohort	CEA	355	530	1,2,4	Unclear	Time to return to work, health care costs, health related quality of life, functional capacity, pain	No	No
Simmons, 1988 ⁴⁷	Multidisciplinary therapy vs situation before	70% clbp, with and without radiculopathy	Cohort	CO	60	-	1,2	Charges	Medical expenses	No	No
Stieg, 1986 ⁵¹	Interdisciplinary pain management vs situation before	Chronic pain	Cohort	CO ⁵	53	-	1,2,4	Unclear	Decreased health care costs to the insurers	No	No
Manipulation and physical therapy											
Meade, 1990 ³⁹	Chiropractic vs conventional hospital outpatient treatment	LBP of mechanical origin	RCT	CEA	384	357	1,4	Average costs	Oswestry pain disability index, straight leg raising and lumbar flexion	No	
Nyiendo, 1991 ⁶¹	Chiropractic vs medical treatment	LBP, back injury	Cohort	CD	94	107	1,2,4	Unclear	Medical expenses and decreased sick days	No	
Skargren, 1997 ⁴⁹	Physiotherapy vs chiropractic therapy	low back and neck pain	RCT	CEA	219 chir.	192 phys.	1,2,4	Charges	Pain intensity, use of pain killers, sick leave, oswestry lbp disability index, general (global) health	no	
Tuchin, 1995 ⁵³	Chiropractic therapy vs conventional medical management	Back pain, spinal injuries	Pilot study (20 AT random.)	CEA	20	20	1,2	Not clear	Days off work	No	

First Author & year	Intervention and reference treatment	Diagnosis/subjects	Study design	Economic design	N exp.	N cont.	Costs *	Valuation costs	Main outcome measures	Dis-counting	Sensi-tivity analyses
Malmivaara, 1995 ³⁶	Bedrest vs exercises vs ordinary activities	acute LBP	RCT	CEA	67 + 52	67	1,2,3,4	Charges actual patient costs	Days of sick leave, intensity and duration of pain, Oswestry back disability index, quality of life, ability to work, lumbar flexion	Yes	No
Mitchell, 1990 ⁴¹	Intensive exercises and educational sessions vs treatment as usual	Back injury	Cohort	CEA ⁵	703	2172	1,2,4	Unclear	Compensation costs for wage loss and health care	No	No
Timm, 1994 ⁵²	Active treatment vs passive treatment	CLBP	RCT	CEA	4 X 5050		1,2	Average costs	Pain status, episode of clbp, length of pain relief, Oswestry pain disability index	No	No
Surgery											
Bell, 1997 ²⁸	Spinal cord stimulation	Failed back surgery syndrome	Cohort	CD	Dif. Samples	Litera-ture	1,2	Estimates and charges	Medical costs of scs, medical resource use	yes	Yes
Devulder, 1997 ³¹	Spinal cord stimulation vs 'chronic maintenance'	Failed back surgery syndrome	Cohort: 13 years	CO	69		1,2, (4)	Unclear	Use of stimulator, return to work, pain relief	no	No
Malter, 1996 ³⁷	Lumbar discectomy	Herniated intervertebral disc	Based on RCTs	CEA/CUA	Dif. Literature samples		1	Cost estimates, tariffs	quality of life, quality adjusted life years	Yes	No
Manucher, 1995 ³⁸	Chemonucleolysis vs laminectomy	Herniated lumbar disc	Cohort	CEA	100	100	1,2	Charges	Patient evaluation, clinical examination, rate of return to work, change in job	No	No
Ramirez, 1985 ⁴²	Chemonucleolysis versus laminectomy	Herniated nucleus pulposus, LBP, neck pain	Cohort	CEA ⁵	40	40	1,2	Charges	Length of hospitalization, unsatisfactory results, repeat operations	No	No
Ray, 1997 ⁴³	Lumbar interbody fusions vs anterior composite interbody grafting	Back pain	Cohort	CMA	25	25	1,2	Charges	Blood loss, surgical time translated into costs	Yes	No
Stevenson,	Automated percutaneous	Contained lumbar disc	RCT	CEA	31	40	1,3	Actual data	Treatment success on 4-	No	No

First Author & year	Intervention and reference treatment	Diagnosis/subjects	Study design	Economic design	N exp.	N cont.	Costs *	Valuation costs	Main outcome measures	Dis-counting	Sensi-tivity analyses
1995 ⁵⁰	lumbar discectomy (APLD) vs microdiscectomy	herniation			(APLD)				point scale		
Tunturi, 1979 ⁵⁴	Posterior fusion	Patients subjected to fusion of Lumbosacral spine	Cohort	CO	118		1,3,4	Charges	Work output	Yes	No
Other therapies											
Liang, 1982 ⁵⁷	Roentgenograms in primary care vs no roentgenograms	Acute LBP	Cohort	CEA	Un-clear	Un-clear	1	Charges	Days of suffering from undiagnosed and untreated disease, radiation exposure, (in millions), dollar Charges for lumbar roentgenograms	No	Yes
Lissovoy, 1997 ³⁵	Intrathecal morphine therapy vs conventional management	Failed back surgery syndrome	Model	CEA	Model	Model	1,2	Charges	Pain relief	Yes	Yes
Maclean, 1993 ⁵⁶	Hospital care given by family physicians or other physicians	Back pain	Cohort	CEA	68	560	1,2	Charges	MedisGroups major morbidity, in-hospital mortality, length of stay	No	Yes
Shekelle, 1995 ⁴⁶	Comparison of provider types of care	Back pain	Cohort	CA	686 ⁶		1,2	Charges	-	Yes	No

* Costs identified

1. program costs (direct)
2. costs of other health care utilization (direct)
3. out of pocket expenses of patients and their families (direct)
4. costs of productivity losses due to absence from work (indirect)

Abbreviations

CMA	Cost Minimization Analysis
CBA	Cost Benefit Analysis
CEA	Cost Effectiveness Analysis
CUA	Cost Utility Analysis
CA	Cost Analysis
CO	Cost Outcome Description
CD	Cost Description
HCA	Human Capital Approach

LBP Low Back Pain

CLBP Chronic Low Back Pain

⁵ some of the consequences in this study are valued in monetary terms

⁶ 686 persons were divided in 5 subgroups depending on back pain care

3.4.2 Economic evaluation designs

The most frequently used economic evaluation design in our review (20 studies) is the cost-effectiveness analysis. The consequences in these studies were valued in natural units, for example pain intensity, pain duration, health care consumption, absenteeism and continuation of activities. In four of the cost-effectiveness analyses some of the consequences were translated into monetary terms.^{33,41,42,45} We did not classify them as cost-benefit analysis because this type of analysis measures all costs and benefits in pecuniary units and computes a net monetary gain or loss. The study of Ray⁴³ was classified as cost-minimization analysis, since effectiveness between two lumbar interbody fusion procedures (Ray Threaded Fusion Cages and 360° fusions) were assumed to be the same. Although Malter *et al.*³⁷ classified their study as a cost-effectiveness analysis, it can also be categorized as a cost-utility analysis, since benefits are finally expressed in quality adjusted life years. Three of the studies^{41,45,46} analyzed only costs, without an economic assessment of outcomes. These are classified as cost analyses. Another four partial evaluation studies^{18,28,30,32}, also restricted their analysis to costs but were classified as cost descriptions because they did not compare the costs of alternatives. Finally five studies could be classified as cost-outcome description, since they evaluated the costs and outcomes of only one intervention.^{27,31,45,51,54}

3.4.3 Alternatives compared

Another important standard in economic evaluation research is the comparison of at least two alternatives. The alternatives with which the various treatments were being compared were not always explicitly mentioned in the articles. Of the 33 studies, 25 used a comparison of two or more alternative interventions, other medical care, or a 'no intervention control'. Six studies compared the intervention to the situation before.^{27,45,51} Two studies did not make a comparison at all.

The studies reviewed evaluated a wide range of different types of treatments. To provide an overview of the studies, we listed the treatments according to Cats-Baril and Frymoyer⁹ into three categories for which cost-effectiveness studies are important: injury prevention and 'back schools' (such as workplace design and screening programs); post-incidence management and secondary prevention programs (return-to-work programs designed to bring the injured worker back to a productive status as soon as possible and to avoid the 'not-working' downward spiral); and back injury and pain remedies which involves medical treatment.

Injury prevention and back schools. Back school programs have been studied extensively during the last few years. Originally developed in Sweden, back schools offer an educational and skills program in a group setting and are directed toward pain management. Two of the studies evaluated the cost-effectiveness of a back school program.^{29,55} Both studies compared the treatment with a no treatment control, which is the most reasonable alternative treatment in prevention programs because omitting the program will not harm the patient in the sense that his current health state will become worse. In accordance with the conclusions from other evaluations of back schools⁵⁸,

these two studies showed that back school programs do not result in different cost-effectiveness compared to no treatment control. In a study of Hochanadel and Conrad³², an on-site industrial physical therapy program was compared to the situation before implementation of the program, and found a reduction in absenteeism and therefore future savings. In Coleman and Hansen's study³⁰ an educational program compared to a no treatment control resulted in a decrease in total costs of back injuries, in terms of lost time and medical expenses. Conclusions on cost-effectiveness cannot be taken because both studies were only partial evaluation studies.

Post-incident management programs and secondary prevention programs. As emphasized above, the indirect costs due to lost working time are substantial for back pain patients. Post-incident management programs, aimed at returning patients to work, are therefore very important in the treatment of back pain. Ten of the studies dealt with such post management and secondary prevention programs.^{27,33,34,40,44,45,48,51,64,69} Although the content of these programs differs enormously, the overall conclusion from these studies is that such programs, generally for acute back pain in a work setting, result in cost savings, especially by reducing the number of lost work days and overall time loss from work. Nevertheless it should be taken into consideration that only five of these studies performed a full economic evaluation.⁶⁰ Two programs^{45,51} compared the new program to the situation at baseline. In these studies the experimental group served as its own control. The other studies compared the treatment of the study group to a no treatment control^{27,40}, to alternatives given to patients who were treated in the community at large^{34,41,59}, with treatment as usual^{33,48} or did not compare the interventions.²⁷

The last intervention category focuses on clinical treatments and was subdivided into 'manipulation and physical therapy', 'exercise therapy', 'surgery' and 'other therapies'. The treatments were usually compared with another treatment, or the usual treatment, rather than with a no treatment control.

Manipulation and physical therapy. Four studies, of which three were cost-effectiveness studies, compared chiropractic treatment with another kind of therapy.^{24,39,49,53} In summary, these studies concluded that chiropractic treatment was cost-effective compared to conventional treatment (but these conclusions were not statistically tested) and equally effective compared to physiotherapy.

Exercise therapy. Three cost-effectiveness studies compared active treatment with several kinds of passive treatment.^{36,41,52} Although also here, statistics were poor, the studies concluded that active treatment reduced absenteeism and increased mobility in acute and chronic low back pain.

Surgery. Eight studies evaluated some type of surgical intervention, of which five were full economic evaluations. Most authors reported positively on the cost-effectiveness of these interventions, although the conclusions were not based on statistical evidence. The economic consequences of spinal cord stimulation for patients with a failed back surgery syndrome were evaluated positively in two recent studies.^{28,31} Ramizer *et al.*⁴² and Manucher and Javid³⁸ concluded that chemonucleolysis was cost-effective compared to laminectomy for the treatment of herniated lumbar disc. The studies of

Malter *et al.*³⁷ and Stevenson *et al.*⁵⁰ concluded that lumbar discectomy was cost-effective compared to no surgical treatment and less cost-effective compared to microdiscectomy. Finally, Ray Threaded Fusion Cage to patients with progressive disabling low back pain⁴³ and posterior fusion of the lumbosacral spine⁵⁴ were evaluated and deemed to result in cost-savings.

Other therapies. Four studies could not be categorized under one of the previous headings and were therefore not comparable to other studies in the review. Two studies comparing different providers of back pain care concluded that care given by the general practitioner was significantly less costly, compared with care provided by the chiropractor, internist, orthopedist⁴⁶ or other caregivers in the hospital.⁵⁶ The study by Liang and Komaroff⁵⁷ on the cost-effectiveness of obtaining a diagnostic roentgenograms in acute low back pain patients concluded that the risks and costs of performing lumbar roentgenograms at the initial visit do not seem to be justified by the relatively small benefit which results. Finally based on a modeling study, Lissovoy *et al.*³⁵ concluded that intrathecal morphine therapy may result in cost savings relative to conventional therapy for patients with failed back surgery syndrome.

The next sections describe the evaluation methods used in the partial and full economic evaluation studies. A summary of the most important findings can be found in table 4.

3.4.4 Methods used in determining costs

Costs identified. Although the perspective of the study determines the costs and consequences that will be examined, only four studies mentioned the perspective explicitly.^{28,35,37,43} All studies included the cost of the intervention itself and 24 studies included the costs of other health care services.^{28-31,35,36,38,40-49,51-53,56,59,61,62} Costs borne by the patients and their families were only included in the studies of Malmivaara *et al.*³⁶, Stevenson *et al.*⁵⁰ and partly in Tunturi *et al.*⁵⁴ Indirect costs were measured by the patient's absenteeism in 21 studies.^{27,29-34,36,39,40,41,45,47-49,51,54, 55,59,61,62}

Data sources. Following the selection of costs relevant to the study, one has to determine how these costs can be measured. To measure medical consumption, some studies use the data-bases of health care suppliers^{29,31,36,38,43,45,49,50}, insurance companies^{27,32,36}, workers compensation boards⁴⁸ or questions to the patient.^{27,32,36,40,47,50} Once, a combination of two or more data sources was used to independently validate medical consumption measures.⁴⁷ Information on absenteeism was obtained from (national) insurance companies^{33,34,61,44,51} and the data-bases of employers^{27,55,62} and from the patient.³¹ It was unclear how some studies obtained their cost data.^{57,41,59,52,54}

Cost valuation. After identifying which costs are to be included in an economic evaluation and how they are measured, costs have to be valued. This means putting currency labels on the volumes found. Only the study of Stevenson⁵⁰ calculated the actual economic costs related to the intervention. Thirteen used charges^{28,29,35,36,42-44,46,47,54,56,57,61}, cost estimates^{27,32-34} or average costs.^{39,52} The remaining studies were vague about how they obtained costs or fees. In general, the Human Capital Approach is the most frequently-used method of valuing indirect costs. In this approach lost production which results from temporary absence from work, disability or premature death is

valued using the average earnings.^{63,64} However, in our sample only Brown *et al.*²⁹ used the Human Capital approach to value indirect costs. Skargren *et al.*⁴⁹ estimated the indirect costs from the mean income in different sex and age groups, including social costs. The other studies which took indirect costs into account were not clear as to how they valued them.

3.4.5 Methods used in determining outcomes

An economic evaluation study should examine not only the costs of the alternatives but also the consequences (or the 'output'). The studies included in the review used a large variety of outcome measures. To facilitate the comparison among the outcomes, we categorized the outcomes into therapeutic measures, quality of life measures and changes in resource use or benefits.²²

Therapeutic measures. In principle, therapeutic outcomes can be measured objectively. Therapeutic outcomes refer only to an individual's ability to function and not to the significance, preference or value attached to this ability by the individual or by others.³⁸ In several studies these therapeutic outcomes are measured in terms of duration of pain and number of back injuries. Because back pain is frequently not an objective demonstrable disorder, therapeutic outcomes can best be combined with other outcomes such as quality of life measures, since quality of life addresses different aspects of the disease.⁶⁵

Quality of life measures. As mentioned earlier, quality of life provides an important outcome measure for back pain treatments. The concept of 'quality of life' in this setting has been limited to 'health-related quality of life' or 'health status'. Measures of these dimensions can generally be classified as either disease-specific or generic. Eight studies used quality of life measures.^{33,34,36,37,39,48,49,52} Seven of these studies used domain-specific questionnaires focusing on those aspects directly related to having pain and on pain-related complaints. Such questionnaires describe changes in domain-specific quality of life. They provide good discrimination between alternative therapies, and may be sufficient as outcome measures, particularly when all aspects change in the same direction and comparisons with other interventions in different medical areas are not required. But since they focus on the specific aspects of a disease they are not applicable across diverse populations. Generic instruments and utility measures are broadly applicable since they cover a broader range of dimensions of health-related quality of life, and aggregate the dimension scores into a single overall score.⁶⁶ Since economic evaluation studies are often directed towards overall comparisons between disease categories, generic and utility measures are considered to be important as outcome measures. Only the studies of Malter *et al.*³⁷ and Sinclair *et al.*⁴⁸ used generic measures. The study of Malter *et al.*³⁷ also used time trade off scores to determine quality adjusted life years. However with regard to this study it should be mentioned that the quality of life data were based on data from another trial.⁶⁷

Changes in resource use or benefits. 'Changes in resource use and benefits' turned out to be the most important outcome measure in back pain studies. Some of the studies even restricted the outcome assessment to these intermediate effects of the intervention.^{27,36,41,45,47,51,53-56,59,61,62} In the studies selected for this review, outcomes

dealing with absenteeism and return to work seemed to be important. This was not a surprise because most studies dealt with back pain-related injury and acute back pain. In many studies the outcomes were measured in volumes and translated into monetary units. For instance, days of absenteeism may be translated into compensation costs for wage loss, or length of hospitalization may be translated in terms of hospital costs. Therefore such changes in resource use and benefits were considered in our review along with the methods of measuring costs, which have been discussed in the previous section.

3.4.6 Discounting and sensitivity analysis

Another important factor in economic evaluation is that different studies may deal with costs and consequences which have different time dimensions. In some programs the benefits and costs may occur immediately, while in others the benefits of the intervention may continue, or only be realized, in the future whereas the costs are incurred immediately. In the latter case one has to discount the figures, to convert future costs and benefits into equivalent present values.⁶⁸⁻⁷³ Only a few studies discounted their data^{27,28,35-37,43,46,54,62}, although other studies collected data over several years.^{29,31,33,39,42,45,47,51,55,61}

Furthermore, a sensitivity analysis has to be performed to adjust for the assumptions made during the study. If the changes in results due to the different assumptions are minor, the results may be treated with greater confidence. If the sensitivity analysis produces significant changes in the results, then greater caution is necessary when interpreting the data and more effort is required to reduce the uncertainty or improve the accuracy of the critical variables. However, a sensitivity analysis was performed in only four studies.^{28,35,56,57}

3.4.7 Some methodological aspects of the studies

A sound epidemiological design is very important for an economic evaluation. The design of the study is one of the main determinants of the quality of the results from economic evaluation studies. The quality of results is especially important for clinical and public policy makers, who use them to make decisions concerning alternative treatments. Economic evaluation studies can be conducted preferably using prospective data (for example with an RCT), (more commonly) using retrospective data, or by modeling.²⁰ Out of the 21 full economic evaluation studies in our review which used comparison groups, only 11 randomized the patients into these groups.^{29,34,36,39,49,50,52,53,55,59,61} Eight studies did not use an alternative therapy as a comparison at all.^{27,30,32,44,45,47,51,62} Several studies gave either ethical or practical reasons for not using an alternative comparison group (intervention).^{51,54,62,74} Drummond⁷⁵ argued that, if no alternative comparison group is used, effectiveness information based on previous good quality studies should be provided. Four of the eight studies^{32,33,45,47} report the conclusions of other studies of the same treatment, but do not say anything about the quality or validity of these earlier treatment results.

Finally it was remarkable that at least twelve studies did not report on statistically significant evidence on the costs outcomes and the target variable for the economic

evaluation.^{37,38,41-43,50, 51,54-57,61} However, these studies did generally provide reasons why their findings can nevertheless be called significant.

3.5 Conclusion and discussion

Despite the rising health care costs of low back pain and the increased interest in economic evaluations, the economic aspects of low back pain interventions have received little attention. A systematic search from 1984 until 1997 through several data-bases found several authors who suggested there was a need for economic evaluation of back pain interventions, but only 30 studies in English were found which actually contained such an evaluation. Of these studies 21 were classified as full economic evaluation studies.

The interventions for back pain have been divided into injury prevention programs, post-incidence management and secondary prevention programs, and clinical treatment. From this review it was impossible to draw definite conclusions on the cost-effectiveness of each category. The small number of comparable treatments, the large differences in and low quality of the costing methodology and the large differences between the back pain populations, limit the comparisons between the estimates in the studies reviewed. Furthermore in the majority of the studies the evidence about differences in costs and effects were made without statistical analysis. Conclusions on cost-effectiveness were therefore sometimes misleading. It should be mentioned as well that in some studies reviewed in this paper the main purpose was not the economic evaluation of the treatment, but the assessment of the effect of the treatment instead. As a result the methodology of performing an economic evaluation was backgrounded. Differences in cost-effectiveness possibly reflect the characteristics of the evaluated programs and the differences between methods, rather than the differences between costs and effects. Nevertheless, conclusions concerning the economic impact of the interventions were still made.

Injury prevention programs, except for back school programs, (for work-related back injury or back pain) seem to result in cost savings due to reduced absenteeism and medical costs. The same is true for the post-incidence management and secondary prevention programs. The results of the studies dealing with chiropractic treatment favored this treatment in comparison to conventional treatment. Furthermore, exercise therapy (compared to passive treatment), spinal cord stimulation and chemonucleolysis were found to produce more favorable results. Mainly due to low costing methodology there was only limited evidence on cost-effectiveness for lumbar discectomy and lumbar interbody fusions.

Most of the studies in the review dealt with acute back pain and back pain related injuries. The cost-effectiveness of the different therapies was therefore mostly expressed in reduced absenteeism. Chronic low back pain interventions were hardly subjected to economic evaluation.

The major type of economic evaluation in the studies reviewed is a cost-effectiveness analysis. In performing an economic evaluation, the relevant cost factors and

consequences should be measured. Almost all included effect measures in terms of changes in resource use, and five used clinical outcome measures. Only eight studies used health-related quality of life measures. This is remarkable, since the acceptance of quality of life measures as valid indicators of whether or not a therapy is beneficial has grown in recent decades, especially in pain-related disorders.

Besides the consequences, the relevant cost factors should be determined, measured and valued in economic evaluation studies. Despite the importance of absenteeism related to back pain, not all studies measured these indirect costs. It was also frequently unclear how costs were measured and valued. Furthermore, most studies used charges as a substitute for true costs, although some used estimates or average costs.

Two other important steps in performing economic evaluation studies we identified include the use of discounting and sensitivity components of the economic estimations. In most of the studies the time-span of the benefits and costs differed by several years, yet only nine of these studies used discounting to account for the difference. Only four studies included a sensitivity analysis to test the robustness of the conclusions to variations in the assumptions, although several authors mentioned limitations in the evaluation of both the costs and effects.

In short, this review has shown that there is a significant need to improve the application of economic evaluation in the area of back pain. By now, the conclusions of this review should be read keeping the shortcomings mentioned in the back of the mind. Decision about the most cost-effective therapy for back pain is therefore not possible at the moment.

One must keep in mind that the poor technical quality of the economic evaluation of the studies is not particular to the field of back pain. Previous studies of the literature on economic evaluation have noted that, despite the growth in the literature on economic evaluation during the 1980s, there had been little improvement in the methods used to perform economic analyses.^{72,76,77} One major reason is that economic evaluation analysis is complex and difficult to perform correctly without specific training in the analytic techniques.⁷⁷

If economic evaluation data is to be used for decision-making, the methods employed in the different studies must be valid and clearly documented. Comparisons between the estimates in the studies are also limited because of large differences in costing methodology. This means that methods of cost and outcome measurement have to be standardized before study results can be compared and used for decision-making and planning. The recommendations established by standardization should include the aspects taken into account in the checklist used for this review.

However some additional difficulties can arise when evaluating studies in the field of back pain. The methodological problems in the definition, classification and diagnosis of low back pain make it difficult to interpret epidemiological studies, which naturally leads to difficulties in measuring costs. Another complication is that many new treatments for back pain or low back pain have been developed, while the effectiveness or lack of effectiveness of the earlier treatments has not been established. It is also known that the success of any treatment for a chronic disease may vary over time.

Despite all these difficulties, the impression is that the quality of economic evaluation in the field of back pain can be improved, especially cost measurement. Using a checklist such as the one used for this review would be an important step towards increasing the validity and comparability of cost-effectiveness studies in the field, and would lead to better decisions in the prevention and treatment of back pain. Such a standardized framework, along with an improvement in the quality of the studies in general, will require that health economists and experts in the field of low back pain work together in any study from its outset.

Acknowledgements. The authors would like to thank Steven Linton, Maureen Rutten, André Ament and Reiner Leidl for reviewing earlier versions of the manuscript. The authors would also like to express their thanks to Jan van Emmerik and Ria Verheyen for their assistance in preparing the article and the Netherlands Health Research Promotion Program (SGO) for their financial support.

References

1. Andersson, G.B.J. The Epidemiology of Spinal Disorders. In: Frymoyer JW. *The Adult Spine: principles and Practice*. New York: Raven Press, 1991: 107-146.
2. Frymoyer, J.W., Cats-Baril, W.L. An overview of the incidences and costs of low back pain. *Orthop Clin North Am* 1991; 22: 263-271.
3. Girolamo de, G. Epidemiology and social costs of low back pain and fibromyalgia. *Clin J Pain* 1991; 7 (Suppl. 1): s1-s7.
4. Borenstein, D. Epidemiology, etiology, diagnostic evaluation, and treatment of low back pain. *Curr Opin Rheumatol* 1992; 4(2): 226-232.
5. Frymoyer, J.W. Back pain and sciatica. *The New England Journal of Medicine* 1988; 318(5): 291-299.
6. Kelsey, J.L., Golden, A.L., Mundt, D.J. Low back pain/Prolapsed Lumbar Intervertebral Disc. *Rheumatic Disease Clinics of North America* 1990; 16(3): 699-711.
7. Abenham, L., Suissa, S. Importance and economic burden of occupational back pain: a study of 2,500 cases representative of Quebec. *J Occup Med* 1987; 29(8): 670-674.
8. Akeson, W.H., Murphy, R.W. Low back pain. *Clinical orthop* 1977; 129(2).
9. Cats-Baril, W.L., Frymoyer, J.W. The economics of spinal disorders. In: *The Adult Spine. Principles and practice*. New York: Raven Press, 1991.
10. Grazier, K.L., Holbrook, T.L., Kelsey, J.L. *et al. The Frequency of Occurrence, Impact, and Cost of Musculoskeletal Conditions in the United States*. Chicago: American Academy of Orthopaedic Surgeons, 1984.
11. Jonsson, E. In: Statens Beredning för Utvärdering av medicinsk metodik, The Swedish Council on Technology Assessment in Health Care. *The problem of back pain, proceedings from a conference*, SBU, Sweden Stockholm, 1989.
12. Klaber Moffett, J., Richardson, G., Sheldon, T.A., Maynard, A. *Back Pain. Its management and cost to society*. The University of York: Centre for Health Economics, York Health Economics Consortium, NHS Centre for Reviews & Dissemination, 1995.
13. Rice, D.P., Hodgson, T.A., Kopstein, A.N. The economic costs of illness: A replication and update. *Health Care Financing Review* 1985; 7: 61-80.

14. Snook, S.H. The cost of back pain in industry. *J Occ Med: State Art Rev* 1988; 3(1): 1-5.
15. Tulder van, M.W., Koes, B.W., Bouter, L.M. A cost-of-illness study of back pain in The Netherlands. *Pain* 1995; 62: 233-240.
16. Snook, S.H., Webster, B.S. The cost of disability. *Clin Orthop* 1987; 221: 77-84.
17. Spengler, D.M., Bigos, S.J., Martin, N.A., Zeh, J., Fisher, L., Nachemson, A. Back injuries in industry: a retrospective study, I Overview and cost analysis. *Spine* 1986; 11(3): 241-245.
18. Webster, B.S., Snook, S.H. The cost of compensable low back pain. *J Occup Med* 1990; 32(1): 13-15.
19. Leavitt, S.S., Johnston, T.L., Robert, D.B. The process of recovery: patterns in industrial back injury, part 1 costs and other quantitative measures of effort. *Industrial medicine* 1972; 40(8): 7-14.
20. Luce, B., Elixhauser, A. Estimating costs in the economic evaluation of medical technologies. *Int. J of Technology Assessment in Health Care* 1990; 6: 57-75.
21. Health Economics. Economic evaluation bibliography. *Health Economics* (Suppl) 1992 December; 1.
22. Drummond, M.F., O'Brien, B., Stoddart, G.L., Torrance G.W. *Methods for the economic evaluation of health care programmes*. Second edition. Oxford, 1997.
23. Finkler, S.A. The distinction between costs and charges. *Annals of Internal Medicine* 1982; 96: 102-109.
24. NHS Centre for Reviews and Dissemination. *NHS Economic evaluation Database*, 1998. University of York, York
25. OHE-IFPMA Database limited. *The Health Economic Evaluation Database (HEED)*. 1998, London.
26. Evers, S.M.A.A., Wijk van, A.S. *Toolkit for reviewing literature*. Department of Health Economics, University of Limburg, Maastricht 1994; working paper: 94-3.
27. Aarås, A. The impact of ergonomic intervention on individual health and corporate prosperity in a telecommunications environment. *Ergonomics* 1994; 37(10): 1679-1696.
28. Bell, G.K., Kidd, D., North, R.B. Cost-effectiveness analysis of spinal cord stimulation in treatment of failed back surgery syndrome. *Journal of Pain and Symptom management* 1997; 13 (5): 286-295.
29. Brown, K.C., Sirles, A.T., Hilyer, J.C., Thomas, M.J. Cost-effectiveness of a back school intervention for municipal employees. *Spine* 1992; 17(10): 1224-1228.
30. Coleman, S., Hansen, S. Reducing work-related back injuries. *Nursing Management* 1994; 25(11): 58-61.
31. Devulder, J., De Laat, M., Van Bastelaere, M., Rolly, G. Spinal cord Stimulation: A valuable treatment for chronic failed back surgery patients. *Journal of Pain and Symptom management* 1997; 13 (5): 296-301.
32. Hochanadel, C.D., Conrad, D.E. Evolution of an on-site industrial physical therapy program. *JOM* 1993; 35(10): 1011-1016.
33. Linton, S.J., Bradley, L.A. An 18-month follow-up of a secondary prevention program for back pain: help and hindrance factors related to outcome maintenance. *Clin J Pain* 1992; 8(3): 227-236.
34. Linton, S.J., Hellsing, A.L., Andersson, D. A controlled study of the effects of an early intervention on acute musculoskeletal pain problems. *Pain* 1993; 54: 353-359.

35. Lissovoy de, G., Brown, R.E., Halpern, M., Hassenbusch, S.J., Ross, E. (1997) Cost-effectiveness of Long-Term Intrathecal Morphine Therapy for Pain Associated with failed Back Surgery Syndrome. *Clinical Therapeutics*; 19 (1); 96-85.
36. Malmivaara, A., Häkkinen, U., Aro, T., Heinrichs, M.J., Koskeniemi, L., Kuosma, E., Lappi, S., Paloheimo, R., Servo, C., Vaaranen, V., Hernberg, S. The treatment of acute low back pain - bed rest, exercises, or ordinary activity? *N Eng J Med* 1995; 332(6): 351-355.
37. Malter, A.D., Larson, E.B., Urban N., Deyo, R.D. Cost-effectiveness of lumbar Discectomy for the treatment of Herniated Intervertebral disc. *Spine* 1996; 21 (9): 1048-1055.
38. Manucher, J., Javid, M.J. Chemonucleolysis versus laminectomy: a cohort comparison of effectiveness and charges. *Spine* 1995; 20 (18): 2016-2022.
39. Meade, T.W., Dyer, S., Browne, W., Townsend, J., Frank, A.O. Low back pain of mechanical origin: randomised comparison of chiropractic and hospital outpatient treatment. *BMJ* 1990; 300(6737): 1431-1437.
40. Mitchell, L.V., Lawler, F.H., Bowen, D., Mote, W., Asundi, P., Purswell, J. Effectiveness and cost-effectiveness of employer-issued back belts in areas of high risk for back injury. *JOM* 1994; 36(1): 90-94.
41. Mitchell, R.I., Carmen, G.M. Results of a multicenter trial using an intensive active exercise program for the treatment of acute soft tissue and back injuries. *Spine* 1990; 15(6): 514-521.
42. Ramirez, L.F., Javid, M.J. Cost effectiveness of chemonucleolysis versus laminectomy in the treatment of herniated nucleus pulposus. *Spine* 1985; 10(4): 363-367.
43. Ray, C.D. Threaded Fusion Cages for Lumbar interbody fusions. An economic comparison with 360 fusions. *Spine* 1997; 22(6) pp 681-685
44. Ryan, W.E., Krishna, M.K., Swanson, C.E. A prospective study evaluating early rehabilitation in preventing back pain chronicity in mine workers. *Spine* 1995; 20(4): 489-491.
45. Ryden, L.A., Molgaard, C.A., Bobbitt, S.L. Benefits of a back care and light duty health promotion program in a hospital setting. *J Community Health* 1988; 13(4): 222-230.
46. Shekelle, P.G., Markovich, M., Louie, R. Comparing the costs between provider types of episodes of back pain care. *Spine* 1995; 20(2): 221-227.
47. Simmons, J.W., Avant, W.S. Jr, Demski, J., Parish, D. Determining successful pain clinic treatment through validation of cost effectiveness. *Spine* 1988; 13(3): 342-344.
48. Sinclair, S.J., Hogg-Johnson, S., Mondloch, M.V., Shields, S.A. The effectiveness of an early active intervention program for workers with soft-tissue injuries. The early claimant cohort study. *Spine* 1997; 22 (24): 2919-2911.
49. Skargren, E.I., Oberg, B.E., Carlsson, P.G., Gade, M. (1997). Cost-effectiveness of chiropractic and physiotherapy treatment for low back pain and neck pain. Six-month follow-up. *Spine* 1997; 22(18): 2167-2177.
50. Stevenson, R.C., McCabe, C.J., Findlay, A.M. An economic evaluation of a clinical trial to compare automated percutaneous lumbar discectomy with microdiscectomy in the treatment of contained lumbar disc herniation. *Spine* 1995; 20 (6): 739-742.
51. Stieg, R.L., Williams, R.C., Timmermans-Williams, G., Tafuro, F., Gallagher, L.A. Cost benefits of interdisciplinary chronic pain treatment. *Clin J Pain* 1986; 1:189-193.
52. Timm, K.E. A randomized control study of active and passive treatments for chronic low back pain following L5 laminectomy. *JOSPT* 1994; 20(6): 276-286.

53. Tuchin, P.J., Bonello, R. Preliminary findings of analysis of chiropractic utilization and cost in the workers compensation system of New South Wales, Australia. *Journal of Manipulative & Physiological Therapeutics* 1995; 18 (8): 503-511.
54. Tunturi, T., Niemela, P., Laurinkari, J., Patiala, H., Rokkanen, P. Cost-benefit analysis of posterior fusion of the lumbosacral spine. *Acta orthop. Scand.* 1979; 50: 427-432.
55. Versloot, J.M., Rozeman, A., Son van, A.M., Akkerveeken van, P.F. The cost-effectiveness of a back school program in industry. A longitudinal controlled field study. *Spine* 1992; 17(1): 22-27.
56. MacLean, D.S. Family practice and the health care system. Outcome and cost of family physicians' care. Pilot study of three diagnosis-related groups in elderly inpatients. *Journal American Board of Family Practice* 1993; 6: 588-93.
57. Liang, M., Komaroff, A.L. Roentgenograms in primary care patients with acute low back pain. A cost-effectiveness analysis. *Archives of Internal Medicine* 1982; 142: 1108-1112.
58. Keijsers, J.F.E.M. *The efficacy of back schools: empirical evidence and its impact on health care practice* (dissertation). Maastricht, 1991.
59. Mitchell, R.I., Carmen, G.M. The functional restoration approach to the treatment of chronic pain in patients with soft-tissue and back injuries. *Spine* 1994; 19(6): 633-642.
60. Turk, D.C. Efficacy of multidisciplinary pain centers in the treatment of chronic pain. In: *Pain Treatment Centers at a Crossroads; a practical and conceptual reappraisal, Progress in Pain Research and Management*. Seattle: IASP Press (7) 257-273, 1996.
61. Nyiendo, J. Disabling low back Oregon workers' compensation claims. Part III: Diagnostic and treatment procedures and associated costs. *J Manipulative Physiol Ther* 1991; 14(5): 287-297.
62. Wiesel, S.W., Boden, S.D., Feffer, H.L. A quality-based protocol for management of musculoskeletal injuries. *Clinical Orthop* 1994; 301: 164-176.
63. Hodgson, T., Meiners, M. Cost-of-Illness Methodology: A Guide to Current Practices and Procedures. *Health and Society* 1982; 60(3).
64. Koopmanschap, M.A., Ineveld van, B.M. Towards a new approach for estimating indirect costs of disease. *Social science and medicine* 1992; 34: 1005-1010.
65. Rice, D.P., Kelman, S., Miller, L.S., Dunmeyer, S. *The economic cost of alcohol and drug abuse and mental illness*. Rockville: National Institute of Mental Health, DHHS pub. no. (ADM) 90-1694, 1990.
66. Spilker, B. *Quality of Life Assessment in Clinical Trials*. New York: Raven Press, 1990.
67. Weber, H. Lumbar disc herniation: A controlled, prospective study with ten years of observation. *Spine* 1983; 8: 131-140.
68. Drummond, M., Brandt, A., Luce, B., Rovira, J. Standardizing methodologies for economic evaluation in health care; practice, problems and potential. *Int. J. of Technology Assessment in Health Care* 1993; 9(1): 26-35.
69. Eisenberg, J.M. Clinical Economics; A guide to the economic analysis of clinical practices. *JAMA* 1989; 262(20): 2879-2886.
70. Fried, B.J., Worthington, C., Deber, R.B. Economic evaluations in the Canadian mental health, System I: Theory behind economic evaluation. *Canadian Journal of Psychiatry* 1989a; 34: 633-636.
71. Fried, B.J., Worthington, C., Deber, R.B. Economic evaluations in the Canadian mental health, System II: From theory to practice in mental health care. *Canadian Journal of Psychiatry* 1989b; 24: 637-640.

72. Gerard, K. Cost-utility in practice: A policy maker's guide to the state of the art. *Health Policy* 1992; 21: 249-279.
73. Department of clinical epidemiology and biostatistics, McMaster University Health Sciences Centre. How to read clinical journals: VII. To understand an economic evaluation. *Can. Med Assoc J* 1984; 130: 1428-1434, 1542-1549.
74. Linton, S.J., Bradley, L.A., Jensen, I., Spangfort, E., Sundell, L. The secondary prevention of low back pain: a controlled study with follow-up. *Pain* 1989; 36: 197-207.
75. Drummond, M.F. *Principles of Economic Appraisals in Health Care*. New York: Oxford University Press, 1980.
76. Rutten-van Mólken, M.P.M.H., van Doorslaer, E.K.A., Rutten, F.F.H. Economic appraisal of asthma and COPD care: a literature review 1980-1991. *Social Science and Medicine* 1992; 35(2): 161-175.
77. Udvarhelyi, S., Colditz, G.A., Rai, A., Epstein, A. Cost-effectiveness and cost-benefit analysis in the medical literature. *Annals Internal Medicine* 1992; 116: 238-244.

4

COST-EFFECTIVENESS OF COGNITIVE-EDUCATIONAL TREATMENT OF FIBROMYALGIA

This chapter is published as:

Mariëlle E.J.B. Goossens¹, Maureen P.M.H. Rutten-van Mölken¹, Reiner M. Leidl¹, Stefan G.P.M. Bos¹, Johan W.S. Vlaeyen^{2,3}, Nicole J.G. Teeken-Gruben². Cognitive-educational treatment of fibromyalgia, a randomized clinical trial. Part II: Economic evaluation. *Journal of Rheumatology* 1996; 23 (6): 1246-1254.

¹ Department of Health Economics, Maastricht University

² Institute for Rehabilitation Research, Hoensbroek

³ Department of Differential and Experimental Psychology, Maastricht University

Summary

In a 3 year randomized clinical trial the cost-effectiveness of a 6 week educational/cognitive intervention (ECO) is compared with an educational/discussion intervention (EDI) and a waiting list condition (WLC). A total of 131 fibromyalgia patients were randomly allocated to the ECO, EDI or WLC intervention. The ECO and EDI groups were followed for 12 months, whereas the WLC group was followed for 6 weeks. Direct health care and non-health care costs, and the indirect costs associated with lost production due to illness, were calculated. The effects were measured in terms of utilities, using rating scale and standard gamble methods. The study estimated the treatment costs to be US \$980 per patient for both ECO and EDI. The total direct health care costs of the ECO treatment were US \$1623 higher than those for EDI. This difference was significant. Indirect costs for the two groups were not significantly different. At 6 weeks there was a significant difference in rating scale utilities between the three groups, caused by a significantly greater improvement in the EDI group as compared to the WLC group. However no significant differences in either rating scale or standard gamble utilities were found between the ECO and EDI groups immediately after treatment, or at the 6 or 12 month follow-ups. This study concluded that the addition of a cognitive component to the educational intervention led to significantly higher health care costs and no additional improvement in quality of life as compared to the educational intervention alone. This conclusion is robust through a range of plausible values used in a sensitivity analysis.

4.1 Introduction

Fibromyalgia is a chronic pain syndrome associated with symptoms of stiffness and fatigue. It has a significant impact on quality of life.¹ Patients frequently suffer from pain, depression, anxiety, decreased participation and pleasure in leisure activities, impaired function in daily life or at work and increased dependence on family and friends.²⁻⁴ Although there are no detailed cost of illness studies for fibromyalgia, it is known that this chronic pain syndrome is associated with extensive health care utilization due to the absence of a clear etiology or effective therapies. A study by Cathey *et al*² of the health service utilisation of 81 fibromyalgia patients showed that their utilization of outpatient medical services was higher than outpatient services utilization of both the control subjects and the national averages. However the health care utilization was similar to that of patients with other chronic pain disorders such as osteoarthritis and low back pain. This study also found very high hospitalisation rates before fibromyalgia was diagnosed. Several studies have shown that job limitations are associated with fibromyalgia.^{1,2,5,6} A study of disability claims paid by a large Canadian insurance company has shown that over 50% of fibromyalgia patients who receive such a benefit are disabled for more than 2 years.⁵ These long-term disability payments and compensation arising from litigation reflect the considerable loss of potentially productive years by fibromyalgia. However, these studies all report on work disability, which may be an underestimate of the total indirect costs, since in most studies more than 85% of the patients are women without a paid job.

Despite the considerable burden and costs associated with fibromyalgia and the increasing need to use resources efficiently, only one economic evaluation of treatments for fibromyalgia was found. This study, conducted by Lightfoot *et al*.⁷ investigated the treatment of fibromyalgia patients with lyme disease using intravenous antibiotics. The drug was directed at the lyme disease. Therapy costs, including the costs of treating side-effects, were found to be more than US \$80 000 for every case of lyme disease which was treated effectively. The cost calculations in this study were limited to the costs of health care utilisation, which is not sufficient to economically evaluate a chronic disease. In a chronic illness such as fibromyalgia, the costs borne by patients and their families and the costs of lost productivity are likely to be substantial and should not be excluded. Full economic evaluations compare alternative treatments with respect to all relevant costs and effects. Since there are not sufficient resources to support every new intervention, the results of such full economic evaluations can be used to inform decision-makers about the most efficient use of the scarce resources, so that the total health gains from the use of these resources can be maximized.

The main question addressed in this study is whether a combined educational/cognitive therapy is effective, and cost-effective, when compared to educational therapy alone. To assess the short term effect of both therapies, a control group of patients on the waiting list was also followed. As far as we know, this is the first full economic evaluation of a treatment modality for fibromyalgia. The societal perspective

has been adopted in this study because of the wide array of social and economic consequences of fibromyalgia, as described earlier.

4.2 Patients and methods

4.2.1 Study design

In a three-year randomized controlled clinical trial, 131 patients aged 18-65 who met the American College of Rheumatology's criteria of fibromyalgia were randomly assigned to an educational-cognitive group (ECO, 49 patients), an educational-discussion group (EDI, 39 patients) and a waiting list condition (WLC, 43 patients). The ECO and EDI treatments were structured so that groups of 6 patients participated in a 6-week program containing 12 half-day treatment sessions, given at the outpatient clinic of a rehabilitation centre.

Educational program. The educational program, which was applied to both ECO and EDI, consisted of 12 2-hour sessions, with each session being conducted by one member of the interdisciplinary rehabilitation staff. The program was intended to provide information about psychosocial factors that influence pain, ergonomic principles as applied to daily activities, and social security legislation. The information provided did not only focus on fibromyalgia specifically, but also on chronic pain in general. Each session ended with a physical exercise, such as swimming or bicycling, excluding systematic physical or fitness training.

Cognitive treatment. The cognitive treatment applied was a coping-skills treatment and was aimed at decreasing distorted pain attributions and at increasing self-efficacy expectations. This group program consisted of 12 sessions of 90 minutes and was conducted by a psychologist. The program was grouped into 3 phases: a reconceptualization phase, a skills acquisition phase and a generalization phase. The goal of the reconceptualization phase was to modify the pain experience in terms that imply self-control and resourcefulness. In the skills acquisition phase the patients learned to use imagery to influence their pain. In addition, applied relaxation was used and supported by EMG biofeedback. During the skills acquisition phase the patient is learned to relax and to use relaxation in situations with personally relevant stress. During the generalization phase the patient is gradually exposed to tension-eliciting stimuli, and encouraged to use the relaxation and imagery skills in the presence of these stimuli.

Group discussion. The group discussion component of the EDI program was intended as an attention-control for possible a-specific effects of the cognitive treatment. During this program patients were requested to read parts of a book about pain, written for pain patients, and to then share the information and their own thoughts with the other group members. In addition, participants listened to various audiotaped music fragments. Each session ended with a homework assignment, consisting of brief relevant reading assignments and listening to audiotaped musical fragments. The group discussion was conducted by the same psychologist who was in charge of the cognitive treatment, and consisted of the same number of sessions. Hence, neither the subjects nor the therapists

of the interdisciplinary rehabilitation staff who are in charge of the educational programma were aware of the difference between the two treatments.

A more detailed description regarding the content of the treatment programs has been given in Vlaeyen *et al.*⁸

All patients were seen at the outpatient clinic of the rehabilitation centre two weeks before the start of the treatment (PRE1), at the start of the treatment (PRE2), after completion of the 6-week (POST) treatment program and 6 (FU1) and 12 months (FU2) after the termination of the treatment. Patients in the WLC group were followed-up for 6 weeks after randomisation. This group was included only to measure the short-term effects of both the ECO and EDI treatments.

Clinical effects were expressed in terms of 5 primary measures (pain control, pain coping, knowledge, tension and relaxation), 4 secondary measures (catastrophizing, pain intensity, pain behavior, activities) and 3 affective measures (fear, depression and obsessive-compulsiveness) described in the preceding paper. This article focuses on the outcomes, in terms of costs and utilities (where utilities are values assigned to the patient's quality of life). The economic evaluation specifically addresses the question of whether the addition of a cognitive therapy to an educational therapy is cost-effective when compared to educational therapy alone.

4.2.2 Costs

To evaluate the economic consequences, the direct health care and non-health care costs were considered, as well as the indirect costs. The direct health care costs included the costs of the ECO and EDI treatment programs and the costs of all other pain-related health care utilisation, which includes GP contacts, outpatient specialist contacts, physiotherapy, alternative health care, hospitalisations, home help, prescribed medications and over-the-counter medications. Direct non-health care costs include costs of paid and unpaid help, transportation costs, out of pocket expenses for pain-related activities and purchases. Indirect costs refer to the value of the production lost due to illness-related absence from work or days lost from house-keeping. The costs in 1993 Dutch Guilders have been converted into US Dollars (\$) at the 1993 Purchasing Power Parities rate* of 2.134:1 (OECD Health Data File, 1995). Table 1 provides an overview of the unit prices of important cost components.

* Purchasing Power Parities (PPP) are rates of currency conversion that equalize the purchasing power of different currencies. This means that the PPP conversion rate eliminates the differences in price levels between countries.

Table 1. Unit prices used in calculations of direct health care and direct non-health care costs (in US \$)

Unit price per contact (US\$)		Unit price per contact (US\$)	
General practitioner	13.44	Hospitalisation (per day)	251.17
<i>Outpatient specialist</i>		<i>Alternative health care</i>	42.17
Rheumatologist	15.02	Homeopath	84.35
Neurologist	17.37	Acupuncturist	19.92
Orthopaedist	11.94	Natural therapist	34.68
Gynaecologist	12.21	Manual physician	40.91
Internal Medicine	12.44	Orthomanual physician	14.06
Psychiatrist	20.89	Reflexologist, magnetiser,	
Rehabilitation physician	11.50	Hypnotiser, paranormal healer etc.	14.06
Urologist	11.94	Home help (per hour)	7.03
Cardiologist	11.07	Unpaid help from family, friends (per hour)	
<i>Physiotherapy</i>		<i>Transportation</i>	
Regular	14.23	Taxi (per km)	
Manual therapy	27.46	Own car (per km)	0.62
Caesar/Mensendieck	15.64	Public transport (per km)	0.27
Ultrasound	16.23		0.09

Program costs. The calculation of total costs per patient of the ECO and EDI programs was based on the volumes in the program protocols, the time schedules of therapists, and detailed cost-accounting studies in the rehabilitation centre using the direct allocation method to allocate service department costs to production departments.

Direct health care costs. Data on pain-related health care utilisation was obtained from patients using a weekly cost diary which was specifically developed for this study. It was covering visits to the general practitioner, outpatient specialist, and physiotherapist, hospitalisations, alternative health care, over the counter (OTC) medications, and prescribed medications. The prescribed medication was later subdivided in anti-rheumatics, analgesics, sleep inducers and tranquilizers, anti-psychotics and anti-depressants. Patients were asked to complete this diary for the entire study period. It was left to the judgement of the patients whether particular items of health care utilisation were related to their fibromyalgia. The costs of GP contacts (US \$13.44 per consultation), outpatient specialist contacts (ranging from US \$11.07 to US \$20.89 per consultation), contacts with a physiotherapist (ranging from US \$14.23 to US \$27.46 per consultation), hospitalisations (US \$251.17 per day) and home help (US \$14.06 per hour) were based on charges. A population-weighted average of the charges to be paid by socially-insured and privately-insured patients was used. The costs of drugs were based on Dutch pharmacy prices. Costs of alternative health care (ranging from US \$19.92 to US \$84.35 per consult) were based on the guidelines of the professional associations for alternative medicine.

Direct non-health care costs. Data on direct non-health care costs related to fibromyalgia was also collected using the weekly cost diary. Patients were asked to report the following: type of transportation, distance travelled, hours of unpaid help from family or friends, house-keeper expenses, out-of-pocket expenses for activities such as swimming, physical exercise etc. and out-of-pocket expenses for equipment, aids, modifications to the home, clothing, etc. necessary to participate in the program

or for self-management. Costs of unpaid help by family or friends were valued using the shadow price for professional help (US \$7.03 per hour). For the other direct non-health care costs, the prices reported by the patient in the weekly cost diary were used.

Indirect costs. To calculate the indirect costs, patients were asked to report in the weekly cost diary how many days they were absent from their paid employment and how many days they were unable to perform their usual daily activities. The indirect cost calculation was based on the Human Capital Approach, which estimates the value of the potential production lost during the entire period of absenteeism using, for both sexes, the national average gross hourly wage of US \$10.78. The Human Capital Approach measures the full potential loss of production but is likely to overestimate the actual loss, because in reality missing workers may be replaced by others, or the absentee may make up the lost production once he or she returns to work. To approximate the actual production loss, it is more realistic to assume that losses occur only during the time needed to replace the sick worker, or to reorganise the production process. This is called the friction time. In the sensitivity analysis we estimate the value of the production loss assuming that no production is lost after this friction time. An estimation of the length of the friction period in the Netherlands, as of our study period, was obtained from Koopmanschap *et al.*⁹ who developed the Friction Cost Approach and calculated that approximately 3 months was, on average, necessary to fill vacancies.

4.2.3 Utilities

The effects are measured in terms of utilities. A utility is a single comprehensive outcome measure that reflects the value or preference that respondents assign to a particular health state. This value is expressed on a scale ranging from 1 (perfect health) to 0 (death) and takes into account both the positive treatment effects and the negative side effects. In this study, utilities were elicited from the patients participating in the trial by means of the Maastricht Utility Measurement Questionnaire. The instrument is a slightly adapted Dutch version of the McMaster Utility Measurement Questionnaire.^{10,11} It is administered as an interview. Utility measurement using this instrument consists of two parts: first, patients are asked to describe their own health state of the past two weeks. They are requested to rate their current functional level using a 5-point scale (1=best level, 5=worst level) for six domains: physical state and mobility, self-care, emotions, leisure activities, pain and other complaints and side-effects of treatment. Second, patients are asked to value both a hypothetical reference state and their own health state by means of rating scale and standard gamble techniques. The hypothetical reference state is a description of a state with medium impaired quality of life, described in terms of the above 6 domains. This state serves as a reference point when patients value their own health state. The rating scale is visualized as a thermometer, with perfect health (100) and death (0) on the extremes. By means of the rating scale, utilities are measured *directly* by asking the patients to place the health states on the thermometer. By means of the standard gamble method utilities are *derived* from the patients' responses to decision situations under risk. In the standard gamble patients are offered a choice between two alternatives: alternative A is

a gamble with chance P of gaining perfect health and chance $1-P$ of dying immediately. Alternative B is the certainty of living in the health state being evaluated (either the patient's own current health state or the reference state) for the remainder of one's life. Probability P is varied in steps of 10% until the patient is indifferent between the two alternatives. According to the axioms of expected utility theory, the value of P at the indifference point is the expected value of this gamble, and thus the utility assigned to the health state being evaluated.¹² To facilitate the standard gamble questions a chance board with a probability wheel on it was used as a visual aid.

4.2.4 Quality adjusted life years

The Quality Adjusted Life Year (QALY) is an effect-measure which includes both the effects in terms of quality of life and effects in terms of survival.¹³ To calculate QALYs, utilities are used as correction factors to adjust years of life to allow for differences in the quality of life. For example, if the quality of life of a patient with a remaining life expectancy of 20 years improves by 0.5 of a utility because of a treatment, then this treatment results in 10 (20×0.5) QALYs gained. QALYs can be related to costs, resulting in a cost per QALY gained ratio. This ratio makes it possible to compare the cost-effectiveness of the current intervention with other programs of all kinds. For this study the QALYs gained by the ECO and EDI interventions will be calculated for illustrative reasons only.

4.2.5 Statistical analysis

To analyze the differences in direct, indirect and total costs between the ECO and the EDI groups, costs per patient-year were calculated.¹⁴ This means that the observed costs of the patients with one or more missing weekly cost diaries were extrapolated to a 1-year period. Since the distribution of costs was skewed to the right, the non-parametric Mann-Whitney test was used to assess the statistical significance of cost differences between the ECO and the EDI groups. Analyses of variance on log-transformed costs per patient-year confirmed the results of the Mann-Whitney test; the p -values of both tests were virtually the same. The differences in the changes in utilities for the ECO, EDI and WLC groups were analyzed using covariance analysis (ANCOVA), with the treatment group used as a factor and the baseline scores and a measure of social desirability as covariates.

4.3 Results

4.3.1 Patient characteristics

One hundred and thirty-one fibromyalgia patients were randomly assigned to one of the three different groups. There were no statistical differences between the treatment groups in clinical characteristics and socio-demographic characteristics.¹⁵ Most of the patients were women (88%) with a mean age of 44 and a pain duration of 10 years.¹⁵ At baseline the patients were asked if they had had contact with an outpatient specialist, a physiotherapist or had had alternative health care contacts in the last 12

months. There was no difference in the health care utilisation among the groups in the year before the trial. Of the whole sample, 98% had at least one outpatient specialist contact, 72% had used physiotherapy at least once, and 30% had at least one alternative health care contact in that year. Utilities did not differ between groups at baseline. The values (divided by 100) were ECO 0.46 (SD 0.12), EDI 0.45 (SD 0.15), WLC 0.47 (SD 0.13) for rating scale utilities and ECO 0.83 (SD 0.15), EDI 0.80 (SD 0.18), WLC 0.78 (SD 0.16) for standard gamble utilities. On the six domains of health, the average scores for all patients were 3.1 (SD 0.67) on physical state and mobility, 1.9 (SD 0.71) on self care, 3.2 (SD 0.73) on emotions, 3.3 (SD 0.64) on leisure activities, 3.9 (SD 0.60) on pain and other complaints and 1.6 (SD 0.75) on side effects from treatment. The classification on the domains of health did not differ between the groups except for the domain 'physical state and mobility'. On this domain there was a difference between the ECO and the EDI group, with the latter reporting a more favourable health state ($p=0.035$). Twenty-six patients did not complete the treatment or the follow-ups at one year (for ECO and EDI patients) or 6 weeks (for WLC patients). As was shown in the preceding paper⁸ there were no differences in demographic and clinical baseline characteristics between the dropouts and those who completed the study. The dropouts were included in the analyses until the moment of dropout.

4.3.2 Costs

Program costs. Using the standard therapy protocols, the costs of the ECO and EDI programs were estimated to be US \$980 per patient (table 2). Although the ECO program is more demanding, since the program component given by the psychologist requires more active participation from the patient, there is no difference in costs between the two programs because the number of therapy hours given by the various therapists are equal.

Table 2. Program costs of the ECO and the EDI treatments per patient (in US \$)

Program components	No. of 30 min therapy sessions	Costs of 30 min therapy	No. of patients in group	Total costs per patient
Rehabilitation specialist*	8	57.51	6	76.68
Rehabilitation specialist	2	57.51	1	115.02
Welfare work	4	43.05	6	28.70
Psychology	36	50.81	6	304.86
Physiotherapy	12	16.42	6	32.85
Sporting and games	14	19.50	6	45.50
Treatment program manager	6	16.42	6	16.42
Ergotherapy	7	25.07	6	29.25
Rehabilitation day treatment (in days)	12	27.56	1	330.76
<i>Total program costs</i>				<i>980.04</i>

*The rehabilitation specialist is mentioned twice in this table. The first row refers to group therapy, and the second row to individual therapy.

Other direct health care costs. As noted before, the other direct costs have been measured by means of the patients' weekly cost diaries. During the entire follow-up period, the patients in the ECO and EDI group completed and returned 84% and 80%,

respectively, of the weekly cost diaries. Table 3 shows the annual volumes per patient for the various categories of health care utilization. Although not all categories show significant differences between the two groups, there is a clear pattern in the direction of the differences. All categories reflect higher health care utilization in the ECO group as compared to the EDI group. It can also be seen that many of the patients in both groups received physiotherapy and called on unpaid help. Of the patients in the ECO and EDI groups, 49% and 36%, respectively, received physiotherapy at least once, and 29% and 13%, respectively, received physiotherapy once or more than once per week during the entire year. Sixty percent and 52% of the patients in the ECO and EDI groups, respectively, received unpaid help from family or friends.

Table 3. Mean (SD) health care utilization and mean (SD) use of other health care services per patient-year by cost category and treatment group

	ECO N= 35	EDI N= 31	P-value*
Number of:			
General practitioner contacts	6.7 (7.2)	5.6 (7.2)	0.131
Outpatient specialist care contacts	6.9 (8.1)	3.3 (4.0)	0.121
Physiotherapy contacts	25.1 (30.4)	11.1 (20.3)	0.037
Alternative health care contacts	4.6 (12.6)	1.3 (4.4)	0.025
Hospitalizations	1.8 (8.87)	0.5 (2.8)	0.376
No. of different medicines prescribed	6.0 (5.3)	5.7 (9.6)	0.162
No. of different OTC** medications bought	9.2 (17.2)	1.3 (2.9)	0.179
Hours of professional home help	20.6 (70.1)	17.9 (68.8)	0.579
Hours of paid housekeeping help	76.5 (105.5)	34.8 (56.8)	0.137
Hours of unpaid help from family or friends	165.8 (267.8)	84.6 (133.0)	0.324
Health activities	55.8 (37.4)	48.1 (34.8)	0.275
Expenses for equipment, clothing etc	0.9 (1.4)	0.3 (0.8)	0.012

* Mann-Whitney test

** OTC: over the counter medications.

The mean total direct costs per patient-year, including program costs, were estimated to be US \$4260 (SD 6510) for the ECO group and US \$2637 (SD 4649) in the EDI group. Figure 1 shows that both the direct health care costs ($p=0.003$), the direct non-health care costs ($p=0.028$) and the total direct costs ($p=0.001$) per patient-year were significantly lower in the EDI group than in the ECO group.

Table 4 shows the various components of direct costs as measured by the patient-year approach. Although not all differences in the subcategories of direct costs reach conventional levels of statistical significance, direct costs were higher in the ECO group than in the EDI group in all categories. The differences in the costs for alternative health care, travel expenses and expenses for equipment, clothing etc. reach statistical significance. In calculating the cumulative distribution of direct costs per patient over time the costs in the ECO group turned out to grow faster than in the EDI group over the entire period.

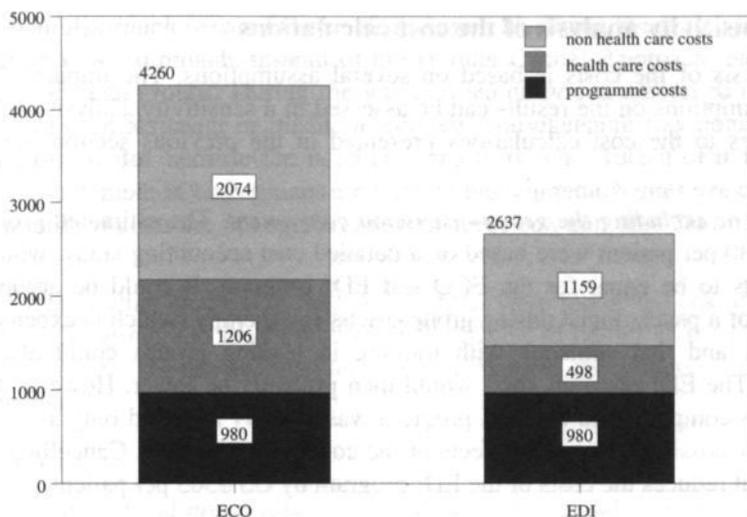


Figure 1. Estimated total costs per patient year per estimated group (US \$)

Table 4. Mean (SD) direct costs per patient-year by cost category and treatment group (US \$).

Cost category	ECO n=35	EDI n=31	P-value ¹
<i>Health care costs</i>			
General practitioner	90 (206)	75 (311)	0.131
Outpatient specialist care	114 (351)	55 (151)	0.133
physiotherapy	373 (1135)	154 (609)	0.083
alternative health care	145 (905)	19 (126)	0.023
hospitalizations	330 (3327)	127 (1502)	0.376
prescribed medications	119 (296)	63 (170)	0.071
OTC medications	35 (160)	5 (27)	0.218
<i>Non-health care costs</i>			
Home help	289 (2104)	252 (2036)	0.579
Paid housekeeping help	353 (1165)	191 (700)	0.279
Unpaid help from family or friends	1165 (4016)	595 (1995)	0.324
Expenses for health activities	65 (203)	51 (190)	0.491
Travel expenses	92 (296)	52 (78)	0.047
Expenses for equipment, clothing etc.	110 (639)	55 (111)	0.017

¹Mann-Whitney test

Indirect costs. Patients in the ECO group reported a mean of 42 (SD 82) days per year of pain-related absence from work and a mean of 50 (SD 56) days in the same year, on which they were not able to perform their usual daily activities. The EDI group reported 34 (SD 71) and 44 (SD 55) days of absenteeism and inactivity, respectively. However the differences between the two groups were not statistically significant on either measure. Consequently, although the estimated indirect costs based on the Human Capital Approach were also somewhat higher for the ECO group (US \$6379, SD 15569) than the EDI group (US \$5817, SD 14303), they were not significantly different ($p=0.872$).

4.3.3 Sensitivity analysis of the cost calculations

The analysis of the costs is based on several assumptions. The impact of changing these assumptions on the results can be assessed in a sensitivity analysis. A number of alternatives to the cost calculations presented in the previous section are discussed below.

Including or excluding the group-discussion component. The estimated program costs of US \$980 per patient were based on a detailed cost accounting study, which showed these costs to be equal for the ECO and EDI program. It could be argued that the presence of a psychologist during group-discussion therapy (which is expensive) is not necessary, and that someone with training in leading groups could also lead the sessions. The EDI program costs would then probably be lower. However the group-discussion component in the EDI program was initially intended only as an attention-control for possible a-specific effects of the cognitive treatment. Cancelling this entire component reduces the costs of the EDI program by US \$305 per patient.

Cost-accounting versus charges. Another assumption was that the program costs were best estimated by an institutional cost-accounting study. However from the rehabilitation centre's perspective, costs are better reflected by charges. Based on the charge for outpatient rehabilitation care, which is US \$106 per day, the program costs would be $12 \times \text{US } \$106 = \text{US } \1272 per patient. This is US \$292 more per patient than was calculated initially. Because there is only one charge for outpatient rehabilitation care, irrespective of its content, there is again no difference between the costs of the ECO and EDI programs, whether or not the cognitive component is included.

Excluding high cost events. The difference in direct health care costs between the ECO and EDI program was heavily influenced by differences in the costs associated with physiotherapy, unpaid help and hospitalisations. In the ECO group, 29% of the patients had physiotherapy on average once or more than once per week for the entire year, compared to 13% in the EDI group. Excluding the costs of physiotherapy decreases the difference in direct health care costs from US \$707 to US \$561, still in favour of EDI. Excluding the costs of unpaid help for the patients with more than 10 hours of unpaid help a week (11% in ECO and 3% in EDI) reduces the difference in direct non-health care costs from US \$916 to US \$522, again in favour of the EDI program. The differences in hospitalisation costs are due to there being 3 hospitalized patients in the ECO group in comparison with 1 hospitalized patient in the EDI group. Excluding the hospitalisation costs reduces the difference in direct health care costs from US \$707 to US \$504, still in favour of EDI. Even if all three high cost components are excluded, the difference in total direct health care costs between the two groups is still significant.

Mean costs versus median costs. The calculations of total direct costs associated with both programs were based on the assumption that the costs are best estimated using mean costs. Estimating the incremental costs using median costs affects the cost differences between the two programs. The direct cost per patient of the ECO treatment is then US \$421 higher than for the EDI program.

Human Capital Approach versus Friction Cost Approach. Using the main assumptions of the Friction Cost Approach, instead of the Human Capital Approach, reduced the indirect costs in both groups. During the whole research period eleven patients were absent from paid employment or unable to perform housekeeping functions for more than three months. But because the number of patients who were absent for longer periods was comparable in both groups the friction costs approach does not change the finding that indirect costs are not significantly different between the EDI and the ECO group.

Overall, no matter what changes in cost assumptions were made, the cost differences between the ECO group and the EDI group were always in favour of the EDI group. This means that the sensitivity analyses strongly support the results found in the baseline analysis.

4.3.4 Utilities

The results of the clinical effect measures are described in detail in Vlaeyen *et al.*⁸ In the first part of the Maastricht Utility Measurement Questionnaire the patients had to classify themselves on six domains of health. In general there was no change over 6 weeks of time and no difference between the groups on these domains, except for significant differences during the treatment in the domain 'leisure activities' in favour of the ECO condition as compared to WLC ($p=0.039$), and in the domain of 'self care' in favour of EDI, as compared to WLC. As for the utilities after the 6 weeks of treatment, there was a significant difference in rating scale utilities between the three groups ($p=0.037$) (table 5). This difference was due to a statistically significant greater improvement in the EDI group as compared to the WLC group ($p=0.012$).

Table 5: Differences in mean change scores after 6 weeks between ECO, EDI and WLC

	ECO (n=32)	EDI (n=24)	WLC (n=33)	P [*]
<i>Rating scale[#]</i>				
M	+ 0.070	+ 0.108	+ 0.018	0.037
SD ^{**}	0.13	0.13	0.13	
<i>Standard gamble</i>				
M	+ 0.003	+ 0.005	- 0.028	0.473
SD	0.12	0.12	0.12	

* P-value from ANCOVA; utilities are adjusted for baseline and social desirability.

** Mean change score

Rating scale utilities were divided by 100.

** Standard deviations equal the $\sqrt{\text{mean square residual}}$. and were divided by 100.

Figure 2 shows a greater improvement in rating scale utilities in the EDI group at the POST measurement (EDI 0.108 vs. ECO 0.070, $p=0.304$) and FU1 (EDI 0.107 vs. ECO 0.057, $p=0.338$) though this trend is not significant. At FU2 the improvement is greater in the ECO group (ECO 0.093 vs. EDI 0.067, $p=0.590$). Again this difference is not significant. The change in standard gamble utilities was not significantly different at either POST, FU1 or FU2.

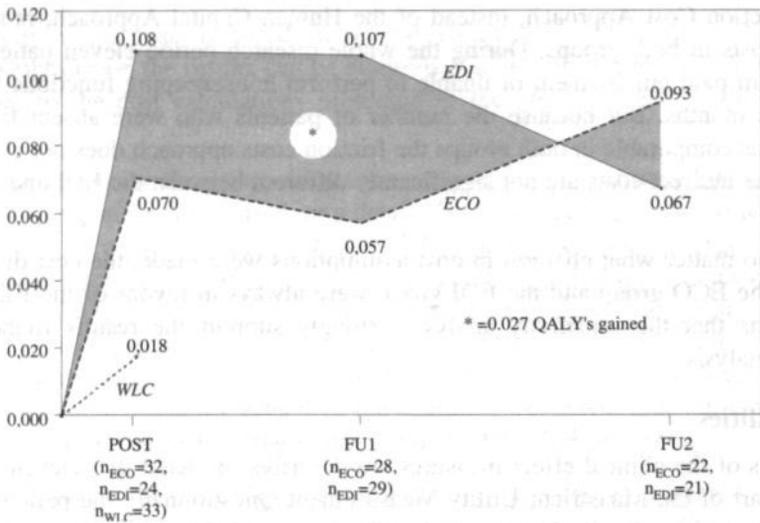


Figure 2 Change in rating scale utilities, adjusted for baseline differences and social desirability. Rating scale utilities were divided by 100.

4.3.5 Quality Adjusted Life Years

The changes in rating scale utilities and standard gamble utilities between baseline and follow-up were used to calculate the QALYs gained by the alternative programs. Based on rating scale results, figure 2 shows that the EDI treatment produced a gain of 0.027 QALYs per patient per year as compared to the ECO treatment. Standard gamble utility measurements showed that the EDI treatment yielded a gain of 0.022 QALYs as compared to the ECO treatment. Because these figures combine the results from different numbers of patients at different moments, the statistical significance of this gain could not be tested. However, the differences when the patients who did not complete the entire year to FU2 were excluded were not significant ($p=0.245$ and $p=0.220$ for changes in rating scale and standard gamble utilities, respectively).

4.4 Conclusion and discussion

This economic evaluation study showed that the intervention combining group education and group discussion (EDI) is to be preferred over the intervention which combined group education and group cognitive therapy (ECO), because the program costs of both interventions were equivalent but both the direct health care and non-health care costs were significantly lower for patients participating in the EDI program. All subcategories of direct costs differed in the same direction as the total costs, though not all of these differences reached statistical significance. This is also true for the

indirect costs. The cost differentials found are robust through a range of plausible values used in sensitivity analysis.

Despite the lower costs of the EDI program, no significant differences in utilities were found, except for the significantly greater improvement in rating scale utilities in the EDI group as compared to the WLC group at the post-treatment measurement. At best, the rating scale utilities suggest a short-term trend towards greater improvement in quality of life in the EDI group as compared to the ECO group; a trend which seems to disappear in the long run. Overall, using the rating scale method to measure utilities, this results in a non-significant estimate of 0.027 QALYs gained due to the EDI program. Such a small increase in QALYs indicates that the EDI program has little advantage over the ECO program, with regard to preferences for the own quality of life.

The observation that direct medical and non-medical costs were lower for the EDI program than for the ECO program is not in accordance with pre-trial expectations. The overall goal of the cognitive component of the ECO program was to increase patients' pain control using techniques such as diverting attention, reinterpreting pain sensations, coping self-statements, etc. As a result, health care resource use was expected to be lower rather than higher in the ECO group compared to the EDI group. One possible explanation may be that the cognitive group therapy increases the patients' awareness of the illness and the problems and weaknesses associated with it. This, in turn, may have increased the demand for both professional and para-professional help. A broader and more intense cognitive program that also addresses individual patients' needs might have prevented the increase in health care utilisation. Another explanation would be that the group-discussion component of the EDI program, which was intended only as an attention placebo, may have given the patients support and understanding. This resulted in less demand for help in the EDI group, from professional care givers, family members or friends, compared to the ECO group. In that case, the group discussions can no longer be seen as a placebo but rather as an active treatment.

A comment worth mentioning in this context is the following. Because we did not measure health care resource use in the year preceding the trial it can be questioned easily whether there was a difference in health care resource use between the groups already at baseline. However we have several arguments to believe that this was not the fact. First of all there was no difference in clinical characteristics between the groups. In addition, there was no difference between the groups proportions in outpatient specialist care, physiotherapy and alternative health care care contacts in the last 12 months. Although we had no information about the frequency and the content of the contacts, this gives us no reason to believe that health care utilization was different between the groups before the trial.

All previous fibromyalgia studies involving costs have been restricted to calculations of changes in health care resource use. In our study, efforts were made to take into account all relevant categories of costs, including the non-medical costs borne by patients and their families. Because of the patients' decisive role in reporting, the identification and measurement of these non-medical costs is somewhat arbitrary.

However in the sensitivity analysis we have shown that different assumptions regarding these costs or even excluding these costs would not have altered the overall conclusion.

Adding utility measures to the battery of pain-specific and domain-specific outcome measures described in the preceding paper allowed us to assess whether the changes in these outcome measures have any impact on the patient's valuation of his or her overall health state. Utilities are meant to express the net effect of the treatment programs, because both the positive and negative effects of the programs are integrated in the overall value a patient assigns to his health state. However, the large difference between the rating scale and standard gamble methods reflect the methodological problems underlying utility measurement. Numerous previous publications, some of them using the same instrument as we did, have reported that rating scale values are lower than standard gamble values.¹⁵⁻²⁰ In this study we found differences between 25% and 35%, depending on the moment of measurement. Several phenomena might explain this difference, the most important one being a patient's risk attitude which plays a role in the standard gamble technique but not in the rating scale technique. Risk-averse behavior is probably reinforced in this study by the fact that our fibromyalgia patients have tried numerous, and usually not very effective, treatments before. This might have caused a reluctance to try a new (although hypothetical) treatment that gives them chance P to gain perfect health. Furthermore, in our study, as in other studies, P was varied in steps of 10%.^{10,17,19} For fibromyalgia patients these steps are probably too large. In the first step patients have either to claim that they are in full health or accept a therapy with a 10% risk of dying; more than 40% of the patients were not prepared to take this risk on at least one measurement. Thus these large steps produce an upward bias in the standard gamble utilities. Smaller changes might have produced a greater variance in utilities. Due to this large ceiling effect, the standard gamble utilities were already so high at baseline that there was little likelihood of recording an improvement. Therefore, it appears justified to conclude that the standard gamble measurement protocol as used here is not suitable for eliciting utilities from certain types of chronic patients, such as those suffering from fibromyalgia.

For decision-making purposes, with regard to the wide adoption of the programs, both the ECO and EDI program should have been compared to the usual care alternative. However, considering the enormous variety of treatments for fibromyalgia, it was not possible to define an 'average' form of medical management within this study. Evaluating the available treatments would require extra studies. 'Usual care' could also not be derived from the health care utilization pattern of the patients on the waiting list, because patients stayed on this list for only a relatively short time and are likely to receive minimal or no care in anticipation of future treatment. Furthermore, the treatment given to patients in the rehabilitation centre does not represent 'usual care' for fibromyalgia in the Netherlands. As a consequence it was not possible to calculate the incremental cost-effectiveness ratios of the ECO and EDI programs as compared to usual care.

In summary, we found that the addition of a group-discussion component to an educational program was more cost-effective than the addition of a cognitive component. The first led to considerable savings compared to the latter, in both direct medical and non-medical costs, whereas the clinical outcome measures showed no clear differences between the two groups. From a cost-utility perspective, the EDI program is preferable to the ECO program. Because neither intervention produced significant long-term improvements in quality of life, the search for an effective therapy of fibromyalgia must go on.

Acknowledgments. We would like to acknowledge Peter Heuts and the Staff of the department of Pain Management of the Lucas Foundation for Rehabilitation Hoensbroek, for the management of the patients included in the study, and Henk Goei The and his colleagues for their continuous support in selecting the patients in this study. We are also grateful to Jan van Emmerik and Hubert Schouten for their statistical advice, Silvia Evers for her valuable comments on an earlier draft of this paper, and Ferd Sturmans and Eddy van Doorslaer for their advice at various stages of the study.

References

1. Wolfe, F. Fibromyalgia: The clinical syndrome. *Rheumatic Disease Clinics of North America* 1990; (16)3: 681-97.
2. Cathey, B.S.N., Wolfe, F., Kleinheksel, S.M., Hawley, D.J. Socioeconomic impact of fibrositis. A study of 81 patients with primary fibrositis. *American Journal of Medicine* 1986; 81: S3A.
3. Hawley, D.J., Wolfe, F., Cathey, M.A. Pain, functional disability and psychological status: A 12 month study of severity in fibromyalgia. *Journal of Rheumatology* 1988;15: 1551-1556.
4. Wolfe, F., Smythe, H.A., Yunus, M.B., *et al.* The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. *Arthritis and Rheumatism* 1990; 33: 160-172.
5. McCain, G.A., Cameron, R., Kennedy, C. The problem of longterm disability payments and litigation in primary fibromyalgia: The Canadian perspective. *Journal of Rheumatology* 1989; 16; supplement 19:174-176.
6. Girolamo, de G. Epidemiology and social costs of low back pain and fibromyalgia. *The Clinical Journal of Pain* 1991; (suppl.1)7: S1-S7.
7. Lightfoot, R.W., Luft, B.J., Rahn, D.W., *et al.* Empiric parenteral antibiotic treatment of patients with fibromyalgia and fatigue and a positive serologic result for Lyme disease. A cost-effectiveness analysis. *American College of Physicians* 1993; 119(6): 503-509.
8. Vlaeyen, J.W.S., Teeken-Gruben, N.J.C., Goossens, M.E.J.B., *et al.* Cognitive-educational treatment of fibromyalgia, a randomized clinical trial. Part 1. Clinical effects. *Journal of Rheumatology* 1996; 23: 1237-1245.
9. Koopmanschap, M.A., Rutten, F.F.H. Indirect costs in economic studies: confronting the confusion. *PharmacoEconomics* 1993; 4: 446-454.
10. Bennett, K., Torrance, G.R., Tugwell, P. Methodologic challenges in the development of utility measures of Health-related Quality of Life in Rheumatoid Arthritis. *Controlled Clinical Trials* 1991; (suppl)12: 118-28.

11. Bakker, C.H., Rutten-van Mólken, M.P.M.H., Van Doorslaer, E.K.A., Bennet, K., Van der Linden S.J. Health related utility assessment by rating scale and standard gamble in patients with ankylosing spondylitis or fibromyalgia. *Patient Education and Counseling* 1993; 20: 145-152.
12. Torrance, G.W., Feeny, D. Utilities and Quality-Adjusted Life Years. *International Journal of Technology Assessment* 1989; 5: 559-575.
13. Drummond, M.F., Stoddard, G.C., Torrance, G.W. *Methods for the evaluation of health care programmes*. Oxford: Oxford University Press, 1987.
14. Rutten-van Mólken, M.P.M.H., Van Doorslaer, E.K.A., Van Vliet, R.C.J.A. Statistical analysis of cost outcomes in a randomized controlled clinical trial. *Health Economics* 1994; 3: 333-345.
15. Read, J.L., Quinn, R.J., Berwick, D.M., et al. Preferences for health outcomes: comparisons of assessment methods. *Medical Decision Making* 1984; 4: 315.
16. Bleichrodt, H. Testing the validity of expected utility theory in health state valuation: Some experimental results. *Institute for Medical Technology Assessment paper no. 93.23*, Rotterdam, Erasmus University of Rotterdam, 1993.
17. Rutten-van Mólken, M.P.M.P., Bakker, C.H., Van Doorslaer, E.K.A., et al. Methodological issues of patient utility measurement: Experience from two clinical trials. *Medical Care* 1995; 33 (9): 922-937.
18. Bakker, C.H., Rutten-van Mólken, M.P.M.H., Van Doorslaer, E.K.A. Feasibility of utility assessment by rating scale and standard gamble in patients with ankylosing spondylitis or fibromyalgia. *Journal of Rheumatology* 1994; 21: 269-274.
19. Bakker, C.H., Rutten-van Mólken, M.P.M.H., Hidding, A. Patient utilities in ankylosing spondylitis and the association with other outcome measures. *Journal of Rheumatology* 1994; 21: 1298-1304.
20. Mulley, A.G. Assessing patients' utilities. Can the ends justify the means? *Medical Care* 1989; 27: S269.

5

HEALTH ECONOMIC ASSESSMENT OF BEHAVIORAL REHABILITATION IN CHRONIC LOW BACK PAIN

This chapter is published as:

Mariëlle E.J.B. Goossens^{1,2}, Maureen P.M.H. Rutten-van Mölken³, Ank M.J. Kole-Snijders^{1,4}, Johan W.S. Vlaeyen^{1,5}, Gerard van Breukelen⁶, Reiner Leidl⁷.
Health Economics 1998; 7: 39-51.

¹Institute for Rehabilitation Research, Hoensbroek

²Department of Health Organisation, Policy and Economics, Maastricht University

³Institute for Medical Technology Assessment, Erasmus University Rotterdam

⁴Lucas Foundation of Rehabilitation, Hoensbroek

⁵Department of Differential and Experimental Psychology, Maastricht University

⁶Department of Methodology and Statistics, Maastricht University

⁷Department of Health Economics, University of Ulm, Germany

Summary

The aim of this cost-effectiveness study was to compare a combined operant program plus cognitive/relaxation program with an operant program plus attention-control and to compare both programs with a waiting-list control group and with operant rehabilitation as usual provided by the same rehabilitation centre. One hundred forty eight patients with chronic low back pain were randomly assigned to the different conditions. The economic endpoints were the costs of the program and other health care utilization, costs for the patient, and indirect costs associated with production losses due to low back pain. The effects were measured in terms of global assessment of change and utilities, using rating scale and standard gamble methods. The 3-year study determined that adding a cognitive component to an operant treatment did not lead to significant differences in costs and improvement in quality of life when compared to the operant treatment alone. Compared to the common individual rehabilitation therapy it can be concluded that the same effects can be reached at the same or lower costs with a shorter, more intense standardized group program. The operant treatment alone is more effective than providing no treatment in the waiting group.

5.1 Introduction

Low back pain has been recognized as a major cause of morbidity, disability and economic loss. Low back pain occurs in most adults at some point in their lives, but in about 95% of the patients it disappears within six weeks. The annual incidence is about 5% and the prevalence lies between 15% and 40% in most Western countries.^{1,2} Nevertheless many studies in different countries have shown that back pain is not only a major health problem but also a tremendous socioeconomic problem. In the Netherlands the total expenses on back pain equaled 1.7% of the GNP³ in 1991. Moreover, musculoskeletal disorders (50% of which back pain) and mental disorders rank first in causes of absenteeism and disablement. Several studies found that the costs of back pain are not normally distributed, but heavily skewed to the right due to a minority of back pain patients (about 10% to 25%) who account for more than 75% of the costs.⁴ These high costs are largely due to production losses because of disability and to a smaller extent due to health care utilisation.

The still unknown pathomechanism of low back pain and the difficulty of making a clear diagnosis are the main reasons for the expansion of treatments in recent decades. Many studies in the field of back pain have led to a better understanding of what influences low back pain. Presently, chronic low back pain is seen as a problem concerning the central nervous system, rather than as a peripheral problem. Behavioral factors such as the impact of cognitions, pain behaviors and tension levels have been demonstrated to be important predictors of back pain disability.⁵⁻⁷ This knowledge has extended the large battery of treatments with behavioral rehabilitation programs. Behavioral rehabilitation is based on the bio-psycho-social model and integrates behavioral science and rehabilitation medicine. It applies cognitive-behavioral treatment in the setting of the rehabilitation centre. Patients are given the opportunity to learn skills through graded training adapted to their physical status.⁷ In general, behavioral rehabilitation programs for chronic low back pain patients consist of at least an operant, a cognitive and a respondent treatment component. The goal of the operant treatments is to increase health behaviors and activity levels step by step, and to decrease pain behaviors and excess disability using operant conditioning principles. Cognitive treatments are aimed at decreasing distorted ways of thinking of pain and increasing feelings of self-control by using cognitive coping skills such as attention diversion and imagery. Respondent treatments are aimed at reducing muscle tension levels by using respondent conditioning principles applied in relaxation and EMG biofeedback training.⁸

Studies evaluating the effectiveness of behavioral rehabilitation of chronic low back pain show that this approach is successful in helping patients reduce their levels of disability. Patients are more active, show less pain behaviors, report better pain coping skills and improved affect, although levels of experienced pain intensity often remain unchanged⁸⁻¹¹. One of the remaining questions is which treatment components are essential and which can be omitted while retaining effectiveness. So far, such component analyses are relatively scarce. This randomized controlled trial was aimed

at examining the surplus value of a cognitive treatment and relaxation when added to an operant treatment. The clinical effects of this study are described by Kole-Snijders *et al.*¹² The full treatment was compared to an operant program with group discussion, a waiting list, and treatment as usual. Compared to the waiting list control, both group treatments resulted in higher activity levels, less pain behaviors, less negative affect, and higher pain control and pain coping. The only difference between the full program and the operant+attention control was better pain coping and pain control for the full program immediately after treatment. Although there were no statistical differences between the experimental treatments and the usual operant rehabilitation, there were significantly less positive responders and more dropouts in the usual operant rehabilitation condition.

The current paper concerns the first study addressing the economical aspects of these type of interventions. Since the prevalence and associated burden and costs of low back pain are high and expected to increase in the near future, the cost-effectiveness of various low back pain interventions are badly needed. It is therefore surprising that economic aspects of low back pain treatments have generally received little attention.¹³ Insight in both the costs and effects of various health care interventions could be used to support decision-making about the allocation of scarce resources and to obtain maximum gain in health from these treatments.

5.2 Patients and methods

5.2.1 Subjects

Criteria for acceptance into the study were: age between 18 and 65 years, low back pain for more than six months, observable pain behavior, discrepancy between objective clinical findings and the pain complaints, and a partner who is willing to participate in a parallel partner program. Exclusion criteria were illiteracy, pregnancy, specific back disorders like ankylosing spondylitis and herniated disc, involvement in a legal procedure because of disability and severe psychopathology preventing the patient from participating in a group therapy. The patients were referred by family practitioners, one university hospital, several general hospitals, and the rehabilitation centre at Hoensbroek.

5.2.2 Study design

The economic analysis was performed alongside a randomized clinical trial examining the surplus efficacy of a cognitive program and relaxation when added to an operant treatment. Patients were randomly assigned to a waiting list or received an operant program with cognitive program and relaxation (OPCON) or an operant treatment with a group discussion treatment (OPDIM). The latter was added as an 'attention-control', to control for the nonspecific effects of the cognitive and relaxation treatment. The waiting-list control condition (WLC) refers to a waiting list period of about 10 weeks, during which no protocolized therapy was given. After completion of the waiting period, patients received individualized operant rehabilitation as usually provided in

the rehabilitation centre (USUAL). Patients were randomized as follows across treatments: Before the first pretreatment measurement each patient was given a number that was written down on a card and folded up. An independent researcher blindly drew a card and assigned it to one of the three conditions. In case a patient dropped out before the first pre-treatment assessment, he or she was replaced by another patient with the same gender and age, who was not yet included in a randomisation. OPCON and OPDIM were given alternated in time, but this schedule was not known to the research assistant who did the measurements and observations nor to the therapists other than the behavior therapist who gave the cognitive treatment and group discussion. If the number of patients screened was too small to fill three groups, OPCON and OPDIM were filled first. The WLC condition was then included in the next randomisation.

The operant program is aimed at increasing health behaviors and activity levels, and to decrease pain behaviors and excess disability.¹⁴ The program always starts with a number of baseline exercises in which the patient exercises to the limit of tolerance. The therapist then sets a quota of exercises to be performed each session. Initial quotas are lower than baseline levels, but are increased systematically toward a preset goal. In situations where patients are not able to actually perform certain activities, shaping procedures can be applied during which the activities are relearned. In contrast to the traditional medical approach, the patients exercise to quota (the patient rests when a certain amount of exercise is performed) and not to tolerance (the patient rests when pain increases). This approach is different from the 'back schools' in that the latter primarily rely on providing information, while the operant approach allows the patient to realise that it is safe to move even when also increasing his or her activity level. An important component in the protocolized operant program is the partner group instruction, during which partners are taught to recognise the difference between pain behavior and healthy behavior, to be more aware of their own responses toward these behaviors, and to attend more to health behaviors than to pain behaviors. Partner involvement facilitates the generalisation of change in the home setting.

The cognitive program with relaxation consisted of 12 sessions of 90 minutes and was conducted by a psychologist. The cognitive program aimed at decreasing distorted pain cognitions and increasing self-expectations. Patients learned to use imagery to influence their pain. Applied relaxation was used to teach patients to relax and to use relaxation in situations with personally relevant stress. At the end of the program patients chose exercises they considered to be the most useful and received an individually tailored tape to use after discharge.

The group discussion treatment contained an equal number of sessions which were given by the same therapists as the cognitive treatment with relaxation. However, since the discussion program was intended to act as an attention-control, the specific techniques and skills used in the cognitive/relaxation treatment were deleted. In groups of five patients, pain-related issues were discussed. They also listened to music, both during the session and at home.

The operant treatment accounts for 87% of the total therapy hours of both the OPCON and OPDIM treatment, whereas the cognitive/relaxation treatment and the discussion

treatment account for 13% of the total therapy hours of the OPCON and OPDIM, respectively. Both OPCON and OPDIM are eight week programs, five of which are inpatient and three of which are outpatient. Most components of both programs are organized as group sessions with five patients per group.

The *USUAL rehabilitation program* is, in contrast to OPCON and OPDIM, an individualized treatment based on the principles of the operant treatment, but excluding cognitive skills training, group discussion and partner group instruction (see table 1). The duration of this treatment and the proportion of inpatient/outpatient treatment depends on the specific patient situation.

Patients participating in OPCON or OPDIM were seen for data collection at five different times: two weeks before starting the program (PRE1), at the start of the program (PRE2), immediately after completing the program (POST), and six and 12 months after completing the program (FU1, FU2). In the analyses the mean score of the two PRE measurements was used. Patients assigned to WLC were also seen at PRE1, PRE2, and POST. WLC patients receiving the USUAL program after the waiting period were asked to participate in measurements immediately after completion of the USUAL program and six and 12 months after USUAL therapy. For these WLC patients the measurement at POST was the PRE measurement for the USUAL therapy. At these measurements patients completed tests, questionnaires, a utility measurement interview, and had their cost diary reviewed.

5.2.3 Costs

To evaluate the economic effects of the treatment, the first step is to identify relevant categories of resource use; secondly the identified volumes of categories must be measured and these volumes multiplied by the resource costs.

Identification of cost categories. As economic endpoints direct health care costs, direct non-health care costs and indirect costs were considered. The direct health care costs included costs of the OPCON, OPDIM, and USUAL programs and the economic consequences of the programs in terms of changes in health care utilisation. This included costs of additional therapies, drug use, hospitalisations, and visits to health care providers (general practitioner, specialist, physiotherapist, alternative medicine). The direct non-health care costs included out-of-pocket expenses, costs of paid and unpaid help, and travel costs of attending a back pain treatment program or visit. The indirect costs refer to the value of production lost to society due to illness-related absence from work or days of inactivity at home.

Measurement of resource utilisation. Information on the resource use associated with OPCON, OPDIM, and USUAL were obtained using data and invoices from the financial department of the rehabilitation centre, time schedules filled in by the therapists, and the therapy protocols. The OPCON and OPDIM program were based on treatment protocols, but the USUAL program was individualized and not protocolized. The volumes of the other direct costs were obtained from the patient-cost diary in which the patients recorded all medical consumption related to their back pain. The patients in OPCON, OPDIM, and USUAL completed this cost diary during the 12

months after the end of the therapy. Since the patients in WLC participate in the USUAL program after the waiting list period, resource use while on the waiting list could only be recorded for a limited period of about 10 weeks. The calculation of indirect costs associated with production losses was based on recordings of the number of days absent from work and the days lost from housekeeping. These were recorded by the patient in the cost diary.

Resource costs. The costs in 1993 Dutch guilders have been converted into US dollars (\$) at the 1993 GDP-based Purchasing Power Parities rate (PPP) of 2.134:1.¹⁵ The costs of the rehabilitation programs were based on a detailed cost-accounting study. The costs of various specialists, the general practitioner and physiotherapist consulted during the follow-up period are based on the 1993 charges of the 'Centraal Orgaan Tarieven Gezondheidszorg' (Central Organisation for Health Care Charges). A population-weighted average of social and private insurance charges was used. The costs of drugs were based on Dutch pharmacy prices. The prices of over-the-counter medication are based on the real costs reported by the patients in a cost diary. Costs of unpaid help by family or friends were valued using the price for professional help (US \$7.03 per hour) as a shadow price. For the other direct nonhealth care costs, the prices reported by the patient in the cost diary were used. Indirect costs were calculated in two ways. The Human Capital Approach estimates the value of the potential production loss during the entire period of absenteeism.^{16,17} The central assumption of the Friction Cost Approach assumes that production is lost only during the friction time. This approach is proposed by Koopmanschap *et al.*¹⁸ because the Human Capital Approach overestimates the value of the lost production since sick employees can be replaced or work can be postponed and made up later. Production is only lost during the friction time, which is the time needed to replace the employee or reorganize the production process. The friction time can be estimated by the time needed to fill vacancies. In this study an average friction time of three months was used.¹⁸ In both approaches lost production was valued using the national average gross hourly wage of US \$10.78 for 1993.

5.2.4 Utilities

The Maastricht Utility Measurement Questionnaire was used to elicit utility values from the patients. The instrument is a translated and slightly adapted version of the McMaster Utility Measurement Questionnaire.^{19,20} Patients were first asked to rate their current functional level using a 5-point scale (1=best level, 5=worst level) for six domains: activities of daily living, self-care, emotions, leisure activities, pain and other complaints, and side effects of treatment. The actual valuation of their health state was measured by means of rating scale and standard gamble techniques. The rating scale is a visual analogue scale ranging from perfect health (100) to death (0). The patients are asked to place their health status on the scale, visualized as a thermometer. The value

¹⁵Purchasing Power Parities (PPP) are rates of currency conversion that equalize the purchasing power of different currencies. The PPP conversion rate eliminates the differences in price levels between countries. It enables international price and volume comparisons of GDP and its component expenditures

placed on the thermometer directly equals the utility. In the standard gamble patients are offered a choice of remaining in their current health state for the rest of their life or undergoing a treatment with a chance of return to perfect health (p) or a chance of dying immediately after the treatment ($1-p$). The probability of treatment success is varied until the patient is indifferent between remaining in his or her current health state or undergoing the risky treatment. At this indifference point the utility equals p . To visualize the term 'probability', the standard gamble method is performed with a probability wheel as a prop.

5.2.5 Global assessment of change

At follow-up patients were asked to give a global assessment of the change in their health state on a 3-point scale (improved, deteriorated, or remained the same compared to the baseline situation). Patients were asked to compare their health state of the last two weeks with their health state at baseline.

5.2.6 Statistical analysis

To calculate the costs the patient-year approach was used, where the observed costs of the patients with one or more missing weekly cost diaries were extrapolated to a 1-year period.²¹ The differences in baseline utility scores were analysed using ANOVA statistics. To test the differences in costs and effects between the OPCON, OPDIM, and USUAL conditions, a multiple linear regression analysis with a hierarchical backward elimination method was performed. Since the distribution of the costs was skewed to the right, the square root transformation of the costs was applied for the regression analyses. The different regression analyses were done with total direct costs, the direct health care costs, and direct nonhealth care costs as dependent variables. Multiple regression analysis was also performed to test differences in utilities. Independent variables in the regression analysis of both the costs and utilities were program condition, gender, age, education, pain duration, disability pension, treatment credibility, biomedical status, activity level, pain behavior, psychopathology, depression, fear, marital problems, baseline standard gamble utilities, baseline rating scale utilities, and the interaction variables age*therapy, gender*therapy and baseline*therapy. To analyze the difference in costs between the conditions, only the OPCON, OPDIM, and USUAL are compared with each other. No costs comparisons have been made with the WLC condition, because the patients in this condition only completed the cost diaries the first 10 weeks of the research period.

At each step of the analysis tests were done to check for high collinearity ($VIF > 10$) and/or outliers (Cook's Distance (Cook D) and Studentized Residual (Sresid)). If Cook $D > 1$ or Sresid < -3 or > 3 the case was removed. By looking at plots of the relationship between each independent variable and the dependent variable a possible curve linear relationship was excluded. It was also checked whether the predictive errors were normally distributed (Zresid).

5.3 Results

5.3.1 Baseline characteristics

A total of 237 patients with chronic low back pain were presented to participate in the study. Of these patients 62 (26.2%) did not fulfill the inclusion and exclusion criteria and 27 (11.4%) did not want to participate in the study. Of the 148 remaining patients, 59 were randomly assigned to the OPCON group, 58 to the OPDIM group and 31 to the WLC/USUAL group. Patients in the three conditions did not differ on demographical data.¹² Of the patients 64% were female, the mean age was 39.8 years and the average duration of pain since the onset was 9 years 10 months. Seventy-eight percent had lower levels of education (i.e. less than 10 years) and 79% received a work disability payment for a mean duration of 3.7 years. Before entering the program 12 patients reported having a job, 39% had received back surgery, and 28% used supportive equipment for ambulation. Data from the compulsory insurance fund were analysed to test for differences in outpatient specialist care and physiotherapy among groups one year prior to the trial. Of the compulsory insured patients 63% had used physiotherapy at least once and 93% had at least one outpatient specialist contact. This was not significantly different among the three groups. In addition, no differences between the groups were found on baseline utility scores measured by means of rating scale ($p=0.222$) and standard gamble ($p=0.291$) method. The scores on the six dimensions of health (1.0=best, 5.0=worst) did not differ between the groups. The score on the dimension 'pain and other complaints' was 4.0 (SD=0.56), which suggests that the back pain patients suffer from frequent pain. They also report serious limitation in activities of daily living (3.6, SD=0.62) and in leisure activities (3.8, SD=0.68). The scores on the other domains of health are 2.0 (SD=0.90) for self care, 3.4 (SD=0.68) for emotions, and 1.6 (SD= 0.79) for side effects from treatment.

During the one year follow-up 41 patients (27.7%) dropped out at several stages of the study. Attrition was not equally divided over the treatment conditions: 24% for OPCON, 24% for OPDIM, 3% for WLC (patients were only followed for 8 weeks), and 43% for USUAL. These patients did not differ from the remaining patients on biographical parameters, age, sex, educational level, and clinical characteristics. The main reasons for dropout were problems in home help during absence of the patients, not willing to participate in the research, family circumstances, and non-back related illness of the patient or another family member.

5.3.2 Costs

Program costs. The costs of OPCON are higher than the costs of OPDIM (respectively US \$9196 and US \$8607 per patient) when excluding the attention-control component of the OPDIM treatment. The USUAL treatment has been offered to all patients on the WLC after an average of eight weeks of waiting time. This treatment is adapted to individual needs. This means that duration, content, and intensity of the treatment differ by patient. The average costs of USUAL are US \$8667 (SD=1014.45), varying from US \$7808 to US \$10229. Table 1 shows that the difference in costs between the protocolized programs and USUAL is mainly because the USUAL program contains

less psychology, more occupational therapy, more treatment days, and almost no interdisciplinary meetings. In the protocolized programs, the costs for interdisciplinary meetings contribute as much as 11% of the total costs.

Table 1. Program costs of OPCON, OPDIM, and USUAL (US \$)

Program components	Costs per hour (days)	Hours (days) of therapy			No. of patients in group			Total costs		
		OPCON	OPDIM	USUAL	OPCON	OPDIM	USUAL	OPCON	OPDIM	USUAL
Rehabilitation therapy	115.02	4.5	4.5	5.25	1	1	1	517.60	517.60	595.24
Social work	86.10	1	1	2.95	1	1	1	86.10	86.10	254.00
Psychology (indiv.)	101.62	4	4	2.65	1	1	1	406.49	406.49	269.30
Psychology (group)	101.62	30.5	-	-	5	-	-	589.40	-	-
Psychology (partner group)	101.62	9	9	-	5	5	-	255.11	255.11	-
Recreational therapy	227	-	-	0.4	-	-	5	-	-	11.35
Occupational therapy (indiv.)	50.14	11.5	11.5	8.75	1	1	1	576.62	576.62	438.73
Occupational therapy (group)	50.14	26	26	29.35	15	15	5	86.91	86.91	294.48
Physiotherapy (group)	32.85	13	13	44	5	5	5	427.04	427.04	289.07
Physiotherapy (indiv.)	32.85	49.5	49.5	12.7	1	1	1	325.21	325.21	417.18
Sport and games therapy (group)	39.00	26	26	36.25	5	5	5	262.78	262.78	238.28
Sport and games therapy (indiv.)	39.00	4.5	4.5	7.8	1	1	1	175.49	175.49	255.45
Interdisciplinary meeting	338.63	3	3	0.27	1	1	1	1015.89	1015.89	91.49
Team meeting	424.73	1	1	1.08	1	1	1	424.73	424.73	456.59
Inpatient day treatment (no. of days)	151.94	25	25	30.7	1	1	1	3798.62	3798.62	4664.70
Day treatment (no. of days)	27.56	9	9	14.2	1	1	1	248.07	248.07	391.40
<i>Total program costs</i>								<i>9196.06</i>	<i>8606.66</i>	<i>8667.26</i>

Other direct health care costs. An average of 82% of the patients completed the weekly cost diaries and returned respectively 78% (OPCON), 76% (OPDIM), and 50% (USUAL) of the total number of diaries. Table 2. shows the annual health care utilisation per patient per condition. Remarkable are the many hours of unpaid help from family and friends. In the OPCON, OPDIM, and USUAL group respectively 54%, 40% and 50% of the patients received help from family or friends.

Table 2. Mean (SD) health care utilisation per patient-year by treatment group

Number of:	OPCON (n=43)	OPDIM (n=40)	USUAL (n=26)
General practitioner contacts	7.0 (8.4)	7.9 (10.9)	6.0 (7.5)
Outpatient specialist care contacts	5.1 (9.2)	4.6 (10.6)	2.8 (6.0)
Physiotherapy contacts	21.5 (45.9)	12.9 (26.6)	10.6 (21.7)
Visits to psychologists/outpatient mental health centre	1.3 (5.9)	1.0 (0.7)	2.1 (5.1)
Alternative medicine contacts	1.6 (5.0)	1.0 (3.5)	5.5 (20.8)
Inpatient hospital days	1.4 (4.6)	0.1 (0.5)	0.0 (0.0)

Number of:	OPCON (n=43)	OPDIM (n=40)	USUAL (n=26)
Different recipes prescribed	2.5 (3.2)	2.4 (3.0)	3.2 (4.4)
Different OTC medications bought	2.2 (7.9)	5.4 (3.0)	6.3 (13.9)
Hours of professional home help	25.8 (29.4)	16.2 (57.7)	22.6 (68.2)
Hours paid housekeeping help	82.1 (163.0)	54.0 (89.5)	35.1 (81.3)
Hours of unpaid help from family or friends	156.0 (330.5)	63.5 (117.0)	83.9 (129.8)
Health activities	81.8 (52.6)	79.2 (56.2)	35.6 (24.8)
Expenses for equipment, clothing, etc	0.8 (1.1)	1.7 (2.9)	2.1 (4.3)

The total direct health care costs are US \$1088 (SD=1812) for OPCON, US \$575 (SD=695) for OPDIM, and US \$651 (SD=848) for USUAL. The total direct nonhealth care costs are US \$2316 (SD=3691), US \$1544 (SD=1636) and US \$1641 (SD=1621) for OPCON, OPDIM, and USUAL, respectively (figure 1). Costs were heavily skewed to the right. When program costs are excluded 20% of the patients accounted for about 56% of the direct costs. Median (interquartile range) total direct health care costs are US \$393.52 (US \$1131) in OPCON, US \$243 (US \$913) in OPDIM and US \$392 (US \$868) in USUAL. Median (interquartile range) total direct nonhealth care costs are US \$999 (US \$3341), US \$1095 (US \$2477) and US \$1338 (US \$2110) respectively in OPCON, OPDIM and USUAL. Non parametric tests indicated that these costs were not significantly different between the groups.

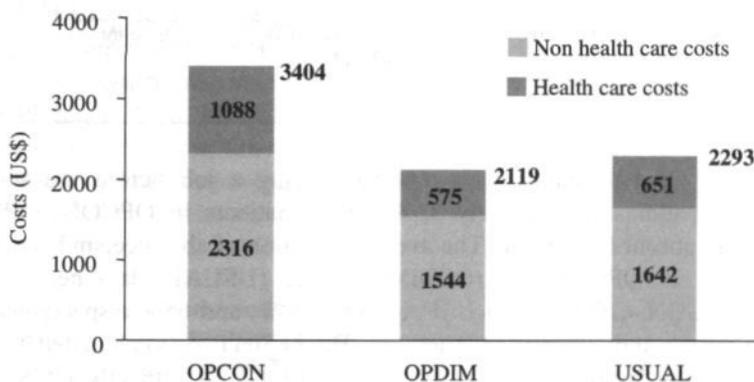


Figure 1. Estimated total direct costs per patient-year per treatment group (US\$)

Regression analyses on the square root of the costs confirmed that there were no differences in costs between OPCON and USUAL and between OPCON and OPDIM. Results from the ordinary least square regression analysis with total direct costs and total direct non-health care costs as dependent variables are shown in the upper part of table 3. It appears that patients with more psychopathology, a lower activity level, more fear, and a higher pain control have higher direct health care costs. Women and patients with lower activity levels appear to have higher direct nonhealth care costs (and therefore also higher total direct costs). Furthermore, patients with lower activity levels and psychopathology have higher costs than patients with higher activity levels and with lower levels of psychopathology.

Table 3. Regression analyses direct costs and utilities (FU2): OPCON vs OPDIM vs USUAL

Dependent variable	N	adj R2	OPCON vs USUAL		OPDIM vs USUAL		Explanatory variables		
			B (95% CI)	P	B (95% CI)	P	B (95% CI)	P	
Direct health care costs	93	.23	4.11 (-6.22; 4.43)	.431	3.19 (-7.18; 13.56)	.543	Psychopathology	.88 (.30; 1.45)	.003
							Activity level	-1.69 (-3.25; -.14)	.033
							Pain control	1.62 (-.02; 3.26)	.053
							Anxiety level	.11 (.008; .21)	.035
Direct non health care costs	98	.24	-1.86 (-18.25; 14.51)	.821	1.63 (-14.82; 18.07)	.845	Gender	28.79 (15.61; 41.97)	.000
							Activity level	3.16 (-5.56; -7.57)	.011
Total direct costs	96	.29	2.87 (-13.22; 8.96)	.724	.67 (-15.37; 16.71)	.934	Gender	29.28 (15.42; 43.13)	.000
							Psychopathology	2.10 (1.18; 3.02)	.000
							Activity level	-2.64 (-4.99; -.29)	.028
Rating scale	91	.25	10.69 (.812; 20.56)	.034	7.06 (-2.73; 16.85)	.155	Baseline	.60 (.343; .851)	.000
							Disability pension	-6.65 (-13.25; -.05)	.048
Standard gamble	96	.49	.06 (-.016; .135)	.122	.05 (-.03; .13)	.184	Baseline	.72 (.56; .88)	.000
							Gender	.06 (.02; .110)	.020

Indirect costs. Of 12 patients who reported having a job before the start of the treatment, 10 patients (respectively 3, 4, and 3 patients in OPCON, OPDIM and USUAL) were absent from work. The average duration of absenteeism because of low back pain was 2 (OPCON), 19 (OPDIM), and 10 (USUAL) days per year. Of the patients in OPCON, OPDIM, and USUAL 60%, 60%, and 65% respectively reported on average 89, 60, and 105 days of inactivity. The indirect costs calculated by means of the Human Capital Method were US \$6522 (SD= 17387) for OPCON, US \$5938 (SD= 16172) for OPDIM, and US \$8213 (SD= 17394) for USUAL. The difference between the groups was not statistically significant.

5.3.3 Utilities

Figure 2 shows the utility scores measured by the rating scale and the standard gamble. Immediately after treatment the patients in OPCON and OPDIM value their own health state significantly higher on the rating scale compared to the no-treatment condition (WLC) (for OPCON vs WLC; 95% CI = 7.85 - 24.18, for OPDIM vs WLC; 95% CI = 7.68 - 24.27). There is no significant difference between the two experimental groups and between the experimental groups and USUAL, either immediately after treatment or after 6 and 12 months. For the standard gamble utilities there are no differences

between the groups. Table 3 shows the least square regression analyses for the standard gamble and rating scale at FU2.

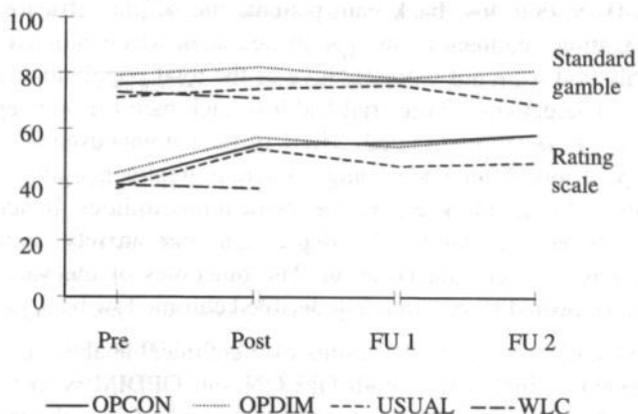


Figure 2. Rating scale and standard gamble utilities per treatment group.

5.3.4 Global assessment of change

Figure 3 shows the percentage of patients who evaluate the change in their health state as improved, deteriorated, or not changed. After 8 weeks there is a significant difference between the program groups and the WLC group in the percentage of patients who report their health state as improved ($p=0.000$). The percentage of patients who evaluate their health state as improved, deteriorated, or equal compared to the situation at baseline is not significantly different between OPCON, OPDIM, and USUAL.

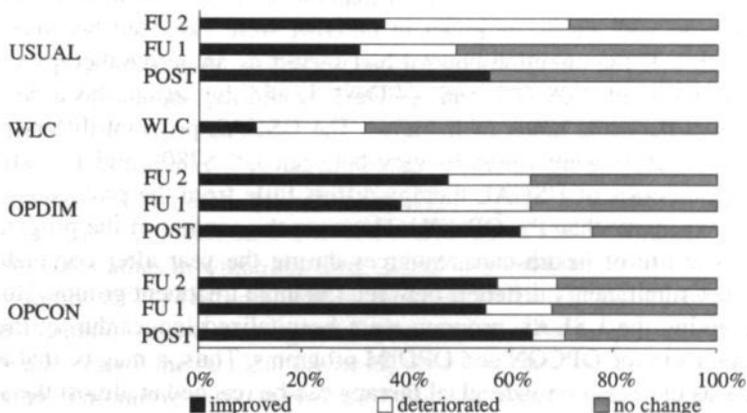


Figure 3. Global assessment of change per program group

5.4 Conclusion and discussion

For a group of 148 chronic low back pain patients the surplus efficacy of adding a cognitive and relaxation treatment to an operant treatment was evaluated. The patients participating in this trial were not representative of the total population of chronic low back pain patients. The patients in our trial had low back pain for on average 10 years and were moderately to severely disabled. 27.7% of the patients used walking aids and only 11% had a paid job. After undergoing numerous treatments, all without lasting improvement, they had reached the end of the medical possibilities. In addition to pain they had various other complaints like depression and anxiety, although severe psychopathology was an exclusion criterion. The outcomes of the various treatment programs might have been different in a less disabled chronic low back pain sample.

The economic evaluation confirms the results of the clinical analysis by showing that the operant treatment as included in both OPCON and OPDIM is more effective in improving utilities than providing no intervention at all which was the quintessence of the WLC.¹² Immediately after treatment patients in both OPCON and OPDIM report a higher valuation of overall health on the rating scale as compared to patients in WLC. In contrast to the expectations, adding a cognitive/relaxation program of about US \$1251 per patient to the operant program did not result in any additional short-term or long-term effects. When compared to USUAL, both OPCON and OPDIM do not seem to have additional effects. All three groups show an improvement, but the improvement is not significantly larger in OPCON and OPDIM. This may be due to a power problem caused by a large number of patients (43%) who were lost to follow-up in the USUAL group. The power for detecting the 10% difference at a significance level of 0.05 between OPCON, OPDIM, and USUAL is 0.58 for rating scale values and 0.79 for standard gamble values. It is not possible to determine the power of the costs, because there is no reference value for the smallest clinically relevant value of the costs.

The program costs of OPCON were higher than those of OPDIM when the costs of the group discussion with music as given in OPDIM were excluded because it was an attention control. If the attention-control had served as an active therapy component, the program costs of OPCON and OPDIM would be equal, because the same disciplines give the same hours of therapy. The USUAL program differs per patient and its program costs were found to vary between US \$7808 and US \$10229. The average program costs of USUAL therapy differs little from the protocolized OPDIM and is more expensive than the OPCON. However the content of the programs differs essentially. The use of health care resources during the year after completion of the program is not significantly different between the three treatment groups. However, all patients receiving the USUAL program were hospitalized longer during the program than the patients in the OPCON and OPDIM programs. Thus, it may be that at least the same effects as the common individual therapy can be reached at almost the same costs by a shorter, more intense, standardized group program.

The absence of an additional effect of the cognitive/relaxation treatment was unexpected. One explanation might be that, as mentioned earlier, the discussion/music

treatment that was intended as an attention control has actually served as an active treatment. The support patients experience from other patients during the group discussions might have had unexpected positive effects. A second explanation might be that most patients were not able to complete their homework assignments from the cognitive program. Cognitive therapy is most successful when patients practice at home on a regular basis. In this study only a small number of patients returned fully completed homework books. A third explanation might be that this cognitive program was too difficult given the relatively low educational level of the patients. The absence of short term improvements induced by a cognitive program might have decreased the patients' motivation. Finally, the contrast between the OPCODE and OPDIM programs might have been too small because in both programs about 87% of the total program time was associated with the operant, which is in itself effective.¹⁰ Only 13% remained for either a cognitive/relaxation program in OPCODE or a discussion/music program in OPDIM. In addition, in order to increase blinding of the patients, the cognitive/relaxation program and the discussion/music program was given by the same behavioral therapist. Contamination bias could have been present. Contamination bias was checked by nine independent behavior therapists who listened to 10 audiotapes of various moments in the cognitive/relaxation program and the discussion/music therapy. After listening to the tapes they had to indicate to which of the two therapies the audiotape referred. Overall, only 18% of the 90 possibilities were incorrectly classified. Thus, the presence of contamination bias is probably limited.

The absolute necessity of five weeks of inpatient treatment can be questioned. Since these patients do not need 24 hour care, outpatient treatment could possibly reduce the costs.²³ However, given the severity of the patient population, the multidisciplinary team feels that it is important for patients to leave their home in order to concentrate on the therapy. Furthermore, a recent study by Pither *et al.*²³ in a comparable group of chronic low back pain patients has shown that the outcomes of inpatient treatment are better than the outcomes of outpatient treatment. Outpatient treatment combined with staying overnight in a motel could fulfill this need and would still reduce the program costs.

Patient derived utilities were used to estimate the net impact of the treatment programs on the quality of life of the patient. Utility measures differ from the other quality of life measures in that the first not only describe but also value the quality of life. In this trial, utilities were measured by means of the rating scale and the standard gamble method. All patients valued their own health state higher on the standard gamble than on the rating scale. This pattern is consistent with many earlier findings^{24,25} and is also found in several studies in which the same utility instruments as in this study were used in similar settings^{26,27,28}. Different phenomena might explain the differences between the standard gamble and the rating scale preferences. The most important is that the difference may reflect the fact that the rating scale does not incorporate a choice with a risk of dying. The rating scale values are measured under certainty, in comparison to the standard gamble utilities which are measured under uncertainty. This means that if patients are not risk neutral, differences can be expected between rating scale and standard gamble scores. Risk-averse behavior of the chronic low back pain patients

might have been reinforced by the fact that these patients have had back pain for an average of 10 years. In this period they have tried many different therapies, without much success. Patients therefore are not likely to try a new risky treatment, which could involve a chance of dying. Of the chronic low back pain patients 53% were not willing to take a 10% chance of dying on at least one of the measurement times, resulting in a utility of 0.95. As a result of the ceiling effect there was little room for improvement. It can be concluded that the standard gamble as used in this setting is not suitable for evaluating a therapy for chronic low back pain patients.

To summarize, the results of the health economic assessment of behavioral rehabilitation in chronic low back pain suggest that the addition of a cognitive/relaxation therapy to an operant therapy as evaluated in this randomized clinical trial resulted in higher costs and no substantial surplus effects.

Acknowledgments: We are grateful to Peter Heuts and the staff of the Department of Pain Management of the Lucas Foundation for Rehabilitation Hoensbroek for the management of patients included in the study and Hugo van Eek, Jos Offermans and the RIAGG Maastricht. We are also grateful to Jan van Emmerink and Jolanda van Haastregt for their statistical advice and thank Ferd Sturmans and Eddy van Doorslaer for advice on the study. The study was supported by a grant from the investigative medicine program of the Health Insurance Executive Board.

References

1. Frymoyer, J.W., Cats-Baril, W.L. An overview of the incidences and costs of low back pain. *Orthopedic Clinics North Am* 1991; 22: 263-271.
2. Andersson, G.B.J. The Epidemiology of spinal disorders. In: Frymoyer J.W. *The adult spine: principles and practice*. New York: Raven Press, 1991: 107-146.
3. Tulder van M.W., Koes B.W., Bouter L.M. A cost-of-illness study of back pain in The Netherlands. *Pain* 1995; 62: 233-240.
4. Nachemson, A.L. Newest knowledge of low back pain. A critical look. *Clinical Orthopedics and Related Research* 1992, 279, 8-20.
5. Philips, H.C. Avoidance behaviour and its role in sustaining chronic pain. *Behaviour Research and Therapy* 1987; 25: 273-279.
6. Waddell, G. 1987 A new clinical model for the treatment of low-back pain. *Spine* 1987; 12: 632-644.
7. Vlaeyen, J.W.S. *Chronic low back pain: Assessment and treatment from a behavioural rehabilitation perspective*. Lisse, Swets & Zeitlinger, 1991.
8. Vlaeyen, J.W.S., Haazen, I.W.C.J., Kole-Snijders, A.M.J., van Eek, H., Schuerman, J.A. Behavioural rehabilitation of chronic low back pain: Comparison of an operant treatment, an operant-cognitive treatment and an operant-responder treatment. *British Journal of Clinical Psychology*, 1996; 34, 95-118.
9. Turner, J.A., Chapman, C.R. Psychological interventions for chronic pain: a critical review. II. Operant conditioning, hypnosis, and cognitive-behavioural therapy. *Pain* 1982; 12: 23-46.
10. Linton, S.J. Behavioural remediation of chronic pain: a status report. *Pain* 1986; 24: 125-141.

11. Cohen, M.J., Naliboff, B.D., McArthur, D.L. Implications of medical and biopsychosocial models for understanding and treating chronic pain. *Critical Reviews in Physical and Rehabilitation Medicine* 1989; 3: 135-160.
12. Kole-Snijders, A.M.J., Vlaeyen, J.W.S., Goossens, M.E.J.B., Rutten van Mólken, M.P.M.H., Heuts, P.H.T.G. Chronic low back pain: what does cognitive treatment add to operant treatment? Results of a randomized clinical trial. *Journal of consulting and clinical psychology* (accepted).
13. Goossens, M.E.J.B., Evers, S.M.A.A. Economic evaluation of back pain interventions. *Journal of Occupational Rehabilitation* 1997; 7 (1); 15-32.
14. Fordyce, W.E. *Behavioural methods for chronic pain and illness*. St. Louis: Mosby 1976.
15. OECD, Statistics Directorate. *Historical statistics 1960-1994*, Paris 1996
16. Rice, D.P., Hodgson, T.A., Kopstein, A.N. The economic cost of illness: a replication and update. *Health Care Financing Review* 1985; 7, 61-80.
17. Koopmanschap, M.A., Rutten, F.F.H. Indirect costs in economic studies; Confronting the confusion. *PharmacoEconomics* 1993; 4(6): 446-454.
18. Koopmanschap, M.A., Ineveld, B.M.van. Towards a new approach for estimating indirect costs of disease. *Social Science and Medicine* 1992; 34: 9, 1005-1010.
19. Bennett, K., Torrance, G.R., Tugwell, P. Methodologic challenges in the development of utility measures of health-related quality of life in rheumatoid arthritis. *Controlled Clinical Trials* 1991; 12: 118-128 (suppl).
20. Bakker, C.H., Rutten-van Mólken, M.P.M.H., van Doorslaer, E.K.A., Bennet, K., van der Linden, S. Health related utility assessment by rating scale and standard gamble in patients with ankylosing spondylitis or fibromyalgia. *Patient Education and Counseling* 1993; 20: 145-52.
21. Jacobson, N.S., Folette, W.C., Revensdorf, D., Baucom, D.H., Hahlweg, K., Margolin, G. Variability on outcome and clinical significance of behavioural marital therapy: a reanalysis of outcome data. *Journal of Consulting and Clinical Psychology* 1984; 52: 497-504.
22. Rutten-van Mólken, M.P.M.H., Doorslaer, E. van, Vliet, R.C.J.A. van. Statistical analysis of cost outcomes in a randomized controlled clinical trial. *Health Economics* 1994; 3; 333-345.
23. Turner, J.A., Romano, J. Evaluating psychological interventions for chronic pain: issues and recent developments. In: Beneditti, editor. *Advances in pain research and therapy*. New York; Raven Press, 1984.
24. Pither, C.E., Williams, A.C. de C., Richardson, P.H., Nicholas, M.K., Harding, V.R., Ridout, K.L., Ralphs, J.A., Richardson, I.H., Justins, D.M. and Chamberlain, J.H. Inpatient versus outpatient pain management: Results of a randomized controlled trial. (submitted).
25. Read J.L., Quinn R.J., Berwick, D.M., Fineberg, H.V., Weinstein M.C. Preferences for health outcomes; comparisons of assessment methods. *Medical Decision Making* 1984; 4: 315.
26. Van Busschbach J. *De validiteit van QALY's*. [The validity of QALYs.] Proefschrift. Sanders Instituut. Gouda Quint-Arnheim, 1994.
27. Bakker, C., Rutten-van Mólken, M., Doorslaer, E. van, Bennett, K., Linden, S. van der. Feasibility of utility assessment by rating scale and standard gamble in patients with ankylosing spondylitis or fibromyalgia. *Journal of Rheumatology* 1994; 21: 269-274.
28. Rutten-van Mólken, M.P.M.H., Bakker, C.H., Doorslaer, E.K.A. van. Methodological issues of patient utility measurement: experience from two clinical trials. *Medical Care* 1995; 33:922-37.
29. Goossens M.E.J.B, Rutten-van Mólken M.P.M.H., Leidl R.M., Bos S.G.P.M., Vlaeyen J.W.S., Teeken-Gruben N.J.G. Cognitive-educational treatment of fibromyalgia: a randomised clinical trial II. Economic Evaluation. *Journal of Rheumatology*; 1996; 23 (6): 1246-1254.

6 THE COST DIARY: A METHOD TO MEASURE DIRECT AND INDIRECT COSTS IN COST EFFECTIVENESS RESEARCH

This chapter is provisionally accepted as:

Mariëlle E.J.B. Goossens¹, Maureen P.M.H. Rutten- van Mólken², Johan W.S. Vlaeyen^{1,3}, Sjeff M.J.P. van der Linden⁴. The cost diary: a method to measure direct and indirect costs in cost effectiveness research. *Journal of clinical epidemiology*, 1999.

¹ Institute for Rehabilitation Research, Hoensbroek

² Institute for Medical Technology Assessment, Erasmus University Rotterdam

³ Department of Differential and Experimental Psychology, Maastricht University

⁴ Department of Internal Medicine/Division of Rheumatology, University Hospital Maastricht

Summary

From a societal perspective long-term clinical trials or follow-up studies should preferably not only include an evaluation of the health effect for the patient, but also an economic evaluation. In order to yield comprehensive medical and non medical resource use data, we at least partly depend on respondents' recall for collecting these costing data. A patient cost diary was developed in order to estimate total resource use, expenses and lost production due to illness and treatment. We applied the cost diary in two randomized clinical trials evaluating the cost-effectiveness of behavioral rehabilitation in fibromyalgia and chronic low back pain patients. The use of the diary was evaluated for 205 patients, studying the feasibility, the influence of the period of data collection on the results and some aspects of validity. The response rate turned out to be 85% and in total 68% of the diaries were returned. Similar outcomes were found for the three alternative periods of data collection: keeping the diary one week every month, two weeks every two months or a full year. Finally, self-reported specialist care contacts were generally in agreement with data from an insurance company. However, for physiotherapy contacts there were differences between the self-reported and insurance data. This study shows how the cost diary might be used successfully in cost-effectiveness studies.

6.1 Introduction

It is widely recommended that cost-effectiveness studies be conducted from the societal perspective. From this broad perspective, all costs and savings are considered. This means that studies designed to investigate cost-effectiveness should base their conclusions upon all relevant medical and non-medical costing data that they are able to collect. However, apart from the productivity costs, the costs for non-medical goods and services associated with an intervention are rarely included, since collecting this information is complex and demands careful scheduling in appraisals covering long-term care.^{1,2} Especially in evaluation studies of treatment for most chronic disorders, data collection should not be limited to the medical costs. Chronic diseases may not only generate a large and growing need of medical care, but they may also have a profound socio-economic impact on the patients, their families and their social and working relations.

In the context of cost-effectiveness analyses linked to a clinical trial, there are various sources to obtain data, such as patient records, data bases from insurance companies, hospital or provider data bases, interviews with patients or providers, questionnaires and diaries. Routine data bases compiled by insurance companies, hospitals and other health care institutions (such as pharmacies and general practitioners) are commonly used. Insurance companies may seem the most reliable data source. However, they are often unable to provide detailed resource use information due to budget or capitation paid systems. Furthermore, the large multiplicity of providers the back pain patient may see, limits the feasibility of the data collection. Even more important is the lack of visibility into medical consumption not covered by the patient's insurance. Moreover, there are no routine records for patients and family resource use, additional expenditures by the patient's or caregiver's household, as well as the input from informal caregivers.³ Because there are no institutional records tracking information on the type, frequency and magnitude of these expenditures, data must be obtained through self-reporting channels such as questionnaires and diaries.^{2,3} This means that studies undertaking a cost-effectiveness analysis for chronic musculoskeletal pain and from the societal perspective must, at least in part, depend upon information provided by the respondents.

Compared to questionnaires, diaries have been reported to offer several advantages in terms of feasibility and validity.⁴⁻⁸ While questionnaires usually rely on momentary recall, diaries provide information prospectively over a period of time, resulting in a minimum recall error (forgetting an event entirely, or remembering an event but forgetting its correct date), and therefore generally provide better and more complete data reporting.^{4-6,9} However, little research has been carried out on the validity of self-reported health care and utilization, primarily because of the scarcity of independent 'true' data sources to compare self-reports against.¹⁰ If valid costing and utilization data is to be collected, either the time period should be reduced to a limited period or a more accurate data collection method, such as the health diary, should be used.^{11,12} However, a short survey period would in general preclude the study of chronic disease

care, since a large number of chronic disease therapies are only effective when administered over a longer period of time or complications of the therapies arise only later.

Because no standardized method to estimate total resource use, expenses and lost production due to illness and treatment appear to be available, we have developed a cost diary. This cost diary allows for the estimation of direct and indirect costs in long-term clinical trials or follow-up studies perceived from a societal perspective. In the present study we examine the feasibility and validity of the cost diary. This paper presents the following issues: 1) the development and content of the cost diary; 2) the feasibility and data collection period of the cost diary in two clinical trials; and 3) aspects of construct validity of the cost diary.

6.2 Patients and methods

6.2.1 Subjects

The data we use were obtained in two randomized clinical trials evaluating the cost-effectiveness of cognitive behavioral treatment for fibromyalgia and chronic low back pain. Patients were enlisted between January 1992 and January 1994. The cost diary has been used in both studies for estimating direct health care costs (focusing on pain-related health care utilization), costs incurred by the patient and the family, and indirect costs. Since the focus of this article is on the methodological issues of the cost diary, it is not our intention to address the effectiveness or cost-effectiveness of the cognitive behavioral therapies. Details on this can be found in Vlaeyen *et al.*¹³, Goossens *et al.*^{14,15} and Kole-Snijders *et al.*¹⁶ For the purpose of this paper, patients involved in the fibromyalgia trial ($n=72$) and the chronic low back pain trial ($n=133$) were pooled into one sample ($n=205$). Only where there is a difference between the two groups of patients will the results be reported separately. Nearly 72% of the patients were women, the mean age was 42 years and the average duration of pain since the onset was 10 years. Seventy-eight percent had lower levels of education (i.e. less than 10 years), 22% of the patients were single at the time, 57.3% received financial disability compensation, 22.1% had a job and 19% used supportive equipment.

6.2.2 The cost diary

Structure of the cost diary. In order to obtain information that was as complete as possible, the layout of the diary was designed to be suitable for self-completion by the patient. We presented the diary in booklet form, containing instructions, an example (one completed booklet), a telephone number to dial in case of questions, and an accompanying letter explaining the objective of the diary. In the letter, as well as in the instruction, we also asked the patient to record only disease-specific resource use. Each diary covered a period of up to four weeks, in which each column represented one week (see Figure 1). The patient received the diaries in sets of 6 booklets covering a total of six months. To encourage a high response, prepaid envelopes were provided in order to send the first 3 booklets back after the 3 months. If the diaries were not

returned, the patients were called with the kind request to return them as soon as possible. The patients were asked to complete the diaries for a period of one year after treatment. Each patient attended three face-to-face interviews (immediately after treatment, 6 months and 12 months after treatment). At the first interview the patient received a verbal detailed instruction to fill in the diary at each moment he/she used one of the services included in the diary. The instruction is repeated in writing in each set of booklets and verbally at each face to face interview, to encourage prospective reporting of the information. The completed diaries were discussed with the patient during these interviews, to minimize partial responses and missing values. The patient was also asked to bring receipts and packaging for any medication purchased and bills of other expenses to check the information recorded in the diary. During intake of the diaries, the patient was encouraged to complete the next set of diaries.

Pre-test. Before widespread distribution, a group of chronic musculoskeletal pain patients at a rehabilitation center was used as a focus group to verify the clarity and completeness of the diary. During this phase, it was important to determine whether the diaries were clear enough to be understood and if the list of physicians, therapists, activities, (informal) help, medication and other aspects was complete. Several changes were made to the diary based on suggestions received, and over a period of two weeks 20 pain patients tested the modified version.

Components of the cost diary. Patients were required to record resource use in a detailed way to allow the multiplication with unit prices. The following information was requested:

- a) Direct health care costs; impact of the interventions on the use of other health care services, such as, visits to the general practice, specialist care, alternative medicine, and physiotherapy, days of hospitalization, and unities applied of each prescribed medication.
- b) Direct non-health care costs; costs incurred by the patient and the family, such as costs of over-the-counter medication, costs of health activities, hours of paid and unpaid household help, transportation and the value of other out-of-pocket expenses.
- c) Indirect costs; these costs refer to the value of production lost due to illness-related absence, such as the number of days absent from (paid and unpaid) work and days lost from housekeeping and other daily activities.

The patient was directed to indicate when, where and how often the activities, consultations, etc., took place within the time period of interest. He or she was also directed to document the name of the health care providers and his/her location, which gave us the possibility to check the information when the written text was unclear or incomplete and to calculate transportation costs. Figure 1 shows the entire cost diary.

Figure 1. components of the cost-diary

Weeknumber:	Week	Week	Week	Week
Date:Monday to Sunday
.....
General practitioner (name)	Number	Number	Number	Number
.....
Specialist (name, specialty)	Number	Number	Number	Number
.....
.....
.....
.....
Physiotherapy (name, type of therapy)	Number	Number	Number	Number
.....
.....
.....
Alternative medicine (name, type of therapy)	Number	Number	Number	Number
.....
.....
.....
Activities	Number	Number	Number	Number
Swimming
Gymnastics
Yoga
Other.....
Other.....
Work and household	Number	Number	Number	Number
Days of sick leave
Days unable to perform usual activities (for example household activities)

Weeknumber:	Week	Week	Week	Week
Date:Monday to Sunday

Prescribed medication (name, dosage, unities)	Number	Number	Number	Number
.....
.....
.....
.....
.....

OTC medication (name, costs)	Costs	Costs	Costs	Costs
.....	f.....	f.....	f.....	f.....
.....	f.....	f.....	f.....	f.....
.....	f.....	f.....	f.....	f.....
.....	f.....	f.....	f.....	f.....
.....	f.....	f.....	f.....	f.....

Assistance	Hours	Hours	Hours	Hours
Assistance from family and friends
Home help
Domestic help
total costs	f.....	f.....	f.....	f.....

Extra expenses, like adaptations at home, assistive devices (sort and costs)				
.....	Costs	f.....		
.....	Costs	f.....		
.....	Costs	f.....		
.....	Costs	f.....		
.....	Costs	f.....		

Hospital				
Date of admission.....				
Date of discharge.....				
Ward.....				
Name of hospital				

Valuation of the diary items. The following prices have been assigned to the resource quantities recorded in the cost diaries. The costs of the general practitioner (US \$13.44 per consultation), various outpatient specialists (ranging from US \$11.07 to US \$20.89 per consultation) and physiotherapists (ranging from US \$14.23 to US\$27.46 per consultation) consulted during the follow-up period, were based on a weighted average of the 1993 charges for privately or socially insured patients. In addition, for the costs of the various outpatient specialist visits, a weighted average of the charges for the first consultation and the follow-up consultations was used. The costs of drugs were based on Dutch pharmacy prices, including a mark-up for pharmacy expenses of US \$4.98. The prices of over-the-counter medication are based on the real costs reported by the patients in the cost diary. Costs of unpaid help by family or friends were valued using the price for professional help (US \$7.03 per hour) as a shadow price. For the other direct nonhealth care costs, prices reported by the patient in the cost diary were used. Lost production was valued using the national average gross hourly wage of US \$10.78 for 1993. For this calculation assumptions were based on the Friction Cost Approach.¹⁷ Information about the number of days the respondent was supposed to work was obtained at baseline during an interview with the patient. The costs in 1993 Dutch guilders have been converted into US \$ at the 1993 GDP-based Purchasing Power Parities rate (PPP) of 2.134:1*.¹⁸

6.2.3 Feasibility

To assess feasibility, the proportion of patients who completed the diaries successfully and the number of completed weeks were both considered. We also assessed the proportion of patients who only completed one diary (four weeks) and the proportion of patients who did not complete any of the diaries. Logistic regression analyses were performed to test whether or not patient characteristics were related to the response rates (completion of at least one diary-booklet and completion of the total set of diary-booklets). Based on findings in previous studies^{8,9,19,20} we hypothesized that more diaries would be completed by older patients, patients who are married, patients who have a bad health-related quality of life score (rating scale) and distressed patients (depression, fear). We also expected a positive relation between response rate and social desirability (scale for social desirability) (for a description of the instruments, see Goossens *et al.*¹⁵ and Kole-Snijders *et al.*¹⁶)

6.2.4 Influence of different periods of data collection

Because we believed that relevant information would be lost if the diaries were completed for only a limited period, we asked the patients to keep the diaries for one complete year. One may assume that keeping a diary for one year is too much effort for the patient. An alternative would be to ask the patient to keep the diary for only a limited period, for example, 2 weeks every 2 months or one month before each follow-

*Purchasing Power Parities (PPP) are rates of currency conversion that equalize the purchasing power of different currencies. The PPP conversion rate eliminates the differences in price levels between countries. It enables international price and volume comparisons of GDP and its component expenditures.

up visit and extrapolate that information to one year. We compared three alternative periods of data collection. First, we calculated total costs based on data from 52 weeks. This was the original lay-out of the cost-diary as used in the studies. In case of one or more missing diaries, costs were extrapolated to a full year period. Second, we calculated total costs using data from the first 2 weeks of each 2 months only (weeks 1, 2, 9, 10, 17, 18, 25, 26, 33, 34, 41, 42, 49 and 50). Third, we calculated total costs based on data from months 1, 6 and 12. Correlation coefficients for the three alternatives were calculated, and the differences in costs between the three alternatives were compared using the Wilcoxon Signed Ranks test.

6.2.5 Validity

The validity of direct health care costs was studied by comparing self-reported information with data obtained from insurance companies. We considered these data to reflect the 'true' or actual utilization because registration of these data is tied to reimbursement, which is an incentive for organizations to establish and maintain good quality registration routines. We have chosen to use the data of the largest sick fund, where 84% of the fibromyalgia patients and 61% of the chronic low back pain patients were insured. Within the scope of the study, we were forced to limit the validation to the number of physiotherapy contacts and specialist contacts, because other services are financed through capitation systems or budget systems. For the validation we were further forced to limit the sample to 40 chronic low back pain patients, first of all because of reasons related to a change in reimbursement system of physiotherapy. From 1994 a distinction of different forms of physiotherapy was no longer made. We could therefore only use the patients who had completed their year of follow-up before the end of 1993. Complete insurance data that match the period of the diary was only available for 40 patients. Contacts with any type of specialist were included as long as the contact was related to low back pain. The hypothesis for this comparison was that the (12-month follow-up) data of the cost diary result in similar estimations of total costs as the data of the compulsory health insurance company. To test differences in estimations, the non-parametric Wilcoxon Signed Ranks test was used due to the non-normal distribution of the data.

6.3 Results

6.3.1 Feasibility

Of the chronic low back pain and fibromyalgia patients, 85% completed at least one diary-booklet (four weeks). In total 68% of the diaries were returned, which corresponds to an average of 32 weeks in one year. Figure 2 shows the frequency distribution of the number of weeks the diaries were completed. Fifty percent (103) of the respondents completed the diary for the entire year of follow-up. Fifteen percent did not return any diary, primarily because they withdrew from the study. Results from the logistic regression analysis with both response rates as dependent variable showed that married respondents and respondents with more fear were more compliant than

non-married respondents and respondents with less fear, but these were only of borderline significance ($p=0.051$ and $p=0.083$). Furthermore the response rates were not significantly related to gender, education, age, pain coping, duration of pain, depression, social desirability and quality of life.

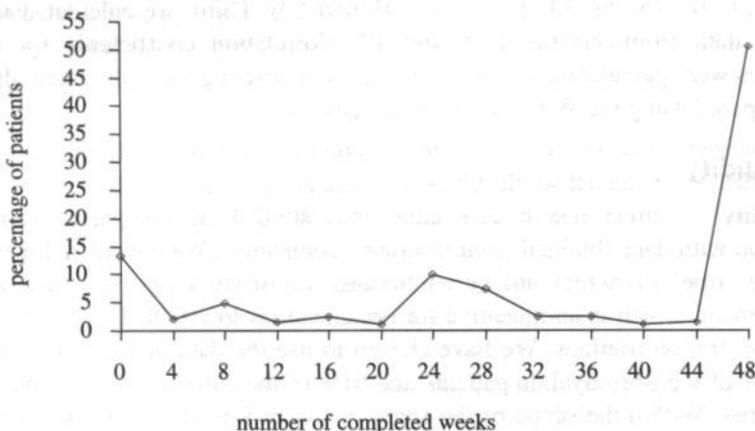


Figure 2. Number of completed weeks per patient

6.3.2 Time perspective

Table 1 shows that there is no significant difference in direct health care costs, direct non-health care costs and indirect costs between data from the entire year and the extrapolation of data from a limited period, either from 2 weeks every two months or from 3 months a year. Spearman correlation coefficients between the cost data in the alternative periods are all significant (for all comparisons $p<0.00$), but show nevertheless only moderate agreement.

Table 1. Comparison of mean (standard deviation) costs per patient per year (in 1993 US dollars) calculated for 3 different periods of data collection

	Alt 1:	Alt 2:	Alt 3:	Alt 1 vs Alt 2		Alt 1 vs Alt 3	
	Complete follow up Mean (SD)	2 weeks each 2 months Mean (SD)	3 months per year (1,6,12) Mean (SD)	P*	Cor.**	P*	Cor.**
Direct health care costs	817 (1348)	707 (896)	721 (1029)	0.550	0.33	0.737	0.32
Direct non-health care costs	1734 (2432)	1860 (2304)	2046 (3267)	0.179	0.31	0.062	0.26
Total direct costs	2551 (2835)	2567 (2607)	2767 (3543)	0.879	0.30	0.687	0.17

* Wilcoxon Signed Ranks test

** Spearman correlation

6.3.3 Validity

The mean number of contacts to a specialist (any specialist as long as the contact was related to the back pain) based on the data from the insurance company was 1.20 (SD

1.49) per year (table 2). Using the cost-diary the estimated number of contacts was 1.25 (SD 2.01) visits per patient per year. Non-parametric tests indicated no significant difference between the two methods ($p=0.930$). The average number of visits to the physiotherapist, as recorded in the diaries (5.40; SD 14.47) was significantly different from the average number indicated by the insurance company (0.55; SD 1.88) ($p=0.021$). The overreporting in the cost diary was due to 50% of the patients consulting a physiotherapist, who registered an average of 1 visit per week during the complete year.

Table 2. Comparison between cost-diary data and insurance company data (of mean contact to specialist care and physiotherapy)

	Cost diary Mean (SD)	Insurance company Mean (SD)	P. value
Specialist care	1.25 (2.01)	1.2 (1.49)	0.93
Physiotherapy	5.40 (14.47)	0.55 (1.88)	0.021

6.4 Conclusion and discussion

In order to perform comprehensive cost-effectiveness studies evaluating treatment for chronic musculoskeletal pain from a societal perspective, it is essential to at least partially obtain information from patients when collecting costing data. In this paper we have described the use of a patient diary developed to collect costing data in long-term clinical trials. We have examined the feasibility, data collection period and aspects of construct validity of the patient cost diary using data from two economic evaluation studies, which were the first attempt at implementing this particular tool.

In this study, the response rate was quite acceptable. Eighty-five percent of the patients returned at least one diary-booklet and in total 68% of the diaries were returned. Fifty percent of the patients completed the diaries over the total follow-up period of one year, which corresponds to data summarized in a previous review of health diaries.⁶ Verbrugge concluded that the largest loss of eligible respondents typically occurs at the very beginning of the survey period, and attrition during the diary period itself tends to be quite low. Diary agreement rate was found to correlate with the length of time that respondents were asked to maintain the diary.⁶ Moreover, no relationship was observed between response rate and relevant characteristics of the patient group. This would suggest that the use of a cost diary as a method to collect data does not demand that extra attention need be paid to these factors in order to yield a high response rate. However, since this cost diary was used in a specific group of pain patients, precaution should be taken in generalizing this finding to other patient populations.

We hypothesized that keeping the diary for part of the year would be a representative alternative to keeping the diary for the whole year. To answer this hypothesis it would have been better to look at two different samples. However in the context of an already large randomized controlled trial, this was not possible. We therefore based our analysis on one sample. If the instruction to complete the diary prospectively is well understood, this should not make any difference. In that case the inter-dependence

between two periods is limited. Our analysis showed that in fact, both alternatives (keeping the diary two weeks each two months or three months a year) show no significant different results to keeping the diary for the whole year. Nevertheless, the Spearman correlations between the alternatives showed only moderate agreement. This could furthermore be explained because the study was not powerful enough to show equivalence. For future economic evaluation studies performed from a societal perspective, important advantages can be realized by collecting information for an abbreviated period rather than for the whole trial period. Shorter diaries are less time-consuming for the patient and also less expensive for the researchers in terms of data entry and analysis time, not to mention correspondence costs (mail and telephone charges). However, when collecting data for only an abbreviated period, several other aspects should be taken into careful consideration. Firstly, agreement between the different time periods need not hold if there are time patterns, for example in case of acute diseases, in case of recent coordinated interventions (as in randomized clinical trials) and in case of chronic diseases with seasonal effects. In these circumstances, great care should be taken to avoid bias. Secondly, collecting data for an abbreviated period will in general increase the variance of the estimate. Therefore, requiring less effort per patient must be weighed against including more patients. The smaller standard deviations we found of alternative 2 and 3, especially for the direct health care costs of alternative 2 could not be explained and must therefore be attributed to random error".

In our paper we made an attempt to validate the self-reported data of the diaries. A commonly used strategy to assess the validity of self-reported data is to compare self-reports with data from routine records that are considered to be accurate.²²⁻²⁴ One problem with this approach is that patients own payments are not easily quantified, since these costs can only be obtained directly from the patient and are usually not registered elsewhere. The validity of the other direct health care costs can be obtained by comparing the self-reported information with data from health care providers and insurance companies. However, in clinical trials with a large number of patients (as is the case in our studies), the number of health care providers is large for it to be practical to gather the necessary information. For this reason we decided to compare the diaries with data from the largest sick fund where most of the patients were insured. These data were deemed to be 'true' or actual utilization data since registration of these data is tied to reimbursement of patient costs.

Within the scope of the study, we compared the number of physiotherapy contacts and specialist care contacts noted in the diary, with data available from the compulsory health insurance company for 40 chronic low back pain patients from whom we had data of comparable periods. It was not possible to study resource utilization other than number of contacts, for instance actions performed. For the physiotherapy visits we found that there was enormously higher reporting of contacts in the diary. At first glance these findings might lead one to incorrect assertions concerning the validity of the cost diary, but when examined in detail, the difference could be explained because half of the patients who visited a physiotherapist went on average once per week (during the full year). These visits were not to be found in the data of the sick fund,

which resulted in much lower physiotherapy costs registered in the data of the sick fund. In accordance with the literature it would have been more likely to find under-reporting of health care utilization rather than over-reporting.^{6,21} However, the question arises as to whether we are really dealing with a case of over-reporting here. It is quite possible that these were visits to a private therapist not registered by the insurance company. Another possibility is that the utilization reported by the patient as being a physiotherapist, was not physiotherapy in the 'strict' definition, but only recognized by the patient as such, for example exercise therapy or fitness therapy provided in a sports center. A different kind of explanation is that the authorization for a certain number of physiotherapy visits was given before the patients entered into the trial and that the patients in question used the authorization only after the trial period. In this case, the physiotherapy visits would be registered in the cost diary but would not be traceable in the data of the insurance company (on the same date). In consideration of these issues, it is likely that the instruction of the diary was not specific enough to distinguish between physiotherapy visits and other related therapy contacts. Further development of the cost diary should take this into account, and one corrective action might be to ask if the patient had personally paid for the physiotherapy visit or had a referral from the general practitioner. Nevertheless, these explanations may still not totally explain the enormous difference between two data sources.

For specialist care visits, it was easier to compare the two methods because it was possible to match the type and exact date of the visit. No significant difference was found between the data reported by the patient and the data in the records of the insurance company. Presuming that the reporting of specialist care visits is no different from the reporting of other health care utilization which can be specified in the same detail, we can conclude that the cost-diary, provided that certain precautions are taken, may be a valid instrument to measure direct medical costs. Although similar results on validity of patient recall data were found in patients with rheumatoid arthritis, we have to bear in mind regarding our conclusions that the diary was used in a setting with specific pain patients. Since this only represents a first attempt to validate the self-reported information, additional research is needed.²²⁻²⁴ It was not our purpose to design one standardized diary applicable to a wide range of diseases. In our opinion, this will be impossible, because of the large variation in health care utilization that exists between different diseases, for example asthma versus schizophrenia. However, the content of the diary can partly be standardized for related diseases, such as different kinds of musculoskeletal diseases. The layout and the instruction of the cost diary can furthermore be adapted and standardized regardless of the disease that is studied.

In our studies, we found that precautions must be taken when relying on patients to record cost data. Inaccuracies and incomplete responses were minimized by asking for specific, detailed information during the interview. Furthermore the interview enabled us to gather information on travel expenses and additional costs for health care visits. These items were missing or not clear enough in the current cost diary. For example, in our diary we requested information regarding visits to alternative medicine. Since it was necessary to translate the data into costs and these costs may vary significantly from one respondent to the next, we should have explicitly asked for expenses. In

further development of the diary these aspects should be taken into account. Effort should be made to describe the requested information in a very detailed manner in order to prevent misunderstandings and incomplete reporting. Taking these supplementations into account the diary method might be successfully adopted as an instrument to answer cost-effectiveness questions in long-term clinical trials.

Acknowledgements; We gratefully acknowledge Silvia Evers and Steven Linton and for their valuable comments on earlier drafts of this paper and thank Jolanda van Haastregt, Gerard van Breukelen, Ingrid Nooyen and Jan van Emmerink for their statistical advice. We also thank Eric Debougnoux (CZ-ziektekostenverzekering) for providing data and Mieke Zemmeling for taking care of a part of the data entry. The study was supported by a grant from the investigative medicine program of the Dutch Health Insurance Executive Board.

References

1. Smith, K., Wright, K. Informal care and economic appraisal. A discussion of possible methodological approaches. *Health Economics* 1994; 3: 137-148.
2. Gold, M.R., Siegel, J.E., Russel, L.B., Weinstein, M.C. *Cost-effectiveness in health and medicine*. New York/Oxford: Oxford University Press, 1996.
3. Drummond, M.F, O'Brien, B., Stoddart, G.L., Torrance, G.W. *Methods for the economic evaluation of health care programmes*. Second edition. Oxford: Oxford University Press, 1997.
4. Freer, C.B. Self-care: A health diary study. *Med Care* 1980; 18 (8): 853-861.
5. Freer, C.B. Health diaries: a method of collecting health information. *J Royal Col Gen Prac* 1980; 30; 279-282
6. Verbrugge, L.M. Health diaries. *Med Care* 1980; 18 (1): 73-95.
7. Carp, F.M., Carp, A. The validity, reliability and generalizability of diary data. *Experimental Aging Research* 1981; 7 (3): 281-296.
8. Norman, G.F., McFarlane, A.H., Streiner, D.L., Neale, K. Health diaries: strategies for compliance and relation to other measures. *Med Care* 1982; 20 (6); 623-629.
9. Kooiker, S.E. Measuring the occurrence of common symptoms of health. Why do health diaries and retrospective questionnaires lead to different results? (Submitted).
10. Carsjo, K., Thorslund, M., Warneryd, B. The validity of survey data on utilization of health and social services among the very old. *J Gerontology: Social Sciences* 1994; 49 (3): S156-S164.
11. Green, S., Kaufert, J., Corkhill, R, Creese, A., Dunt, D. The collection of service utilisation data: a research note on validity. *Soc Sci & Med* 1979; 13A: 231-234.
12. Cleary, P.D., Jette, A.M. The validity of self-reported physician utilization measures. *Med Care* 1984; 22 (9): 796-803.
13. Vlaeyen, J.W.S., Teeken-Gruben, N.J.C., Goossens, M.E.J.B., Rutten van Molken, M.P.M.H. Cognitive-educational treatment of fibromyalgia: A randomized clinical trial. I Clinical effects. *J Rheumatol* 1996; 23: 1237-1245.
14. Goossens, M.E.J.B., Rutten- van Mólken, M.P.M.H., Leidl, R.M., Bos, S.G.P.M., Vlaeyen, J.W.S., Teeken-Gruben, N.J.G. Cognitive-educational treatment of fibromyalgia: a randomised clinical trial II. Economic Evaluation. *J Rheumatol* 1996; 23 (6); 1246-1254.

15. Goossens, M.E.J.B., Rutten-van Mólken, M.P.M.H., Kole Snijders, A.M.J., Vlaeyen JWS, van Breukelen, G., Leidl, L. Health economic assessment of behavioural rehabilitation in chronic low back pain; A Randomised clinical trial. *Health Economics* 1998; 7: 39-51.
16. Kole-Snijders, AMJ, Vlaeyen, JWS, Goossens, MEJB, Rutten van Mólken, MPMH, Heuts, PHTG, van Breukelen, G, van Eek, H. Chronic low back pain: what does cognitive coping skills training add to operant behavioral treatment? Results of a randomized clinical trial. *Journal of consulting and clinical psychology* (accepted).
17. Koopmanschap, M.A., Rutten, F.F.H. Indirect costs in economic studies; Confronting the confusion. *Pharmacoeconomics* 1993; 4(6): 446-454.
18. OECD, Statistics Directorate. *Historical Statistics 1960-1994*. Paris: OECD; 1996.
19. Roghmann, K.J., Haggerty, R.J. The diary as a research instrument in the study of health and illness behavior; Experiences with a random sample of young families. *Med Care* 1972; 10 (2): 143-163.
20. Liang, M.H., Larson, M., Thompson, M. Costs and outcomes in rheumatoid arthritis and osteoarthritis. *Arthritis Rheum* 1984; 27 (5): 522-529.
21. Glandon, G.L., Counte, M.A., Tancredi, D. An analysis of physician utilization by elderly persons: systematic differences between self-report and archival information. *J Gerontology* 1992; 47 (5): S245-S252.
22. Meenan, R.F., Yelin, E.H., Henke, C.J. The costs of rheumatoid arthritis. A patient-oriented study of chronic disease costs. *Arthritis Rheum* 1978; 21 (7): 827-833.
23. Lubeck, D.B., Spitz, P.W., Fries, J.F. A multicenter study of annual health service utilization and costs in rheumatoid arthritis. *Arthritis Rheum* 1986; 29 (4): 488-493.
24. Thompson, M.S., Read, J.L. The cost-effectiveness of auranofin: results of a randomized clinical trial. *J Rheumatol* 1988; 15 (1): 35-42

7 PATIENT UTILITIES IN CHRONIC MUSCULOSKELETAL PAIN: HOW USEFUL IS THE STANDARD GAMBLE METHOD?

This chapter is published as:

Mariëlle E.J.B. Goossens¹, Johan W.S. Vlaeyen^{1,2}, Maureen P.M.H. Rutten- van Mölken³, Sijf M.J.P. van der Linden⁴. Patient utilities in chronic musculoskeletal pain: How useful is the Standard Gamble method? *Pain* 1999; 80: 365-375.

¹ Institute for Rehabilitation Research, Hoensbroek

² Department of Medical, Clinical and Experimental Psychology, Maastricht University

³ Institute for Medical Technology Assessment, Erasmus University Rotterdam

⁴ Department of Internal Medicine/Division of Rheumatology, University Hospital Maastricht

Summary

The main goal of current pain management approaches is to increase the patients' quality of life by improving pain coping skills and by reducing the levels of disability in daily life, often despite persistent pain. Direct measurement of quality of life is of crucial importance in economic evaluation research, in which not only is the estimation of financial costs and benefits included, but so is the evaluation of costs and benefits in terms of changes in health states. The purpose of this study is to compare the psychometric qualities of two instruments for assessing patients' utilities, the rating scale (RS) and the standard gamble (SG). Such instruments are designed for their application in economic evaluation research, but have seldomly been used in chronic pain trials. Both methods provide a single measure between 0 and 1. The relationship between these utility measures and descriptive and domain-specific quality of life measures was examined in 133 fibromyalgia patients and 148 patients with chronic non-specific low back pain. Mean utility score at baseline was .43 with the RS and .78 for the SG. The correlation between both methods was found to be poor ($r=0.21$). Both measures appeared to be fairly stable in a two-week test-retest period ($ICC=.74$ and $.77$). Scores on the description of patient's own health on 6 domains, global assessment of change and domain specific measures correlated moderately with the RS scores and low with the SG. Multiple regression analyses demonstrated that 32% of the variance in RS values and only 13% of the variance in SG utilities could be explained by domain-specific measures. The results suggest acceptable construct validity for the RS but insufficient construct validity for the SG. Valuations of own health appear only partially to be related to the assessment of the pain-specific measures and measures of distress. It can be concluded that the RS and domain-specific measures assess partly different, but nevertheless complementary aspects of health-related quality of life. It is therefore recommended to include in economic evaluation studies both domain-specific measures and valuation measures. Finally, in chronic musculoskeletal pain patients, RS scores were found to be more responsive in detecting significant changes in preferences than SG scores. For use in patients with chronic musculoskeletal pain, the RS is preferred to the SG for establishing accurate decisions about the impact of new interventions on their health outcomes.

7.1 Introduction

Health related quality of life is conceptually an important outcome measure when comparing alternative programs for the treatment of chronic disorders. Especially in chronic musculoskeletal pain, treatment effect often means improvement of the patient's physical function, social function and psychological well-being, rather than pain reduction. A wide array of measurement instruments have been used within the musculoskeletal pain field to assess quality of life. These measures can be divided into two major classes: specific instruments and generic instruments and they differ from each other in their measurement properties, range of application and practicability. Specific instruments focus on aspects of health status that are relevant to chronic musculoskeletal pain patients, and specific to the area of primary interest for the program, such as pain, pain coping, pain beliefs, pain control, and emotional distress. Since they focus on aspects specific to a domain or group of patients these instruments are highly sensitive in detecting small but important changes. For clinicians and researchers whose main goal is to determine the impact of the program on how patients feel and function, specific instruments are of major importance. Examples of specific instruments are McGill Pain Questionnaire¹, Coping Strategies Questionnaire², Pain Cognitions List³ and Fibromyalgia Impact Questionnaire.⁴ In spite of their specificity, the disadvantage of these measures is that they can not be used to compare treatment effect across conditions or programs. When the applicability of treatment effects among different domains of health, populations and interventions is needed, the second class of the generic instruments are eligible.

Generic instruments are designed to sample the complete spectrum of function, disability and distress that are relevant to quality of life and aggregate them into a single summary score. Apart from their broad applicability, they may be able to pick up unexpected beneficial and adverse effects that can easily be missed by specific measures. Limitations of these instruments are that they may not be responsive to small but clinically important differences between treatments for specific conditions. To generic instruments belong health profiles and utility measures. Health profiles such the MOS 36-item short-form health survey (SF-36)⁵ and the Nottingham Health Profile (NHP)⁶ are specially aimed at measuring and describing (changes in) quality of life.⁷ Utilities differ from health profiles in that they reflect how a patient values a state of health and not just how he or she describes the characteristics of the health state. Knowledge of these aspects makes them useful in health care decision making and cost-effectiveness analysis. The utility approach assigns numerical values (utilities or preferences) to a health state on a very generalizable 0 (anchored as death) to 1 (anchored as complete health) scale. Utilities are described as numbers that represent the strength of an individual's preference for a particular health state under uncertainty. Preferences are the values people assign to different health outcomes when uncertainty is not a condition of measurement.⁸ Three techniques have been extensively used to elicit preferences/utilities; the rating scale (RS), the time-trade-off (TTO) and the standard gamble (SG). RS, TTO and SG are methods to measure preferences for health

states on a scale from 0 (death) to 1 (full health). The RS typically asks respondents to place health states on a line with the clearly defined endpoints. The TTO presents the respondent with the task of determining what amount of time they would be willing to give up to be in a better versus a poorer health state.⁹ The SG is the classical method of measuring utilities and is based directly on the axioms of the expected utility theory of Von Neumann and Morgenstern.¹⁰ Although this theory prescribed how a rational individual 'ought' to make decisions when faced with uncertain outcomes, it is not based on any psychological theory, but purely economical. The SG is in fact the direct application of one of these axioms. The SG technique begins with asking the respondent to consider a hypothetical choice between certainty of continued life in the health state of interest and a gamble with two possible outcomes; full health (probability p) and death (probability $1-p$). The probability p is varied until the respondent is indifferent between the two alternatives. The indifference point resembles the utility of the health state. The greater the risk of death the subject is willing to consider, the lower the indifference probability and thus the lower the utility of the health state of interest. For a more detailed description of the SG method and the other methods we refer to Gold *et al.*⁹ and Drummond *et al.*¹¹

Although considerable differences between the methods have been observed^{12,13}, the typical pattern is that the SG yields the highest scores, followed by the TTO and the RS. Not only does the literature report on differences in mean scores between methods, values for certain health states also seem to differ in their relationship to health status measures. In their systematic review, Revicki and Kaplan¹³ summarized selected evidence pertaining to the relationship between health status and utility measures. In general, their review suggests that health status scales are poorly to moderately correlated with SG and TTO scores, and are more closely associated with RS preferences.

In contrast to the rather ease of choosing the specific instruments best fitting the study objectives, the above-mentioned differences between utility methods indicate that it is not easy to decide what measure will be most appropriate. To date, no consensus has been reached on which method should be used. SG is strongly preferred by economists, because it is the only measure with a theoretical foundation and therefore the only instrument measuring 'true' utilities. However, this does not make it obvious to assume that the SG method is most suitable to apply in chronic musculoskeletal pain. It would be most favorable to incorporate the utility measure which is the most reliable and responsive and which correlates most strongly with the concepts of the other health status measures. Within chronic musculoskeletal pain utility measures have not been applied extensively. In their review, Revicki and Kaplan¹³ reported on 3 studies dealing with rheumatoid arthritis¹⁴⁻¹⁶, using both health status and preference-based instruments. Liang *et al.*¹⁴ and Kaplan *et al.*¹⁵ used the Quality of Well-Being Questionnaire (QWB)¹⁷ to measure preferences. More recently, Bakker *et al.*¹⁸⁻²⁰ applied the Maastricht Utility Measurement Questionnaire (MUMQ) to patients with ankylosing spondylitis and fibromyalgia. The MUMQ consists of the description of health related quality of life by six dimensions of health, and the valuation of health status with SG and RS methods. For both patient groups, RS values correlated better

with health status measures and disease activity measures than SG utilities. As far as we know, there are no other studies that examined the relationship between health status and utility in chronic musculoskeletal pain patients.

The purpose of this paper is threefold. Firstly, in a sample of fibromyalgia and chronic low back pain patients we want to examine the psychometric properties of the MUMQ (consisting of health description, SG and RS) including reproducibility, validity and responsiveness. Secondly, we want to examine the extent of agreement between the RS and the SG. Finally, we address the question whether these utility measures can be satisfactorily used in chronic musculoskeletal pain patients to interpret treatment results in terms of cost-effectiveness.

7.2 Patients and methods

7.2.1 Subjects

Data were obtained in two randomized clinical trials on cognitive behavioral rehabilitation for fibromyalgia and chronic low back pain respectively. The sample included 281 patients who were all referred for behavioral rehabilitation; 133 patients with fibromyalgia according to the ACR criteria²¹ and 148 chronic pain patients with chronic non-specific low back pain, according to Spitzer *et al.*²² In the fibromyalgia trial, patients were randomly assigned to an educational program (including information about ergonomics, legislation, psycho-social factors affecting pain) plus cognitive coping skills training, an educational program plus attention-control, and a waiting list control group. In the chronic low back pain study, patients were assigned to operant behavioral treatment plus cognitive-coping skills training, an operant-behavioral treatment plus attention-control or a waiting list control group. The cognitive coping skills training (including applied relaxation, attention diversion, imagery, self-task) aimed at decreasing distorted pain cognitions and increase perceived control over pain and its consequences.²⁷ The operant-behavioral treatment (including graded activity, activity pacing, spouse training) is aimed at increasing health behaviors and activity levels, and at decreasing pain behaviors and excess disability.^{28,29} Of the total sample 75% were female, the mean age was 42 years, and the average duration of pain since the onset was 10 years. Seventy-four percent had lower levels of education (i.e. less than 10 years). There were no significant differences on these characteristics between the two conditions.

7.2.2 Measures

Socio-demographics. Prior to the treatment provided in the trial, a general questionnaire was administered which included the following variables: gender, age, duration of pain, education, sick pension, and marital state.

Utilities. The MUMQ was used to elicit utility values from the patients. The instrument is a translated and slightly adapted version of the McMaster Health Utility Index.³⁰ Patients were first asked to describe their current health status using a five-point scale for six domains: activities of daily living, self-care, emotions, leisure activities, pain

and other complaints, and side effects of treatments. The dimensions of health are constructed out of five ordinal descriptions where '1' indicates the best level and '5' the worst level. The combination of the levels indicated by the patient, one for each domain was used to define the patient's own health state. The actual valuation of their health state was performed by means of RS and SG techniques. In the MUMQ, the RS is a visual analog scale ranging from perfect health on top (showing the value 100) to death on the bottom of the scale (showing 0). The patients are asked to place their health status (as indicated earlier) on the scale, visualized as a thermometer. The value placed on the thermometer directly equals the utility. The RS is an interval scale on the response level. This means that we assume that each individual difference between for example 40 and 50 equals the difference between 60 and 70. The RS scores are transformed into utilities by dividing them by 100. As opposed to the RS method where risk is not included in the measurement process, the SG technique is directly based on the Von Neumann-Morgenstern utility theory and is valued under the assumption of risk.¹⁰ In the SG, patients are offered a choice of remaining in their current health state for the rest of their life or undergoing a treatment with a chance p of perfect health or a chance of $1-p$ of dying immediately after the treatment (see figure 1).

At follow-up, patients were asked to globally rate their change in health status on a three-point scale (improved, deteriorated, or remained the same compared to the baseline situation). Then, a final SG question was asked, which provided the relative change score of the patient. To enable the comparison a patient was confronted with his prior description of his baseline health state (conceptualization of the own health state by the six dimensions). If the health state of the patient had improved, the patient was required to choose between two alternatives: a) remaining in the own health state for the rest of his or her life, or b) undergoing a treatment with a chance of p to return to perfect health or a chance of $1-p$ to relapse into his or her own health state as it was at baseline. If the patient's current health was worse, then the current health state would replace 'death'. The patient then had to choose between return to own health state as it was at baseline or undergoing a treatment with a chance of p to return to perfect health or a chance of $1-p$ to remain in his or her own current health state. Of course, if the health state of the patient at follow-up was equal to the health state at baseline, the last SG question is irrelevant. In the SG questions, the probability of treatment success is varied in steps of 10% until the patient is indifferent between remaining in his or her current health state or undergoing the risky treatment. At this indifference point, the utility equals p . To visualize the term 'probability', the SG method is performed with a probability wheel as a prop, as shown in figure 1.

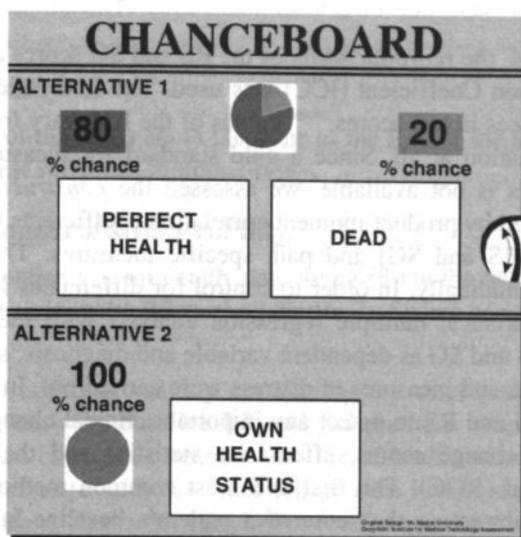


Figure 1. Chanceboard*

*The Chanceboard is translated and slightly adapted version of the McMaster Health Utility Index of Bennett *et al.*²⁶ (with permission).

Pain-related variables. *Pain intensity:* The Pain Rating Index of the Dutch version of the McGill Pain Questionnaire (MPQ-DV)^{1,31} was included as a measure of current pain intensity. *Pain behavior:* The Pain Behavior Scale (PBS)^{32,33} is a widely used observation scale for the quantification of observable pain behaviors. The total score was included in this study. *Activity tolerance:* the Behavioral Approach Test (BAT)³⁴, during which a pain-eliciting stimulus is placed in a standardized environment, and the patient is instructed to engage with a set of standardized physical activities as long as possible. In this study the walking distance in meters was used. *Pain cognitions:* The Pain Cognition List (PCL-e)³ is a 77-item questionnaire aimed at the assessment of distorted pain cognitions and experienced self-control. For this study the scales 'Pain Impact' (PCL-im) and 'Catastrophizing' (PCL-cat) were used. *Pain control:* A Dutch version of the Coping Strategies Questionnaire (CSQ)^{2,35} measures different ways of coping with pain, such as diverting attention, ignoring pain, praying, positive self-talk, increasing activities, relaxation and pain control. For this study, the scale 'pain control' was used.

Distress. *Depression:* In this study, the Dutch translation of the Beck Depression Inventory (BDI)³⁶ was selected. This version is a 21-item self-report questionnaire designed to measure the severity of depression and was suggested by the Dutch Committee for the Standardization of Depression Questionnaires.³⁷ *Fear:* The total score of the Dutch version of the Fear Survey Schedule (FSS-III-R)³⁸ was used. The FSS-III-R is a 76-item questionnaire consisting of clusters of phobic complaints including social phobia, agoraphobia, fear of bodily injury, illness and death, fear of sex and aggression and fear of living organisms.

7.2.3 Analysis

For the assessment of the *reproducibility* of the RS and SG scores after two weeks, the Intra Class Correlation Coefficient (ICC) was used. The ICC reflects both systematic and random differences in test scores.^{39,40} Values of the ICC vary from -1 to +1, with 0 indicating no association at all. Since a gold standard for measuring preferences in chronic pain patients is not available, we assessed the *construct validity*. Construct validity was examined by product-moment correlation coefficients between utilities (as measured with the RS and SG) and pain specific measures. This was done cross-sectionally and longitudinally. In order to control for differences in diagnostic groups and patient characteristics, multiple regression analysis with hierarchical backward elimination, with RS and SG as dependent variable and diagnosis, socio-demographics, pain-related variables and measures of distress were carried out. In order to investigate the ability of the SG and RS to detect any important clinical changes, three strategies were used: use of change scores, effect size statistics and the receiver operating characteristic method (ROC). The first and most common method of demonstrating *responsiveness* is subtracting the score of a patient's baseline health state from the score of a patient's follow-up health.⁴⁰ Change scores were tested by use of paired t-test statistics. Relative change in SG was calculated by combination of the patient's global assessment of change and scores on the last SG question. The utility received from this SG question combined with the utility of a patient's baseline health state resulted in a relative change score*. Deyo *et al.*⁴⁰ proposed the use of effect size statistics for comparing responsiveness. Effect size is calculated by dividing the within-patient mean change score by the standard deviation of the mean change score of the stable patients.³⁹ Global assessment of change was used as an external criterion to distinguish between changed and stable patients. The third method of demonstrating responsiveness is the receiver Operating Characteristic Method (ROC). The ROC curve synthesizes information on the sensitivity and specificity for detecting improvement by an external criterion.^{41,42} The area under the ROC curve can be interpreted as the probability of correctly identifying the improved from the non-improved patients.⁴¹ The area ranges from 0.5 (no accuracy in discriminating improved from unimproved) to 1.0 (perfect accuracy). Also here the global assessment of change was used as an external criterion. Since ROC analysis requires the external criteria to be dichotomous, the three-point scale was collapsed to a dichotomous variable (improved/non-improved: stable and deteriorated).

* If a patient indicates an improvement, the relative change score is calculated as follows: The utility at follow-up is calculated as $U_f = p + (1-p)U_b$, where U_b is the known utility at baseline. Then the relative change score is calculated as $U_f - U_b$. In case of a perceived deterioration in health status, the relative change score is calculated as follows: The utility at baseline is calculated as $U_b = p + (1-p)U_f$, where U_f is the known utility at follow-up. Accordingly, the relative change score is calculated as $U_f - U_b$.

7.3 Results

7.3.1 Reproducibility

The two-week ICC of the SG is about the same as the ICC of the RS; 0.77 compared to 0.74, but both suggest sufficient reproducibility.

7.3.2 Cross-sectional construct validity

Median utilities of patients' own health state assessed via the SG method were higher than the values assessed via the RS method (figure 2). Scores between the two methods differed by least 25%.

Cross-sectional and construct validity correlations are displayed in Tables 1 and 2. Pearson correlation between baseline RS and SG methods was low (0.21; $p=0.05$). Since the six domains of health constitute the core of health-related quality of life, we expected them to correlate with the valuation of own health state, composed of the same constructs. Table 1 shows the correlation coefficients between the RS scores and the six domains of health, ranging from -0.034 to -0.43 (with only two domains being less than 0.30). The SG correlations were even poorer, all being less than 0.25. For both methods, the correlations were strongest with the health domains 'activities of daily living' and 'pain'.

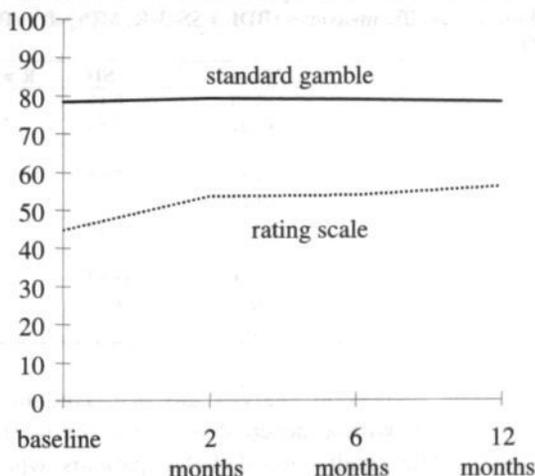


Figure 2. Mean scores on rating scale and standard gamble at baseline, 2 months, 6 months and 12 months.
(Standard gamble scores were multiplied by 100)

Table 1. Pearson correlations at baseline between the six domains and the RS and SG

6 domains of health [#]	RS	SG
Activities of daily living	-.434**	-.161*
Self-care	-.257**	-.122*
Emotions	-.394**	-.143*
Leisure activities	-.379**	-.138*
Pain and other complaints	-.360**	-.214**
Side effects of treatment	-.034	-.013

** = $p < 0.01$, * = $p < 0.05$

[#] Contradictory direction with rating scale and standard gamble scores (lower score indicates better health)

Cross-sectional construct validity correlations are displayed in table 2. Pearson product-moment correlation coefficients ranged between 0.03 and 0.33 for the SG scores, with only one of 9 correlation coefficients being higher than 0.25. Correlation coefficients for the RS were somewhat higher and ranged between 0.14 and 0.40, with 5 of 9 correlation coefficients being higher than 0.25. The RS correlated statistically significantly with all the selected pain-related variables and measures of distress. The highest correlation was found between RS and the measures of depression, pain impact, and catastrophizing, and the lowest between SG and fear, activity tolerance and pain behavior.

Table 2. Means, standard deviations (SD) and pearson correlation coefficients (R) among SG and RS and domain-specific measures (BDI, FSS-3-R, MPQ-DV, PBS, BAT, PCL-im, PCL-cat, CSQ).

Variables	Mean	SD	R with RS	R with SG
Standard gamble (SG)	0.78	0.17	.21	--
Rating Scale (RS)	42.94	13.83	--	.21
Activity tolerance (BAT)	294.84	176.27	.27	.03
Depression (BDI)	13.57	7.51	-.40	-.22
Coping (CSQ)	9.14	7.09	.15	.08
Fear (FSS-3-R)	120.49	38.26	-.20	-.05
Pain intensity (MPQ)	25.47	12.67	-.25	-.33
Pain behavior (PBS)	16.48	95.62	-.14	-.09
Catastrophizing (PCL-cat)	48.9	11.5	-.36	-.17
Pain impact (PCL-im)	50.61	9.15	-.40	-.21

* $p < .05$, ** $p < .01$, *** $p < .001$

The results of the multiple regression analysis are displayed in table 3. Variation inflation factors (VIFs) were within acceptable range (between 1.05 and 1.50), suggesting no collinearity. The results revealed that patients who valued their own health state higher on the RS, were more active, less depressed, and less disabled by the pain. Also patients who were single tended to rate their own health state as better. A model with six independent variables explained 32% of the variance in the RS scores. Only 13% of the variation in SG utilities could be explained. Moreover, only one of the selected variables significantly predicted SG utilities. Patients showing less pain intensity had higher SG scores. Socio-demographics such as age, gender, education and sick pension did not explain the variation in RS and SG utility scores.

Table 3 Multiple regression analysis of RS and SG scores.[#]

Independent variables	RS			SG		
	Adj R	R ²	Beta	Adj R ²	R ²	Beta
Activity tolerance (BAT)	.30	.32	.29	.14	.13	
Depression (BDI)			-.23			-.14 (.051)
Pain behavior (PBS)			-.13			
Coping (CSQ)			.14			
Pain impact (PCL-im)			-.26			
Marital status			-.26			
Pain intensity (MPQ)						-.28

p < .05, *p < .01, ***p < .001,

[#]Backward elimination method

7.3.3 Longitudinal construct validity.

Longitudinal construct validity correlation coefficients for RS and SG change scores are given in table 4. Of all the domain-specific instruments the change in activity tolerance correlated best with the two preference-based instruments. Again the correlations are higher with the RS change scores. These scores correlated strongly with depression and pain impact.

Table 4. Longitudinal construct validity (12-month follow-up) of RS and SG scores (Pearson correlation coefficients)

Variables	RS	SG
Rating scale	-	.18*
Standard gamble	.18*	-
Activity Tolerance (BAT)	.38***	.18*
Depression (BDI)	-.38***	-.11
Coping (CSQ)	.18*	.14
Fear (FSS-3-R)	.04	.04
Pain (MPQ)	-.11	.02
Pain behavior (PBS)	.08	.02
Catastrophizing (PCL-cat)	-.06	-.002
Pain Impact (PCL-im)	-.29**	-.15

p < .05, **p < .01, ***p < .001

In the regression analysis using RS score as dependent variable a model of four variables explained 30% of the change. Also here there was no matter of collinearity (VIFs range between 1.0 and 1.01). Patients with increased activity tolerance, less depression, less fear, and less pain impact (table 5) showed higher RS changes, indicating more improvement in health-related quality of life. Only 11% of the variance in SG scores could be explained. Women and patients with higher activity levels and less pain impact showed more improvement on the SG.

Table 5. Multiple regression analysis of change in RS and SG scores.[#]

Independent variables	RS			SG		
	Adj R	R ²	Beta	Adj R ²	R ²	Beta
	.27	.30		.09	.11	
Activity Tolerance (BAT)			.33***			.20*
Depression (BDI)			-.25**			
Fear (FSS-III-R)			.18*			
Pain impact (PCL-im)			-.21*			-.17*
Gender						.18*

p < .05, **p < .01, ***p < .001

[#]backward elimination method

7.3.4 Responsiveness

Absolute change score. Figure 2 provides information on the mean RS and SG scores at the different measurement points. Graphical inspection reveals an improvement between baseline and 2, 6 and 12 months, when measured with the RS method. The improvement measured on the RS was statistically significant at all periods (paired t-statistics; $p < 0.001$ at 2 months ($t=8,26$), 6 months ($t=7,31$) and 12 months ($t=8,54$)). Significant change was not found for the SG. Baseline scores for the SG and RS are shown in table 2.

Absolute versus relative change SG scores. Two different methods were used to calculate change in SG utility due to treatment. The first is the conventional method of subtracting the utility of a patient's baseline health state from the utility of a patient's follow-up health state. The second method is to calculate the change using the last SG question. The utility received from this SG question combined with the utility of a patient's baseline health state results in a relative change score. Table 6 compares the responsiveness of both methods separately for patients who indicated an improvement and for patients who indicated a deterioration when asked directly. By self-assessment, 43% (at 2 months and 39% and 40% at 6 and 12 months respectively) of the patients rated themselves as improved, 15% (22% and 19% respectively) as deteriorated and 42% (38% and 41% respectively) as not changed. Table 6 shows that for both patients who indicated an improvement and patients who indicated a deterioration, the relative change score is significantly higher than the absolute change score. This indicates that when patients are asked directly to value improvements or deteriorations compared to baseline and when patients are informed about their baseline health state, the responsiveness to small changes improved. Table 6 also shows a decrease in standard deviations in relative change scores, indicating that the variations between patients diminish when measured in a relative compared to an absolute manner.

Table 6. Comparing absolute with relative (SG) change scores of the patient's own health state.

	N	Absolute change SG		Relative change SG	
		Mean (SD)		Mean (SD)	
Improved 0-2 months	94	0.031 (0.145)		0.128 (0.105) [*]	
Deteriorated 0-2 months	33	-0.026 (0.185)		-0.146 (0.117) [*]	
Improved 0-12 months	53	0.0005 (0.150)		0.134 (0.127) [*]	
Deteriorated 0-12 months	25	-0.072 (0.132)		-0.173 (0.143) [*]	

^{*}p < 0.001 (Wilcoxon matched pairs signed rank test)

Effect size statistics. Patients' global assessment of change was also used as external criterion to calculate effect sizes. The mean change score in SG and utility measures differed between patients who classified themselves as improved and as non-improved (table 7). The effect size statistic in both methods was higher in the improved group compared to those in the non-improved group. By comparing the effect size of the RS and the SG, it is evident that the RS shows more change than the SG. This means that if patients indicated themselves as improved, the degree of improvement detected by the SG was small, indicating low sensitivity to change.

Table 7 Responsiveness; baseline scores, mean changes (12 months), standard deviations (SD), effect size statistics and area under ROC-curve in the improved, non-improved and deteriorated patients

Instrument (min - max)	N	Baseline score (SD)	Mean changes (SD)	Effect size [*]	Area under ROC-curve
<i>RS (0-100)</i>					0.79
Entire sample	135	44.6 (14.2)	11.9 (16.7)	-	
Improved (12 mo)	54	65.8 (15.1)	20.9 (16.9)	1.71	
Non-improved (12 mo)	55	52.6 (17.9)	6.9 (12.2)	0.57	
Deteriorated (12 mo)	26	43.1 (14.6)	1.5 (12.6)	0.12	
<i>SG (0-1)</i>					0.51
Entire sample	135	0.78 (0.16)	-0.007 (0.14)	-	
Improved (12 mo)	54	0.80 (0.16)	0.0005 (0.15)	0.004	
Non-improved (12 mo)	55	0.78 (0.17)	0.0071 (0.12)	0.06	
Deteriorated (12 mo)	26	0.73 (0.17)	-0.072 (0.13)	-0.62	

^{*}Effect size is calculated as mean change score divided by the standard deviation of the mean change score of the non-improved patients

ROC-curves. Figure 3 shows the ROC plots of the RS and the SG, using global assessment of change as the external criterion. The best discrimination occurs with a curve that reaches the upper left corner. For the RS the curve was to the left above the diagonal, showing moderate discriminative ability. The curve for the SG was close and nearly parallel to the diagonal, indicating very low discriminative ability. This was also shown by the area under the ROC curve, which was only 0.51 for the SG (table 7).

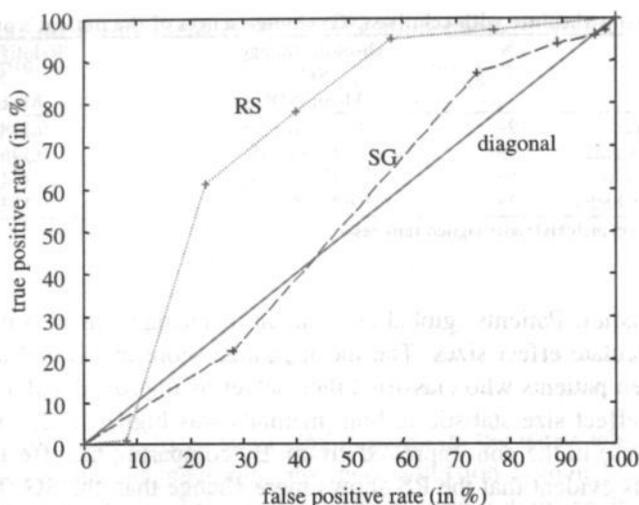


Figure 3. ROC curves of the change scores for SG and RS

7.4 Conclusion and discussion

In this paper a psychometric analysis of the RS and SG as contained in the MUMQ was performed, by assessing its reliability, validity and responsiveness. The sample population in the study were 281 chronic musculoskeletal pain patients, including fibromyalgia and chronic non-specific low back pain patients.

Two-week test-retest reproducibility of the RS and SG method was satisfactory and comparable to other studies.⁹ The correlation between the two methods was poor ($r=0.21$), whereas the SG method resulted in substantially higher scores than the RS method. These findings are similar to a previous study with fibromyalgia patients, which used the SG and RS in a similar context²⁰ and other studies dealing with other chronic conditions, such as asthma^{43,44}, chronic obstructive lung disease⁴⁵, chronic renal disease^{46,47}, ankylosing spondylitis¹⁹ and chronic stable angina.⁴⁸

In terms of validity, the RS appeared to be superior to the SG. The relationship between the six domains of health and the health state utilities were in the expected direction. The highest correlation was found between the RS and the domains of health. The pattern of associations between the domain-specific variables and the RS scores were in the expected direction. The highest associations were found for depression, impact of pain on daily life, catastrophizing, activity tolerance and pain intensity. A different pattern was found for the SG method which shared most variance with pain intensity. When change scores in these variables are considered, a slightly different pattern of association emerged for the RS. Changes in depression, activity tolerance, and pain impact were significantly related to changes in RS utilities. Only changes in activity tolerance appeared to be related to changes in SG utilities. The results indicate a higher (cross-sectional and longitudinal) construct validity for the RS,

which is consistent with other studies.^{8,19,20,43,44,45} Also in line with other studies, in which the variance in valuations measures was explained by a combination of health status measures^{8,44,46,48} the explanatory power of the regression equation for the RS (cross-sectional 32% and longitudinal 30%) was higher than for the SG (cross sectional 13%, longitudinal 11%). Overall, this suggests that, even though a large part of the variation could still not be explained, for RS at least one third of the variance was explained by identifiable health state attributes. These findings indicate that RS preferences are more closely related to other measures of health status than are the SG values. RS and health status instruments are both based on similar measurement methods and are the result of direct evaluation of health.⁸ However, although these two types of instruments may be measuring a common construct, the relationship is too weak to make them interchangeable.⁴⁶

The lack of responsiveness of the SG is likely to be associated with the high baseline utilities of 0.95, leaving very little room for improvement. There are a number of possible explanations for this. First of all, in chronic musculoskeletal pain, the possible risk of dying after treatment, as brought to the patient's attention in the SG question, is rather hypothetical and far from realistic for this patient-group. Twenty-eight percent of the chronic musculoskeletal pain patients reached the same indifference probability of 0.95, on all measurements, although the dimensions of health changed. Apparently, these patients were never willing to take a risk of dying larger than 10%, irrespective of their own health state. It might be that the description of the levels of the health state loses some value when the risk-element is brought into prominence. Methods that involve uncertainty and risk may make sense in clinical applications, where patients may be at risk due to condition (for example cancer) or treatment (for example surgery)⁹, but not in situations as described in this paper, where risk of dying is not a very realistic option. Secondly, the disease history of chronic musculoskeletal pain patients might have reinforced their risk-aversity. In an average period of 10 years of persistent pain they have tried many different therapies, usually without much success. Further, after long periods of pain, patients may have adapted themselves to the situation of living with chronic pain. A treatment with a chance of dying would therefore not be a realistic option for these chronic patients. The last two explanations suggest that SG scores might not only be affected by attitudes regarding risk taking, but that patients' willingness to take risk also partly depends on how SG is presented. It is possible that the alteration of the probability in steps of 10% (100/0, 90/10 etc) may have been too large. On the other hand, smaller probabilities may be difficult for people to handle.⁴⁹ It still is unknown how much smaller these changes should be and whether smaller changes would give better results. Finally, the presence of SG turned out to be more responsive (and more closely related to RS scores) when patients were directly asked to value improvements or deteriorations compared to baseline. An explanation for this finding might be that this situation is more direct and less imaginary. People are directly confronted with an improvement or deterioration in their health state and probably take this change in health state into account when answering the SG-question.

Some investigators, have interpreted the weak correlation between the SG and health status measures as supportive for validity of the instrument, because in their opinion it indicates that the utility measure is differentiating adequately between health status and health state preference measures.⁵⁰ However, several arguments concerning this adequacy of differentiation between these types of measures makes this interpretation questionable. To start with, these investigators refer to the discriminant validity of the SG, indicating a logically weaker agreement with measures of different attributes. Determining discriminant validity however requires specifying the constructs that account for the variance in the measures as well as the directions and if possible the strength of the expected relationship.⁷ Many assumptions have been made regarding the possible constructs that influence the SG scores, such as attitude to risk, time preference, cognitive task and other variables related to patient circumstances.¹³ So far, these assumptions have never been tested. The possibility that the discrepancy between SG and health utility measures is generated by these non-health variables should not be neglected. What exactly are these (health) factors? Should they be incorporated into an overall valuation of a treatment outcome, and if so, how? When it is unclear what constructs exactly explain the SG scores (and how), perhaps the most parsimonious explanation is that the SG represents not more than just 'measurement error'. Besides that, if SG does not capture most relevant aspects of a condition and non-health factors would explain most of SG scores (in our sample between 87% and 89%) is it then still fair to presume that what the SG values is actually a health state?

In sum, health-related quality of life measures and utility-based methods are important in evaluating interventions for chronic musculoskeletal pain. This study shows that domain-specific (pain-related and distress) measures and valuation methods assess different but complementary and related aspects of health related quality of life. Overall, the rather high correlation between the RS, different domain specific measures and quality of life dimensions, the higher responsiveness to detect change, and the ease of administering the RS would favour the use of a RS to evaluate treatment of chronic musculoskeletal pain. On the other hand, the RS method has a weaker theoretical foundation than the SG method, and it has therefore been argued by many economists and decision scientists that only the SG measures true preferences. However, the poor validity and the lack of responsiveness of SG utilities in chronic musculoskeletal pain patients is troublesome. The use of an insensitive measure like the SG may lead to inaccurate decisions about the impact of new interventions on health outcomes. As long as the determinants of SG scores are not fully understood, the use of the SG method in a setting with chronic musculoskeletal pain should not be recommended.

Although we found complementary and related aspects between pain specific measures and RS, it should be beared in mind that these measures were originally designed for different purposes and have therefore different measurement properties. Specific measures contain more accurate descriptions, whereas utility measures are based on patients' preferences and allow for cross-program comparisons. Dependent on the perspectives of the studies undertaken, (e.g. physician's and patient's perspective, and health care policy perspective) conclusions can be focused on particular instruments. In order to collect health-related quality of life data in clinical trials for economic

evaluation purposes, it seems advisable not to rely on just one type of quality of life measure. Both, pain specific and generic utility instruments, with satisfactory psychometric characteristics to best suit the study objectives should be included.

Acknowledgements: The authors thank Geert Crombez and an anonymous reviewer for their valuable comments on an earlier draft of this paper and Jolanda van Haastregt and Riekje de Vet for their statistical advice. The study was supported by grants from the investigative medicine program of the Health Insurance Executive Board (nr OG 91/088) and from the Dutch Praevention Fund (nr 28-2055).

References

1. Melzack, R. *Pain measurement and Assessment*. New York: Raven Press, 1983: 33-37.
2. Rosenstiel, A.K., Keefe, F.J. The use of coping strategies in chronic low back pain patients: Relationship to patient characteristics and current adjustment. *Pain* 1983; 17: 38-44.
3. Vlaeyen, J.W.S., Geurts, S.M., van Eek, H., Kole-Snijders, A.M.J., Schuerman, J.A., Groenman, N.H., What do chronic pain patients think of their pain? Towards a pain cognition questionnaire. *British J of Clinical Psychology* 1990b; 29: 95-118.
4. Burchardt, C., Clark, S., Bennett, R., The fibromyalgia impact questionnaire; Development and validation. *J Rheumatol* 1991; 18: 728-733.
5. Ware, J.E., Sherbourne, C.D. The MOS 36-item short-form health status survey: 1. Conceptual framework and item selection. *Medical Care* 1992; 30 (6): 473-481.
6. Kind, P., Carr-Hill, R. The Nottingham health profile: A useful tool for epidemiologists? *Soc Sci Med* 1987; 25: 905-910.
7. Patrick, D.L., Erickson, P. *Health status and health policy. Quality of life in health care evaluation and resource allocation*. New York Oxford: Oxford University Press, 1993.
8. Revicki, D.A., Wu, A.W., Murray, M.I. Change in clinical status, health status and health utility outcomes in HIV-infected patients. *Med Care* 1995; 33: AS173-AS182.
9. Gold, M.R., Patrick, D.L. Torrance, G.W., Fryback, D.G., Hadorn, D.C., Kamlet, M.S., Daniels, N., Weinstein, M.C. Identifying and valuing outcomes. In: M.R. Gold, J.E. Siegel, L.B. Russel, M.C. Weinstein, editors. *Cost-effectiveness in Health and Medicine*. New York Oxford: Oxford University press, 1996:82-134.
10. Neumann, Von J., Morgenstern, O. *Theory of games and economic behavior*. Princeton University Press, 1947.
11. Drummond, M.F., O'Brien, B., Stoddart, G.L. Torrance G.W. *Methods for the economic evaluation of health care programmes*. Second edition. Oxford, Oxford University Press, 1997.
12. Nord, E. Methods for quality adjustment of life years. *Soc Sci Med* 1992; 34: 559-569.
13. Revicki, D.A., Kaplan, R.M. Relationship between psychometric and utility-based approaches to the measurement of health-related quality of life. *Quality of Life Research* 1993; 2: 477-487.
14. Liang, M.H., Larson, M.G., Cullen, K.E., Schwartz, J.A. Comparative measurement efficacy and sensitivity of five health status instruments for arthritis research. *Arthritis and Rheumatism* 1985; 28: 542-574.
15. Kaplan, R.M., Anderson, J., Wu, A., Mathews, W.C., Korin, F.J., Orenstein, D. The quality of well-being scale; applications in AIDS, cystic fibrosis, and arthritis. *Med Care* 1989; 27: S27-S43.

16. Bombardier, C., Ware J., Russell I., Larson, M., Chalmers, A., Read, J.L. Auranofin therapy and quality of life in patients with rheumatoid arthritis: results of a multicenter trial. *Am J Med* 1986; 8: 683-690.
17. Kaplan, R.M., Bush, J.W. Health related quality of life measurement for evaluation research and policy analysis. *Health Psychology* 1982; 1: 61.
18. Bakker, C.H., Rutten-van Mólken, M.P.M.H., van Doorslaer, E.K.A., Bennet, K., van der Linden, S. Patients' utilities in ankylosing spondylitis and the association with other outcome measures. *J of Rheumatol* 1994; 21: 1298-1304.
19. Bakker, C., Rutten-van Molken, M., Hidding, A., Doorslaer, E. van, Bennett, K., Linden, S. van der. Feasibility of utility assessment by rating scale and standard gamble in patients with ankylosing spondylitis or fibromyalgia. *J. of Rheumatol.*, 21 (1994) 269-274.
20. Bakker, C.H., Rutten-van Mólken, M.P.M.H., van Santen-Hoeufft, M., Bolwijn, P., van Doorslaer, E.K.A., Bennet, K., van der Linden, S. Patients' utilities in fibromyalgia and the association with other outcome measures. *J. of Rheumatol.*, 22 (8) (1995) 1536-1543.
21. Wolfe, F., Smythe, H.A., Yunus, M.B., Bennett, R.M., Bombardier, C., Goldenberg, D.L., Tugwell, P., Campbell, S.M., Abeles, M., Clark, P., Fam, O.G., Farber, S.J., Fiechtner, J.J., Franklin, C.M., Gatter, R.A., Hamaty, D., Lessard, J., Lichtbroun, A.S., Masi, A.T., McGain, G.A., Reynolds, W.J., Romano, T.J., Russel, I.J. and Sheon, R.P. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. *Arthritis Rheum* 1990; 33: 160-172.
22. Spitzer, W.O., LeBlanc, F.E., Dupuis, M. Scientific approach to the assessment and management of activity-related spinal disorders: report of the Quebec Task Force on spinal disorders. *Spine* 1987; 12 (suppl. 7): S1-S59.
23. Vlaeyen, J.W.S., Teeken-Gruben, N.J.C., Goossens, M.E.J.B., Rutten-van Molken, M.P.M.H., Cognitive-educational treatment of fibromyalgia: A randomized clinical trial. I Clinical effects. *J Rheumatol* 1996; 23 (6): 1237-1245.
24. Goossens, M.E.J.B., Rutten-van Mólken, M.P.M.H., Leidl, R.M., Bos, S.G.P.M., Vlaeyen, J.W.S., Teeken-Gruben, N.J.G. Cognitive-educational treatment of fibromyalgia: a randomised clinical trial II. Economic Evaluation. *J Rheumatol* 1996; 23 (6): 1246-1254.
25. Goossens, M.E.J.B., Rutten-van Mólken, M.P.M.H., Kole-Snijders, A.M.J., Vlaeyen, J.W.S., van Breukelen, G., Leidl, L. Health economic assessment of behavioural rehabilitation in chronic low back pain; A Randomised clinical trial. *Health Econ* 1998; 7: 39-51.
26. Kole-Snijders, A.M.J., Vlaeyen, J.W.S., Goossens, M.E.J.B., Rutten-van Mólken, M.P.M.H., Heuts, P.H.T.G. Chronic low back pain: what does cognitive treatment add to operant treatment? Results of a randomized clinical trial. *Journal of consulting and clinical psychology* (accepted).
27. Turk, D.C., Meichembaum, D. and Genest, M. *Pain and behavioral medicine, a cognitive-behavioral approach*. New York: Guilford Press, 1983.
28. Fordyce, W.E. *Behavioral methods for chronic pain and illness*. Mosby, St. Louis, 1976
29. Roberts, A.H. The operant approach to the management of pain and excess disability. In: A.D. Holzman and D.C. Turk, editors. *Pain management. A handbook of psychological treatment approaches*. New York: Pergamon Press, 1986.
30. Bennett, K., Torrance, G.,R., Tugwell, P. Methodologic challenges in the development of utility measures of health-related quality of life in rheumatoid arthritis. *Controlled Clinical Trials* 1991; 12: S118-S128.
31. Kloot, W.A. van der & Vertommen, H. *De MPQ-DLV: een standaard Nederlandstalige versie van de McGill Pain Questionnaire. Achtergronden en handleiding*. Lisse: Swets & Zeitlinger, 1989.

32. Richards, J.S., Nepomuceno, C., Riles, M. & Suer, Z. Assessing pain behavior: the UAB pain behavior scale. *Pain* 1982; 12: 393-398.
33. Vlaeyen, J.W.S., Pernot, D., Snijders, A.M.J., Eek H. van, Schuerman, J.A., Groenman, N.H. Betrouwbaarheid en validiteit van een Nederlandse versie van de Pain Behavior Scale (PBS). [Reliability and validity of a Dutch version of the Pain Behavior Scale] *Nederlands Tijdschrift voor Psychologie* 1990a; 45: 184-189.
34. Kole-Snijders, A.M.J., Vlaeyen, J.W.S., Eek, H. van, Schuerman, J.A. and Groenman, N.H. Behavioral assessment of residual health behavior in chronic low back pain patients through the pain-rest contingency principle. Paper presented at the IASP VIth World Congress. *Pain* 1990; 5: S394.
35. Spinhoven, P., Kuile, M.M. ter, Linssen, A.C.G., Ganzendam, B. Pain coping strategies in a Dutch population of chronic low back pain patients. *Pain* 1989;37: 77-83.
36. Beck, A., Rush, A.J., Shaw, B.P., Emery, G. *Cognitive therapy of depression*. New York: Guilford Press, 1979.
37. Zitman, F.G., Griez, E.J.L., Hooijer, Chr. Standaardisering depressievragenlijsten. [Standardizing depression questionnaires]. *Tijdschrift voor psychiatrie* 1989;31: 114-135.
38. Wolpe, J., Lang, P.L. A fear Survey Schedule for use in behaviour therapy. *Behav Research Ther* 1964; 2: 27-30.
39. Guyatt, G., Walter, S., Norman, G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chron Dis* 1987; 40(2): 171-178.
40. Deyo, R.A., Diehr, P., Patrick, D.L. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Controlled Clin Trials* 1991; 12(suppl): 142-158.
41. Deyo, R.A., Centor, R.M. Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. *J Chron Dis* 1986; 39: 897-906.
42. Hanley, J.A., McNeil, B.J. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983; 148: 839-43.
43. Rutten-van Mólken, M.P.M.H., Custers, F., Van Doorslaer E.K.A., Jansen M.C.M., Heurman, L., Maesen F.P.V., Smeets, S.J., Bommer, A.M., Raaijmakers, J.A.M. Comparison of performance of four instruments in evaluating the effects of salmeterol on asthma quality of life. *Eur. Res. Journal* 1995; 8 (6): 888-898.
44. Blumenschein, K., Johannesson, M. Relationship between quality of life instruments, health state utilities, and willingness to pay in patients with asthma. *Ann Allergy Asthma Immunol* 1998; 80: 189-94.
45. O'Brien, B., Viramontes, J.L. Willingness to pay: a valid and reliable measure of health state preference? *Med Dec Making* 1994; 16: 289-297.
46. Revicki, D.A. Relationship between health utility and psychometric health status measures. *Med. Care* 1992; 30: MS274-MS282.
47. Hornberger, J.C., Redelmeier, D.A., Petersen, J. Variability among methods to assess patients' well-being and the consequent effect on a cost-effectiveness analysis. *J Clin Epidemiol* 1992; 5: 505-512.
48. Nease, R.F., Kneeland, T., O'Connor, G.T., Sumner, W., Lumpkins, C., Shaw, L., Pryor, D., Sox, H.C. Variation in patient utilities for outcomes of the management of chronic stable angina. Implications for clinical practice guidelines. *JAMA* 1995; 273: 1185-1190.
49. Kahneman D. Tversky A. Prospect theory. An analysis of decision under risk. *Econometrica* 1979; 47:263.

50. Bartman, B.A., Rosen, M.J., Bradham, D.D., Weissman, J., Hochberg M., Revicki D.A. Relationship between health status and utility measures in older claudicants. *Quality of Life Research* 1998; 7: 67-73.

CONCLUSION, GENERAL DISCUSSION AND RECOMMENDATIONS

8

In the first part of the study, the effect of the concentration of the solution on the rate of the reaction was investigated. It was found that the rate of the reaction increased with the concentration of the solution. This is because the higher the concentration, the more particles there are in a given volume, and therefore the more collisions between particles occur. This leads to a higher rate of reaction.

The second part of the study investigated the effect of temperature on the rate of the reaction. It was found that the rate of the reaction increased with temperature. This is because the higher the temperature, the more particles have enough energy to overcome the activation energy barrier and undergo a successful collision. This leads to a higher rate of reaction.

The third part of the study investigated the effect of the surface area of the solid reactant on the rate of the reaction. It was found that the rate of the reaction increased with the surface area of the solid reactant. This is because the larger the surface area, the more particles are exposed to the other reactant, and therefore the more collisions occur. This leads to a higher rate of reaction.

Based on the results of the study, the following recommendations are made:

- To increase the rate of the reaction, the concentration of the solution should be increased.
- To increase the rate of the reaction, the temperature should be increased.
- To increase the rate of the reaction, the surface area of the solid reactant should be increased.

8.1 Introduction

This thesis reports on two economic evaluations of cognitive behavioral rehabilitation in chronic musculoskeletal pain and discusses the psychometric analysis of the cost diary and two utility methods applied in the outcome studies. This chapter presents the main findings of the different studies reported here and discusses the principle conclusions. The chapter concludes with some recommendations for policy and research.

8.2 Methodological considerations

In chapter 2 we summarized the most important methodological issues in economic evaluation research and applied these issues to chronic musculoskeletal pain. We showed that in chronic musculoskeletal pain some outcomes are less accessible than they may be in other disorders, such as therapeutic outcomes, and that some outcomes and costs need to be given higher priority, such as quality of life and patient and family costs. It has furthermore to be kept in mind that for the measurement of both costs and outcomes there is no single method or approach. Therefore, we emphasized the importance of the transparent reporting of methods and study results in order to facilitate the interpretation of the results from economic evaluations, and to assess the possibilities to generalize results from one setting to another. Multidisciplinary involvement in the economic evaluation process will contribute to the final interpretation and implementation of the study results.

Chapter 3 reviews economic evaluations of interventions in back pain, and shows that there is an urgent need for sound economic assessment in this area. From the 31 studies published until 1997, there is no firm basis for conclusions about the cost-effectiveness of the interventions. We argued that this is mainly due to the small number of studies of comparable treatments, the large differences in and the low quality of costing methodologies, and the lack of evidence based on statistical analysis. Furthermore, the studies included in the review dealt mainly with acute back pain and back pain related injuries. In chronic low back pain, as well as for any form of pain management, cost-effectiveness studies are seldomly carried out or even totally lacking. Furthermore, there is a complete lack of economic evaluations for fibromyalgia interventions.

The conclusions taken from the review are not surprising. This is partly because, in general, the primary purpose of the studies included in the review was not economic evaluation, but assessment of clinical effectiveness. Economic evaluations were additional and subsidiary components, and therefore the performance of the economic evaluation was placed in the background. Also the complexity of assessing the cost-effectiveness of pain (and in particular chronic pain) partially explains the low priority given to the performance of economic evaluation. Ferrel¹ contended that the subjective nature of pain (and discomfort), the complexity of assessing quality of life, and little historical weight given to costs for carers and families make cost evaluation in pain

treatment a complex undertaking. In our review, too, cost-effectiveness studies were, in general, limited to the measurement of traditional costs and outcomes, such as direct medical costs (program costs) and indirect costs (productivity losses). However, in line with Ferrel¹, we also urge that alongside these more or less tangible costs and outcomes, economic evaluation of the management of pain will only be complete when quality of life and the burdens imposed on the patient and carers are included.

The review indicates the need for good economic evaluation studies in the field of back pain. This implies a role for funding agencies to incorporate economic evaluation as a part of study proposals. Furthermore, these agencies should assist researchers with specific guidelines and promote transparency in the use and reporting of the methodology applied. This would increase the validity and comparability of the evaluations and may facilitate decision making in the prevention and treatment of musculoskeletal pain.

Our studies described in chapters 4 and 5 can be considered as the first cost-effectiveness studies in the field of behavioral medicine. The 'fibromyalgia trial' in chapter 4 examined the additional effectiveness of cognitive coping skills when added to an educational intervention, by comparing it to group discussion (intended as an attention control) added to an educational intervention, or to a waiting-list condition. The cognitive coping skills program was protocolized and organized as 12 outpatient group sessions of 4 hours and lasted 6 weeks in total. The analyses failed to confirm the supplemental benefits of the coping skills training in terms of reductions in direct health care costs, direct non-health care costs, indirect costs and increase in valuations of quality of life. The addition of a group discussion component to the group education resulted in significantly lower direct costs compared to the addition of coping skills training to the group education, whereas the health state valuations showed no significant differences between the two interventions, immediately after treatment and at 6 and 12 months. However, although there was no difference between the interventions, the valuation of quality of life improved significantly in both groups immediately after treatment and after one year. Furthermore patients in the educational-discussion program valued their quality of life significantly higher compared to the waiting-list control group.

The second study examined the cost-effectiveness of the addition of coping skills training to an operant-behavioral program compared to the addition of a group discussion component (intended as attention control) to the operant-behavioral program for chronic low back pain patients (chapter 5). The treatments were protocolized as 5 weeks of inpatient and 3 of outpatient treatment. Both programs were further compared to a waiting list control group and to an individualized operant treatment, as usually given in the rehabilitation center. The intervention groups did not lead to differences in direct health care costs, direct non-health care costs, indirect costs and quality of life valuations between the interventions. Nevertheless, just as in the fibromyalgia trial (chapter 4), this study concluded that immediately after treatment the interventions resulted in significantly higher quality of life valuations (measured by the rating scale method) as compared to the waiting-list control. The quality of life values in the intervention groups were significantly higher during the one-year follow-

up compared to pre-treatment. Compared to the statistical significance, further analysis on the clinical significance showed that almost no patients in the usual treatment group improved immediately after treatment or at one-year follow-up.² These findings suggest that the comprehensive and interdisciplinary nature of the treatments, rather than the specific components that are added to it, may be the important factors for improvement.

The expected supplemental efficiency of coping skills training when added to a multidisciplinary operant-behavioral rehabilitation program was not found in either trial. Vlaeyen *et al.*³ and Kole-Snijders *et al.*² gave several explanations for the fact that the addition of coping skills training was less effective than expected. First, the cognitive coping skills training may not have been specific enough to the individual needs of the patient. Second, the treatment requires great effort and self-discipline from the patients and this may have been too difficult given the relatively low educational level of the patients. Third, patients were not sufficiently compliant to do their homework assignments as planned. The effects of the cognitive therapy can only be expected when patients practice at home on a regular basis. A final reason might be the relatively small part (only 13%) the coping-skills training plays in the total treatment.

Besides these explanations for the small contrast between the treatment interventions, there are some other factors which should be considered when interpreting the results from both trials. To start with, the *choice of comparator* may have influenced the relevance of both economic evaluations to policy makers. In the fibromyalgia and chronic low back pain trial the supplemental efficiency of coping skills training was compared to a group discussion component. This group discussion was originally added as an *attention control*, to control for nonspecific effects of the coping skills training. However, in both trials the addition of the group discussion component showed equal (in chronic low back pain) or better (in fibromyalgia) results when compared to the addition of the coping skills training. This indicates that the group discussion may have served as a new, active treatment and can no longer be seen as an attention control: the addition of coping skills training was compared to the addition of another untested component. Alternatively, one can conclude that the addition of a cognitive coping skills training was not effective and differences with the waiting-list control and differences in time may be due to the constant treatment factor (present in all groups): the operant behavioral treatment (for chronic low back pain) and the group education (for fibromyalgia). Our design did not allow for direct assessment of the pure effect of the coping skills training, nor the pure effect of the operant, educational or the discussion program could be assessed.

Using an *attention control* in chronic pain outcome studies may be questionable. Its use has long been a subject of debate in psychotherapy and behavioral research^{4,6} because in this type of research it is difficult to tell whether the condition merely serves as an attention control. Provider and patient factors may also influence this process. In order to enhance treatment credibility and blinding for patients, both the attention control condition and the cognitive coping skills training were provided by the same behavioral therapists. However, a drawback to this is that there may have been

contamination bias between these conditions. This has been tested in both trials showing that contamination seems to be absent in the fibromyalgia trial and only some degree of contamination cannot totally be excluded in the chronic low back pain trial. For example, the group discussion component of the attention control might have worked as a form of exposure and perhaps influenced outcomes (e.g. fear reduction).²

Furthermore, in both trials, we compared the interventions to a *waiting-list control* group and concluded that the interventions resulted in short-term (quality of life) improvements. We, however, doubt the reliability and usefulness of the waiting-list controls with regard to the cost analysis. For ethical reasons it was not appropriate or feasible to keep the patients on the waiting list for a longer period. Patients stayed only up to 10 weeks on that list after which they were offered active treatment as well. This short period prevented us from estimating costs for controls. The completed cost diaries showed that in the waiting period, the patients received minimal or no care in anticipation of the promised treatment. Therefore, this data does not give us a reliable picture of the chronic musculoskeletal pain patient in average daily practice. Including these findings into the cost analyses may lead to an underestimation of the costs and may consequently bias the calculation of a cost-effectiveness ratio in favor of no rehabilitation treatment.

To overcome the problem concerning the choice of comparator, we agree with Parloff's⁵ and Schwartz *et al.*⁷ advice that future behavioral studies (to estimate both effectiveness as well as cost-effectiveness) should compare two alternative interventions rather than comparing an experimental therapy with an attention-control or a waiting-list control. Alternatively, in our trials the interventions could have been compared to the standard practice, or the 'best-available' alternative. However, considering the enormous variety of treatments for chronic musculoskeletal pain and little proven effectiveness, it is difficult to define current practice or best alternative. Neither could 'usual care' be derived from assessing the patterns of health care use in the waiting-list controls of both studies. As mentioned previously, these patients stayed only for a maximum of 10 weeks on the list, in which period few sought other interventions. In the chronic low back pain trial, a partial solution to the problem came from constructing a 'usual rehabilitation care' control for patients initially on the waiting-list. After completion of the waiting period these patients received a reduced form of operant-behavioral program as usually given in the rehabilitation center. However, the results of the comparison with 'usual' treatment may only be relevant from the perspective of the rehabilitation center under study, and may not provide a useful societal perspective on the problem.

In this context, comparison with observational studies might be worthwhile to overcome the lack of a best-available alternative. These studies might enable the real world estimation of costs and outcomes and could be used to measure total cost savings. Furthermore, incorporating economic evaluation in studies evaluating chronic musculoskeletal pain treatment at different stages of the disorder would enhance the possibility of life-time calculations of costs of medical consumption. This could also be valuable when modeling the cost-effectiveness of interventions for chronic musculoskeletal pain.

Apart from the choice of comparator, the second methodological constraint in both studies relates to the *generalizability* of the findings. Both economic evaluations were conducted alongside randomized controlled clinical trials (RCTs) which may have limited the *generalizability* of the findings. Notwithstanding its high internal validity, results that have emerged from RCT-based economic evaluation studies have often been criticized for having a limited scope and being unrepresentative for the entire patient population.⁸ Also it can be questioned if our conclusions can be generalized to other situations with chronic musculoskeletal pain patients. First of all, in our studies patients were treated in a rehabilitation center, which is not the usual treatment setting for these patients. Treatment in the rehabilitation setting may also have influenced the total costs of the programs. In the chronic low back pain trial for example, the direct medical costs were largely driven by the costs of the 5 week inpatient treatment. This is indeed a debatable aspect of the treatment. Most chronic low back pain patients do not need 24 hours of inpatient care. Outpatient treatment for the majority of the patients would reduce the costs and might turn out to be more cost-effective.

Furthermore, patients participating in both trials were not *representative* of the total population of chronic musculoskeletal pain patients. The patients had been in pain for on average 10 years, had several psychological complaints, such as depression, anxiety and phobia, and had undergone several treatments, all without lasting improvements. Many of these patients had no other medical intervention options still open to them. Therefore, the results from our RCTs may not provide solutions to a 'real world' question. The findings might have been different in a population less disabled due to chronic musculoskeletal pain.

In summary, several factors, related to the choice of comparator and the study design, decreased the generalizability of the results of both studies within a broader context of chronic musculoskeletal pain. Two other factors which merit consideration in interpreting the study results are the measurement of the costs and the quality of life valuations (chapters 6 and 7).

A *patient cost diary* was developed for both studies to measure direct health care costs, direct non-health care costs and indirect costs. In chapter 6, the usefulness of the patient cost diary is evaluated for its feasibility, aspects of data collection and construct validity. With certain additions regarding content, the analysis led to the conclusion that the proposed cost diary might be a useful instrument for the assessment of health care resource use and patient-related costs. Feasibility of the patient cost diary was generally satisfactory and response was not related to certain patient characteristics. We tested the length of the period of data collection and found no significant differences in total costs between diaries kept for approximately a quarter of a year and a whole year. For long term clinical trials this may indicate that health care utilization and patient costs may not have to be collected during the entire trial period, requiring less time from the patient and lowering research costs. Because the analysis for the period of data collection was done on one sample only, a more comprehensive comparison of two different samples with different periods of data collection may be necessary to substantiate this finding. Finally, construct validation of the self-reported information turned out to be difficult. Patients co-payments are not registered

elsewhere and some utilization data cannot be obtained at a patient level. Content validity was only studied by comparing specialist care and physiotherapy data recorded in the diary with that obtained via an insurance company. For specialist care data the two databases revealed no significant differences. In contrast, the number of physiotherapy contacts was higher in the cost diaries. Reasons for this may be the following. First, the activities involving physical activity (such as fitness exercises and aerobics) may have been misrecorded as physiotherapy. Alternatively, insurance company data may have been underreported, due to a backlog in declarations. This has been tested, showing no reason for underreporting of insurance company data. Second, our analysis was only based on a small sample of patients. During the period of data collection there was a change in the reimbursement system, which forced us to concentrate on the patients who had completed one year of follow-up before this change came into action (end of 1993). Furthermore, complete insurance data that fully matched the period of the diary was only available for 40 patients. To strengthen the findings on the construct validity, a more extensive analysis with more variables (reported in extensive detail) and a larger sample would be necessary. Further studies are needed to establish the diary's applicability in economic evaluation studies for other settings.

The patient cost diary was specially designed for our studies because a gold standard to estimate use of health care resources, patient and family costs and indirect costs was lacking. The most up-to-date guidebooks^{9,10} agree that there are no routine records for resource use and that these costs should be obtained through self-reporting channels, such as surveys, telephone interviews or diaries. However, they do not provide guidance as to the most valid instruments to use or how they should be applied. In practice, patient and family costs are rarely included in economic evaluation studies (as was shown in chapter 3) and even when these costs are included, a clear description of how to measure is usually lacking. Future economic evaluation studies dealing with chronic disorders, particularly those that seek to address patient and societal perspectives, should incorporate these costs into the analysis and should report transparently on the methodology used.

The question arises as to whether it is possible to accept the patient cost diary as a standard to measure direct health care resources used, indirect costs and patient and family costs in clinical trials. For the cost diary a core set of questions, a standard design and layout can serve as a template for application in different diseases. The variables in the cost diary need to be adjusted to suit specific situations, because for example sufferers of different diseases vary in their pattern of health care utilization.

Chapter 7, which deals with patient *utilities* in chronic musculoskeletal pain, illustrates another dilemma in economic evaluation, namely lack of consensus in the measurement of utilities. In chapters 4 and 5, utilities have been assessed using the rating scale (RS) and standard gamble (SG) methods. In chapter 7, the psychometric qualities of the RS and SG are assessed by analyzing the reliability, validity and responsiveness of both instruments. The study shows that the RS is to be preferred above the SG for establishing accurate decisions about the impact of interventions on the health state of chronic musculoskeletal pain patients. Regression analyses reveal

that in chronic musculoskeletal pain the preferences measured with the RS method are more closely related to other domain specific measures than the SG values are. RS and domain-specific measures assess partially different, but nevertheless complementary, aspects of health-related quality of life. Therefore, in the evaluation of cognitive behavioral interventions for chronic musculoskeletal pain, utility instruments and pain specific measures should be used to complement each other. Contributing to the unresponsiveness of the SG method as observed in our studies may be the unrealistic assumptions in the SG questions, the long disease history of chronic musculoskeletal pain patients and the indirect presentation of the SG method. Regardless of the many assumptions made about the possible (non-health) constructs that influence the SG scores, clarity on this issue is still lacking. If SG does not capture most relevant aspects of a condition and if non-health related factors explain most of the SG scores, then in our opinion it may not be correct to presume that SG actually values health-related quality of life. Future research into the different concepts that influence the SG scores is needed. The development of methods to measure risk attitude may be a step in this direction. At this moment standardization of one specific utility instrument for *all* cost-effectiveness analyses may therefore not yet be recommended.

8.3 Consequences for different parties

In chapter 1 of this thesis the process of Medical Technology Assessment (MTA) is fitted into a conceptual framework, which shows that after the evaluation of effectiveness and cost-effectiveness the next step is implementation of study results into practice.

In this chapter we have pointed out that some methodological dilemmas have hindered efforts to reach firm conclusions concerning national policy implications about cost-effectiveness of cognitive-behavioral rehabilitation for chronic low back pain and fibromyalgia. Both trials failed to show the expected supplemental efficiency of the coping skills training when added to a multidisciplinary operant-behavioral rehabilitation program. It seemed that the comprehensive and interdisciplinary nature of the treatment are more important than the inclusion of cognitive skills training or group discussion. However, while policy implications must remain the reference point for economic evaluations, to enhance the 'real world' relevance of our studies, these policy implications must be accompanied with implications for the different parties involved. Although the phase of implementation is beyond the scope of this thesis, we have tried to briefly formulate the implications of our findings for a variety of perspectives: patients, providers, rehabilitation center, and finally the insurance companies.

From the viewpoint of the *patients*, those patients included in trials may not be representative for the entire population of patients with chronic musculoskeletal pain. Nevertheless, for the specific group of patients in *these* trials, the interdisciplinary treatments resulted in significant improvements in quality of life. Better quality of life (the primary effect measure in the economic evaluation) may not be the only achievement. Studies on the (clinical) effectiveness on the same samples^{2,3} have also

indicated significant improvements in domain-specific measures. Patients show higher activity tolerance and improved mood, and seem to have better control of pain and to suffer less from their pain, as can be observed from decreased pain behavior. Because the control group was only followed up for 10 weeks, we lack data to confirm that the difference in improvements in favor of the intervention groups was still significant after one year. Nevertheless, it is very unlikely that the control group would have improved as well. In these studies we dealt with moderately to severely disabled and long-term (on average 10 years) chronic patients for whom 'spontaneous' improvements are very unlikely. Furthermore, the cost analysis shed some light on the economic impact of the pain on the individual patient. Patient and family costs, including over-the-counter medication, home help, health activities, and other out of pocket costs, accounted for about 69% of the total health care expenditures by the patients.

For the provider and the rehabilitation center, the study findings show that shorter, but more intensive and protocolized group programs result in equivalent effects as compared to the usual individualized rehabilitation program. First, this indicates the importance of continuing to work with protocols. For the rehabilitation center the use of protocolized treatments gives more transparency in the construction of the treatments, the hours involved for each discipline and the total costs. In future, this transparency may facilitate therapy planning and may contribute to the most efficient distribution of the available resources. Second, the specific improvements highlight the importance of integrating those treatment ingredients which were not yet provided in the usual rehabilitation treatment, such as frequent interdisciplinary meetings, group programs, additional individual counseling, and patient shared responsibility in the process of improvement.^{2,3} For the rehabilitation center it is also worth reconsidering the necessity of inpatient treatment for the chronic low back pain patients. A subset of patients may not necessarily need 24 hours of inpatient care as is currently usual practice in the rehabilitation center.¹¹ In future this may result in fewer hospitalizations and more rehabilitation day treatments for a selection of the patients and consequently in cost savings.

For the purchasing authorities, such as the health insurance companies, the results of the studies are also valuable. At the moment insurers are often hesitant to pay for multidisciplinary treatment of pain, an unwillingness which, according to Frederico,¹² is influenced by factors such as the lack of familiarity with chronic pain, skepticism about positive benefits of multidisciplinary treatment of pain, the diversity of disciplines involved in this treatment and the inconsistency between centers. In our studies, the establishment of treatment protocols gives more insight into the composition of the costs of programs, such as the different disciplines involved and the number of hours spent per discipline. Second, our studies inject more transparency into the (positive) effects resulting from multidisciplinary pain treatment, and provide more understanding of complexity, the possible areas for improvement, and the needs of the severe chronic musculoskeletal pain patient. Third, as mentioned in the previous part, an increase in the number of patients treated in an outpatient basis and a decrease in the number of inpatient treatments may result in a reduction of the mean costs per insured

patient. Taking these implications into consideration may provide payers with an incentive to reimburse chronic pain treatment. Nevertheless, we are aware that for insurers a 'before-and-after' comparison between health care utilization of different interventions would have been of interest. However, because this was not incorporated into our study design, such as change in health care utilization due to the interventions could not be analyzed. The inclusion of a 'before-and-after' comparison would have been a valuable supplement to our data.

8.4 Recommendations for health policy

We emphasize that the implications of the studies for the various parties involved might call for a change in current policy and require specific recommendations with regard to rehabilitation programs for this group of chronic musculoskeletal pain patients. To implement these changes and implementations as efficiently as possible the following activities are of crucial importance.

Adjustment of the national guidelines for the management of chronic low back pain and fibromyalgia. For low back pain, it might be beneficial to update to the national guidelines for the management of low back pain, published in 1996 by the Dutch College of General Practitioners (NHG).¹³ For chronic low back pain, only in case of high health care use and increase of dysfunctioning do the guidelines recommend referral to a multidisciplinary team in a rehabilitation center. The content of this treatment is further left up to the providers' own interpretation. In this regard the NHG-standard could advise a multidisciplinary treatment with a minimum of operant treatment components, such as determining patient treatment goals, team meetings, spouse training, group discussions and individual counseling.² In addition, it will be beneficial to develop specific guidelines on the management of non-specific chronic low back pain for a multidisciplinary team in a rehabilitation setting. Special attention should be paid to the psychosocial treatment principles. With regard to the treatment of fibromyalgia, in the Netherlands no official standard exists yet. More systematic research would be beneficial both in underpinning current practice and so that official guidelines can be developed.

Dissemination of the study results to the different actors in the field. Insurers must become aware that in chronic pain cure is not an outcome. Attention needs to be directed towards psychosocial outcomes. The medical community needs to learn the language of payers by demonstrating cost-effectiveness. We advise insurers to focus not simply on treatment outcomes involving monetary aspects. Currently, insurers tend to consider 'return to work' the most important criterion for treatment success. In general 'return to work' is rather easy to measure and can be translated into savings. For chronic musculoskeletal pain successful 'return to work' can be influenced by numerous factors.¹⁴ For patients who are off work for more than 1-2 years, return to work will often not be a realistic outcome: 'they are virtually unemployable, irrespective of the physical state of their back or further health care'.¹⁵ Even when return to work is still an option, the previous job may have changed or may not be available anymore. Furthermore, finding a new job may be difficult, because an

employer may be unwilling to hire someone with a history of back pain.¹⁴ These factors can bias the final treatment result, when return to work is seen as the only primary effect measure. For multidisciplinary pain treatment, main effects, such as pain coping, pain behavior and domain-specific quality of life can hardly be expressed into monetary units.

Inform patients and the public of treatment options. Evidence-based information on treatment options (linked to the national guidelines) might be provided to the patient through an educational booklet. This might supplement verbal information from providers, prevent unfamiliarity and uncertainty about the variety of different types of care available and might reduce demand of health care use. In the Netherlands, a short pamphlet linked to the NHG-guideline has been developed, which serves as a self-management tool for acute back pain. More thorough patient information materials based on current evidence have been produced in the US,¹⁶ the UK¹⁷ and New Zealand.⁸ We are currently developing such a booklet for the Netherlands, in cooperation with a multidisciplinary team, including general practitioners, physiotherapists, rehabilitation specialists, clinical psychologists, company doctors, medical advisers of insurance companies and health scientists.

8.5 Recommendations for future research

In this thesis we hoped to clarify the application of the principles of cost-effectiveness studies to chronic musculoskeletal pain. We are fully aware that with the different studies addressed in the thesis new or related questions have come up. In the next and final section we give our views on specific issues most worthy for further research in this field.

First, further research regarding cost-effectiveness should match the specific needs of the different actors involved (for example patient satisfaction for the patients, improvement in clinical effects for providers and change in medical consumption for insurers) and should incorporate these different needs into the research questions. To address all these viewpoints, cooperation between a multidisciplinary team of experts in the area in question is important in any cost-effectiveness study from its outset.

Second, the accepted principles of economic evaluation should be followed when designing cost-effectiveness studies. Only properly carried out and reported economic evaluations that include all relevant costs and outcomes allow the study results to be useful in informing policy makers. This indicates that in addition to the clinical effectiveness, the cost-effectiveness needs to be included as a main objective.

The third recommendation involves the choice of alternatives in economic evaluation studies. To estimate the cost-effectiveness, one intervention needs to be compared to another intervention (preferably the best alternative), rather than to an attention-control or a relatively short waiting-list control group. This further leads us to say that observational studies will also serve a purpose to view current practice for chronic musculoskeletal pain.

Fourth, economic evaluation studies in this field need to incorporate patient and family costs and these costs should be obtained through self-reporting, for example by the proposed cost-diary. Since our study is the first evaluation of the cost diary, further studies are required to establish the diary's applicability in cost-effectiveness studies in other settings.

Fifth, we recommend utility measures be used in combination with domain-specific measures, because utility and valuation measures use different measurement concepts compared to domain-specific measures. Further research is needed into the measurement properties of utility and valuation methods applied in chronic musculoskeletal pain.

To conclude, we want to stress that strong recommendations concerning treatment are beyond the scope of this thesis. We nevertheless want to draw attention to the need for further research for multidisciplinary pain treatment. Studies are needed that answer the question what kind of treatment modalities are most effective for chronic musculoskeletal pain patients who have certain characteristics. From a cost-effective point of view, it might be beneficial to develop a multidisciplinary treatment that is targeted at the special needs of the individual patient and is less dependent on patient's compliance and input of time. Furthermore, we may add to the recent recommendations by Waddell¹⁵ and Linton.¹⁹ In their point of view, pain rehabilitation should not only be available as a tertiary service after medical treatment is complete or has failed. Secondary prevention activities may prevent or reduce unnecessary suffering, chronicity and associated health care as well as absenteeism.¹⁹ The knowledge available from our studies can be applied to secondary prevention studies, but also attention should be paid to the cost-effectiveness.

References

1. Ferrell, B.R. How patients and families pay the price of pain. In: M.J.M. Cohen and J.N. Campbell, editors. *Pain treatment centers at a crossroads: A practical and conceptual reappraisal*, Vol. 7, IASP Press, Seattle, WA, 1996: 229-237.
2. Kole-Snijders, A.M.J., Vlaeyen, J.W.S., Goossens, M.E.J.B., Rutten-van Mólken, M.P.M.H., Heuts, P.H.T.G. Chronic low back pain: what does cognitive coping skills training add to operant-behavioral treatment? The results of a randomized clinical trial. *Journal of consulting and clinical psychology* (accepted).
3. Vlaeyen, J.W.S., Teeken-Gruben, N.J.C., Goossens, M.E.J.B., Rutten-van Molken, M.P.M.H. Cognitive-educational treatment of fibromyalgia: A randomized clinical trial. I Clinical effects. *Journal of Rheumatology* 1996; 23: 1237-1245.
4. O'Leary, K.D., Borkovec, T.D. Conceptual, methodological and ethical problems of placebo groups in psychotherapy research. *American Psychologist* 1987; 33: 821-830.
5. Parloff, M.B. Placebo controls in psychotherapy research: A sine qua non or a placebo for research problems? *Journal of Consulting and Clinical Psychology* 1986; 54: 79-87.
6. Strayhorn, J.M. Control groups for psychosocial intervention outcome studies. *American Journal of Psychiatry* 1987; 144 (3): 275-282.

7. Schwartz, C.E., Chesney, M.A., Irvine, M.J., Keefe F.J. The Control groups dilemma in clinical research: Applications for psychosocial and behavioral medicine trials. *Psychosomatic Medicine* 1997; 59: 362-371.
8. Baltussen, R.M.P.M. *Real World Designs in Economic Evaluation. Bridging the gap between clinical research and policy making*. Dissertation. Maastricht: University pers, 1998.
9. Gold, M.R., Torrance P.D.L. Identifying and valuing outcomes. In: Gold, M.R., J.E. Siegel, L.B. Russel, M.C. Weinstein, editors. *Cost-effectiveness in Health and Medicine*. Oxford University Press: New York Oxford, 1996: 82-134.
10. Drummond, M.F., O'Brien, B., Stoddart, G.L., Torrance G.W. *Methods for the economic evaluation of health care programmes. Second edition*. Oxford University Press: Oxford New York Toronto, 1997.
11. Loeser, J.D. *Pain management at the University of Washington*. Satellite conference of the 7th World Congress on Pain, Vancouver, August 23, 1996.
12. Federico, J.V. The cost of pain centers: Where is the return? In: M.J.M. Cohen and J.N. Campell, editors. *Pain treatment centers at a crossroads: A practical and conceptual reappraisal*. Vol. 7, IASP Press, Seattle, WA, 1996: 249-256.
13. Faas, A., Chavannes, A.W., Koes, B.W., van den Hoogen, J.J.M., Mens, J.M.A., Smeel, L.J.M., Romeijnders, A.C.M., van der Laan, J.R. NHG-standaard Lage Rugpijn. [NHG-standard on low back pain.] *Huisarts en Wetenschap* 1996; 39:18-31.
14. Turk, D.C., Efficacy of multidisciplinary pain centers in the treatment of chronic pain. In: M.J.M. Cohen and J.N. Campell, editors. *Pain treatment centers at a crossroads: A practical and conceptual Reappraisal*, Vol. 7, IASP Press, Seattle, WA, 1996, pp. 257-273.
15. Waddell, G. *The back pain revolution*. Edinburgh: Churchill Livingstone , 1998.
16. Back in Action. *Understanding acute low back problems*. US Agency for Health Care Policy and Research, 1994.
17. *The Back Book*. Department of Health, UK, 1996.
18. *Acute Low Back Pain Management. Patient guide*. National Health Committee, New Zealand. 1998.
19. Linton, S.J. The socioeconomic impact of chronic pain: is anyone benefiting? *Pain* 1998; 75: 163-168.

SUMMARY

This thesis deals with the assessment of the economic evaluation of treatment interventions in chronic musculoskeletal pain. In particular, the research was directed at two 3-year randomized clinical trials studying the efficiency of different types of cognitive behavioral treatments for chronic low back pain and fibromyalgia. In further studies, some methodological aspects of measuring costs at a patient level and health-related quality of life were analysed .

Chapter 1 gives an outline of the concept of 'Medical Technology Assessment' (MTA) assessed in the field of chronic musculoskeletal pain. Several steps in the MTA process are described, such as the etiology and diagnostics, the burden of illness and the economic consequences of chronic musculoskeletal pain, with an emphasis on fibromyalgia and chronic low back pain. In addition an overview is provided of the state of the art about the effectiveness and cost-effectiveness of cognitive-behavioral therapy. It appeared that the majority of the treatments do not result in complete health or long-term pain-reduction. Therefore emphasis has shifted more and more from 'cure' to 'care' with increasing attention for a behavioral science approach to chronic pain . Previous studies showed a positive effect of these treatments on several psychological and functional variables. Conclusions are difficult to draw because of inadequate comparisons and methodological shortcomings. Moreover these types of interventions lacked economic evaluations.

Chapter 2 addresses important methodological issues in economic evaluation research, such as the techniques for economic evaluation studies and the analytic perspective. In addition, attention is paid to the inclusion of costs and outcomes, sensitivity analysis, discounting, incremental analysis and ratios, and collecting of data in economic evaluation research. The chapter notes that in chronic musculoskeletal pain some outcomes are less accessible than they may be in other disorders, such as quality of life and patient and family costs. Further emphasis is placed on the transparent reporting of methods and study results. Clear reporting facilitates the interpretation of the results of published studies and may enable one to assess the generalizability of results from one setting to another.

Chapter 3 consists of a review of economic evaluation studies of back pain interventions looking at costs, consequences and other methodological aspects. The chapter gives a detailed summary of the economic evaluation data given by these studies and provides insight into the quality of the economic evaluations in the field of back pain. Four of the studies dealt with injury prevention programs, 10 with post-

incidence management and secondary prevention programs and 16 studies with back injury and pain remedies (clinical treatment). The injury prevention programs (except for back school programs), post-incidence management and secondary prevention programs appeared to produce cost savings due to reduced absenteeism, generally for (acute) low back pain in the workplace. There appeared to be a lack of economic evaluations for the management of chronic low back pain and for any form of pain management. Little attention was further paid to the measurement of patient and family costs and health related quality of life. This chapter shows that the small number of comparable treatments, the large differences between and low quality in the costing methodologies used render it impossible to draw conclusions in favor of any one type of program for the treatment of low back pain.

Chapter 4 reports on a 3-year study on the comparison of the cost-effectiveness of a 6-week educational/cognitive coping skills training with both an educational/discussion intervention and a waiting list condition. A total of 131 fibromyalgia patients were randomly allocated to one of the three conditions. The both treatment programs were followed for 12 months, whereas the waiting-list control condition was followed for 6 weeks. Direct health care and non-health care costs, and the indirect costs associated with lost production due to illness, were calculated. The effects were measured in terms of utilities, using rating scale (RS) and standard gamble (SG) methods.

Treatment costs were estimated to be the same per patient for both interventions. Total direct health care costs of the educational/cognitive coping skills training were significantly higher than those for educational/discussion intervention. Indirect costs for the two groups were not significantly different. At 6 weeks there was a significant difference in RS scores between the three groups, caused by a significantly greater improvement in the educational/discussion group as compared to the waiting-list control group. However, no significant differences in either RS or SG utilities were found between the two treatment conditions immediately after treatment, or at the 6 or 12 month follow-ups.

The chapter concludes that the addition of a cognitive coping skills training to the educational intervention led to significantly higher health care costs and no additional improvement in quality of life as compared to the educational intervention combined with group-discussion. This conclusion is robust through a range of plausible values used in a sensitivity analysis.

In chapter 5 a cost-effectiveness study is presented, in which a combined operant program plus cognitive coping skills training is compared with an operant program plus group discussion (as attention-control) and both programs are compared with a waiting-list control group and with operant rehabilitation as usual provided by the same rehabilitation center. One hundred and forty-eight patients with chronic low back pain were randomly assigned to the different conditions. The economic endpoints were the costs of the program and other health care utilization, costs for the patient, and indirect costs associated with production losses due to low back pain. The effects were measured in terms of global assessment of change and utilities, using RS and SG methods.

The 3-year study determined that adding the cognitive coping-skills training to an operant treatment did not lead to significant differences in costs and improvement in quality of life when compared to the operant treatment combined with group-discussion. Compared to the common individual rehabilitation therapy it was concluded that at least the same effects could be reached at the same or lower costs with a shorter, more intense standardized group program. Treatment is more effective than providing no treatment in the waiting group.

In chapter 6 the development and assessment of the patient cost diary is evaluated. The cost diary was applied in two cost-effectiveness studies on behavioral rehabilitation in 205 fibromyalgia and chronic low back pain patients, as described in chapters 4 and 5. The use of the diary was evaluated on grounds of the feasibility, the influence of the period of data collection on the results, and some aspects of validity. The response rate turned out to be 85% and in total 68% of the diaries were returned. Similar outcomes were found for the three alternative periods of data collection: keeping the diary two weeks every two months, three months per year, or a full year. Finally, self-reported specialist care contacts were generally in agreement with data from an insurance company. However, for physiotherapy contacts there were differences between the self-reported and insurance data. The study in this chapter shows how the cost diary might be used successfully in cost-effectiveness studies.

Chapter 7 reports on a study on the comparison of the psychometric qualities of two instruments for assessing patients' health state valuations, the RS and the SG. Such instruments are designed for their application in economic evaluation research, but have seldomly been used in chronic pain trials. Both methods provide a single measure between 0 and 1. The relationship between these utility measures and descriptive and domain-specific quality of life measures was examined in 133 fibromyalgia patients and 148 patients with chronic non-specific low back pain. Mean utility score at baseline was 0.43 with the RS and 0.78 for the SG. The correlation between both methods was found to be poor ($r=0.21$). Both measures appeared to be fairly stable in a two-week test-retest period ($ICC = 0.74$ and 0.77). Scores on the description of patient's own health on 6 domains, global assessment of change and domain specific measures revealed moderate correlation with the RS scores and low correlation with the SG. Multiple regression analyses demonstrated that 32% of the variance in RS values and only 13% of the variance in SG utilities could be explained by domain-specific measures. The results suggest acceptable construct validity for the RS but insufficient construct validity for the SG. Valuations of own health appear to be only partially related to the assessment of the pain-specific measures and measures of distress. It may be concluded that the RS and domain-specific measures assess partly different, but nevertheless complementary aspects of health-related quality of life. It is therefore recommended to include in economic evaluation studies both domain-specific measures and valuation measures. Finally, in chronic musculoskeletal pain patients, RS scores were found to be more responsive in detecting significant changes in preferences than SG scores. The chapter concludes that for use in patients with chronic musculoskeletal pain, the RS is preferred to the SG for establishing accurate decisions about the impact of new interventions on their health outcomes.

Chapter 8 reports on the most important findings of the different studies addressed in this thesis and discusses some conclusions in the overall context of the thesis. Several methodological constraints are identified that may be considered when interpreting the costs and outcomes of both cost-effectiveness studies presented in this thesis. The chapter briefly formulates the implication of the studies in this thesis for the perspective of the patients, the providers, the rehabilitation center and the insurance companies. Furthermore some recommendations for health policy and future research are given. It is suggested that more attention be paid to the choice of alternatives and to the measurement of patient and family costs, and health-related quality of life, when conducting economic evaluation studies on the (behavioral) management of chronic musculoskeletal pain. It might be beneficial to include cost-effectiveness into study proposals as a primary study objective.

Co-operation between a multidisciplinary team of experts in the area is important to incorporate their specific needs into the research questions. This will facilitate potential changes and implementations of study implications into practice.

SAMENVATTING

De belangstelling voor economisch evaluatieonderzoek op het gebied van chronische pijn in het houdings- en bewegingsapparaat is de laatste jaren enorm gegroeid. De stijgende gezondheidszorgkosten en productieverliezen en het groeiend aantal behandelvormen met vaak teleurstellende resultaten, heeft eraan bijgedragen behandelingen niet alleen op werkzaamheid en effectiviteit te beoordelen, maar ook de economische consequenties ervan te evalueren. Door middel van economische evaluaties worden kosten en effecten van behandelingen met elkaar vergeleken om zo te bepalen welke behandelingen voor chronische pijn in het houdings- en bewegingsapparaat de beste opbrengsten geven in verhouding tot de kosten. Dit proefschrift beschrijft de toepassing van economische evaluatie op cognitieve gedragsmatige behandelingen voor chronische pijn in het houdings- en bewegingsapparaat. Centraal staan twee gerandomiseerde klinische studies waarin de efficiency van een cognitief gedragsmatige behandeling voor patiënten met chronische lage rugpijn en fibromyalgie bestudeerd wordt. Verder wordt een aantal methodologische aspecten met betrekking tot het meten van kosten via de patiënt en gezondheidsgerelateerde kwaliteit van leven geanalyseerd.

In hoofdstuk 1 wordt het concept 'Medical Technology Assessment' (MTA) toegelicht en in de context geplaatst van chronische pijn in het houdings- en bewegingsapparaat. Een aantal stappen in het MTA proces worden besproken, zoals de aetiologie en diagnose, de ernst van de aandoening en de economische gevolgen van chronische pijn in het houdings- en bewegingsapparaat. Hierbij ligt de nadruk op fibromyalgie en (chronische) lage rugpijn. Vervolgens wordt een overzicht gegeven van de laatste stand van zaken op het gebied van effectiviteit en kosten-effectiviteit van (cognitieve gedrags)therapie voor beide chronische pijnsyndromen. Het blijkt dat de meerderheid van de behandelingen niet bijdragen aan genezing of pijnreductie op lange termijn. Hierdoor is de aandacht steeds meer verschoven van 'cure' naar 'care', met een toenemende aandacht voor chronische pijn vanuit de gedragswetenschappen. Uit de voorhanden zijnde studies blijkt dat deze behandelingen effect hebben op verschillende psychologische en functionele variabelen. Conclusies zijn echter moeilijk te trekken vanwege onvoldoende vergelijkingsmogelijkheden en methodologische tekortkomingen. Tevens ontbreken economische evaluaties van deze behandelingsvormen.

Hoofdstuk 2 belicht diverse methodologische aspecten die belangrijk zijn in economische evaluatie studies, zoals de verschillende technieken en het analytische perspectief van de studie. Daarnaast wordt aandacht besteed aan de includering van

kosten en resultaten, sensitiviteits analyse, discontering, incrementele analyse en ratio's. Het blijkt dat een aantal effecten bij aandoeningen in het houdings en bewegingsapparaat moeilijker te bepalen is, zoals gezondheidsgerelateerde kwaliteit van leven en patient-gerelateerde kosten. In dit hoofdstuk wordt verder de nadruk gelegd op het belang van een inzichtelijke weergave van toegepaste methoden en verkregen resultaten bij de uitvoering van economische evaluatie studies. Een duidelijke rapportage vergemakkelijkt de interpretatie van gepubliceerde studies en geeft mogelijkheden om de resultaten te generaliseren naar verschillende studie settings.

In hoofdstuk 3 worden de resultaten beschreven van een literatuurstudie van economische evaluaties van rugpijn interventies. Hierbij wordt gekeken naar de kosten, effecten en andere methodologische karakteristieken. Het hoofdstuk geeft een gedetailleerde samenvatting van de economische evaluaties en de methodologische kwaliteit van de rugpijn studies. In deze review worden 30 economische evaluatie studies betrokken waarvan 4 studies primaire preventie programma's evalueren, 10 studies post-incidentie management en secundaire preventie programma's evalueren en 16 studies de efficiency van (klinische) behandelingen onderzoeken. De economische evaluaties tonen aan dat primaire preventie (behalve de rugscholen) en secundaire preventie programma's in het algemeen kosten besparend zijn, als gevolg van onder andere een daling van het ziekteverzuim van (acute) lage rugpijn op het werk. Economische evaluaties van behandelingen van chronische lage rugpijn en van elke vorm van 'pijn management' blijken te ontbreken. Daarnaast wordt ook weinig aandacht besteed aan het meten van patient-gerelateerde kosten en gezondheidsgerelateerde kwaliteit van leven. Binnen dit hoofdstuk wordt aangetoond dat vanwege de uiteenlopende behandelingsvormen en de diversiteit tussen en lage kwaliteit van methodologieën om kosten-effectiviteit te meten, het niet mogelijk is om conclusies te trekken over de meest kosten-effectieve behandeling voor lage rugpijn.

Binnen de behandelprogramma's van chronische pijnklachten wordt de laatste jaren overwegend aandacht besteed aan cognitief-gedragstherapeutische interventies. Het gaat hierbij steeds vaker om combinaties van operante en cognitieve behandelingen. Een operante behandeling gericht op motorische activiteit, heeft tot doel gezond gedrag opnieuw aan te leren en pijngedrag af te leren. Een cognitieve behandeling is gericht op pijnattributies en verwachtingen met betrekking tot de eigen controle over pijn. Een combinatie van deze twee benaderingen sluit goed aan bij de gedachtengang dat een behandeling betere resultaten oplevert wanneer bij patiënten het gevoel van controle over de pijn is toegenomen en negatieve verwachtingen zijn gecorrigeerd. Dit betekent dat bij toepassing van uitsluitend een operant programma, effecten kunnen worden verwacht op de motorische component van de pijn, maar dat pijn cognities niet automatisch gecorrigeerd worden. Binnen de gerandomiseerde kosten-effectiviteitsstudies beschreven in de hoofdstukken 4 en 5 wordt dan ook getoetst of een combinatie van educatie met cognitieve pijn-coping meer oplevert dan educatie alleen, en een combinatie van operant met cognitieve pijn-coping meer oplevert dan operant alleen.

Hoofdstuk 4 betreft een economische evaluatie van een cognitief-gedragsmatige behandeling voor patienten met fibromyalgie. De resultaten worden beschreven van

een drie jaar durende studie waarin de kosten-effectiviteit van drie interventies wordt vergeleken: een educatieve-cognitieve pijn coping vaardigheidstraining, een educatief-discussie programma en een wachtlijst controle groep. In totaal zijn 131 fibromyalgie patiënten gerandomiseerd over deze drie programma's. De patiënten in de educatieve-cognitieve pijn coping vaardigheidstraining en in het educatief-discussie programma zijn gedurende 12 maanden gevolgd, de wachtlijst-controle conditie is daarentegen alleen over een periode van 6 weken gevolgd. Kosten zijn berekend in termen van directe gezondheidszorg kosten, directe niet gezondheidszorg kosten en indirecte kosten, gerelateerd aan productie verliezen. De effecten zijn gemeten in termen van waarderingen voor kwaliteit van leven, waarbij gebruik is gemaakt van de rating scale (RS) en de standard gamble (SG) methode. De behandelkosten per patiënt zijn gelijk voor beide interventies. De totale directe gezondheidszorg kosten van het educatieve-cognitieve pijn-coping vaardigheidstraining zijn significant hoger dan van het educatieve-discussie programma. De indirecte kosten zijn niet significant verschillend tussen de drie groepen. Wat betreft de effecten is na 6 weken een significant verschil voor de RS utiliteiten te zien tussen de drie groepen. Dit verschil is veroorzaakt door een significante verbetering in kwaliteit van leven van de patiënten in de educatief-discussie groep, vergeleken met de wachtlijst controle groep. Echter, tussen de patiënten in de educatief-discussie groep en de patiënten in de educatieve-cognitieve pijn coping vaardigheidstraining is direct na de behandeling, alsook na 6 en 12 maanden follow-up geen verschil gevonden op de RS en SG utiliteiten. Uit bovenstaande resultaten kan geconcludeerd worden dat de toevoeging van een cognitieve pijn coping vaardigheidstraining aan een educatieve interventie vergeleken met een educatief-discussie programma, leidt tot hogere gezondheidszorgkosten en geen extra verbetering in kwaliteit van leven. Dit is in tegenstelling tot de verwachtingen! Deze conclusie blijkt robuust te zijn nadat verschillende kostenvariabelen in een sensitiviteitsanalyse zijn getest.

In hoofdstuk 5 worden aan de hand van een driejarige gerandomiseerde klinische studie de kosten en de effecten vergeleken van drie interventies voor mensen met chronische lage rugpijn. De eerste interventie betreft een operant programma met toevoeging van een cognitieve pijn coping vaardigheidstraining Het tweede programma betreft een operant programma dat gecombineerd wordt met een groepsdiscussie (geincludeerd als een aandachts controle voor de pijn coping vaardigheidstraining). Beide programma's worden verder vergeleken met een wachtlijst-controle groep en met de standaard operante behandeling die wordt gegeven in hetzelfde revalidatie centrum. Op basis van toeval worden honderdachtenveertig chronische lage rugpijn patiënten toegewezen aan een van de condities. De kosten worden gemeten in termen van programmakosten en andere gezondheidszorgkosten, patient-gerelateerde kosten en indirecte kosten, gebaseerd op verloren gegane productie door ziekteverzuim van betaalde en onbetaalde arbeid. De effecten worden gemeten in termen van ' global assessment of change', en RS en SG waarderingen.

De studie heeft aangetoond, wederom in tegenstelling tot de verwachtingen, dat de toevoeging van cognitieve pijn coping vaardighheids training aan een operant programma, in vergelijking tot een operant programma gecombineerd met een

groepsdiscussie niet leidt tot een significant verschil in kosten en verbetering in kwaliteit van leven. In vergelijking tot de standaard operante behandeling kan worden geconcludeerd dat tenminste dezelfde effecten kunnen worden bereikt voor ongeveer gelijke (of waarschijnlijk) lagere kosten met een in verhouding korter, intensiever en geprotocolleerd groepsprogramma. Daarnaast blijkt voor chronische lage rugpijn behandelen effectiever dan niet behandelen.

In hoofdstuk 6 wordt de ontwikkeling en het gebruik van het kostendagboek geëvalueerd. Hierbij wordt gekeken naar de bruikbaarheid, de invloed van de periode van dataverzameling op de resultaten, en naar een aantal aspecten van validiteit. Het kostendagboek is toegepast in twee kosten-effectiviteitsstudies om over de periode van 1 jaar de totale gezondheidszorg kosten, uitgaven en productie verliezen te meten van 205 patiënten met fibromyalgie of chronische lage rugpijn. Binnen dit kostendagboek is gevraagd naar het aantal bezoeken aan de huisarts, specialisten, paramedische hulpverleners en alternatieve genezers, en verder naar het aantal ziekenhuisopnamen, activiteiten, uren betaalde en onbetaalde hulp in de huishouding, dagen verzuim (van betaald en onbetaald werk), dagen van inactiviteit en extra uitgaven. Van de patiënten heeft 85% het dagboek minimaal een maand bijgehouden en is 68% van het totaal aantal dagboeken geretourneerd. De drie verschillende perioden van dataverzameling leveren geen verschillende resultaten op: de totale kosten verschillen niet significant wanneer het dagboek twee weken per twee maanden, drie maanden per jaar of een volledig jaar wordt bijgehouden. Tevens verschilt het aantal door de patiënt gerapporteerde fysiotherapie consulten met de gegevens van de ziektekostenverzekeraar. Hiervoor zijn een aantal verklaringen gegeven, waaronder een mogelijke achterstand in declareren en onduidelijkheid van de vraagstelling in het dagboek. De twee data bronnen verschillen niet voor het aantal specialisten consulten. De studie in dit hoofdstuk geeft weer op welke wijze een kostendagboek succesvol gebruikt kan worden in kosten-effectiviteits studies. Om de bruikbaarheid en construct validiteit van het kostendagboek te bevestigen is een gedetailleerde rapportage van de afzonderlijke variabelen noodzakelijk.

In hoofdstuk 7 worden diverse psychometrische kwaliteiten van de RS en SG met elkaar vergeleken. Deze instrumenten worden vaak toegepast in economische evaluatie studies, maar zijn nog maar zelden toegepast in het gebied van de chronische pijn. De relatie tussen deze utiliteits instrumenten en beschrijvende en ziekte specifieke kwaliteit van leven instrumenten is bestudeerd bij 133 fibromyalgie patiënten en 148 chronische lage rugpijn patiënten. Tijdens de baseline waardeerden deze patiënten hun kwaliteit van leven gemiddeld met 0.43 op de RS en 0.78 op de SG. De correlatie tussen beide methoden bleek laag te zijn ($r=0.21$). De tweeweekse test herstest betrouwbaarheid voor de waardering van de kwaliteit van leven van de patiënt, uitgedrukt als intraclass correlatie coëfficiënt (ICC) was goed te noemen (ICC=0.74 voor RS en ICC=0.77 voor SG). De beschrijvingen van de patiënt aan de hand van zes dimensies van gezondheid, de 'global assessment of change' en de domein specifieke instrumenten correleren matig met de RS en laag met de SG. Multiple regressie analyses laten zien dat de 32% van de variantie in RS scores en slechts 13% van de variantie in SG scores verklaard kan worden door de scores op de domein specifieke

kwaliteit van leven instrumenten. Deze resultaten doen vermoeden dat de waarderingen verkregen met de RS methode de domeinspecifieke kwaliteit van leven dichter benaderen dan SG waarderingen. Daarnaast blijkt maar een gedeeltelijke relatie te bestaan tussen de waardering voor de eigen gezondheidstoestand en de score op de pijnspecifieke maten en maten van emotioneel welbevinden. Hieruit kan geconcludeerd worden dat de RS en domeinspecifieke instrumenten gedeeltelijk verschillende, maar tevens overeenkomende aspecten van kwaliteit van leven beoordelen. Het is daarom aan te bevelen om binnen economische evaluatie studies zowel domeinspecifieke instrumenten als ook waarderings instrumenten te gebruiken. Tenslotte blijkt voor patiënten met chronische pijn in het bewegingsapparaat de RS methode gevoeliger om verandering te meten dan de SG methode. Wanneer beleidsbeslissingen genomen moeten worden omtrent de invloed van een nieuwe therapie op de algehele gezondheidstoestand van patiënten met chronische pijn in het houdings- en bewegingsapparaat, kan geconcludeerd worden dat de voorkeur wordt gegeven aan de RS boven de SG.

Hoofdstuk 8 beschrijft op een beknopte manier de meest belangrijke bevindingen van de studies die in dit proefschrift worden beschreven. Bij het meten en interpreteren van de kosten en effecten van de beide kosten-effectiviteits studies die in dit proefschrift zijn geëvalueerd, zijn verschillende methodologische beperkingen geïdentificeerd. Dit betreft het design van de studies, de keuze van de te vergelijken alternatieven en het meten van patientgebonden kosten en waarderingen voor kwaliteit van leven. Tevens wordt ingegaan op de mogelijke implicaties van de studies, bekeken vanuit het oogpunt van de patiënten, de behandelaars, het revalidatie centrum en de zorgverzekeraars. Verder worden aanbevelingen gedaan voor gezondheidsbeleid en toekomstig onderzoek. Zo wordt onder meer aanbevolen om voor deze specifieke groep patiënten met chronische pijn in het houdings- en bewegingsapparaat, gedragsgeoriënteerde revalidatie met minimaal een aantal operante principes op te nemen als nationale richtlijn. Met betrekking tot economisch evaluatie onderzoek op het gebied van gedragsgeoriënteerde revalidatie van chronische pijn in het houdings- en bewegingsapparaat, wordt aanbevolen meer aandacht te besteden aan de keuze van de te vergelijken alternatieve interventies en de meting van patiënt-gerelateerde kosten en kwaliteit van leven. Het is belangrijk om de kosten-effectiviteits analyse in een onderzoeksvoorstel op te nemen als één van de primaire vraagstellingen. Ook reeds in de fase van onderzoek is samenwerking tussen een multidisciplinair team van experts in dit gebied belangrijk om ieders specifieke belangen in de onderzoeksvragen op te nemen. Uiteindelijk zal dit bijdragen aan een soepelere implementatie van de studie resultaten naar de praktijk.

DANKWOORD

Het dankwoord is het enige onderdeel in dit proefschrift dat niet gebaseerd is op voor anderen toetsbare feiten. Dit maakt het dankwoord onbelangrijk voor de wetenschap, maar van moreel significant belang voor mezelf. Het mogen schrijven van het dankwoord betekent dat mijn proefschrift bijna is voltooid. Bijna, want rest mij nog al die mensen te bedanken zonder wiens aandacht en steun ik nooit aan het schrijven van dit dankwoord zou zijn toegekomen. Sommige van hen wil ik graag in willekeurige volgorde noemen.

Johan Vlaeyen, dat ik mijn eerste woord van dank aan jouw richt is echter niet willekeurig. Je inzet als co-promotor, is voor mij van grote waarde geweest. Door de prettige collegiale sfeer, en je enthousiasme en gedrevenheid wist je me keer op keer de stimulans te geven die ik op dat moment nodig had. Ook het feit dat je altijd even op mijn kamer kwam binnen lopen om te vragen hoe het er voor stond, is voor mij zeer waardevol geweest. Misschien moest je me tijdens de eindspurt wel eens te vaak positief bekrachtigen. Johan, bedankt!

Maureen Rutten, jij hebt mij als eerste geïntroduceerd in het gebied van de MTA en 'standard gambles' en werd vanzelf co-promotor. Dat onderzoek doen heel leuk en ontspannen kon zijn werd me tijdens onze gezamenlijke congressen snel duidelijk. Maureen, bedankt voor je inzet en dat je ondanks het feit dat je naar Rotterdam ging (en Brian erbij kwam) betrokken bent gebleven. Gelukkig was Stein minder ver!

Sjef van der Linden, in de zomer van 1997 kwam ik voor het eerst bij je met mijn promotieplan. Vanaf die tijd is (bijna) alles volgens plan verlopen. Beste Sjef, ik bedank je voor het vertrouwen dat je in me hebt gesteld. Veel heb ik geleerd van onze discussies en je kritische commentaren. Kritischer ben ik in ieder geval geworden.

De leden van de beoordelingscommissie: Prof. Onno van Schayck, Prof. Hans Maarse, Prof. Frans Rutten, Dr. Ir. Riekje de Vet en Dr. Steven Linton dank ik voor het beoordelen van mijn proefschrift. Prof. Dr. Eddy van Doorslaer leverde een belangrijke bijdrage aan de voorbereiding van de kosten-effectiviteitsstudies. Prof. Reiner Leidl, voormalig voorzitter van de vakgroep Economie van de Gezondheidszorg (EGZ, nu BEOZ) bedank ik voor zijn inbreng tijdens de uitvoering en verslaglegging van de studies.

De medewerking van de behandelaars en de patiënten was van onmisbaar belang in dit project. Het behandelteam "Reuma en Pijn" van de SRL Limburg, onder leiding van

Peter Heuts bedank ik voor de goede uitvoering van de behandelingen. Peter Heuts, bedankt voor de prettige samenwerking en voor je hulp bij de kostprijsberekeningen. De patiënten in beide projecten bedank ik voor hun bereidheid om tot vijf keer toe door mij geïnterviewd te worden en hun gezondheidstoestand te waarderen op het kansbord en de thermometer. Daarnaast hebben zij de moeite genomen om gedurende 12 maanden het kostendagboek bij te houden.

Het klinische deel van de chronische lage rugpijn en fibromyalgie studie werd uitgevoerd door respectievelijk Ank Kole-Snijders en Nicole Teeken-Gruben. Ank en Nicole, jullie wil ik bedanken voor de samenwerking en het samen schrijven van de eind-rapporten en de artikelen hieromtrent. Stefan Bos, leverde als stagiaire een belangrijke bijdrage aan de kostprijsberekeningen. Sylvia Jansen op de Haar voerde samen met mij de vele metingen uit. Bedankt voor alle gezelligheid Sylvia!

Silvia Evers, mijn beste maatje en collega, mijn vraagbaak en uitlaatklep. Vanzelfsprekend ben jij mijn 'nimpf'. Ons werk en de vooruitzichten waren vaak onderhevig aan onzekerheid en de nodige ups en downs, onze vriendschap en collegialiteit werd hierdoor alleen nog meer versterkt. Lieve Silvia, een betere vriendin had ik mij in deze periode niet kunnen wensen. Jammergenoeg denkt niet iedereen zoals wij, 'lama-makara'. Loek, bedankt voor al die lekkere ijsjes die je speciaal voor mij maakte!

Jan van Emmerik, ook jij bent ondanks alle veranderingen tijdens mijn loopbaan continu aanwezig geweest. In mijn tijd bij EGZ wist je me altijd uit de brand te helpen, was het nou voor statistiek, een nieuw programma, de computer die weer 'kuren' had of voor de auto die het 's avonds laat begaf, je stond altijd klaar. Jan, jouw kennis en handigheid was goud waard. Speciaal bedank ik je ook voor het 'lay-outen' van het proefschrift. Enne... ik heb mijn belofte gehouden, je bent paranimpf.

Steven Linton, thank you for paying so much interest to my work during all these years and for your willingness to comment on various drafts of my papers. It has been a great opportunity and pleasure to work with you. Very special to me was the visit I made to your institute and the possibility you gave me to give a workshop. I will never forget the painting-lesson with your family, my first time ever! I still wonder if you ever sold my painting. Steven, thanks for your collegiality and friendship.

Thanks to Egon Jonsson of the Swedish Council on Technology Assessment in Health Care (SBU) in Stockholm and Alf Nachemson who made it possible to update the back pain review. It is a great opportunity for me to be a member of the SBU Back Pain group. Gordon Waddell, Anders Norlund and Lars Ake Marke of the Economics subgroup thanks for your valuable comments and discussions and taking interest in the progress of my thesis. The rest of the group I thank a lot for the lively meetings and the nice atmosphere during the rest of the, mostly, 2-day rendezvous somewhere near Stockholm. I look forward to the presentation of our book! Jessica, thank you for your friendship. Maurits, mijn Nederlandse collega in het gezelschap dank ik voor het samen schrijven van de 'chronische nek en rugpijn' hoofdstukken en voor de gezellige 'pre-en-post-meetings' activiteiten. Een Nederlands onderonsje was op z'n tijd wel zo prettig!

In addition, I thank Bob Wilkinson, Frank McGuillan and David Gilbert for the valuable English corrections of some parts of this thesis. Gerard van Breukelen dank ik voor de deskundige statistische adviezen.

Gedurende ruim vier jaar heb ik met veel plezier gewerkt bij de vakgroep Economie van de Gezondheidszorg (het huidige BEOZ) van de Universiteit Maastricht. Gezelligheid en vriendschap stonden voorop, wat bleek uit de vele gezamenlijke sportactiviteiten, avondjes uit en verlengde congresbezoeken. Na het verlaten van de vakgroep beseftte ik pas hoe veel dit voor mij betekende. Speciaal wil ik Raymond, Rob (bedankt voor het commentariëren van de discussie), Grant, Ria en Mieke bedanken. Jullie zorgden voor de juiste ontspanning op het juiste moment. Ria, ik heb de oplossing, 'Sweet reserve' importeren (en thuis een aardbeieveldje)! Mieke, als ik drie paranimfen had mogen hebben, dan had ik niet lang hoeven na te denken. Jouw invulling van het korte-termijn-assistentschap heeft voor ons een lange-termijn-vriendschap opgeleverd. Bedankt voor je betrokkenheid en luisterend oor! Vastenaovend houden we erin.

Mijn huidige iRv-collega's van het thema 'revalidatie-processen' hebben ervoor gezorgd dat het pijltje op de 1 meter grote thermometer tegen mijn kast altijd een hoge kwaliteit van leven aanwees. Bedankt voor de geweldig goede sfeer en de warme belangstelling die ik altijd heb gekregen tijdens de eindspurt. Jolanda en Ingrid, jullie waren altijd bereid om te hulp te schieten wanneer 'SPSS weer eens vreemd deed'. Bedankt! Jolanda, ook fijn dat je wekelijks als 'carpool-uitlaatklep' wilde dienen. Binnenkort poolen van bij jou (ik stel voor vanaf de winter)! Henk, Caroline en Geert, bedankt voor jullie verhalen en adviezen over promoveren. Nu ben jij aan de beurt Yvonne! Jeanine en Anja, mijn congresteam; op naar Wenen. Verder bedank ik Luc de Witte die mij de ruimte heeft gegeven om het proefschrift te voltooien en Josien Houben voor het onvermoeibaar veranderen van de referenties.

Graag wil ik ook mijn vrienden bedanken voor de gezelligheid, de vele gesprekken en het gewoon aanwezig zijn. Sommigen van jullie kwamen er misschien niet meer wijs uit wanneer klaar nu *echt* definitief klaar betekende. Nu is het echt af! Dorit, toen ik wist dat jij mijn voorkant zou ontwerpen was dit een enorme geruststelling voor me. Het is echt uniek geworden! Toda raba, voor het enthousiasme waarmee je dit hebt gedaan, en tevens voor alle Hebreeuwse lessen, die mij een goede afleiding bezorgden en voor onze gezellige en vaak diepe gesprekken. Mijn 'Loopgroep Harmsen' (of toch maar LH-97), de vele kilometers hardlopen, het koffiedrinken en de uitstapjes hebben ruimschoots bijgedragen aan mijn geestelijk welbevinden. Gezamenlijk hebben we heel wat vet verbrand en gestapeld. Theo, Bianca, Monique en Cita bedankt voor jullie echte vriendschap in makkelijke, maar vooral ook in moeilijke perioden. Ik heb weer meer tijd!

Michal, our friendship is sometimes difficult to comprehend for outsiders, a large part of whom, if I am right now, has been gathered during our fourteen years of friendship. We grew up together, not only because of our letters and many trips all over the world, but also by being there, for each other and for ourselves. The distance between Tel Aviv and Maastricht shortened over the years. Michal, friend for ever, we are a perfect team! Guido, jouw enthousiasme voor de wetenschap heeft voor mij al vroeg als

voorbeeld kunnen dienen. Misschien ben ik niet voor niets ook dezelfde weg ingeslagen, alhoewel...wetenschap?

Pap en mam, jullie zijn mijn roots. Het boekje was er dan ook nooit gekomen zonder jullie. Bedankt voor jullie voortdurende steun en het vertrouwen in mijn kunnen. Fijn dat jullie mijn ouders zijn! En mam, bedankt dat er altijd een haneboutje voor me klaar stond om me weer op te pepen! Je bent de beste!

En tenslotte Jeroen, huhu! Jij hebt de eindspurt van zeer dichtbij beleefd. Jouw vermogen om te relativieren en me aan het lachen te krijgen, je rust en je warmte kwamen hierbij precies op het juiste moment. Het is af en ik zal (*nee ik kan*) niks meer veranderen. Het is tijd voor een nieuw hoofdstuk!

Het dankwoord is klaar en ik kan nu met veel overtuiging zeggen dat de waardering voor mijn eigen kwaliteit van leven significant is toegenomen. Het boek kan gesloten worden. Op naar het feest!

Mariëtte

L IST OF PUBLICATIONS

- Teeken-Gruben, N.J.G., Goossens, M.E.J.B., Nooyen-Haazen, I.W.C.J., Vlaeyen, J.W.S., Rutten-van Mólken M.P.M.H. *Tertiare preventie van chronische pijn in het houdings-en bewegingsapparaat: Het effect van een gecombineerd educatief programma en een cognitieve behandeling. Toepassing bij mensen met het fibromyalgiesyndroom.* Eindrapport, IRV 1995; (94)/200/220/mh.
- Nooyen-Haazen, I.W.C.J., Vlaeyen, J.W.S., Kole-Snijders, A.M.J., Teeken-Gruben, N.J.G., Goossens, M.E.J.B. *Fibromyalgie en chronische lage rugpijn: Twee verschillende syndromen? Een psycho-sociale vergelijking.* 1995; IRV (95)/199/802/mh.
- Teeken-Gruben, N.J.G., Vlaeyen, J.W.S., Goossens, M.E.J.B., Rutten-van Mólken, M.P.M.H., Eek van, H., Pelt R.A.G.B., Heuts, P.H.T.G. Effectiviteit van een cognitief-educatief behandelingsprogramma voor mensen met het fibromyalgie syndroom. *GEDRAG & GEZONDHEID* 1995; 23 (5); 191-205.
- Vlaeyen, J.W.S., Teeken-gruben, N.J.C., Goossens, M.E.J.B., Rutten-van Mólken, M.P.M.H., Pelt, R.A.G.B., Eek van, H., Heuts, P.H.T.G. Cognitive-educational treatment of fibromyalgia, a randomized clinical trial. Part I: Clinical effects. *Journal of Rheumatology* 1996; 23 (6) 1238-1245.
- Goossens, M.E.J.B., Rutten-van Mólken, M.P.M.H., Leidl, R., Stefan B., Vlaeyen J.W.S., Teeken-gruben N.J.C. (1996) Cognitive-educational treatment of fibromyalgia, a randomized clinical trial. Part II: Economic Evaluation. *Journal of Rheumatology* 1996; 23(6): 1246-1254.
- Nooyen-Haazen, I.W.C.J., Vlaeyen, J.W.S., Goossens, M.E.J.B., Teeken-Gruben, N.J.C. Cognitief-educatieve behandeling van fibromyalgie: Voorspellende waarde van therapieresultaat. 1996; IRV (96)/05-09/199/0113/in.
- Goossens, M.E.J.B., Evers, S.M.A.A. *Economic evaluation in back pain interventions.* Journal of Occupational Rehabilitation 1997; 7(1): 15-32.
- Goossens, M., Rutten, M., Vlaeyen, J., Kole-Snijders, A. Utility measurement in chronic low back pain patients; a comparison between domain specific and preference based instruments. *Quality of life research* 1997; 6/7: 650-651.
- Goossens, M.E.J.B., Rutten-van Mólken, M.P.M.H., Kole-Snijders, A.M.J., Vlaeyen, J.W.S., Health economic assessment of behavioural rehabilitation in chronic low back pain; A randomized clinical trial. *Health Economics* 1998; 7: 39-51.

- Vlaeyen, J.W.S., Nooyen-Haazen, I.W.C.J., Goossens, M.E.J.B., Breukelen van, G., Heuts, P.H.T.G., Goei The, H. *The role of fear in the cognitive-educational treatment of fibromyalgia*. Proceedings of the 8th World Congress on Pain. IASP Press 1997; 693-704.
- Goossens, M., van der Heijden, G., de Witte, L. *Groepsoefentherapie bij CVA-gehandicapten en hun centrale verzorgers*. Hoensbroek : iRv, Kenniscentrum voor Revalidatie en Handicap, Eindrapport, 1998.
- Goossens, M.E.J.B., Vlaeyen, J.W.S., Rutten M.P.M.H., Linden van der, J.M.J.P. Patient utilities in chronic musculoskeletal pain. How useful is the standard gamble method? *Pain* 1999; 80: 365-375.
- Goossens, M. 'Evidence based' patienten informatie. *Nederlands Tijdschrift voor Pijn en Pijnbestrijding* 1999; 19 (1): 22.
- Kole-Snijders, A.M.J., Vlaeyen, J.W.S., Goossens, M.E.J.B., Rutten van Mólken, M.P.M.H., Heuts, P.H.T.G. Chronic low back pain: what does cognitive treatment add to operant treatment? Results of a randomized clinical trial. (accepted in *Journal of consulting and clinical psychology*).
- Goossens, M.E.J.B., Rutten, M.P.M.H., Vlaeyen, J.W.S., Linden van der Sij. The cost-diary; a method to measure direct and indirect costs in cost-effectiveness research. (provisionally accepted for publication in *Journal of Clinical Epidemiology*).
- Goossens, M.E.J.B., Evers, S.M.A.A., Vlaeyen, J.W.S., Rutten- van Molken, M.P.M.H., van der Linden, J.M.J.P. Principles of economic evaluation for interventions of chronic musculoskeletal pain. (accepted).
- Van Tulder, M., Goossens M. *Chronic neck pain: non-surgical treatment*. Chapter 10. In: A. Nachemson, E. Jonsson, editors. *Evidence based treatment for back pain*. Stockholm, SBU/ Philadelphia, Lippincott (accepted).
- Van Tulder, M., Goossens, M., Waddell, G., Nachemson, A. *Chronic low back pain: non-surgical treatment*. Chapter 13. In: A. Nachemson, E. Jonsson, editors. *Evidence based treatment for back pain*. Stockholm, SBU/ Philadelphia, Lippincott (accepted).
- Goossens, M., Evers, S. *Cost-effectiveness of treatment of neck and low back pain*. Chapter 18. In: A. Nachemson, E. Jonsson, editors. *Evidence based treatment for back pain*. Stockholm, SBU/ Philadelphia, Lippincott (accepted).
- Van Tulder, M.W., Goossens, M.E.J.B., Hoving, J.L. *Non-surgical treatment of chronic neck pain: a systematic review*. (submitted).

About the author

Mariëlle Goossens was born on March 22nd 1969 in Heerlen. In 1987 she obtained her secondary school diploma at the 'Scholengemeenschap Groenewald' in Stein and started her study at the Faculty of Health Sciences from the Maastricht University. After graduation in 1992 she worked till 1996 as a research-fellow at the Department of Health Economics from the Maastricht University on the clinical trials presented in this thesis. In this period she also lectured at several summer university courses on Medical Technology Assessment (MTA) of the Maastricht University and gave workshops on MTA in Sweden. From May 1996 till May 1997 she worked as an MTA researcher at the Department of Medical Informatics, Epidemiology and Statistics of the University of Nijmegen. Since 1996 she has been participating in the 'SBU back pain group' at the Swedish Council on Technology Assessment in Health Care (Stockholm, Sweden), which is a working group with international invited experts. In 1997 she started her current job as an MTA researcher at the Institute for Rehabilitation Research in Hoensbroek and since October 1998 has been a part-time researcher at the Pain Management and Research Centre from the University Hospital of Maastricht. Besides her research activities she is involved as an MTA advisor in several national and international research projects directed at economic evaluation of chronic (musculoskeletal) pain, which is her primary field of interest.

About the author

Mariëlle Goossens was born on March 22nd 1969 in Heerlen. In 1987 she obtained her secondary school diploma at the 'Scholengemeenschap Groenewald' in Stein and started her study at the Faculty of Health Sciences from the Maastricht University. After graduation in 1992 she worked till 1996 as a research-fellow at the Department of Health Economics from the Maastricht University on the clinical trials presented in this thesis. In this period she also lectured at several summer university courses on Medical Technology Assessment (MTA) of the Maastricht University and gave workshops on MTA in Sweden. From May 1996 till May 1997 she worked as an MTA researcher at the Department of Medical Informatics, Epidemiology and Statistics of the University of Nijmegen. Since 1996 she has been participating in the 'SBU back pain group' at the Swedish Council on Technology Assessment in Health Care (Stockholm, Sweden), which is a working group with international invited experts. In 1997 she started her current job as an MTA researcher at the Institute for Rehabilitation Research in Hoensbroek and since October 1998 has been a part-time researcher at the Pain Management and Research Centre from the University Hospital of Maastricht. Besides her research activities she is involved as an MTA advisor in several national and international research projects directed at economic evaluation of chronic (musculoskeletal) pain, which is her primary field of interest.



ISBN 90-9012767-4