

# **An integrated approach to dementia** a clinical and economic evaluation

© Copyright C.A.G. Wolfs, Maastricht 2007

Layout | C.A.G. Wolfs

Production | Datawyse BV | Universitaire Pers Maastricht

Publisher | NeuroPsych publishers

ISBN | 978-90-75579-36-9

NeuroPsych Publishers is a non-profit organisation, which aims at promoting the science of 'Brain and Behaviour' and improving the application of the products of this science in health care and education. NeuroPsych Publishers accomplishes these aims by publishing books, dissertations and other products of scientific activity, by disseminating educational material and publication of tests, assessment scales and other psychometric instruments in the field of Neuropsychology, Neuropsychiatry and other areas within the domain of Brain and Behaviour.

NeuroPsych Publishers, Maastricht University, PO Box 616, NL-6200 MD Maastricht

# **An integrated approach to dementia a clinical and economic evaluation**

## **PROEFSCHRIFT**

ter verkrijging van de graad van doctor aan de Universiteit Maastricht,  
op gezag van Rector Magnificus, Prof. mr. G.P.M.F. Mols,  
volgens het besluit van het College van Decanen, in het openbaar te verdedigen  
op donderdag 6 december 2007 om 14:00 uur

door

**Claire Aldegonda Gerarda Wolfs**

Geboren te Eijsden op 20 mei 1977

**Promotores**

Prof. dr. F.R.J. Verhey

Prof. dr. J.L. Severens

**Co-promotor**

Dr. C.D. Dirksen

**Beoordelingscommissie**

Prof. dr. J. Metsemakers (voorzitter)

Dr. W.B.F. Brouwer

Dr. L. Jönssen

Dr. A.F.G. Leentjens

Prof. dr. J.M.G.A. Schols

The research described in this thesis was performed at the department of Clinical Epidemiology and Medical Technology Assessment, Maastricht University Hospital, the Netherlands, and at the department of Psychiatry and Neuropsychology, Maastricht University, the Netherlands.

The research described in this thesis was funded by the Dutch Research Institute for Care–Medical Sciences (Zorgonderzoek Nederland–Medische Wetenschappen, Zon-Mw), grant number 945-02-055.

The publication of this thesis was financially supported by:

**Novartis Pharma B.V., Internationale Stichting Alzheimer Onderzoek  
Alzheimer Nederland, Lundbeck B.V.**



JANSSEN-CILAG B.V.

Inte]on[



hago

Veur pap en mam



## Contents

Chapter 1	Introduction	9
<b>PART ONE</b>		
Chapter 2	Aims, design and methods	17
Chapter 3	The added value of a multidisciplinary approach in diagnosing dementia: a review	23
<b>PART TWO      Methodological issues</b>		
Chapter 4	Performance of the EQ-5D and the EQ-5D+C in elderly patients with cognitive impairments	41
Chapter 5	The measurement and valuation of informal care	55
<b>PART THREE      Findings of the MEDICIE-study</b>		
Chapter 6	Effectiveness of an integrated multidisciplinary diagnostic approach for psychogeriatric patients: a randomized controlled trial	93
Chapter 7	Economic evaluation of an integrated diagnostic approach for psychogeriatric patients: Results of a randomized controlled trial	107
Chapter 8	Informal caregiving in dementia: results of a survey	125
Chapter 9	GP concordance with advice for treatment following a multidisciplinary psychogeriatric assessment	141
Chapter 10	General discussion	153
	Samenvatting	163
	List of abbreviations	167
	References	171
	Dankwoord	183
	Curriculum vitae	187



# Chapter 1

## Introduction



## Epidemiology

The proportion of elderly in the total population is growing fast and this increase will continue into the next decades. In 2000, 14% of the Dutch population was 65 years or older, and prognostic calculations show that in 2015 and 2035 these percentages will be 17% and 24% (Centraal Bureau voor Statistiek: Statistisch jaarboek, 2000). In parallel, the number of elderly who have a chronic disease is also increasing fast. GP morbidity registrations reported rates for dementia ranging from 0.7% to 2.2% (Hout, 1999) and similar rates for other psychogeriatric problems, such as depression (Blazer et al., 1980; Girling et al., 1995; Verhey et al., 1997). Psychiatric disorders other than dementia, occur in much higher frequencies in the somatically elderly patient, with, for instance, prevalence rates up to 30% in patients with stroke or Parkinson's disease. These figures show the coexistence of somatic and psychological disorders in the frail elderly, and the necessity to create and evaluate integrated multidisciplinary health care services for the elderly.

## Integrated care

An integrated multidisciplinary approach to diagnosing and managing the multiple aspects related to dementia is generally recommended given that no profession is adequately equipped to single-handedly deal with the complex range of mental, physical and social problems that accompanies dementia (APA, 1997; CBO, 2005; NICE, 2006; Olde Rikkert et al., 2006). In the last decades, new types of services that deliver comprehensive assessment for the elderly have been initiated and evaluated. Several randomised controlled trials of geriatric assessment units have been done (Applegate et al., 1990; Cole et al., 1991; Schrijnemakers et al., 1995a; Schrijnemakers et al., 1995b; Toseland et al., 1997). Some of these have demonstrated that a geriatric evaluation can lead to improved functional and mental status, more appropriate placement of disabled older persons, decreased use of health care services and reduced mortality, although other studies failed to find any beneficial effects (Schrijnemakers et al., 1995a; Schrijnemakers et al., 1995b; Stuck et al., 1993; Turpie, 2000). Most of these initiatives, however, were monodisciplinary as they had only input from geriatric medicine.

Furthermore, multidisciplinary memory clinics (MCs) for patients with dementia have been established in many places over the world. MCs play an important role in establishing a possible diagnosis of dementia, performing etiological diagnostics and discussing the outcomes with the patients and their relatives. They provide information on the implications of the diagnosis, on the course of the illness and give advice how to deal with this. However, these services have recently been criticized for being principally hospital-based, for their apparently limited focus that entails only medical diagnostics and pharmacotherapy and for thus diverting resources from community mental health teams (CMHTs) (Pelosi et al., 2006).

On the other hand, CMHTs are mainly concerned with the provision and organization of integrated care, and unlike MCs, CMHTs generally have limited access to hospital-based

diagnostics such as brain imaging. Studies into the efficacy of multidisciplinary psychogeriatric units combining the expertise of geriatric medicine, MCs and CMHTs have not yet been performed (Wolfs et al., 2006).

Recently, an outpatient diagnostic facility, the Diagnostic Observation Centre for PsychoGeriatric patients (DOC-PG), was established in Maastricht, the Netherlands. This facility combines the hospital-based approach of a memory clinic with the care-oriented approach of a regional community mental health team and aims to provide general practitioners (GPs) with detailed diagnostic and therapeutic advice for patients with cognitive disorders.

## **Evaluating a new intervention in dementia**

### **Clinical evaluation**

The evaluation of a new intervention that has the management of complex care in geriatric patients with dementia should apply a broad approach, in accordance to the complexity of the problems accompanying dementia. Outcomes should include effects on patients regarding cognition, behavioural and psychological symptoms, quality of life, global assessments and activities of daily living (Katona et al., 2007). Quality of life is an important outcome measure as it reflects the multi-dimensionality of dementia. Assessment of quality of life also calls attention to positive states and 'personhood' in dementia, as opposed to most other measures of dementia that focus on deficits and pathology (Ready et al., 2003). Declines in memory and cognitive abilities are a consequence of ageing and it needs to be assessed whether this decline is due to normal ageing. Assessment of cognition is highly standardized in dementia research (Rockwood, 2007). Behavioural and psychological symptoms are important as they are very common in dementia and are known to cause considerable suffering in patients and their caregivers (Aalten et al., 2005). Clinical global measures are often utilized because they have proven useful in efficacy trials and in the study and assessment of interventions in dementia in all severity stages (Reisberg, 2007). Finally, activities of daily living are considered important to assess the functional independence of the patients (NICE, 2006).

In addition, outcomes should encompass effects on caregivers. The majority of dementia patients is living at home and are usually cared for by family or friends. Although caregivers do not suffer from the pathology, they are also affected by the illness (Brodaty, 2007; de Vugt et al., 2005a; de Vugt et al., 2005b; de Vugt et al., 2004). Caregivers have high risks for depression and other mental health problems, an increased physical morbidity, higher health-care related costs, and feel more isolated than non-caregivers (Schulz et al., 1995). Furthermore, caregivers are crucial to various aspects in the dementia process as they are the gatekeepers to services, they are essential for conducting research and they are likely to influence outcome (Brodaty, 2007).

**Economic evaluation**

Obviously, an elaborate assessment by an integrated multidisciplinary facility as DOC-PG will have important economic implications which need to be taken into account when studying the effects. The pervasive scarcity of resources due to the increasing number of older adults who are diagnosed with dementia led to a growing interest in analyzing the economic consequences of new interventions. Mental illnesses, which constitute 22% of the total, represent the most expensive diagnosis group.

In absolute numbers, 12.7 billion euro was spent on mental illness in 2003 in the Netherlands, including costs for those afflicted by dementia (Slobbe et al., 2006).

Economic evaluations are part of Medical Technology Assessment (MTA), also referred to as Health Technology Assessment (HTA) which identify, measure, value, and compare costs and outcomes of alternative health care interventions in order to inform decision making (Oortwijn, 2000).

Five different techniques of economic evaluations can be distinguished: cost-minimization analysis, cost-benefit analysis, cost-consequence analysis, cost-utility analysis and cost-effectiveness analysis (Drummond et al., 1997). A cost-minimization analysis is performed when it is assumed that two interventions are equally effective. A cost-benefit analysis expresses both the costs and effects in monetary terms, whereby the benefits of an intervention are evaluated in terms of the willingness to pay for them. To date, this type of analysis is not used as the sole basis for health care decision-making, as assigning monetary values to health outcomes is controversial. A cost-consequence analysis is an analysis in which costs and effects are calculated but not aggregated into quality-adjusted-life-years (QALYs) or cost-effectiveness ratios (Mauskopf et al., 1998). It is a listing of all relevant costs, outcomes and consequences which enables decision-makers to make adequate resource allocation decisions (Mauskopf et al., 1998). A cost-effectiveness analysis (CEA) and a cost-utility analysis (CUA), which are rather similar, are most commonly used and both involve a comparison of alternative interventions regarding their costs and effects. In a CEA, clinically relevant outcomes such as average number of life-years gained, morbid events averted or disease specific outcome parameters are measured. In a CUA the measure of effectiveness is a quality-adjusted-life-year (QALY), which combines quality of life with duration of life. Because CEA and CUA have identical analytical methods, they are often both referred to as CEA. Preferences for health states (health utilities) are used as weights in the calculation of QALYs. To date, the EQ-5D (Brooks, 1996; "EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group", 1990) and the SF-6D (Brazier et al., 2002b) are the only two utility-generating instruments recommended by NICE (NICE, 2006) for use in dementia. However, concern has been raised they are insufficiently sensitive to pick up (clinical) changes in dementia patients (Knapp, 2007).

It is recommended that economic evaluation studies should be conducted from a broad societal perspective, including all relevant costs and health effects regardless of who bears the costs or who obtains the effects. Informal care is however rarely included in economic evaluation studies, since the collection and valuation of these data is complex. Furthermore, the time horizon is crucial and, in dementia, should be at least 12 months

to capture all effects and costs. This is important in dementia, as major cost-driving events such as institutionalization will rarely occur in periods shorter than 12 months.

This thesis describes the findings of the MEDICIE study (Maastricht Evaluation of a Diagnostic Intervention for Cognitively Impaired Elderly), a randomized controlled trial comparing the efficacy and efficiency of DOC-PG and usual care.

## Aims and outline of this thesis

The general objective of this thesis is to evaluate the effectiveness and cost-effectiveness of an integrated multidisciplinary approach to the diagnosis and management of dementia. This thesis consists of three parts, each addressing different aspects of the objectives. The following research questions have been posed concerning these aspects:

### Part one

Chapter 2 describes the aims, design and methods of the MEDICIE-study (no research question)

#### *1. What is the added value of a multidisciplinary approach to dementia?*

The aim of a systematic review presented in chapter 3 is to give an overview of the reported empirical studies on the added value of a multidisciplinary approach to diagnosing dementia, i.e. whether differences can be found between diagnostic multidisciplinary approaches, and/or monodisciplinary ones with respect to the diagnostic outcomes, the impact of diagnostics on treatment decisions and their effects on quality of life and costs.

### Part two: Methodological issues

#### *2. Is the EQ-5D a suitable measure for evaluating Health Related Quality of Life (using proxy ratings) in a population with cognitive impairments?*

The aim of this explorative study (chapter 4) was to compare the performance of the EQ-5D and the EQ-5D+C (i.e. the EQ-5D extended with a cognitive dimension) by assessing their construct validity and responsiveness in patients aged 55 and older with cognitive impairments. We hypothesized that the construct validity and responsiveness of the EQ-5D+C were higher than the construct validity and responsiveness of the EQ-5D.

#### *3. Which factors are associated with the amount of informal care in dementia?*

#### *4. Which of the factors identified in research question 3 are associated with the different methods to measure and value informal care in our study (construct validity)?*

#### *5. To what extent are the methods to measure informal care sensitive to changes in these factors (responsiveness)?*

#### *6. Based on the findings from research questions 4 and 5, which of the methods is most suitable to measure and value informal care in this patient population and their caregivers?*

In chapter 5, several methods for the measurement of informal care are applied. For this, a questionnaire developed by the iMTA in Rotterdam, the Netherlands, is used. Results of three methods for measuring informal care are described (two monetary methods and one non-monetary method) and both construct validity and responsiveness of each of these methods is examined.

### **Part three: Findings of the MEDICIE-study**

#### *7. Is DOC-PG an effective facility for the diagnosis and management of dementia?*

We hypothesized that DOC-PG would have beneficial effects on health-related quality of life (HRQoL) compared with usual care, based on the assumption that both diagnosis of the cognitive disorder according to specialist guidelines and appropriate assessment of the patient's social circumstances are prerequisites for the best possible care for the patient and his/her family. The findings of this study are reported in chapter 6.

#### *8. Is DOC-PG a cost-effective facility for the diagnosis and management of dementia?*

Chapter 7 reports on the economic evaluation of a multidisciplinary integrated approach to the diagnosis and management of dementia alongside a randomised controlled trial. A cost-effectiveness analysis was performed in which DOC-PG was compared to usual care in the Netherlands.

#### *9. Are there differences between DOC-PG and usual care regarding several caregiver aspects?*

A questionnaire, developed by the iMTA in Rotterdam, the Netherlands was used to measure and value informal care, to assess caregiver characteristics (relationship with the patient, education, income, self-rated health, HRQoL, and burden) and to assess social context (caregiver network, formal care, other daily activities of the patients (that support the caregiver), respite care and paid leave from work. Chapter 8 describes differences in these caregiver outcomes between the intervention group and the control group.

#### *10. What is the level of GP concordance with advice for treatment after a multidisciplinary psychogeriatric assessment?*

#### *11. Are the GPs satisfied with the services provided by the DOC-PG?*

The objective of chapter 9 was to examine GP's level of concordance following a multidisciplinary psychogeriatric assessment by the DOC-PG. In particular we were interested in factors influencing concordance. Furthermore, a short satisfaction questionnaire was developed to identify the level of satisfaction regarding the services provided by the DOC-PG.

The main findings and implications of the various chapters are discussed in chapter 10.

# **PART ONE**



# Chapter 2

Aims, design and methods



## **Introduction**

The increasing number of elderly and increasing coexistence of somatic and psychological disorders in these elderly stress the need to create and evaluate integrated multidisciplinary health care services. We conducted a longitudinal study regarding the effectiveness and cost-effectiveness of an integrated approach to the diagnosis and management of dementia, entitled the Maastricht Evaluation of a Diagnostic Intervention for Cognitively Impaired Elderly (MEDICIE) study. In this chapter the aims, design and methods of the MEDICIE-study are described.

## **Aims of the study**

The main aim of the MEDICIE-study is to evaluate the effectiveness and cost-effectiveness of an integrated approach regarding the diagnosis and management of dementia. In this longitudinal study, 230 patients and their caregivers were assessed three times during the course of one year (i.e. baseline, six months and 12 months). The study is a combined effort of the departments of Psychiatry & Neuropsychology and Internal Medicine of the University Hospital of Maastricht, and the department of Old Age of the Community Mental Health Service (CMHT) of Maastricht.

## **Design**

The MEDICIE study was approved by the Medical Ethics Committee of Maastricht University Hospital. Patients were recruited from July 2002 to August 2004 from 60 general practices in the Maastricht region, 7 practices in the Sittard region, and 3 practices in the east Heerlen region (all three regions are in the province of Limburg, in the south of the Netherlands). GPs in these practices were asked to refer all patients with possible dementia or a cognitive disorder. The inclusion and exclusion criteria are given in the Box below.

## **Randomization**

Randomization took place at the general practice level. Randomization on patient level may result in contamination between groups because of a learning effect of PG-consultations in previously referred patients on the GP's diagnostic skills, and was therefore not preferred. Likewise, randomization on GP level would raise methodological problems because contamination within a practice cannot be precluded, and because GP's working on a part-time basis share responsibility of the same patients. In order to control for effects related to differences in general practices, all practices were asked to supply information about the practitioner's experience, demographic characteristics of the practice population, and the practitioner's affinity with geriatric problems. On the basis of these data, two groups of practices were formed, and the patients from these

**Trial inclusion and exclusion criteria****Inclusion criteria:**

Aged 55 years or older;  
 Suspected of having dementia or a cognitive disorder;  
 No referral to other local/regional services in the last two years; and  
 Availability of a proxy (visiting the patient at least once a week).

**Exclusion criteria:**

Acute disorders that need a prompt therapeutic intervention; and  
 Living in a nursing home.

practices were randomly assigned (by means of a computer program, randols) to either the intervention group (DOC-PG) or the control group (usual care). GPs were initially blinded to this procedure and the random allocation sequence was concealed for most of the participants.

**Interventions***Diagnostic Observation Centre for PsychoGeriatric patients (DOC-PG)*

The DOC-PG integrates a disease-oriented diagnostic approach from hospital departments with a care-oriented approach of the department of Old Age of the Regional Institute of Community Mental Health (CMHT). The use of a day hospital facility is chosen because it provides an adequate venue for a programmed investigative package, which might be otherwise impractical to perform in the outpatient clinic. For instance, a detailed somatic and neurological examination, laboratory and CT scan could be arranged through the outpatient departments, but it is altogether easier for an elderly patient to attend the day hospital on two occasions. The DOC-PG has expertise in the fields of old age psychiatry, geriatric medicine, neuropsychology, physiotherapy, occupational therapy, geriatric nursing, and mental health nursing and hence enables multidisciplinary assessment of patients, covering aspects such as somatic screening, psychogeriatric assessment, and evaluation of the required levels of care for the patient and his/her carer. GPs can refer patients to the DOC-PG if a cognitive disorder is suspected. During a 2-week diagnostic screening procedure, patients are first visited at their homes by the community nurse of the CMHT to assess the level of present and needed care, and to assess the capacities of the patient's carers. Subsequently, patients (and their carers) are asked to visit the University Hospital departments of geriatric medicine (for intake and somatic and neurological assessment) and geriatric psychiatry (for neuropsychological assessment, and assessment of functioning). In addition, a CT scan and various blood tests are performed. The results are then discussed at a weekly interdisciplinary meeting in which a definite diagnosis is made and a treatment plan is formulated. The patient's GP is sent a summary of the assessments, the multi-axis diagnosis, and recommendations for treatment and management. Thereafter the GP is responsible for his/her patients even though further investigations may have been recommended.

### *Usual care*

In the control group, GPs provided care as usual. This means that the patients were not referred to the DOC-PG and that either the diagnosis was made by the GP or the patient was referred to one of the separate regional services, e.g. the Maastricht Memory Clinic, geriatric medicine, or the department of mental health for the elderly at the mental health community service (Verhey et al., 1993).

## **Methods**

Informed consent was obtained from all participants. Trained interviewers assessed participants at their homes at baseline (within 2 weeks of the DOC-PG or usual care intervention) and at 6 and 12 months after the baseline measurement. All outcome measures, except the Mini Mental Scale Examination (MMSE), were collected through personal interviews with the patient's proxy. The assessments took approximately two hours. The measures used are described in the following paragraph.

### **Measures**

#### *Health-Related Quality of Life*

Health-related quality of life (HRQoL) was the primary outcome of this study for which we chose to use the EQ 5D. This instrument has been validated in a number of European countries including the Netherlands (Brooks, 1996) and provides a simple descriptive profile and a single index value for health status. It is widely used in cost-utility analyses (Dolan, 1997; Kind, 2003). The EQ-5D describes health status according to five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has three levels, namely, "no problems", "some problems" and "severe problems". Besides the five dimensions, the EQ-5D consists of a visual analogue scale ranging from 0 (worst imaginable health state) to 100 (best imaginable health state).

We also used the SF-36, a 36-item generic questionnaire used to measure nine relevant aspects of the health-related functioning of patients: physical functioning, social functioning, physical problems, emotional problems, mental health, pain and general health perception (Brazier et al., 1992; VanderZee et al., 1996; Ware et al., 1992). Higher scores reflect better functioning. The scale has been translated in Dutch and validated in a population ageing from 18- 89 years in the Netherlands (Zee vd et al., 1993).

#### *Cognitive functioning*

The Mini Mental State Examination (MMSE) was used to detect cognitive impairment, to assess its severity and to monitor cognitive changes over time (Folstein et al., 1975b). The MMSE has a maximum score of 30 points, with different domains being assessed: orientation in regard to time and place (10 points), registration of three words (3 points), attention and calculation (5 points), recall of three words (3 points), language (8 points), and visual construction (1 point). Scores below 24 are considered abnormal and this is the cut-off used for dementia. Scores in the MMSE are often classified into different categories: 26-30 (normal ageing), 21-25 (mild dementia), 15-20 (moderate dementia),

10-14 (moderately severe dementia) and 0-9 (severe dementia). The MMSE has demonstrated validity and reliability in geriatric, psychiatric, neurological and other medical populations (Wagner et al., 1995), also in the Netherlands (Kok et al., 2002). The Global Deterioration Scale (GDS) (Reisberg et al., 1982) evaluates the global level of deterioration in patients with a primary degenerative dementia such as Alzheimer's disease. It consists of seven stages of global functioning ranging from no cognitive deterioration at all (stage 1) to severe cognitive and functional deterioration (stage 7).

#### *NeuroPsychiatric Disturbances*

Patient's behavioural and psychological problems were measured by the Neuropsychiatric Inventory (NPI) (Cummings, 1997). This questionnaire assesses 12 neuropsychiatric disturbances common in dementia: delusions, hallucinations, agitation, dysphoria, anxiety, apathy, irritability, euphoria, disinhibition, aberrant motor behaviour, nighttime behaviour disturbances, and eating disturbances. The NPI distinguishes symptom frequency and severity. It facilitates rapid behavioural assessment by using screening questions. Information for the NPI is obtained from a caregiver familiar with the patient's behaviour by means of a structured interview (Cummings et al., 1994). The scale has been translated in Dutch and has been validated (Kat et al., 2002).

#### *Functional capabilities*

The instrumental Activities of Daily Living scale (IADL) measures seven areas of more complex activities required for optimal independent functioning, with scores reflecting whether patients are completely independent, in need of assistance, or are completely dependent on others for the performance of specific activities (Lawton et al., 1969).

#### *Emotional functioning*

The Cornell Scale of Depression in Dementia (CSDD) is a 19-item depression scale that was developed specifically to measure the severity of depressive symptoms in older adults with dementia. Each item is rated for severity on a scale of 0-2 (0=absent, 1=mild or intermittent, 2=severe). The item scores are added. Scores above 10 indicate a probable major depression. Scores above 18 indicate a definite major depression. Scores below 6 are associated with absence of significant depressive symptoms (Alexopoulos et al., 1988; Kurlowicz et al., 2002).

#### *Informal care*

Informal care was measured by means of a questionnaire, developed by the iMTA in Rotterdam, the Netherlands. This questionnaire included three methods to measure and value informal care. The proxy-good-method (van den Berg et al., 2006) (also called the market cost method or replacement cost method) values the measured time spent on providing informal care at the (labour) market price of a close substitute. In the opportunity-cost-methods (van den Berg et al., 2006), we valued the lost time as the opportunity cost (i.e. hours of paid work, unpaid work or leisure time forgone), or the value of the time in its current best alternative use (hours of paid work, unpaid work or leisure time). The CarerQoL-method (Brouwer et al., 2006) provides us with a non-

monetary outcome regarding the measurement and valuation of informal care. It is aimed at measuring the care-related quality of life of informal caregivers.

### *Caregiver burden*

We used the self-rated burden scale (SRB) to measure subjective burden (Brouwer et al., 2004; van Exel et al., 2005; van Exel et al., 2004). The SRB consists of a single question: “How burdensome is the provision of informal care to you at this moment?” and caregivers are asked to indicate their burden on a horizontal visual analogue scale ranging from 0 (“not at all straining”) to 10 (“much too straining”). The SRB was part of the informal care questionnaire mentioned above. Subjective burden was also assessed using the Perceived Stress caused by Informal Caregiving scale (Dutch translation: *Ervaren Druk door Informele Zorg (EDIZ)*) (Pot et al., 1995). The EDIZ is a 9-item measurement to assess self-perceived pressure from informal care.

### **Data-management**

Personal information of the patients was entered in an Access database and each patient was assigned an individual number. This database was password protected. Data from the collected measurements were entered in SPSS. These data were checked randomly to ensure data integrity. SPSS, STATA and Excel were used for statistical analyses.

# Chapter 3

The added value of a multidisciplinary approach  
in diagnosing dementia: a review



Published in *International Journal of Geriatric Psychiatry*, 2006; 21: 223-232

Claire A.G. Wolfs, Carmen D. Dirksen, Johan L. Severens and Frans R.J. Verhey

## Abstract

**Background:** The impact of dementia constitutes a major public health challenge to our society. A multidisciplinary approach to dementia is generally recommended.

**Objective:** The objective of this review is to give an overview of the reported empirical studies on the added value of a multidisciplinary approach to diagnosing dementia, i.e., whether differences can be found between diagnostic multidisciplinary approaches, and/or monodisciplinary ones with respect to the diagnostic outcomes, the impact of diagnostics and their effects on quality of life and costs.

**Method:** Critical review of studies which involve a multidisciplinary assessment of dementia. Literature was systematically searched in a number of international databases (Medline, PsychInfo, Pubmed and EconLit). A first selection was based on screening titles by one author. Interrater reliability was determined by scoring all abstracts by two authors. The reliability of selecting full articles was based on scoring a random sample by two authors. Kappa values were calculated. The Kappa statistic ( $k$ ) was also used as an indicator of the diagnostic agreement between single disciplines and multidisciplinary teams.

**Results:** Five studies were identified as relevant for the purpose of this review. The most important outcome was the level of diagnostic agreement between single disciplines and multidisciplinary teams (MTs). Overall, there was substantial agreement (kappa: 0.37-0.76) on the diagnosis dementia (syndrome), but not on the subtypes (aetiology) (AD: 0.26-0.60; VD: 0.35-0.52; depression: 0.30-0.46). Diagnoses also differed between the disciplines.

**Conclusions:** Merely studies on diagnostic accuracy were found, all using different reference standards. Kappa values were chosen as outcome measure and every diagnosis was considered of equal value. Therefore, the added value could not definitively be demonstrated. However, the potential added value of a multidisciplinary approach to diagnosing dementia probably lies in the ability to differentiate among the subtypes of dementia and to detect comorbid psychiatric conditions such as depression. Disciplines within an MT complement each other in establishing a specific diagnosis, which is considered to be the added value of such a multidisciplinary team.

## Introduction

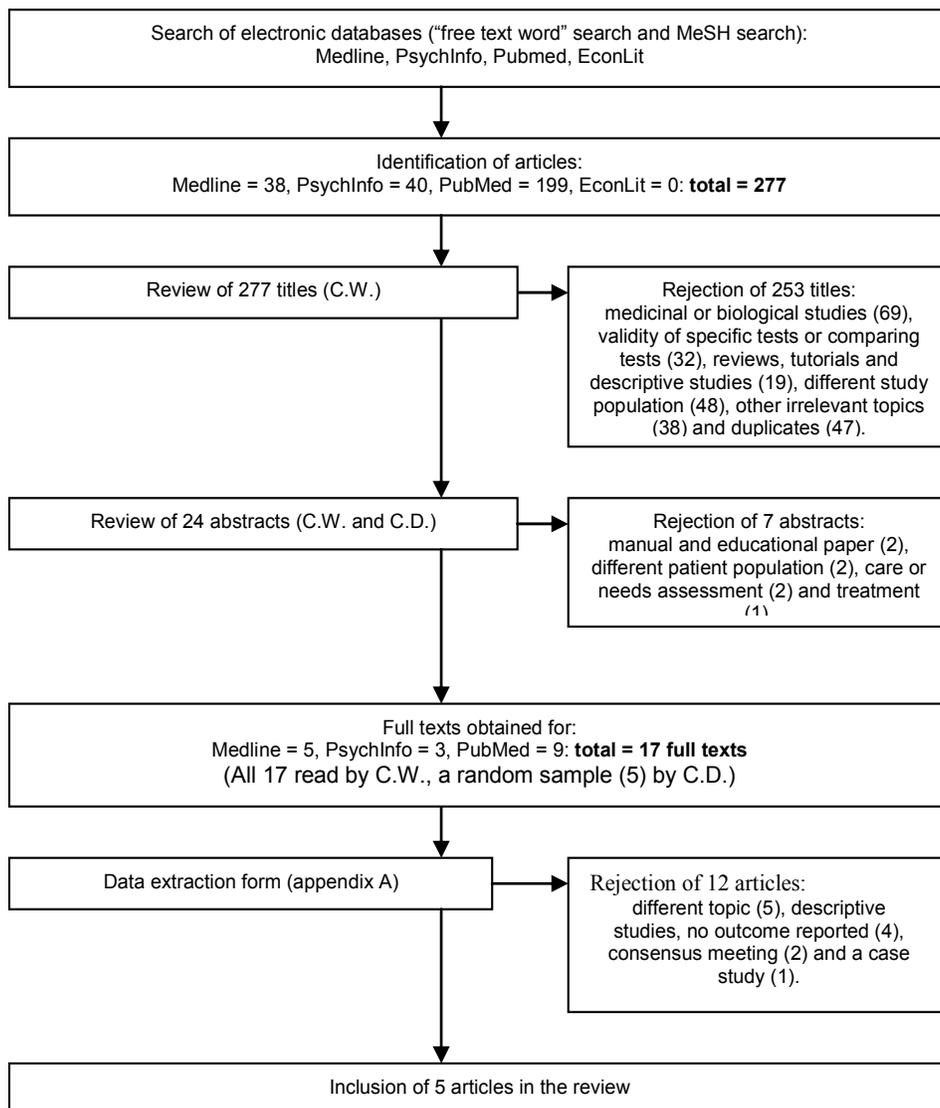
Elderly people with cognitive impairments suffer from a complex range of mental, physical and social problems. Early detection of dementia is a complicated challenge because of this multiplicity of underlying aetiologies (Hogh et al., 1999). A multidisciplinary approach to diagnosing and managing dementia is generally recommended (APA, 1997; CBO, 2005), as there is no single profession which is adequately equipped to deal with all of these on its own (Collighan et al., 1993). Multidisciplinary memory clinics (MCs) constitute a relatively new concept in psychogeriatrics. They can play an important role in the diagnosis, treatment and scientific research on the early and differential diagnosis of dementia (Bayer et al., 1987; Van der Cammen et al., 1987; Verhey et al., 1993). Particularly, the MCs are involved in a variety of clinical studies (Hejl et al., 2002; Hogh et al., 1999; Vraamark Elberling et al., 2002). Although the validity of certain diagnostic tests and techniques used in the multidisciplinary settings has been amply investigated (Baillon et al., 2003; Schramm et al., 2002; van Hout et al., 2001; Walstra et al., 1997), data on the added value of a multidisciplinary approach as such with respect to dementia appear to be sparse.

The aim of this systematic review is to give an overview of the reported empirical studies on the added value of a multidisciplinary approach to diagnosing dementia, i.e., whether differences can be found between diagnostic multidisciplinary approaches, and/or monodisciplinary ones with respect to the diagnostic outcomes, the impact of diagnostics on treatment decisions and their effects on quality of life and costs.

## Methods

### Search strategy

A systematic literature search was conducted in Medline, PsychInfo, PubMed and Econlit to identify articles published in medical journals reporting on research regarding multidisciplinary diagnostic assessments in the elderly. An advanced search was performed using the free text words “multidisciplin\*” or “memory clinic”, as well as the terms “diagnos\*” or “assessment” or “evaluation”. Furthermore the terms “dement\*” or “psychogeriatr\*” or “cognitive disorder” or “cognitive impairment” or “memory disorder” or “memory impairment” were utilised to include those subjects relevant to the present study. The free text words dementia and diagnosis were also used as Medical Subject Headings. Subsequently, the results were limited to studies that were comparative, published in English and included an abstract. The search eventually resulted in a total of 277 articles. Before reviewing the abstracts, criteria were established to determine whether a study was relevant to the purpose of the present review: 1) the study had to use a multidisciplinary approach regarding assessment, diagnosis or evaluation in people who are suspected of suffering from dementia 2) the study had to consider some kind of outcome in which effects (particularly the added value) of a multidisciplinary approach were described.

**Box 1: Flowchart showing the process of inclusion and exclusion of articles****Selection of studies**

Box 1 shows the process of the inclusion and exclusion of articles. One reviewer (C.W.) assessed all articles for inclusion or exclusion based on the title. In case of doubt she also read the abstract of the article. Of the 277 references found, the reviewer excluded 206 as non-relevant in terms of subject matter. The main reason for this large number of excluded articles was that, while many studies used participants from a (multidisciplinary) memory clinic population, these studies involved topics that were not relevant to the purpose of this review (see box 1). A total of 47 duplicates were also

excluded. Two reviewers (C.W. and C.D.) independently assessed the relevance of the remaining 24 articles by reading the abstracts. They selected those abstracts for which full text paper retrieval was appropriate. Both reviewers excluded 7 studies. A good interrater agreement (kappa) of 0.75 was reached. In case of discrepancies between the reviewers, the articles (4) were read in full. One reviewer (C.W.) read all 17 articles in full, while the second reviewer (C.D.) read a random sample (5) of the 17 full articles. To increase the objectivity and consistency of the decision to include or exclude a paper based on reading the full text, a data-extraction form was developed (appendix A). Besides selection, this form was used to score all relevant items with regard to study characteristics, composition of the teams, diagnostic assessments, results and conclusions. There was full agreement on the random sample of articles read between the two reviewers. A total of 12 articles were excluded from the review.

### **Analysis**

From the results section of the data extraction form, various measures (namely sensitivity, specificity and percentage agreement) describing the level of diagnostic agreement were converted into a Kappa statistic ( $k$ ). This measure was chosen as the unit of diagnostic agreement, because different reference standards were used in the various articles. We therefore decided to ignore all reference standards and to consider every diagnosis of equal value, irrespective of who formulated this diagnosis (a nurse or a multidisciplinary memory clinic). The unweighted kappa statistic was used because our diagnostic categories were merely nominal. Furthermore, we pooled the kappas for the individual articles to enhance the precision of the overall estimate using the following formula:  $K_{all} = \sum K_i * w_i / \sum w_i$  (Rothman, 1986) in which  $K$  is the Kappa value and  $w$  is the weight ( $1 / \text{variance}$ ). We also calculated the confidence intervals;  $R^{\wedge}D + 1.96 S^{\wedge}D(R^{\wedge}D)$  in which  $R^{\wedge}D$  is the point estimate (pooled kappa) and  $S^{\wedge}D(R^{\wedge}D)$  is the estimated standard deviation of the point estimate (Rothman, 1986).

## **Results**

### **General remarks on methodological aspects**

Five studies were identified as relevant for the purpose of this review i.e. Collighan et al., 1993 (Collighan et al., 1993); Dennis et al., 1998 (Dennis et al., 1998); Ferran et al., 1996 (Ferran et al., 1996); van Hout et al., 2000 (van Hout et al., 2000) and Verhey et al., 1993 (Verhey et al., 1993). The selected articles consisted of four prospective case series and one retrospective cohort study. In all of the selected studies comparisons of diagnostic assessments between multidisciplinary teams and/or single disciplines were made. Tables 1a and b describe the main characteristics of the selected studies.

**Table 1a:** Main characteristics of the selected studies.

	<b>Collighan et al.</b>	<b>Dennis et al.</b>	<b>Ferran et al.</b>
<b>Research question</b>	What is the level of agreement between diagnoses made by multidisciplinary teams (MTs) (professionals other than doctors) compared with diagnoses made by psychiatrists?	Is a Community Psychiatric Nurse (CPN) able to diagnose early dementia and differentiate dementia subtypes in a memory clinic population?	What are the clinical characteristics and level of service use in a cohort of 200 patients suspected of having presenile dementia (assessed by a multidisciplinary team)?
<b>Study design:</b>			
Type of study	Prospective case series	Prospective case series	Retrospective cohort study
Comparative study	Yes	Yes	Yes
Duration study	11 months	Not stated	36 months
Follow-up	If necessary	No	3 to 23 months
Care setting	Inner London health district	Outpatient hospital and patient's home	Outpatient hospital
<b>Study population:</b>			
Number of participants	100 (30 male, 70 female)	61 (27 male, 34 female)	200 (116 male, 84 female)
Mean age (SD)	78.2 (6.6)	67.7 (10.8)	56.0 (9.3)
Patient characteristics	People in need of psychogeriatric assessment.	People suspected of having dementia (e.g. Alzheimer's disease) or Mild Cognitive Impairment (MCI).	People suspected of having dementia with onset before the age of 65 (presenile or early onset dementia).
In- and exclusion criteria	Not stated	Not stated	Not stated
<b>Main outcome</b>	Agreement between diagnoses MT and psychiatrist (who consulted 2 psychogeriatricians).	Agreement between diagnoses by nurse (who consulted a psychiatrist), ICD-10 and MC.	Agreement between diagnoses by referrer and MT.
Outcome measure	% Agreement, sensitivity, specificity and kappa	% Agreement, sensitivity, specificity and kappa	Sensitivity and specificity
Reference standard	Diagnoses by psychiatrists.	ICD-10 diagnoses	Diagnoses by MT
Source of referral	Open referral system	GPs, Psychiatrist, CPNs and consultant physicians	GPs, Psychiatrists, Neurologists
Country of origin	United Kingdom	United Kingdom	United Kingdom

**Table 1b** Main characteristics of the selected studies.

	van Hout et al.	Verhey et al.
<b>Research question</b>	Are General Practitioners (GPs) able to diagnose dementia and identify the type of dementia accurately and confidently?	What are the differences in diagnostic outcome between a systematic multidisciplinary approach and a monodisciplinary one in common clinical practice?
<b>Study design:</b>		
Type of study	Prospective case series	Prospective case series
Comparative study	Yes	Yes
Duration study	16 months	62 months
Follow-up	No	No
Care setting	Outpatient hospital	Outpatient hospital
<b>Study population:</b>		
Number of participants	107 (42 male, 65 female)	430 (242 male, 188 female)
Mean age (SD)	74.0 (8.0)	61.7 (not stated)
Patient characteristics	People suspected of having dementia.	People suspected of having dementia or having memory complaints associated with aging.
In- and exclusion criteria	Inclusion: >55 years, cognitive impairment. Exclusion: Not stated	Inclusion: Referred to memory clinic. Exclusion: Not stated
<b>Main outcome</b>	Agreement between diagnoses by GPs and memory clinic (MC).	Agreement between diagnoses by referrer and memory clinic.
Outcome measure	Kappa	Sensitivity and specificity
Reference standard	Diagnoses by MC	Diagnoses by MC
Source of referral	GPs	GPs, Neurologists, Psychiatrists, Geriatricians and others
Country of origin	The Netherlands	The Netherlands

### Multidisciplinary teams

The different disciplines involved in the multidisciplinary teams are summarized in table 2. In 3 of the 5 selected studies a psychiatrist, a neuropsychologist and/or a psychologist and a neurologist were part of the multidisciplinary team. A geriatrician belonged to the team in 2 studies. Furthermore some teams comprised of a registrar, a social worker, a psychiatric nurse, an occupational therapist, a speech therapist, a neurosurgeon, a consultant, a clinical assistant and a senior house officer. On the whole, the

multidisciplinary teams consisted of at least 3 disciplines (Dennis et al., 1998; Ferran et al., 1996; van Hout et al., 2000; Verhey et al., 1993) to a maximum of 11 disciplines (Collighan et al., 1993) involved in the diagnostic assessment of dementia. However, in Collighan's and Ferran's study, not every patient was assessed by the complete multidisciplinary team. Collighan (1993) described the cooperation of professionals other than doctors involved in the diagnostic process. A team member carried out initial assessments, while diagnoses were made during multidisciplinary team meetings.

**Table 2:** Disciplines involved in the multidisciplinary assessment teams (% of patients assessed).

	Collighan	Dennis	Ferran	van Hout	Verhey
Psychiatrist		100%	64%		100%
(Neuro)psychologist			92%	100%	100%
Neurologist			88%	100%	100%
Geriatrician		100%		100%	
Occupational therapist	18%				
Social worker	25%				
(Psychiatric) nurse	24%				
Registrar	20%				
Other	13%	100%			

### Multidisciplinary diagnostic assessments

The diagnostic assessments used in the selected studies are outlined in tables 3a and b. The assessments performed by the multidisciplinary teams, as well as the assessments in the comparative (monodisciplinary) conditions, were described if possible. All multidisciplinary diagnostic procedures took place in an outpatient care setting. The comparative assessments were conducted at home (Dennis et al., 1998), at the GP's practice (Ferran et al., 1996; van Hout et al., 2000; Verhey et al., 1993) as well as in an outpatient setting (Collighan et al., 1993; Ferran et al., 1996; Verhey et al., 1993). The patient's history, as part of the assessment procedure, was described in all studies, although the data were collected in different ways. Standardized physical, neurological, psychiatric, and neuropsychological examinations were all performed in two studies (Dennis et al., 1998; Verhey et al., 1993). The three remaining studies (Collighan et al., 1993; Ferran et al., 1996; van Hout et al., 2000) carried out at least two of these examinations. More examinations were done whenever it was considered necessary. Imaging techniques (CT and MRI), as well as routine blood tests, were commonly used. Multiple criteria were used to diagnose dementia, i.e., DSM-III, DSMIII-R and DSM IV (APA, 1980, 1987, 1994), ICD-10 (W.H.O., 1993), NINCDS-ADRDA (McKhann et al., 1984), Newcastle criteria for Lewy body dementia (McKeith et al., 1992) and the Dutch Dementia Guidelines (De Bruyne et al., 1991.). Collighan used unclear broad diagnostic categories. Every team held a multidisciplinary meeting to reach consensus on the (etiological) diagnosis.

**Reported outcomes of the selected studies**

The most important outcome measure in the selected articles were differences and similarities in diagnoses between multidisciplinary teams and/or single disciplines. Tables 4a and b (outcomes) & 5 (pooled kappa values and confidence intervals) summarize the results of the studies. The strength of agreement was described according to the Landis and Koch (Landis et al., 1977) classification. Dennis et al. (1998) demonstrated that there was substantial agreement on the diagnosis of the dementia syndrome between the community psychiatric nurse (CPN) and the Memory Clinic (MC) ( $k=0.76$ ) and between the CPN and the ICD-10 ( $k=0.75$ ), but fair to moderate agreement on the aetiology (AD;  $k=0.60$ ,  $k=0.46$  respectively; VD;  $k=0.43$ ,  $k=0.29$  respectively).

Three studies (Ferran et al., 1996; van Hout et al., 2000; Verhey et al., 1993) compared the referrer's diagnoses with those of multidisciplinary teams. Again, there was fair to substantial agreement on the syndromal diagnosis of dementia between the teams and the referrers (kappa: 0.37-0.63), but only slight to moderate agreement on the aetiology (kappa: AD: 0.12-0.73; VD: 0.35-0.52; depression: 0.07-0.67). Agreement on these etiological diagnoses varied among the single disciplines and this variation was different for the three studies. There were no data available on the syndrome diagnosis in Ferran's study and it should also be noted that the patients in this study suffered from a distinct segment of dementia (i.e. presenile dementia). Therefore the patients of this study were different from those in the other studies. Collighan et al. (1993) demonstrated substantial to almost perfect agreement between a multidisciplinary team (consisting of professionals other than doctors) and a psychiatrist (dementia;  $k=0.88$ , depression;  $k=0.93$  delirium;  $k=0.66$ ), and moderate agreement between a team of doctors and a psychiatrist (dementia;  $k=0.55$ , depression;  $k=0.73$  delirium;  $k=0.47$ ). However, no aetiological diagnoses were made, merely syndromal ones (dementia).

In order to achieve an overall picture of the levels of agreement, the data were pooled where possible (table 5). The kappa values of Collighan et al., (1993) were left out since the multidisciplinary team differed too much from the other teams. There was substantial agreement between monodisciplinary diagnosticians and multidisciplinary teams on the dementia diagnosis (kappa: 0.37-0.76), but not on the aetiology (AD: 0.26-0.60; depression: 0.30-0.46). Overall, neurologists and the multidisciplinary teams agreed most on the AD diagnosis, whereas most agreement with psychiatrists and GPs was noticeable for depression and VD.

Since the required raw data in the study of Ferran et al. (1996) were no longer available, we calculated standard deviations, variances and weights based on the disease prevalence of the total study population (all referrers) instead of the prevalence for each referrer. An additional result in Ferran et al's study (1996) was a higher admission rate and level of service use in the group assessed by the multidisciplinary team because of the high prevalence of severe behavioural symptoms in this group.

**Table 3a:** Diagnostic assessments in the selected studies

	Collighan et al.	Dennis et al.
<b>Diagnostic assessment MT:</b>		
History	Yes (patient and other)	Yes
Physical examination	Yes	Yes
Neurological examination	No	No
Neuropsychological examination	No	Yes
Psychiatric examination	Yes	Yes
CT/MRI	No	Yes (both, % not stated)
Blood tests	Yes	Yes
Additional	No	No
Diagnostic criteria	Not stated (some broad diagnostic categories)	ICD-10 and Newcastle criteria for Lewy body dementia
Duration of assessment	Not stated	Not stated
<b>Diagnostic assessment (comparison)</b>	Half of the patients assessed before team assessment, the other half after team assessment by research psychiatrist. Psychiatrist presented findings to 2 psychogeriatricians to reach consensus. Assessment by means of a research interview	Every patient assessed by CPN by means of semi-structured interview: Reasons for referral; Geriatric Depression Scale; Checklist for anxiety, psychosis and mania; Physical checklist; Checklist daily and social functioning; Mental examination. Discussion with psychiatrist to agree on diagnosis. Based on notes from MC, diagnoses were made according to ICD-10 criteria by another, (independent) psychiatrist.
Diagnostic criteria	Not stated	Not stated
Duration of assessment	Not stated	0.7 to 1 hours

**Table 3b:** Diagnostic assessments in the selected studies

	Ferran et al.	van Hout et al.	Verhey et al.
<b>Diagnostic assessment MT:</b>			
History	Yes (when available)	Yes (patient and other)	Yes (patient and other)
Physical examination	No	Yes	Yes
Neurological examination	Yes (88%)	Yes	Yes
Neuropsychological examination	Yes (92%)	No	Yes
Psychiatric examination	Yes (64%)	Yes	Yes
CT/MRI	Yes (96% CT or MRI, 9% both)	Yes CT, no MRI	Yes ( CT, % not stated)
Blood tests	No	Yes	Yes
Additional	No	Yes: Cognitive screening and functional assessment	Yes, ECG (100%), chest X-ray (100%) and EEC (when epilepsy is suspected),
Diagnostic criteria	ICD-10	DSM-IV, NINCDS-ADRDA	DSM-III, DSM-III-R and NINCDS-ADRDA
Duration of assessment	Less than one day	4.5 hours (within 3 days)	4 to 5 hours
<b>Diagnostic assessment (comparison)</b>	Every patient diagnosed by referrer. The referrer's assessment procedure was not stated. The initial diagnosis was revised after team assessment.	Every patient assessed by GP. Assessment by means of the Guideline on Dementia of the Dutch College of General Practitioners (DDG).	Every patient diagnosed by referrer. The referrer's assessment procedure was not stated.
Diagnostic criteria	Not stated	D.D.G.	Not stated
Duration of assessment	Not stated	0.6 hours	Not stated

## Discussion

The aim of this review was to give a systematic overview of the reported studies on the added value of a multidisciplinary approach to diagnosing dementia compared with a monodisciplinary approach.

Even though a multidisciplinary policy with regard to dementia is generally recommended in several practice parameters (APA, 1997; CBO, 2005), empirical evidence is very scarce. Only five studies were found on this topic, none of which were randomised controlled trials comparing the diagnostic outcomes, as well as the impact of diagnostics on treatment decisions and their effects on quality of life and costs. A first issue of interest was the composition of the MTs, especially since diagnoses made by individual

**Table 4a:** Outcomes in the different studies

	Collighan et al.	Dennis et al.
<b>Diagnoses</b>	<p><b>Team diagnoses, other than doctors (non-doctors) (n=65):</b> 49% dementia; 34% depression; 3% delirium; 14% other</p> <p><b>Team diagnoses, doctors (n=38):</b> 61% dementia; 32% depression; 5% delirium; 2% other</p> <p><b>Research diagnoses, psychiatrist (n=107):</b> 59% dementia; 29% depression; 3% delirium 9% other</p>	<p><b>CPN's diagnoses:</b> 60% dementia (all); 43% AD; 13% VD; 4% other dementia; 39% other</p> <p><b>MC diagnoses:</b> 56% dementia (all); 43% AD; 11% VD; 2% other dementia; 44% other</p> <p><b>ICD-10 diagnoses:</b> 66% dementia (all); 39% VD; 20% VD; 7% other dementia; 34 % other</p>
<b>Level of agreement: (Kappa (k))</b>	<p><b>dementia;</b></p> <p>non-doctors and psychiatrist :k=0.88 doctors and psychiatrist: k=0.55 team (all) and psychiatrist k=0.80</p> <p><b>depression;</b></p> <p>non-doctors and psychiatrist: k=0.93 doctors and psychiatrist: k=0.73 team (all) and psychiatrist k=0.88</p> <p><b>delirium</b></p> <p>non-doctors and psychiatrist: k=0.66 doctors and psychiatrist: k=0.47 team (all) and psychiatrist k=0.56</p>	<p><b>dementia (all);</b></p> <p>CPN and ICD-10: k=0.75 MC and ICD-10: k=0.73 CPN and MC: k=0.76</p> <p><b>AD;</b></p> <p>CPN and ICD-10: k=0.46 MC and ICD-10: k=0.80 CPN and MC: k=0.60</p> <p><b>VD;</b></p> <p>CPN and ICD-10: k=0.29 MC and ICD-10: k=0.69 CPN and MC: k=0.43</p>
<b>Other</b>	Greater experience, better performance	

disciplines also varied. Four studies in this review described the operation of a memory clinic team (Dennis et al., 1998; van Hout et al., 2000; Verhey et al., 1993) and an early onset dementia service (Ferran et al., 1996) that consisted mainly of psychiatrists, neuropsychologists and/or psychologists, neurologists and geriatricians. The team reported on in Collighan et al's study primarily comprised other professionals although a psychiatrist and a geriatrician were involved to assist in difficult cases. The second question addressed by this review concerned the assessments used in the selected studies. In general, the diagnostic assessments employed were very broad. Physical, neurological, neuropsychological and psychiatric examinations, neuroimaging and additional tests, including blood tests were all used to distinguish between the various types of dementia. Although the implementation of guidelines (AAN, 1994.; APA, 1997; CBO, 2005) is recommended, no standardized model based on particular guidelines was explicitly used.

Finally, we were interested in the major outcomes of the studies. Initially, our goal was to study diagnostic outcomes, the impact of diagnostics on treatment decisions as well as their effects on quality of life and costs. We aimed to achieve this by reviewing RCT's on

**Table 4b:** Outcomes in the different studies

	Ferran et al.	van Hout et al.	Verhey et al.
<b>Diagnoses</b>	<p><b>Initial diagnosis (referrer):</b> 27% AD; 7% VD; 4% depression; 28% unspecified dementia; 34% other</p> <p><b>After follow-up (team):</b> 27% AD; 17% VD; 18% depression; 9% unspecified ; 29% other</p>	<p><b>GP's diagnoses:</b> 67% dementia (all); 46% AD; 21% other dementia; 33% no dementia</p> <p><b>MC diagnoses:</b> 63% dementia (all); 37% AD; 26% other dementia; 37% no dementia</p>	<p><b>Referrer's diagnoses:</b> 26% dementia (all); 7% AD 7% VD; 12% other dementia; 10% depression; 64% no dementia (all)</p> <p><b>MC diagnoses:</b> 35% dementia (all); 20% AD 10% VD; 5% other dementia; 20% depression; 45% no dementia (all)</p>
<b>Level of agreement: (Kappa (k))</b>	<p><b>AD:</b> GP and MT: k=0.51 Psychiatrist and MT: k=0.60 Neurologist and MT: k=0.73</p> <p><b>VD;</b> GP and MT: k=0.52 Psychiatrist and MT: k=0.35 Neurologist and MT: k= -</p> <p><b>depression;</b> GP and MT: k=0.31 Psychiatrist and MT: k=0.07 Neurologist and MT: k=0.25</p>	<p><b>dementia (all);</b> GP and MC: k=0.50</p> <p><b>AD;</b> GP and MC: k=0.32</p> <p><b>other;</b> GP and MC: k=0.29</p>	<p><b>dementia (all);</b> GP and MC: k= 0.63 Psychiatrist and MC: k=0.37 Neurologist and MC: k=0.61</p> <p><b>AD;</b> GP and MC: k= 0.22 Psychiatrist and MC: k=0.12 Neurologist and MC: k=0.36</p> <p><b>depression;</b> GP and MC: k=0.49 Psychiatrist and MC: k=0.67 Neurologist and MC: k=0.31</p>
<b>Other</b>	> rates of service use, a high level of dependence.	GP's diagnostic confidence in 59% of all cases, 50% for AD	

this topic, since they provide the highest level of evidence. However, merely studies on diagnostic accuracy were found. Most important was the level of agreement between the MTs and single professions (GP, neurologist, psychiatrist or CPN), which we indicated by means of the kappa statistic (k). There was substantial agreement on the syndrome diagnosis (dementia) but only fair to moderate agreement on the etiological diagnoses between the single disciplines and the multidisciplinary teams. Several limitations of this review also need to be discussed. Each of the studies had a different design using different patient groups and, most importantly, different reference standards on which the diagnoses were based. It was therefore difficult to compare the studies and find an appropriate indicator of agreement. The kappa statistic enabled us to compare the levels of agreement between the diagnosticians, but this approach did not allow valid judgments about the correctness of the diagnoses. Therefore, we were not able to demonstrate an explicit added value of using a multidisciplinary approach. However, as a multidisciplinary approach to dementia is generally recommended, one may assume that diagnoses formulated by such multidisciplinary teams are more accurate than

monodisciplinary ones. Under these assumptions, the results of the present review indicate that the potential added value of a multidisciplinary assessment does not lie in

**Table 5:** Pooled kappa's [confidence intervals]

Agreement with		Dementia (all)	Alzheimer's dementia	Vascular dementia	Depression
GP	MT	3 & 4;	2 & 3 & 4;	2;	2 & 4;
		k=0.65 [0.53-0.77]	k=0.29 [0.20-0.38]	k=0.52 [0.31-0.72]	k=0.41 [0.30-0.52]
Psychiatrist	MT	4;	2 & 4;	2;	2 & 4;
		k=0.37 [0.17-0.57]	k=0.26 [0.12-0.40]	k=0.35 [0.10-0.60]	k=0.37 [0.23-0.51]
Neurologist	MT	4;	2 & 4;	-	2 & 4;
		k=0.61 [0.46-0.76]	k=0.53 [0.40-0.66]		k=0.30 [0.20-0.40]
CPN	MT	1;	1;	1;	-
		k=0.76 [0.52-1.01]	k=0.60 [0.35-0.85]	k=0.31 [0.07-0.57]	

1. Dennis et al.
2. Ferran et al.
3. van Hout.
4. Verhey et al.

< 0.00 poor  
.00 -.20 slight  
.21 -.40 fair  
.41 -.60 moderate  
.61 -.80 substantial  
.81 -1.00 almost perfect

(Landis & Koch, 1977)

diagnosing the dementia syndrome, but in the ability to differentiate the aetiology of dementia. Every specialist is competent in his or her own domain and this review underlines the notion that a multidisciplinary team is helpful in preventing diagnoses from being biased by the medical profession of the diagnostician (Plugge et al., 1991). Given the necessity of an etiological diagnosis of dementia for current pharmacotherapy (e.g. cholinesterase inhibitors), this finding is relevant for clinical practice.

A second limitation, again concerns the utilization of the kappa statistic. Since every article used a different reference standard, it was difficult to find a measurement of agreement suitable for making comparisons between the articles. The kappa statistic gives some indication as to whether the agreement is better than that predicted by chance alone, but it is difficult to interpret kappa across different circumstances or studies. When comparing heterogeneous data, multilevel analyses need to be used. However, to our knowledge, this technique is not available with respect to kappa values. This means that the pooled data should be interpreted cautiously. Nevertheless, when looking at the kappa values for each study separately (i.e. without pooling) the results are similar.

A third limitation is that only a few studies were identified and the nature and operation of the multidisciplinary teams differed among themselves. The heterogeneity among the multidisciplinary teams, as well as the variety among the disciplines involved with the index diagnoses, did not allow for fair comparisons between the studies. Diagnoses made by the various disciplines differed substantially within the studies. Therefore, the diagnostic input of each of these disciplines should be investigated more thoroughly. Furthermore, given an increasing level of sub-specialization within most disciplines in

general, it may be difficult to compare psychiatrists, neurologists and geriatricians with each other. Besides, although the outcomes in terms of diagnostic accuracy were described in depth, little information was provided on the relevancy of the diagnostic outcome for further treatment and on effects on health status and well-being.

The data available in the studies did not allow us to give satisfactory answers regarding the aim of this review. The focus was merely on the common diagnostic categories, whereas using a broad multidisciplinary approach, other factors contributing to impaired cognitive functioning, such as anxiety, depressive mood or alcohol abuse might be revealed (Verhey et al., 1993). Furthermore, none of the studies addressed the potential disadvantages of having diagnostic assessments done by multiple disciplines. Possible disadvantages are overlapping information, higher costs, unclear responsibilities and organizational complexity.

We conclude that there is a need for an efficient, standardized, multidisciplinary diagnostic service for the diagnosis and treatment for cognitively impaired elderly, especially in view of the increasing number of aged people. Therefore, we propose an interpretative diagnostic model that combines a disease-oriented approach with a care-oriented one, in which specific diagnostic guidelines on dementia such as those of the American Academy of Neurology (AAN, 1994.), American Psychiatric Association (APA, 1997) and the Dutch Institute for Healthcare Improvement or CBO (CBO, 2005) are implemented. Diagnostics should be based on input from various disciplines, using assessments on a variety of variables (containing mental, physical and social aspects) to classify the aetiology and reveal all relevant factors contributing to excess disability. This team should combine expertise from psychiatry, neuropsychology, neurology and geriatrics. Furthermore, the model should include regular re-evaluations of patients, caregivers and support systems so changes can be monitored (Patterson et al., 1990).

Future research should focus on clinical effects (amongst which excess disability) as well as other effects (quality of life), the cost-effectiveness of multidisciplinary diagnostic models regarding cognitive impairment in the elderly and potential disadvantages of having diagnostic assessments done by multiple disciplines. Results of the ongoing MEDICIE-study (Maastricht Evaluation of a Diagnostic Intervention for Cognitively Impaired Elderly), a randomised clinical trial addressing these matters, will be available in 2006 (Wolfs et al., 2005a).

#### *Acknowledgement*

The authors wish to thank Alfons Kessels for his statistical and methodological assistance.



# **PART TWO**

Methodological issues



# Chapter 4

Performance of the EQ-5D and the EQ-5D+C in elderly patients with cognitive impairments



Published in *Health and Quality of Life Outcomes*, 2007, 5:33 doi:10.1186/1477-7525-5-33

Claire AG Wolfs, Carmen D Dirksen, Alfons Kessels, Daniëlle CM Willems, Frans RJ Verhey and Johan L Severens

## Abstract

**Background:** The EQ-5D is a reliable tool for measuring Health-Related Quality of Life (HRQoL). However, concern has been expressed that it may ignore elements of HRQoL, particularly cognition. In response to this concern, the EQ-5D has been extended with a cognitive dimension (EQ-5D+C). The aim of this study was to compare the performance of the EQ-5D and the EQ-5D+C in elderly patients with cognitive impairments by assessing their construct validity and responsiveness.

**Methods:** Data from the MEDICIE study (n=196) were used, in which all questionnaires were rated by proxies.

**Results:** Regarding construct validity, we found similar correlations between the EQ-5D and the Mini Mental State Examination (MMSE) and between the EQ-5D+C and the MMSE. Furthermore, both the EQ-5D and the EQ-5D+C were responsive to changes in the MMSE, with the EQ-5D performing slightly better.

**Conclusion:** We conclude that the EQ-5D performs well for evaluating HRQoL in a population with cognitive impairments. Based on the results of this explorative study, it does not seem necessary to adjust the current classification system by adding a cognitive dimension. However, in order to compare both instruments regarding utility values, it is necessary to develop a new scoring algorithm for the EQ-5D+C by conducting a general population study. Considering the explorative nature of this study, it is recommended that more aspects of the validity of both the EQ-5D and the EQ-5D+C be explored in patients with cognitive impairments using a more tailored study design.

## Background

The increasing number of older adults who are diagnosed with dementia has far-reaching implications for health service delivery and expenditures (Holland et al., 2004b). Economic evaluations are performed more often to assist decision-makers in setting priorities, especially with regard to resource allocation (van Velden et al., 2005). A central component of economic evaluations in health care is the use of preference-based instruments (also called value-based instruments) to measure changes in Health-Related Quality of Life (HRQoL). Preference-based measures, such as the EQ-5D (Brooks, 1996), “EuroQol—a new facility for the measurement of health-related quality of life. The EuroQol Group”, 1990), the SF-6D (Brazier et al., 2002a) and the HUI (Feeny et al., 1995), are standardized multi-dimensional health state classifications (Brazier et al., 2004). For each of these instruments, health states have been valued using techniques such as standard gamble (SG) or time trade-off (TTO) (Brazier et al., 1999). These valuations were used for each instrument to generate a scoring algorithm of which a single utility score for each health state can be deduced.

The EQ-5D is commonly used to measure HRQoL and has been shown to be responsive, internally consistent and reliable in the normal population and other patient groups (Konig et al., 2002; Nowels et al., 2005) as well as in patients with dementia (Ankri et al., 2003; Selai, 1998). However, concern has been raised that it may ignore elements of HRQoL of specific relevance to the elderly such as vision and hearing (Holland et al., 2004b) and in particular cognition (Bryan et al., 2005; Hoeymans et al., 2005; Holland et al., 2004b; Jonsson et al., 2006a; Krabbe et al., 1999b; Selai, 2001). It is known that cognitive problems have an impact on personality, mood, behaviour and global functioning (APA, 1997), which are domains covered by the EQ-5D, but cognition might also be regarded as a separate dimension.

In response to the concern that the EQ-5D ignores cognition, the EQ-5D has been extended with a cognitive dimension (EQ-5D+C) (Krabbe et al., 1999a). In this study of Krabbe et al. (which was an adapted Dutch replication of the Global Burden of Disease study commissioned by the World Bank) (Murray et al., 1996), valuations (by means of a rating scale) elicited from EQ-5D+C descriptions were compared empirically with parallel EQ-5D descriptions in Dutch faculty members (i.e. scientific staff members and management members of the Department of Public Health, the Department of Clinical Epidemiology and Biostatistics and the Institute of Social Medicine).

The EQ-5D+C generated different values compared with the EQ-5D. Whereas the content validity of the EQ-5D improved by adding cognition, both versions evoked equally reliable values. Based on these results, the authors emphasized the importance of considering the inclusion of a cognitive dimension. Furthermore, the EQ-5D+C was used to describe the health status of the Dutch population and to investigate sociodemographic differences (Hoeymans et al., 2005).

In this study, the content validity also improved through the addition of the cognitive dimension, while the reliability remained unaltered. It was concluded by the authors that the EQ-5D+C is an efficient tool for establishing the health status in the community. Another way to examine if the EQ-5D should contain a cognitive dimension is to investigate the performance of the EQ-5D and the EQ-5D+C in a population with cognitive impairments. The aim of this explorative study was to compare the performance of the EQ-5D and the EQ-5D+C by assessing their construct validity and responsiveness in patients aged 55 and older with cognitive impairments.

## Methods

### Study population and data collection

Data were derived from the MEDICIE (Maastricht Evaluation of a Diagnostic Intervention for Cognitively Impaired Elderly) study. The MEDICIE study is a randomized controlled trial (RCT) comparing the effects of a multidisciplinary diagnostic observation centre for psychogeriatric patients (DOC-PG) with usual care on HRQoL, mental and physical health, and the costs and use of health care facilities by patients with psychogeriatric problems (Wolfs et al., 2005a). The DOC-PG is an outpatient facility, providing multidisciplinary assessment by somatic screening, psychogeriatric assessment, and evaluation of the required levels of care for the patient and his (her) carer. The main aim of the DOC-PG is to improve or maintain the HRQoL of patients.

In the MEDICIE study, a total of 234 patients and their caregivers agreed to participate and were included between July 2002 and October 2004. Randomization occurred at the level of general practices. The experimental group visited the new diagnostic facility (DOC-PG), whereas the control group was treated as usual, i.e. the GP made the diagnosis or referred the patient to a specialist facility, namely the Maastricht Memory Clinic (MMC) or the Department of Old Age Psychiatry of the Community Mental Health Service (RIAGG). Patients were followed up after 6 months and 12 months.

All outcome measures, except the Mini Mental State Examination (MMSE), were collected through personal interviews with the patient's proxy. After initial assessment by the aforementioned health care professionals, the baseline MMSE scores were gathered from the patient records. The researchers (C.W. and D.W) were trained to assess patients using the MMSE at the 6 and 12 month follow-up. When possible, follow-up scores by the professionals were used. Sociodemographic data of the patients (gender, age, living arrangements) and proxies (gender, age, relationship to patient) were collected at baseline. Diagnosis was established by the multidisciplinary teams working at the DOC-PG or the MMC/RIAGG respectively, and was based on the DSM-IV criteria or other regular criteria (McKhann et al., 1984). In this study, the baseline data and the data at the six and 12 month follow-up for the entire group were used since, for the purpose of this paper, it was not necessary to analyze the data of the control group and the experimental group separately.

## Instruments

### *Mini Mental State Examination (MMSE)*

The MMSE is used to detect cognitive impairment, to assess its severity and to monitor cognitive changes over time (Folstein et al., 1975b). The MMSE has a maximum score of 30 points, with different domains being assessed: orientation in regard to time and place (10 points), registration of three words (3 points), attention and calculation (5 points), recall of three words (3 points), language (8 points), and visual construction (1 point). Scores below 24 are considered abnormal and this is the cut-off used for dementia. Scores in the MMSE are often classified into different categories: 26-30 (normal ageing), 21-25 (mild dementia), 15-20 (moderate dementia), 10-14 (moderately severe dementia) and 0-9 (severe dementia). The MMSE has demonstrated validity and reliability in geriatric, psychiatric, neurological and other medical populations (Wagner et al., 1995), also in the Netherlands (Kok et al., 2002).

### *EQ-5D and EQ-5D+C*

The EQ-5D is a generic instrument to measure HRQoL. The instrument was developed and validated in a number of European countries including the Netherlands (Brooks, 1996; Lamers et al., (in press); Lamers et al., 2005). The EQ-5D describes health status according to five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has three levels, namely, “no problems”, “some problems” and “severe problems”. This yields 243 potential combinations of health states across the five dimensions. Dolan et al. (Dolan, 1997) have presented 42 of these health states to members of a representative sample of the UK general population, which were valued using the TTO method. Based on these valuations, utility scores can be deduced by means of an additive function. These are now widely used in cost-utility analyses (Kind, 2003). Utility scores can vary between -0.59 (worst health) and 1.00 (perfect health). Besides the five dimensions, the EQ-5D consists of a visual analogue scale (VAS<sub>5D</sub>) ranging from 0 (worst imaginable health state) to 100 (best imaginable health state).

The EQ-5D+C is an extended version of the EQ-5D that includes “cognitive functioning” (memory, concentration, coherence, IQ) as an additional dimension, with a similar operationalization of three levels (as described above) [15]. The EQ-5D+C also includes a VAS<sub>5D</sub>.

In this study, the EQ-5D was administered completely first, that is the five dimensions followed by the VAS<sub>5D</sub>. Subsequently, the proxies were asked to answer the sixth dimension concerning cognitive functioning, whereupon the VAS<sub>5D</sub> was valued a second time (in this study referred to as the VAS<sub>5D+C</sub>). Therefore, in this study, the EQ-5D+C refers to the additional cognitive functioning dimension and the VAS<sub>5D+C</sub>.

## Data analysis

### **Construct validity**

Construct validity, the extent to which an instrument correlates with other measures which it should be related to (Solberg et al., 2005), was estimated by studying

correlations between the EQ-5D and the MMSE and between the EQ-5D+C and the MMSE at baseline and follow-up measures. Although it was expected that effects of cognitive impairment were implicitly expressed in the EQ-5D dimensions concerning self-care, usual activities and anxiety/depression (APA, 1997; CBO, 2005), we hypothesized that the correlations between the EQ-5D+C and the MMSE were strongest.

### **Responsiveness**

In this study, responsiveness was defined as the correlation of the changes in an instrument to changes in other measures which it should be related to, using an anchor-based approach (Norman et al., 1997; Terwee et al., 2003). We evaluated whether changes in the EQ-5D and changes in the EQ-5D+C correlated with changes in the MMSE, the so-called anchor (external standard). Again, it was hypothesized that the correlations between the EQ-5D+C and the MMSE were stronger than the correlations between the EQ-5D and the MMSE.

### **Statistical analysis**

The software used for the analyses was SPSS version 12.0.1 and STATA version 8.2. Background characteristics of the participants (both the patients and their proxies) were summarized using descriptive statistics. Response distributions of the instruments (EQ-5D, EQ-5D+C and MMSE) are given. Missing data of the participants were imputed using multiple imputation (MI). MI provides a useful strategy for dealing with data sets with missing values. Instead of filling in a single value for each missing value, Rubin's (Rubin et al., 1991) multiple imputation procedure replaces each missing value with a set of plausible values (5) that represent the uncertainty about the right value to impute. This results in statistically valid inferences that properly reflect the uncertainty due to missing values.

Kolmogorov-Smirnov tests were used to test for normality. Non-parametric tests for comparisons were selected. Associations between the instruments were analyzed with Spearman rank correlations.

## **Results**

### **Sample characteristics**

Table 1 summarizes sample characteristics. Of the 234 patients that were included in the MEDICIE-study, 64.1% were females. Of the proxies, 66.7% were females and mostly children (-in-law) or spouses of the patient (90.2%). In most cases, dementia (present in 70.1% of the patients) was associated with Alzheimer's disease (41.5%). Patients whose etiological diagnoses could not be determined were assigned to the "other" groups (i.e. other dementia or other cognitive impairment). After six months, 16 patients (6.8%) had died and 11 patients (4.7%) and their caregivers had dropped out of the study. Four patients (1.7%) did not attend the six month follow-up because of personal reasons. After 12 months another 11 patients (4.7%) had died and two more patients (0.9%) and their

**Table 1:** Sample characteristics at baseline

<b>Total n = 234</b>	
<b>Gender of patient:</b>	
Female (%)	150 (64.1%)
<b>Age of patient:</b>	
Mean (SD)	77.8 (6.7)
Range	[55 – 94]
<b>Relationship proxy:</b>	
Spouse	88 (37.6%)
Child (-in-law)	123 (52.6%)
Other	23 (9.8%)
<b>Gender of proxy:</b>	
Female (%)	156 (66.7%)
<b>Age of proxy:</b>	
Mean (SD)	59.8 (13.9)
Range	[30 – 91]
<b>MMSE:</b>	
Mean (SD)	20.18 (5.8)
<b>Dementia:</b>	
	164 (70.1%)
Alzheimer's Disease (AD)	97 (41.5%)
Vascular Dementia (VD)	26 (11.1%)
Mixed Dementia	21 (9.0%)
Other Dementia	20 (8.5%)
<b>No Dementia:</b>	
	70 (29.9%)
Cognitive Impairment / MCI	40 (17.1%)
Other Cognitive Impairment	30 (12.8%)

caregivers had dropped out of the study. The 27 patients (11.5%) who had died were excluded from the analyses as well as the 11 study dropouts (4.7%) who completed only 1 measurement. Missing data of the remaining 196 patients were imputed using MI. To make sure the imputations did not influence our results, separate analyses were performed on the 5 imputed datasets. The results were highly comparable (data not shown). Therefore, the first imputed dataset was used for analysis.

**Table 2:** Responses in the EQ-5D and EQ-5D+C (items, utility scores and VAS scores)

EQ-5D	N = 196	Baseline	6 months	12 months
<b>Mobility:</b>				
Mean (SD)		1.65 (0.48)	1.65 (0.51)	1.67 (0.51)
% no problems		24.7%	36.7%	34.7%
% some problems		65.3%	61.7%	63.3%
% severe problems		0%	1.5%	2.0%
<b>Self-Care:</b>				
Mean (SD)		1.59 (0.65)	1.71 (0.73)	1.97 (0.80)
% no problems		50.0%	45.4%	33.2%
% some problems		40.8%	38.3%	36.7%
% severe problems		9.2%	16.3%	30.1%
<b>Usual Activities:</b>				
Mean (SD)		1.92 (0.72)	2.02 (0.75)	2.17 (0.78)
% no problems		29.6%	27.0%	23.0%
% some problems		48.5%	44.4%	36.7%
% severe problems		21.9%	28.6%	40.3%
<b>Pain/Discomfort:</b>				
Mean (SD)		1.59 (0.65)	1.51 (0.60)	1.54 (0.62)
% no problems		49.5%	54.1%	53.1%
% some problems		41.8%	40.8%	40.3%
% severe problems		8.7%	5.1%	6.6%
<b>Anxiety/Depression:</b>				
Mean (SD)		1.66 (0.67)	1.53 (0.59)	1.54 (0.68)
% no problems		44.9%	52.6%	56.6%
% some problems		43.9%	42.3%	33.2%
% severe problems		11.2%	5.1%	10.2%
<b>Utility score:</b>				
Mean (SD)		0.54 (0.31)	0.57 (0.32)	0.47 (0.34)
Range		-0.35 – 1.00	-0.35 – 1.00	-0.43 – 1.00
<b>VAS5D:</b>				
Mean (SD)		58.84 (19.00)	57.78 (17.67)	56.38 (20.29)
Range		10-100	20-100	10-100
<b>EQ-5D+C</b>				
<b>Cognition:</b>				
Mean (SD)		2.30 (0.58)	2.35 (0.64)	2.47 (0.61)
% no problems		6.1%	8.7%	6.1%
% some problems		58.2%	47.4%	40.8%
% severe problems		35.7%	43.9%	53.1%
<b>VAS5D+C:</b>				
Mean (SD)		49.45 (19.23)	48.43 (18.06)	45.41 (18.90)
Range		0 - 100	10 - 90	0 - 95

### Responses in the EQ-5D, EQ-5D+C and MMSE

The responses in the EQ-5D and the EQ-5D+C at baseline and follow-up measurements are summarized in table 2. At baseline as well as at the 6 month follow-up, most patients had problems with cognitive functioning, usual activities and mobility. At the 12 month follow-up, most patients had problems with cognitive functioning, usual activities and self-care. The mean VAS5D+C scores were significantly lower than the mean VAS5D scores at all measurements (Wilcoxon Signed Ranks Tests,  $p = 0.000$ ).

The responses in the MMSE at baseline and follow-up measurements are summarized in table 3. Most patients had mild to moderate dementia at all measurements.

**Table 3: Responses in the MMSE**

MMSE: N = 196	Baseline	6 months	12 months
Mean (SD)	20.21 (5.78)	18.89 (7.34)	17.61 (8.16)
<b>Severity of dementia:</b>			
26 – 30 (normal ageing)	18.4%	21.9%	21.9%
21 – 25 (mild)	35.7%	26.0%	19.4%
15 – 20 (moderate)	29.6%	25.0%	25.0%
10 – 14 (moderately severe)	10.2%	14.3%	15.3%
0 – 9 (severe)	6.1%	12.8%	18.4%

### Construct validity

Table 4 summarizes the results of the correlations between the MMSE and the EQ-5D and between the MMSE and the EQ-5D+C. At baseline, significant correlations were found between the MMSE and the utility score, and more specifically the self-care dimension and the usual activities dimension, and the VAS5D of the EQ-5D. Correlations were also found between the MMSE and the cognitive dimension and the VAS5D+C of the EQ-5D+C. At the six month follow-up, correlations were found between the MMSE and the utility score, and more specifically all five dimensions, and the VAS5D of the EQ-5D. Correlations were also found between the MMSE and the cognitive dimension and the VAS5D+C of the EQ-5D+C. At the 12 month follow-up, correlations were found between the MMSE and the utility score, and more specifically all dimensions except for the pain/discomfort dimension, and the VAS5D of the EQ-5D. Correlations were also found between the MMSE and the cognitive dimension and the VAS5D+C of the EQ-5D+C.

### Responsiveness

Table 5 summarizes the results of the correlations between the change scores of the EQ-5D and the EQ-5D+C and the change scores of the MMSE. In table 6, the means for the change scores are outlined. Regarding the difference between the six month measurement and the baseline measurement, correlations were found between changes in the utility score, more specifically the self-care dimension, of the EQ-5D and the change scores of the MMSE. Correlations were also found between changes in the cognitive dimension of the EQ-5D+C and change scores of the MMSE. Regarding the difference between the 12 month measurement and the baseline measurement,

correlations were found between changes in the utility score, more specifically the mobility dimension, the self-care dimension and the usual activities dimension, and the VAS<sub>5D</sub> of the EQ-5D and the change scores of the MMSE. Correlations were also found between changes in the cognitive dimension and the VAS<sub>5D+C</sub> of the EQ-5D+C and the change scores of the MMSE.

**Table 4:** Spearman correlations between the EQ-5D and the MMSE and between the EQ-5D+C and the MMSE at baseline, six months and 12 months (construct validity).

N = 196	Baseline	6 months	12 months
<b>EQ-5D:</b>			
Utility score	0.19**	0.45**	0.50**
Mobility	-0.02	-0.17*	-0.23**
Self-Care	-0.28**	-0.50**	-0.55**
Usual activities	-0.34**	-0.42**	-0.52**
Pain/Discomfort	-0.01	-0.14*	-0.05
Anxiety/Depression	-0.05	-0.17*	-0.22**
VAS5D	0.22**	0.23**	0.23**
<b>EQ-5D+C:</b>			
Cognition	-0.35**	-0.52**	-0.54**
VAS5D+C	0.37**	0.47**	0.48**

\*\* Correlation is significant at the 0.01 level (2-tailed)

\* Correlation is significant at the 0.05 level (2-tailed)

## Discussion

The aim of this explorative study was to compare the performance of the EQ-5D and the EQ-5D+C by assessing their construct validity and responsiveness in patients aged 55 and older with cognitive impairments.

Based on our results it can be concluded that the construct validity of the EQ-5D and the EQ-5D+C is comparable in our study population, except for the VAS<sub>5D</sub>. Results regarding construct validity of the EQ-5D are in line with the recent findings of Jönsson et al. (Jonsson et al., 2006a). Contrary to our expectations, correlations between the cognitive dimension and the MMSE were almost similar to the correlations between the self-care and the usual activities dimensions and the MMSE. The presence of more and stronger correlations of both the EQ-5D and EQ-5D+C with the MMSE at the 12 month follow up measurement was possibly due to the fact that the dispersion of the scores using these instruments increased with time. Three studies also showed that cognitive function was positively related to HRQoL in cardiac rehabilitation patients (Cohen et al., 1999), in patients with progressive supranuclear palsy (Schrag et al., 2003b) and in patients with hypertension (Degl'Innocenti et al., 2002). Another study (Banerjee et al., 2006) failed to find a relationship between HRQoL and cognition in patients with dementia.

**Table 5:** Spearman correlations between the change scores of the EQ-5D and the MMSE, and between the EQ-5D+C and the MMSE, i.e. change score of baseline and six months and baseline and 12 months (longitudinal)

Change scores N = 196	Baseline & 6 months	Baseline & 12 months
<b>EQ-5D:</b>		
Utility score	0.16*	0.30**
Mobility	-0.11	-0.17*
Self-Care	-0.18*	-0.35**
Usual activities	-0.12	-0.29**
Pain/Discomfort	-0.04	0.03
Anxiety/Depression	0.00	-0.12
VAS <sub>5D</sub>	0.01	0.17*
<b>EQ-5D+C:</b>		
Cognition	-0.21**	-0.28**
VAS <sub>5D+C</sub>	0.08	0.23**

\*\* Correlation is significant at the 0.01 level (2-tailed)

\* Correlation is significant at the 0.05 level (2-tailed)

With regard to responsiveness, the EQ-5D performed slightly better than the EQ-5D+C, which is also in line with the findings of Jönssen (Jonsson et al., 2006a). An important finding, again contrary to our expectations, is that changes in the MMSE corresponded better with changes in the self-care dimension and the usual activities dimension than with changes in the cognitive dimension.

However, no judgments were made about the strength of the correlations, which would provide us with a stricter criterion regarding the performance of the EQ-5D and EQ-5D+C. In the literature, different classifications were found (Haywood et al., 2002; Landis et al., 1977; Marra et al., 2005; Wu et al., 2002) a clear gold standard being absent. We therefore ignored the classifications and merely described our results. However, it is possible to compare our results with other studies. Our results were in line with correlations between the EQ-5D and clinical measures found in other studies involving diseases such as progressive supranuclear palsy (PSP) (Schrage et al., 2003a), rheumatoid arthritis (RA) (Marra et al., 2005) and stroke (Pickard et al., 2005).

The majority of authors ((Haywood et al., 2002; Marra et al., 2005; Pickard et al., 2005) and others) considered a Spearman's correlation of >0.50 to be strong, a correlation of 0.30 / 0.35 – 0.50 to be moderate and a correlation < 0.30 / 0.35 to be weak. Using these classifications in our study, it can be concluded that both versions performed well with respect to construct validity, as indicated by strong correlations with the MMSE. Regarding responsiveness, it can be concluded that the EQ-5D performed moderately, whereas the EQ-5D+C did less well as indicated by weak correlations with the MMSE. When the more stringent classification of Landis and Koch (Landis et al., 1977) is used (i.e. < 0.00 poor; 0.00-0.20 slight; 0.21-0.40 fair; 0.41-0.60 moderate; 0.61-0.80

substantial and 0.81-1.00 almost perfect), it can be concluded that the EQ-5D and the EQ-5D+C performed moderately with regard to the construct validity.

Regarding responsiveness, fair correlations were found between changes in the EQ-5D and EQ-5D+C and changes in the MMSE. The relatively low responsiveness of the EQ-5D in this study could be due to the, on average, small changes in cognition in a year, or to a ceiling effect because there are only three levels for each dimension of the EQ-5D. Patients' health may improve or decline but not enough to go up or down one level. Instruments that have a greater number of possible responses may be more responsive. Furthermore, it is possible that adaptation to illness on the part of the proxy leads to a lack of responsiveness, especially with a chronic condition such as dementia (Marra et al., 2005). It should also be noted that a lack of clarity exists with regard to the definition and adequate approach for evaluating responsiveness. Some authors argued that there is no need for an additional concept like responsiveness, since it can be viewed as either longitudinal validity or magnitude of the treatment effect (Lindeboom et al., 2005; Streiner, 2003; Terwee et al., 2003). The definition and approach used in this study has also been referred to as longitudinal validity (Terwee et al., 2003).

There are several limitations to this study that need to be recognized. An important limitation of this study concerns our study design. The origin of this study, the MEDICIE trial, was designed to compare the effects of a multidisciplinary diagnostic observation centre for psychogeriatric patients (DOC-PG) with usual care on HRQoL, mental and physical health, and the costs and use of health care facilities by patients with psychogeriatric problems. Therefore, studying the usefulness of the EQ-5D+C in this patient population was framed in this RCT. The EQ-5D was administered first, that is the five dimensions followed by the VAS<sub>5D</sub>. Subsequently, the proxies were asked to answer the sixth dimension concerning cognitive functioning, whereupon the VAS<sub>5D+C</sub> was valued. It would have been better to administer the EQ-5D+C completely as well in order to make valid comparisons between the 2 versions. However, considering the explorative nature of this study, we did not want to burden the participants of the MEDICIE trial by administering a similar questionnaire twice.

Second, regarding the assessment of the EQ-5D+C, the proxies may have focused their attention on the cognitive dimension when scoring the VAS<sub>5D+C</sub>, even though they had been instructed to rate the VAS<sub>5D+C</sub> again based on the overall health. This effect is called a framing effect, which suggests that how something is presented (the 'frame') influences the choices people make (Tversky et al., 1981). Hence, it is possible that the higher correlations of the VAS<sub>5D+C</sub> with the MMSE are due to a framing effect. However, according to Parkin et al. (Parkin et al., 2004), the framing bias also exists when assessing the EQ-5D, meaning that values of the VAS<sub>5D</sub> are affected by end-state descriptors (last named dimensions).

Another possible limitation is the use of proxies to complete the questionnaires. Previous research indicated that there is generally fairly good proxy-patient agreement for observable items such as mobility, self care and usual activities, but poor agreement for non-observable items such as pain and affect (Selai, 2001). Others have found agreement to be poor for the domains most affected by dementia (self-care and usual

activities) (Jonsson et al., 2006a). In the light of the longitudinal nature of our study, the complex health problems of our study population and their progressive global deterioration of intellect and personality, the method of proxy rating had been chosen. It is generally acknowledged that in the later stages of dementia proxy measures are required since patients are no longer capable of making an adequate evaluation of their HRQoL (Jonsson et al., 2006a; Selai, 1998). Furthermore, the use of proxy reports throughout the course of a longitudinal study, rather than substituting them only when the person with dementia becomes unable to report his or her HRQoL, reduces bias over time (Smith et al., 2005). The overall picture of previous research is that rating by proxy is a valid alternative for assessing HRQoL in the presence of dementia (Albert et al., 1996; Jonsson et al., 2006a; Magaziner, 1997; Smith et al., 2005), although it is possible that the scores in the EQ-5D and the EQ-5D+C were biased because of perceived caregiver burden (Logsdon et al., 2002).

A final limitation also concerns the design of our study. Comparisons between the EQ-5D and the EQ-5D+C were merely based on the dimensions and VAS-scores of both versions and not on the utility scores since these are not available for the EQ-5D+C. It should be noted that an algorithm has been developed for EQ-5D+C health states, based on Dutch disability weights (Mathers et al., 1999; Mathers et al., 2001). In the Dutch disability weights study, a comprehensive set of disease-specific disability weights for 175 disease stages associated with 52 disease categories was obtained (Stouthard et al., 1997; Stouthard et al., 2000). Based on these disability weights, an EQ-5D+C regression model was fitted. However, the origins of the EQ-5D+C disability weights and the EQ-5D utility scores differ significantly. First, the algorithm is based on valuations of health experts instead of valuations of the general public. Second, EQ-5D+C health states were valued by means of the person trade-off (PTO) method, whereas EQ-5D health states were valued by means of the time trade-off (TTO) method (Dolan, 1997). PTO differs from TTO in that subjects are required to trade-off person years lived healthy against person years lived with some defined disability, thus making choices in the context of a decision involving other people rather than themselves. Whether the PTO technique is able to reflect actual preferences is still under debate (Green, 2001; Pinto-Prades, 1997). Finally, besides the EQ-5D+C health state description, subjects were given specific information with respect to the disease, which differs from the EQ-5D valuation procedure (Dolan, 1997). Therefore, in our opinion, no valid comparison of EQ-5D+C disability weights with EQ-5D utility scores can be made. In order to develop a new scoring algorithm of which utility scores for the EQ-5D+C can be deduced, a valuation procedure similar to the one used for the EQ-5D should be applied. Presenting EQ-5D+C health states to members of the general population should reduce the framing effect described earlier, as the cognitive dimension will then be 'just' one of the six dimensions in the health states. Furthermore, although in the descriptive part of the EQ-5D a proxy effect may still be present, by using a utility score based on valuations of the general population, possible proxy effects are expected to decrease.

### **Conclusion**

In this explorative study, the construct validity and responsiveness of the EQ-5D and the EQ-5D+C were assessed and compared in patients aged 55 and older with cognitive impairments. We conclude that the EQ-5D performs well for evaluating HRQoL in our population with cognitive impairments using proxy ratings. Therefore, based on the results of this study and given its (serious) limitations, it does not seem necessary to adjust the current classification system by adding a cognitive dimension. However, in the absence of a gold standard for measuring HRQoL, a general population study to obtain valuations for the EQ-5D+C health states could provide a better insight into whether cognition has a separate and significant effect on utility values, and would enable us to compare the utility values deduced from both versions in a correct manner.

### *Competing interests*

The authors declare that they have no competing interests.

### *Authors' contributions*

FV was the principal investigator of the larger trial and is guarantor. FV, AK, CD, and JS designed this trial. CW and DW carried out the outcome measures. AK and CW were responsible for statistical analysis and interpretation of the study, and all authors contributed to interpretation. CW drafted the manuscript, and all authors critically revised it for scientific content and approved the final version.

### *Acknowledgements*

The MEDICIE study is funded by the Dutch Research Institute for Care- Medical Sciences (ZorgOnderzoek Nederland-Medische wetenschappen, ZON-Mw), grant no. 945-02-055.

# Chapter 5

The measurement and valuation of informal care



**Will be submitted as two separate articles**

Claire AG Wolfs, Carmen D Dirksen, Alfons Kessels, Frans RJ Verhey, and Johan L Severens

## Abstract

**Background/relevance:** More and more people are aware of the importance of informal care. Informal care plays a substantial role in the provision of total care and is often called the principle of health care. In the Netherlands, it was estimated in 2001 that 3.75 million people provided informal care. This number is expected to increase even further due to demographic and social-economic developments. More attention to informal care, the position of informal caregivers, the consequences of providing informal care and the inclusion of informal care in economic evaluations is therefore essential.

**Aim:** The aim of the proposed study is to apply several methods for the measurement of informal care in the study "Randomized trial on the effects of a psychogeriatric diagnostic day hospital". For this, a questionnaire developed by the iMTA in Rotterdam, the Netherlands, was used. Results of three methods for measuring informal care are described (two monetary methods and one non-monetary method) and both construct validity and responsiveness of each of these methods is examined.

### Research question(s):

1. Which factors are associated with the amount of informal care in dementia?
2. Which of the factors identified in research question 1 are associated with the different methods to measure and value informal care in our study (construct validity)?
3. To what extent are the three methods to measure informal care sensitive to changes in these factors (responsiveness)?
4. Based on the findings from research questions 2 and 3, which of the methods is most suitable to measure and value informal care in this patient population and their caregivers?

**Study design:** The proposed study is part of a prospective controlled randomised trial into the effects of the multidisciplinary DOC-PG on somatic and mental health, and the costs and use of health care facilities compared with usual care. In the current study, three methods for measuring and valuing informal care are described i.e. the proxy-good-method, the opportunity-cost-method (both monetary methods) and the CarerQol-method (non-monetary method). The opportunity-cost-method was divided in the opportunity-cost A method (i.e. hours of paid work, unpaid work or leisure time forgone), and the opportunity-cost B-method (i.e. preferred number of hours of paid work, unpaid work or leisure time). A literature search is conducted to identify factors that are associated with the amount of informal care in dementia. Based on factors emerging from the review, it was investigated for each of the three methods by means of multivariate regression analysis, which of these factors are associated with informal care (construct validity). Furthermore, the responsiveness was examined for these three methods, by testing whether changes over time in these factors were associated with changes in informal care. Construct validity was assessed by using cross-sectional data from the baseline measurement, whereas responsiveness was assessed by using longitudinal data (changes at 6 and 12 months compared to baseline). With respect to the monetary methods, respectively hours and costs were used as dependant variables in the regression analyses. Judgments regarding construct validity and responsiveness

were based on comparisons of the explained variances (R-Squares) of the regression models.

**Results:** The literature search revealed various factors, which were associated with the amount of informal care, underlining the complex nature of informal care. With regard to the construct validity, the proxy-good method performed best based on the multivariate cross-sectional models, with the highest explained variance of more than 60%. The CarerQol method had the lowest explained variances in the cross-sectional models based on the identified factors. Regarding responsiveness, the multivariate longitudinal models revealed that the proxy-good (hours as dependent variable) method was most responsive to changes after six months, as indicated by the highest explained variance of 19.7%. After 12 months, the opportunity-cost B (costs as dependent variable) method was most responsive with an explained variance of 23.3%, followed by the proxy-good (hours as dependent variable) method with an explained variance of 13.1%.

It should be noted that overall, the R-Squares of the longitudinal models were substantially lower than the R-Squares of the cross-sectional models.

**Conclusions:** Based on the results regarding both construct validity and responsiveness, the proxy-good method is the best method to measure and value informal care in patients suffering from dementia or another cognitive disorder and their caregivers.

## BACKGROUND

More and more people are aware of the importance of informal care as it plays a substantial role in health care. Family members, spouses and friends have become increasingly responsible for providing caregiver assistance to diseased or disabled persons. In the Netherlands, it was estimated in 2001 that 3.75 million people provided informal care (Boer de et al., 2003). This number is expected to increase even further due to demographic and socio-economic developments. As the population is ageing, more people will suffer from dementia in the coming decades. Furthermore, the number of single elderly is increasing, there is an increasing shortage of personnel in health care and institutional care is more and more replaced with community care. Additionally both policymakers and third-party payers place increasing pressure on families to provide care (Robinson, 1997). These developments cause tension between the availability of, and demand for informal care. As there is no cure for dementia, and the majority of service provision falls on care rather than treatment, there is a large demand on informal care (NICE, 2006). From an economic evaluation perspective informal care is an important issue. Studies on dementia (Chiu et al., 1999; van der Roer et al., 2000), low back pain (Maniadakis et al., 2000) and stroke (Anderson et al., 2000) for example have pointed out that the costs of informal care can be substantial compared to formal health care. In economic evaluation studies the total costs and consequences of an intervention are weighted. It is now widely recommended that economic evaluation studies should be conducted from a broad societal perspective, indicating that all relevant costs and health effects should be included, regardless of who bears the costs or who obtains the effects. By incorporating both formal and informal care in the analysis, economic evaluation studies will be more comparable both nationally as well as internationally. By this, it is recognized that informal (unpaid) caregiver's assistance is an important issue in economic evaluation. Yet, systematic reviews (Evers et al., 2000; Evers et al., 1997; Goossens et al., 1997, 2000) show that informal care is rarely included in economic evaluation studies, since the collection of these data is often complex.

Informal caregivers reduce or give up normal activities in order to provide (health) care to a patient and/or their relatives, for which they normally do not receive payment. Informal care may involve financial costs like hiring or buying certain attributes and travel costs, but informal caregiver "time" is noted to be the major input in the treatment of the patient. The time invested in informal care may normally have been invested in paid work, household activities or leisure and should be valued in a cost-effectiveness analysis (Luce et al., 1996; Posnett et al., 1996). In order to incorporate time cost into the cost-effectiveness analysis, both the exact quantity of time spent on informal care and the valuation of that time must be determined. With respect to the quantity of time, an underestimation of true caregiver activities may occur when, through the years, these activities have slightly become part of the daily routine. On the other hand, true caregiver costs may be overestimated when some of the earmarked activities were already part of normal home role functioning, like cooking or shopping, or when a combined use of time occurs, for example in the case of surveillance (Smith et al., 1994). The most practical method to solve these problems is to elicit from the caregivers how much time they actually spend each day providing informal care by means of a structured questionnaire.

Regarding the valuation of time, the time input can be valued by means of the market- or proxy-good method, the contingent valuation method and the opportunity cost method (Busschbach et al., 1998; Smith et al., 1994). In the market-price method the time input of the informal caregiver is valued at the costs of hiring a professional caregiver. In the contingent valuation method informal caregivers are asked to state their willingness-to-pay or willingness to accept to respectively no longer having-, or having to perform caregiver activities. The opportunity cost of time spent on informal care is equal to the value of the best alternative use of time, which is assumed to be the normal use of time. The latter method is consistent with the theoretical foundation of cost-benefit analysis (Luce et al., 1996; Posnett et al., 1996; Smith et al., 1994). The normal use of time can be divided in paid work time, unpaid work time and leisure time. When for example time spent on caregiving would normally have been spent on paid work, that time should be valued as being equal to the wage rate. The valuation however of unpaid working time and leisure by means of adjusted wage rates is a point of discussion. Since caregivers for patients with Alzheimer's disease are mostly elderly persons for which wage rates do not seem to be the appropriate method to capture the opportunity cost, it is often recommended to use the market- or shadow price as the valuation method (Brouwer et al., 1999; Busschbach et al., 1998; Luce et al., 1996).

Informal caregivers may encounter positive aspects in the process of caring for a loved one, but more often they experience a number of negative outcomes such as emotional strain, financial losses, feelings of isolation and health declines (Schulz et al., 1999; van den Berg et al., 2004). Caregiver's burden, which is a complex multi-component construct, that is influenced by characteristics of the patient (i.e. severity of the disease and behavioural problems), characteristics of the caregiver (i.e. perceived stress, psychological wellbeing, lack of knowledge, immature coping), and the social context (i.e. finances, social support, family help, number and nature of the caregiving tasks) (Brouwer et al., 2004; Verhey et al., 2007b) has been extensively described in the literature (Dunkin et al., 1998; Kim et al., 2006; Torti et al., 2004). In order to combine the information density of burden with a valuation component for use in economic evaluations, the CarerQol was recently developed (Brouwer et al., 2006).

However, research regarding the amount of informal care and factors that are associated with this amount is limited (Nordberg et al., 2005). Methods to measure and value informal care are crude and standardization and improvement of these methods is necessary (van den Berg et al., 2004). This is especially important in the light of the recommendation to include informal care in economic evaluations as an important outcome measure (Drummond et al., 1997; Gold et al., 1996). Furthermore, since interventions (that are evaluated in economic evaluations) are usually aimed at improving clinical status, quality of life and/or other circumstances of patients, it is desirable that possible effects of the interventions also influence (the amount of) informal care (Verhey et al., 2007b).

## **Objective(s) of the study**

The objective of the proposed study is to apply different methods to measure and value informal care (by means of a questionnaire, developed by the iMTA in Rotterdam, the Netherlands) in the study “Randomized trial on the effects of a psychogeriatric diagnostic day hospital”. A literature search is conducted to identify factors that are associated with the amount of informal care in dementia. Three methods to measure and value informal care (i.e. the proxy-good-method, the opportunity-cost-method and the CarerQoL-method) are applied and validated. Based on factors emerging from the review, it was investigated for each of the three methods which of these factors are associated with informal care (construct validity). Furthermore, responsiveness was examined for these three methods, by testing whether changes over time in these factors were associated with changes in informal care.

## **Research questions of the study**

1. Which factors are associated with the amount of informal care in dementia?
2. Which of the factors identified in research question 1 are associated with the different methods to measure and value informal care in the current study (construct validity)?
3. To what extent are the three methods to measure informal care sensitive to changes in these factors (responsiveness)?
4. Based on the findings from research questions 2 and 3, which of the methods is most suitable to measure and value informal care in this study population and their caregivers?

## **METHODS**

### **Study design and outcome parameters**

The proposed study is part of a prospective controlled randomised trial into the effects of the multidisciplinary psychogeriatric diagnostic observation centre (DOC-PG) on somatic and mental health, and the costs and use of health care facilities of patients with psychogeriatric problems (dementia, somatic problems with comorbid psychiatric disorders). In the current study, an informal care questionnaire, which was developed by the iMTA in Rotterdam, to measure and value informal care was adapted and extended to make it suitable for use in the MEDICIE study. It was administered by means of an interview with the primary caregiver of the patient at baseline and at 6 and 12 months follow-up. The results are categorized in patient characteristics, characteristics of the caregiver, social context and methods for measuring and valuing informal care. These different methods are investigated for their sensitivity to changes. A literature study was conducted to identify factors associated with the amount of informal care in dementia.

### **Patient characteristics**

Patient characteristics described in this study are gender, age, diagnosis, severity of the dementia, independent functioning and behavioural problems (which have been collected in the MEDICIE-study), number of (self-reported) comorbidities, health, physical function and health related quality of life (HRQoL). The severity of the dementia was assessed by means of the Mini Mental State Examination (MMSE) and the Global Deterioration Scale (GDS). The MMSE (Folstein et al., 1975b; Kok et al., 2002) assessed the severity of cognitive decline. The GDS (Reisberg et al., 1982) evaluated patients on seven stages of global functioning for those suffering from a primary degenerative dementia such as Alzheimer's disease. The Instrumental Activities of Daily Living Scale (IADL) (Lawton et al., 1969) measured seven areas of more complex activities required for optimal independent functioning. The scoring for this scale indicated whether the patient was completely independent, in need of assistance or was completely dependent on others for the performance of specific activities. Finally, the NeuroPsychiatric Inventory (NPI) (Cummings, 1997; Cummings et al., 2001; Cummings et al., 1994) appraised patients' behavioural and psychological problems. Higher scores on all instruments, except on the MMSE, are indicative of more severe problems.

Health and physical function were assessed by means of the Visual Analogue Scale (VAS) of the EQ-5D (Health Related Quality of Life) ("EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group", 1990) and physical functioning aspect of the SF36 (Brazier et al., 1992; VanderZee et al., 1996; Ware et al., 1992) (also measuring Health Related Quality of Life) respectively.

### **Caregiver characteristics**

Caregiver characteristics described in this study are gender, age, relationship with the patient, education, income, self-rated health, health-related quality of life (HRQoL), and burden. Self-rated health and HRQoL were assessed by means of the EQ-5D (Health Related Quality of Life) ("EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group", 1990). Regarding burden, a distinction is made between objective burden and subjective burden. Objective burden concerns the number of caregiving tasks, and whether the caregiver shares a house with the care recipient. The subjective burden of informal care indicates how the caregiver encounters the impact of objective burden (Brouwer et al., 2004). This subjective burden may change over time as a result of changes in caregiving situation, the deterioration of the patient's health or coping strategies. We used the self-rated burden scale (SRB) to measure subjective burden (Brouwer et al., 2004; van Exel et al., 2005; van Exel et al., 2004). The SRB consists of a single question: "How burdensome is the provision of informal care to you at this moment?" and caregivers are asked to indicate their burden on a horizontal visual analogue scale ranging from 0 ("not at all straining") to 10 ("much too straining"). Subjective burden was also assessed using the Perceived Stress caused by Informal Caregiving scale (Dutch translation: *Ervaren Druk door Informele Zorg (EDIZ)*) (Pot et al., 1995). The EDIZ is a 9-item measurement to assess self perceived pressure from informal care.

### **Social context**

The social context of both the patient and the caregiver are the caregiver network (i.e. the availability and number of other caregivers and the number of hours spent on informal care by these caregivers) and formal care (home care, day care and admission). Furthermore, other daily activities of the patients (that support the caregiver), respite care and paid leave from work were considered.

### **Measurement and valuation of informal care**

In the informal care questionnaire applied in the MEDICIE study, three different methods (two monetary methods (1 and 2) and one non-monetary method (3)) for the measurement and valuation of informal care were described. Costs of informal care were expressed in 2005 Euros.

1). The proxy-good-method (van den Berg et al., 2006): This method (also called the market cost method or replacement cost method) values the measured time spent on providing informal care at the (labour) market price of a close substitute. This value can differ per task. Caregivers were asked to indicate the average time spent on 18 different informal care tasks per week. This time can be valued using different methods, resulting in a monetary outcome (Oostenbrink et al., 2004; van den Berg et al., 2004; van den Berg et al., 2006). We subdivided the 18 tasks into household activities of daily living (HDL), activities of daily living (ADL) and instrumental activities of daily living (IADL) according to the classification of van den Berg et al (van den Berg et al., 2006). Household (HDL) activities were valued against the wage rate of a housekeeper (€ 8,54). Personal care (ADL and IADL) was valued against the wage rate of a nurse (€ 32,67) (Oostenbrink et al., 2004; van den Berg et al., 2006). Measuring time spent on informal care is the essence of the proxy-good-method.

2). The opportunity-cost-method (van den Berg et al., 2006): In this method, we valued the lost time as the opportunity cost (i.e. hours of paid work, unpaid work or leisure time forgone), or the value of the time in its current best alternative use (hours of paid work, unpaid work or leisure time). We asked the caregivers:

a) what types and amount of time they actually gave up in order to provide informal care (i.e. number of hours of paid work, number of hours of unpaid work and/or number of hours of leisure time).

b) which activities (paid work, unpaid work or leisure time) they would rather perform, instead of giving informal care and to indicate the preferred amount of these activities.

Hours of paid work were valued against the hourly wage of € 49,21 for men and € 37,47 for women, based on a mean age of caregivers in this sample of 59 years. For unpaid work and leisure time, the hourly wage of a housekeeper was used as shadow price (€ 8,54) which was again in accordance with van den Berg et al (Oostenbrink et al., 2004; van den Berg et al., 2006) .

3). The CarerQol-method (Brouwer et al., 2006): This method provides us with a non-monetary outcome regarding the measurement and valuation of informal care. The CarerQol is aimed at measuring the care-related quality of life of informal caregivers (Brouwer et al., 2006). It combines seven important burden dimensions with a valuation component (a visual analogue scale (VAS) for happiness. The seven burden dimensions

are 1) fulfilment; 2) relational problems; 3) mental problems; 4) problems with daily activities; 5) financial problems; 6) support; and 7) physical problems. The CarerQoI-VAS ranges from 0 (“completely unhappy”) to 10 (“completely happy”). The CarerQoI has been validated in a Dutch sample of heterogeneous caregivers.

### **Study population**

The MEDICIE-study was approved by the medical ethical committee at Maastricht University Hospital. Patients were recruited from July 2002 to August 2004 from 70 GP practices in the Maastricht region and surroundings. GPs were asked to refer all patients that may have dementia or a cognitive disorder.

Patients were included who 1) were older than 55 years; 2) were suspected by the GP of having dementia, or were suspected of having both a somatic and a psychological disorder; 3) had not been referred in the two years prior to this study to one of the participating services; 4) had the availability of a proxy, i.e. a person visiting the patient at least once a week).

Patients were excluded when 1) they suffered from acute disorders that needed a prompt therapeutic intervention; or 2) they lived in a nursing home, or received care that is comparable to that of a nursing home (‘e.g. substitution of care in an elderly home);

## **Data selection and analysis**

### **Missing data**

In this study, data of the informal care surveys were analyzed for those patients who completed the survey at all three measurements (i.e. at the baseline measurement, at the 6 months follow-up and at the 12 months follow-up) and who had the same primary caregiver during the course of the study. Missing items within the survey were imputed by means of a regression model (missing value analysis in SPSS version 12.0.1).

### **Data analyses**

#### *Review of the literature*

First, a systematic literature search was conducted in Medline to identify articles published in medical journals reporting on research regarding (the amount of) informal care in dementia. An advanced search was performed using the free text words “informal care\*” and “cost\*” or “amount” or “hour\*” or “quantity” or “intensity”, as well as the terms “predict\*” or “explain\*” or “regress\*” or “model\*” or “influence\*” or “relat\*”. Furthermore the terms “dement\*” or “psychogeriatr\*” or “cognitive disorder” or “cognitive impairment” or “memory disorder” or “memory impairment” were utilised to restrain the search to those subjects relevant to the present study. Subsequently, the results were limited to studies that were published in English and included an abstract. Before reviewing the abstracts, a criterion was established to determine whether a study was relevant to the purpose of the present review: the study had to use regression analysis to identify factors associated with the amount of informal care. In case of doubt (i.e. if factors were

significant in some studies but not in other studies) these factors were also included in our regression analyses.

### *Construct validity and responsiveness*

Construct validity, defined as the extent to which an instrument corresponds with other measures which it should be related to, was estimated by studying associations between the informal care methods and factors emerging from the literature as being associated with informal care. Responsiveness was defined as the associations of the changes in an instrument to changes in other measures which it should be related to, which is also referred to as 'external responsiveness (Husted et al., 2000) or using an 'anchor-based' approach (Terwee et al., 2003). Responsiveness was investigated by studying the associations between changes in the informal care methods and changes in the factors emerging from the literature review.

### *Construct validity (cross-sectional data)*

For examination of construct validity, the baseline data were used for analysis. First, based on the significant factors found in the literature, univariate regression analyses were performed to explore the association of each factor separately on the informal care methods. Second, multivariate forward stepwise regression models were developed (which was assumed to be most robust) to examine the joint influence of these factors on the different methods to measure and value informal care in our study. The dependent variables in the models were the number of hours spent on informal care and the associated costs respectively (both measurement method 1), the number of hours of time forgone and the associated costs respectively (both measurement method 2a), the number of hours of preferred time and the associated costs respectively (both measurement method 2b) and the valuation component of the CarerQoL. So, in total 7 regression analyses were performed with respect to examination of the construct validity. Judgments regarding the construct validity of each of the three methods to measure and value informal care were based on a comparison of the R-Squares of the multivariate regression models.

### *Responsiveness (longitudinal data)*

First, difference scores were calculated for the methods to measure and value informal care and for the factors which can change over time (such as cognition) by subtracting the baseline scores from the 6 and 12 months scores. Based on these changes scores, univariate regression analyses were performed for each factor separately, in order to explore the influence of a change in each of the factors on a change in informal care. Subsequently multivariate forward stepwise regression models were developed (again, which was assumed to be most robust) to study the responsiveness of the three methods to the joint influence of changes in the factors. The dependent variables in the models were the change in number of hours spent on informal care and the associated change in costs respectively (both measurement method 1), the change in number of hours of time forgone and the associated change in costs respectively (both measurement method 2a), the change in number of hours of preferred time and the associated change in costs respectively (both measurement method 2b) and the change

in the valuation component of the CarerQoI. So, in total 7 regression analyses were performed with respect to the examination of responsiveness.

Again, judgments regarding the responsiveness of each of the three methods to measure and value informal care were based on a comparison of the R-Squares of the multivariate regression models.

## RESULTS

### Description of the study population

In the MEDICIE study, 230 patients were included. In the course of the study, 27 patients (11.7%) died and 13 patients (5.7%) dropped out. Of the study completers, 170 patients completed all three informal care questionnaires and were included in the analyses.

Table 1 summarizes demographic and clinical patient characteristics. Of the 170 patients, 68.3% were female. In most cases, dementia (present in 68.2% of the patients) was associated with Alzheimer's disease (40.0%). Patients whose etiological diagnoses could not be determined were assigned to the "other" groups (i.e. other dementia or other cognitive impairment). The clinical characteristics of the patients showed a progressive deterioration over time.

**Table 1:** Patient characteristics (intervention group and control group) at the three measurements

Total group (n=170)	Baseline	6 months	12 months
<b>Sex patient</b> Female (%)	68.3%	n.a.	n.a.
<b>Age patient</b> Mean (SD)	77.3 (6.7)	n.a.	n.a.
Range	55 - 94	n.a.	n.a.
<b>Dementia:</b>	<b>116 (68.2%)</b>	n.a.	n.a.
Alzheimer's Disease (AD)	68 (40.0%)	n.a.	n.a.
Vascular Dementia (VD)	17 (10.0%)	n.a.	n.a.
Mixed Dementia	17 (10.0%)	n.a.	n.a.
Other Dementia	14 ( 8.2%)	n.a.	n.a.
<b>No Dementia:</b>	<b>54 (31.8%)</b>	n.a.	n.a.
Cognitive Impairment / MCI	26 (15.3%)	n.a.	n.a.
Other Cognitive Impairment	28 (16.5%)	n.a.	n.a.
<b>MMSE</b> <sup>1</sup>	20.2 (5.7)	19.0 (7.2)	18.1 (7.5)
<b>GDS</b> <sup>2</sup>	4.2 (1.1)	4.5 (1.3)	4.7 (1.2)
<b>IADL</b> <sup>3</sup>	16.7 (6.1)	18.6 (6.5)	20.2 (6.6)
<b>NPI</b> <sup>4</sup>	22.1 (15.0)	24.8 (18.1)	28.0 (20.3)
<b>Number of comorbidities</b>	3.0 (2.0)	3.0 (1.8)	3.1 (1.8)
<b>Health</b> <sup>5</sup> Mean (SD)	59.7 (19.2)	58.3 (17.8)	57.3 (19.9)
<b>Physical function</b> <sup>6</sup> Mean (SD)	50.9 (28.9)	49.1 (29.3)	42.4 (29.7)

<sup>1</sup> Mini Mental State Examination; <sup>2</sup> Global Deterioration Scale; <sup>3</sup> Instrumental Activities of Daily Living Scale; <sup>4</sup> NeuroPsychiatric Inventory; <sup>5</sup> Visual Analogue Scale EQ-5D; <sup>6</sup> Physical functioning part of the SF36; n.a.: characteristics which do not change over time

Table 2 summarizes the caregiver characteristics. Of the caregivers, 67.3% were female and most often a child (-in-law) or spouse of the patient (91.2%). Most caregivers (57.1%) had a high education (i.e. they finished secondary school) and an income above average (52.9%). Overall health (as measured with the EQ-5D) slightly deteriorated, but the burden (objective and subjective) remained stable over time.

**Table 2:** Caregiver characteristics (intervention group and control group) at the three measurements

Total group (n=170)	Baseline	6 months	12 months
<b>Sex caregiver Female (%)</b>	67.3%	n.a.	n.a.
<b>Age caregiver Mean (SD)</b>	58.8 (13.0)	n.a.	n.a.
Range	30 – 87	n.a.	n.a.
<b>Relationship patient:</b>			
Spouse	32.4%	n.a.	n.a.
Child (in law)	58.8%	n.a.	n.a.
Other	8.8%	n.a.	n.a.
<b>Education:</b>			
Low	49 (42.9%)	n.a.	n.a.
High	51 (57.1%)	n.a.	n.a.
<b>Income (net):</b>			
€ 545 - € 726	5 ( 2.9%)	n.a.	n.a.
€ 726 - € 908	5 ( 2.9%)	n.a.	n.a.
€ 908 - € 1135	10 ( 5.9%)	n.a.	n.a.
€ 1135 - € 1589	31 (18.2%)	n.a.	n.a.
€ 1589 - € 2269	29 (17.2%)	n.a.	n.a.
> € 2269	90 (52.9%)	n.a.	n.a.
<b>VAS EQ-5D Mean (SD)</b>	77.2 (16.2)	73.1 (16.1)	73.6 (15.8)
<b>Utility score EQ-5D Mean (SD)</b>	0.87 (0.19)	0.87 (0.18)	0.86 (0.20)
<b>Objective burden:</b>			
Number of care giving tasks	6.0 ( 3.0)	6.3 ( 3.1)	6.4 ( 2.9)
Share house with care recipient (%yes)	37.1%	35.3%	33.5%
<b>Subjective burden:</b>			
SRB <sup>1</sup>	3.5 (2.8)	4.0 (2.7)	4.1 (2.5)
EDIZ <sup>2</sup>	4.5 (2.7)	4.5 (2.8)	4.4 (2.9)

<sup>1</sup> Self-Rated Burden scale; <sup>2</sup> Perceived Stress caused by Informal Caregiving scale;  
n.a.: characteristics which do not change over time

Table 3 describes the social context of the patients and their caregivers. Besides the primary caregiver, other caregivers were available for most patients (61.5%), with a mean of 1.5 caregivers extra per patient, providing an additional 4 hours a week on caregiving. At baseline 58.8% of the patients made use of formal care and this percentage increased to 78.8% at the 12 months follow-up measurement. Home care

was used the most (mean 44.3%). Hardly any of the caregivers used services in order to support them (i.e. paid leave from work or respite care).

**Table 3:** Social context (intervention group and control group) at the three measurements

Total group (n=170)	Baseline	6 months	12 months
<b>Caregiver network:</b>			
Other caregivers? (%yes)	61.0%	61.6%	61.8%
Number of other caregivers	1.3 (1.5)	1.5 (2.0)	1.5 (2.0)
Number of hours per week	3.8 (5.8)	3.7 (5.7)	5.0 (8.4)
Other day care (e.g. family, % yes)	4.5%	5.2%	8.6%
Hours per month (mean, SD)	0.4 (2.1)	0.9 (4.9)	1.4 (6.6)
Monthly activities* (% yes)	20.2%	20.2%	20.6%
Hours per month (mean, SD)	4.8 (12.5)	4.9 (9.6)	3.9 (8.0)
<b>Formal care (patient): (%yes)</b>			
On waiting list home care (%yes)	1.8%	0.6%	0.6%
Home care (%yes)	47.1%	45.3%	40.6%
Hours per week (mean, SD)	2.7 (4.8)	3.0 (5.2)	2.8 (5.1)
On waiting list day care (%yes)	4.7%	2.9%	5.3%
Formal day care (%yes)	19.4%	35.9%	40.6%
Days per week (mean, SD)	0.5 (1.4)	1.2 (1.9)	1.6 (2.3)
On waiting list elderly home (% yes)	10.6%	9.4%	6.5%
Live in elderly home (% yes)	8.8%	11.2%	16.5%
On waiting list nursing home (% yes)	1.2%	3.5%	4.1%
Live in nursing home (% yes)	0.0%	9.4%	13.5%
<b>Support (caregiver):</b>			
Paid leave from work (% yes)	0.0%	0.0%	0.6%
Respite care (% yes)	0.6%	1.2%	2.4%

Table 4 presents the results of the proxy-good method to measure informal care (number of hours per week per activity and the percentages of caregivers performing the activities). Most caregivers spent time on IADL and HDL tasks (>90%). The percentage of caregivers performing these tasks remained stable over time. At baseline, 38.5% of the caregivers spent time on ADL tasks and this percentage increased to 50.0%. Overall, most time was spent on the HDL activities (4.6 hours a week). In total, a mean of 9.1 hours per week was provided on caregiving. The total amount of time spent on informal care increased with 2 hours (which is an increase of 24%) over the course of one year. The total costs increase with 41% after a year. This higher increase in costs (than in

hours) is due to the substantial increase in ADL and IADL activities (71% and 48% respectively), which are most expensive compared with HDL activities.

**Table 4:** Proxy-good method (1); informal care time in mean hours (and mean total costs) per week at baseline, 6 months follow-up and 12 months follow-up (mean, SD, percentage of people performing the task)

Total group (n=170)	Baseline	6 months	12 months
<b>HDL activities (total):</b>	<b>4.8 (6.3): 88.0%</b>	<b>4.3 (5.6): 92.4%</b>	<b>4.7 (5.9): 90.0%</b>
Preparation of food and drinks	1.4 (3.1): (41.8%)	1.3 (2.9): (40.0%)	1.6 (3.3): (43.5%)
Cleaning the house	0.8 (2.4): (28.8%)	0.7 (1.6): (39.4%)	0.6 (1.5): (37.6%)
Washing, ironing or sewing	0.6 (1.1): (31.2%)	0.5 (1.0): (35.9%)	0.7 (1.2): (42.4%)
Shopping, groceries	1.0 (1.4): (67.1%)	0.8 (1.3): (63.5%)	0.9 (1.2): (64.1%)
Chores, gardening, maintenance	0.4 (1.1): (29.4%)	0.3 (1.2): (27.1%)	0.2 (0.7): (25.3%)
Other (inspection fridge, garbage etc.)	0.6 (1.5): (47.6%)	0.7 (1.6): (58.8%)	0.7 (2.2): (49.4%)
<b>ADL activities (total):</b>	<b>0.7 (2.3): 38.2%</b>	<b>1.2 (3.3): 45.9%</b>	<b>1.2 (2.9): 50.0%</b>
Personal care ((un)dressing, wash etc.)	0.5 (1.3): (30.0%)	0.8 (2.4): (36.5%)	0.7 (1.7): (39.4%)
Visiting the toilet	0.1 (0.3): (5.9%)	0.2 (1.2): (11.2%)	0.2 (1.1): (14.1%)
Moving around within the house	0.1 (0.3): (9.4%)	0.2 (1.2): (14.1%)	0.2 (0.5): (15.3%)
Eating and drinking	0.0 (0.6): (8.2%)	0.0 (0.2): (10.0%)	0.1 (0.6): (14.7%)
<b>IADL activities (total):</b>	<b>3.1 (3.3): 95.9%</b>	<b>2.9 (4.6): 94.7%</b>	<b>4.6 (18.2): 93.9%</b>
Moving around outside the house	0.4 (1.1): (30.0%)	0.4 (1.0): (32.4%)	0.4 (0.9): (33.5%)
Excursions, visits to friends/family	0.8 (1.4): (51.8%)	0.7 (1.4): (44.1%)	0.6 (1.2): (42.4%)
Contacts with health care services	0.8 (1.2): (79.4%)	0.6 (0.9): (78.2%)	0.5 (0.7): (71.2%)
Organizing home adaptations	0.2 (0.4): (35.3%)	0.2 (0.3): (40.6%)	0.2 (0.4): (34.7%)
Financial matters (insurance, rent)	0.4 (0.6): (55.3%)	0.4 (0.5): (57.1%)	0.4 (0.8): (62.9%)
Medication	0.2 (0.7): (28.2%)	0.2 (0.5): (31.8%)	0.1 (0.3): (28.8%)
Other (mail, remembering)	0.3 (1.1): (17.1%)	0.5 (3.4): (14.1%)	0.2 (0.5): (18.2%)
<b>Total activities:</b>	<b>8.5 (9.0): 100%</b>	<b>8.4 (10.8): 100%</b>	<b>10.5 (21.8): 100%</b>
<b>Total costs:</b>	<b>€ 163,26 (173.7)</b>	<b>€ 170,90 (240.7)</b>	<b>€ 229,85 (645.1)</b>

Table 5a presents the results of the opportunity-cost-A method regarding types and time forgone. Most caregivers indicated that they used leisure time in order to care for the patients. In total, a mean of 5.8 hours per week were forgone in order to provide caregiving. The total amount of time forgone increased with 2.7 hours over the course of one year, which is an increase of 62%. The total costs increase with 66% after one year.

Table 5b presents the results of the opportunity-cost-B method regarding the preferred usage of time in case no more care needed to be given to the patient. Most patients (67.1%) preferred to spent freed time on leisure at the three measurements. The mean number of hours amounted to 5.5 hours, which was similar to the number of time forgone (opportunity-cost method A). The total amount of preferred time increased with 0.5 hours

over the course of one year, which is an increase of 9%. However, the costs decreased with 11% over the course of one year, due to a change in preferences i.e. after 12 months the number of caregivers who would like to spent freed time on paid work had decreased whereas the number of caregivers who would like to spent freed time on leisure (which is cheaper) had increased.

**Table 5a:** Opportunity cost method (2A): Changes in paid work, unpaid work and leisure time (percentage, mean number of hours, SD) caused by providing care

Total group (n=170)	Baseline	6 months	12 months
<b>Paid work forgone (%yes):</b>	4.7%	5.3%	2.9%
Hours forgone (mean, SD)	0.5 (3.1)	0.8 (2.9)	0.9 (1.9)
No change	41.2%	41.8%	45.3%
Not applicable (no paid work)	54.1%	52.9%	51.8%
<b>Unpaid work forgone (%yes):</b>	14.1%	15.3%	10.6%
Hours forgone (mean, SD)	0.9 (3.1)	1.3 (2.6)	1.7 (2.4)
No change	84.7%	82.9%	89.4%
Not applicable (no unpaid work)	1.2%	1.8%	0.0%
<b>Leisure time forgone (%yes):</b>	47.1%	38.8%	34.3%
Hours forgone (mean, SD)	3.1 (5.3)	3.6 (6.3)	4.7 (7.0)
No change	52.9%	60.6%	64.7%
Not applicable (no leisure time work)	0.0%	0.6%	0.0%
<b>Total hours forgone (mean, SD)</b>	4.5 (7.7)	5.7 (10.8)	7.3 (14.3)
<b>Total costs</b>	€ 52,86 (138.3)	€ 70,99 (222.2)	€ 87,86 (251.0)

**Table 5b:** Opportunity cost method (2B): Preferences regarding paid work, unpaid work and leisure time (percentage, mean number of hours, SD)

Total group (n=170)	Baseline	6 months	12 months
<b>When I no longer have to provide informal care, I would spent this freed time on:</b>			
More paid work (%yes)	5.9%	5.3%	4.1%
Hours more (mean, SD)	0.8 (3.9)	0.7 (4.0)	0.4 (2.4)
More unpaid work (%yes)	15.9%	15.9%	17.1%
Hours more (mean, SD)	0.8 (2.4)	0.8 (2.3)	0.8 (2.1)
More leisure time (%yes)	60.0%	71.2%	70.0%
Hours more (mean, SD)	3.7 (5.1)	3.8 (4.1)	4.6 (4.9)
<b>Total number of hours (mean, SD)</b>	<b>5.3 (7.4)</b>	<b>5.3 (6.5)</b>	<b>5.8 (5.9)</b>
<b>Total costs</b>	<b>€ 70,00 (169.6)</b>	<b>€ 67,31 (170.4)</b>	<b>€ 61,93 (108.5)</b>

Table 6 summarizes the results of the CarerQoL, used as a non-monetary outcome for the measurement of informal care. Physical problems (of the caregiver) were most often reported (45.3%) at the three measurements, followed by relational problems (with the patient) and mental problems (44.7% and 43% respectively). Furthermore, 38.6% of the caregivers indicated having problems with carrying out their daily activities. Most caregivers found fulfilment in caring for their loved one. The mean baseline score on the CarerQoL VAS was 7.2 and this score slightly decreased to 6.9 at the 12 months follow up measurement. In total, there was a decrease of 0.3 (4%) on the CarerQoL.

**Table 6:** CarerQoL method (3): Subjective care related quality of life questionnaire (non-monetary outcome regarding the measurement and valuation of informal care).

Total group (n=170)	Baseline	6 months	12 months
<b>Fulfilment (%)</b>			
No	8.8%	7.6%	13.5%
Some	36.5%	33.5%	36.5%
A lot of	54.7%	58.8%	50.0%
<b>Relational problems (%)</b>			
No	54.1%	60.0%	51.8%
Some	35.9%	31.2%	38.2%
A lot of	10.0%	8.8%	10.0%
<b>Mental problems (%)</b>			
No	58.2%	59.4%	55.3%
Some	35.4%	28.8%	32.9%
A lot of	9.4%	11.8%	11.8%
<b>Problems with daily activities (%)</b>			
No	61.2%	61.2%	61.8%
Some	31.2%	34.1%	31.8%
A lot of	7.6%	4.7%	6.5%
<b>Financial problems (%)</b>			
No	98.2%	94.7%	93.5%
Some	1.8%	4.7%	5.9%
A lot of	0%	0.6%	0.6%
<b>Support (%)</b>			
No	21.8%	21.2%	19.4%
Some	33.5%	28.8%	39.4%
A lot of	44.7%	50.0%	41.2%
<b>Physical problems (%)</b>			
No	53.5%	55.3%	55.3%
Some	33.5%	28.2%	27.1%
A lot of	12.9%	16.5%	17.6%
<b>VAS (mean, SD)</b>	7.2 (1.7)	7.0 (1.7)	6.9 (1.6)

In conclusion, the largest relative change over one year was noticed in the opportunity-cost A method, whereas the smallest relative change was noticeable in the CarerQol method.

### **Research question 1**

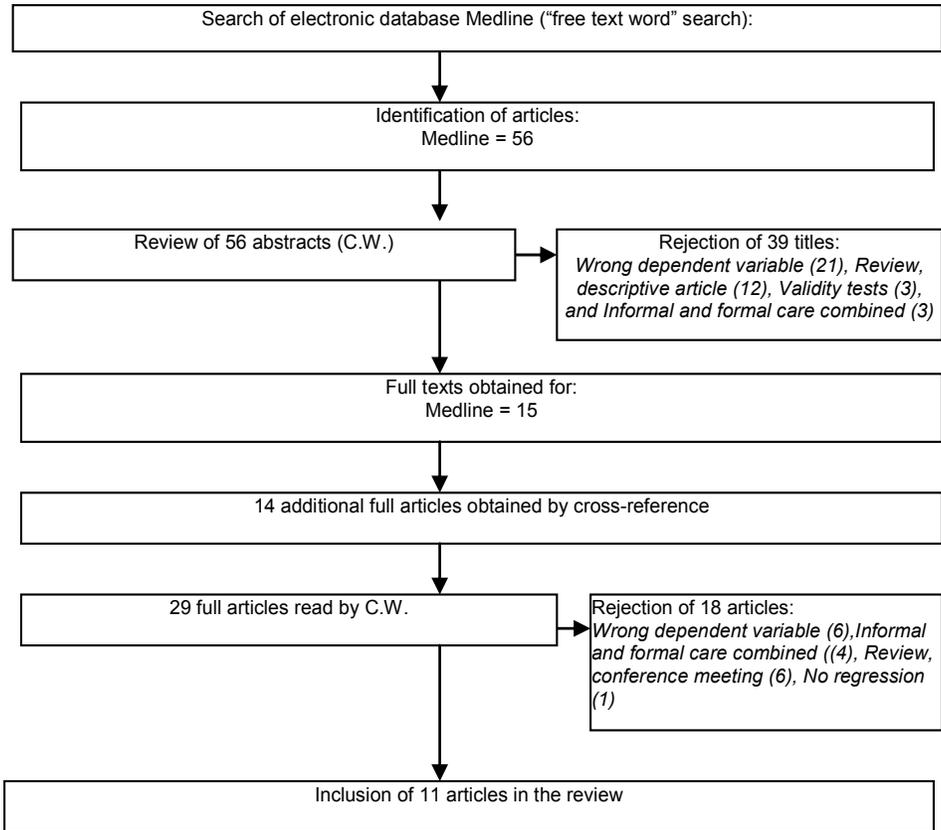
#### *Factors associated with the amount of informal care: Review*

Box 1 shows the process of the inclusion and exclusion of articles. One reviewer (C.W.) assessed all articles for inclusion or exclusion based on the abstract. In case of doubt she also read the full text of the article. Of the 56 references found, the reviewer excluded 39 as non-relevant in terms of subject matter (see box 1). An additional 14 full texts were obtained by cross-references. Eventually, 11 studies were identified as relevant for the purpose of this review (Albert et al., 1998; Feldman et al., 2005; Harrow et al., 2004; Jonsson et al., 2006b; Kemper, 1992; Langa et al., 2001; Nordberg et al., 2005; Schneider et al., 2002; Wimo et al., 2002; Zhu et al., 2006a, 2006b). In all of the selected studies regressions were used to determine predictors of the amount of informal care.

Tables 7a-c describe the research questions of the studies and the dependent and independent variables used in their regression analyses. Gender and age of the patient, living together, severity of dementia, cognitive impairments, behavioural disturbances, (I)ADL impairments, comorbidity, health, physical function and use of formal services were identified as the patient characteristics as possible predictors of the amount of formal care (using total hours of informal care as dependent variable). Of the very few caregiver characteristics that were used in the studies, none were significant (again using total hours of informal care as dependent variable). It should be noted that two studies identified ethnicity and time since diagnosis as predictors of informal care. However, since merely two patients were of Indonesian origin and the fact that we studied a diagnostic facility (i.e. of all patients who were included in MEDICIE study, the time since diagnosis was equal), these variables were not included in our regression models.

As the CarerQol is, in essence, another type of measure to value informal care compared to the monetary measures, a quick search of the literature was additionally conducted on predictors of caregiver burden as a proxy for care-related quality of life of the caregiver (Kim et al., 2006; Torti et al., 2004), to check whether any potentially relevant factors were missing. Similar patient characteristics were found to be significant, and in addition some caregiver characteristics were found to have an association with caregiver burden: age, gender, duration of caregiving, number of caregiving tasks, education, income and self rated competence. No data on self rated competence was available in the MEDICIE study, but the other caregiver characteristics were included in our regression analyses.

**Box 1:** Flowchart to show the process of inclusion and exclusion of articles.



**Table 7a:** Research questions of the studies that are included in the review

Study:	Research question:
1. Wimo, 2002 (Wimo et al., 2002)	To measure, describe and compare formal services with different types of informal care provided to elderly individuals suffering from dementia.
2. Langa, 2001 (Langa et al., 2001)	Obtain nationally representative estimates of the time and associated cost of informal caregiving for the elderly with dementia.
3. Feldman, 2005 (Feldman et al., 2005)	To investigate the longitudinal 1-year decline in cognition and functional ability and the amount of caregiving.
4. Nordberg, 2005 (Nordberg et al., 2005)	To describe the amount of formal and informal care given to non-demented and demented elderly persons living at home.
5. Zhu, 2006 (Zhu et al., 2006a)	Provide an estimation of caregiving hours and cost trajectories.
6. Jönssen, 2006 (Jonsson et al., 2006b)	To estimate the cost of medical care, community care and informal care in different stages of AD in Sweden. . Second aim was to estimate the relationship between costs and disease severity and to identify important cost drivers.
7. Zhu, 2006 (Zhu et al., 2006b)	Estimate effects of patient's clinical characteristics on disease costs.
8. Albert, 1998 (Albert et al., 1998)	1) Derive estimates of hourly informal and formal care, 2) Examine the extent to which dementia severity and presence of comorbidities predict care hours.
9. Schneider, 2002 (Schneider et al., 2002)	1) Explore the configurations of formal and informal care received by people living in their homes, as well as in residential or nursing homes; 2) What services do people use?; 3) What factors are associated with service use?; 4) What services do informal carers provide?; 5) How do the answers to these questions vary according to living arrangements or other factors?
10. Harrow, 2004 (Harrow et al., 2004)	1) What is the total cost of care for informal and formal services that support the activities of daily living (ADLs) and instrumental activities of daily living (IADLs) for community residing AD care recipients; 2) Is there geographic variation in the cost of this care after controlling for caregiver and care recipient characteristics?
11. Kemper, 1992 (Kemper, 1992)	To analyze factors associated with the amount of formal and informal care.

**Table 7b:** Hours and costs of informal care as dependent variables (horizontal) and patient characteristics as independent variables (vertical) used in the review

	Hours informal care (total)	Hours ADL	Hours IADL	Hours supervision	Incremental hours informal care	informal care (yes/no)	Costs informal care
<b>Patient:</b>							
Gender	1ns, 3ns, 4+, 5ns, 8+, 11ns	1ns	1ns	1ns	2ns,	5ns	6ns, 7ns, 10ns
Age	1+, 3ns, 4+, 5ns, 8ns, 11ns	1ns	1ns	1+	2ns,	5ns	6ns, 7ns, 10ns
Marital status	5ns				2+	5ns	
Education	8ns, 11ns						
Living together / relationship	1+, 4+, 8+, 9+	1+	1ns	1+			10ns
Availability of caregivers	11-						
City size (not applicable)	11ns						
Health	11-						
Severity of dementia (stages)	1ns, 3+, 4+, 8ns, 9+	1+	1ns	1ns	2+,		10-
Cognitive impairment	4+, 5ns, 11+						6+
Behavioural disturbances	1+, 5ns, 11+	1+	1ns	1+		5+	7ns, 10+
(I)ADL-impairments	1r, 3ns, 4r, 5+, 11+	1r	1r	1r	2+,	5ns	6+, 7ns
Physical function	9-					5+	7+
Cholinesterase inhibitors							6r
Comorbidity	4ns, 5ns, 8ns, 11+				2+		6ns
Diagnosis	8ns					5+	6+, 7ns
Communication capacity	1ns	1ns	1ns	1ns			
Medicaid eligibility	8-, 11ns						
Amount of formal care	1+, 9-, 11+/-	1ns	1ns	1+	2+		
Race	11+				2ns		
Net worth	11ns						
Time since diagnosis (n.a.)	3+						6+
Time since symptoms							7ns
Year (long term effects)	5+					5ns	

1-11, Studies included in review (see table 7); + = positively associated; - = negatively associated; ns = not significant; r = removed due to high correlation with other variable

**Table 7c:** Hours and costs of informal care as dependent variables (horizontal) and caregiver characteristics as independent variables (vertical) used in the review

	Hours informal care (total)	Hours ADL	Hours IADL	Hours supervision	Incremental hours informal care	Informal care (yes/no)	Costs informal care
<b>Caregiver:</b>							
Age	1ns	1-	1ns	1ns			
Gender	1ns	1ns	1ns	1ns			
Health (0poor, 5excellent)	1ns, 9ns	1ns	1+	1ns			10ns
Coping / burden	1ns	1+	1ns	1ns			10+
Wish to care	1ns	1+	1ns	1ns			
Paid remuneration	1ns	1ns	1ns	1ns			
Hospitalizations	1ns	1ns	1ns	1ns			
Drug use	1ns	1ns	1ns	1ns			
Work time lost	1ns	1ns	1ns	1ns			10ns
Income	1ns	1ns	1ns	1ns			10+
Employment (full, part, no)							
Primary caregiver	8ns						10ns
Other caregivers							

1-11; Studies included in review (see table 7); + = positively associated; - = negatively associated; ns = not significant; r = removed due to high correlation with other variable

**Research questions 2 & 3***Factors associated with the different methods to measure and value informal care (regression models)*

Table 8a shows the variables included in our regression analyses (i.e. in the cross-sectional models, in the longitudinal models or both). Inclusion of these variables was based on the significant variables found in the review.

**Table 8a:** Description of the variables used in the regression models

	Variable		Regression model
<b>Patient:</b>			
Gender	Dichotomous	Male/female	Cross-sectional
Age	Continuous		Cross-sectional
Health	Continuous	VAS EQ-5D	Cross-sectional and longitudinal
Comorbidity	0, 1, > 2	Number of comorbidities	Cross-sectional and longitudinal
Dementia severity	Continuous	GDS <sup>1</sup>	Cross-sectional and longitudinal
Cognitive impairment	Continuous	MMSE <sup>2</sup>	Cross-sectional and longitudinal
Behavioural disturbances	Continuous	NPI <sup>3</sup>	Cross-sectional and longitudinal
(I)ADL impairments	Continuous	IADL-scale <sup>4</sup>	Cross-sectional and longitudinal
Physical function	Continuous	Physical function SF36	Cross-sectional and longitudinal
Formal care	Dichotomous	Yes/no	Cross-sectional and longitudinal
<b>Caregiver:</b>			
Gender	Dichotomous	Male/female	Cross-sectional
Age	Continuous		Cross-sectional
Income	Continuous	1-7	Cross-sectional
Education	Dichotomous	High/low	Cross-sectional
Number of tasks	Continuous		Cross-sectional and longitudinal
Duration of caregiving	Continuous	Months	Cross-sectional
Relation	0, 1, 2	Spouse, child, other	Cross-sectional
Living together	Dichotomous	Yes/no	Cross-sectional and longitudinal
Caregiver network	0, 1, > 2	Number of caregivers	Cross-sectional and longitudinal

<sup>1</sup> Global Deterioration Scale; <sup>2</sup> Mini Mental State Examination; <sup>3</sup> NeuroPsychiatric Inventory; <sup>4</sup> Instrumental Activities of Daily Living Scale

Based on the identified factors (of research question 1) and the directions of the associations with informal care found in the literature, it was hypothesized that the amount of informal care is: higher in men, higher in older patients, higher when the caregiver is a co-resident of the patient, higher when health is poorer, higher when dementia is more severe, higher in case of more impairments (cognitive, behavioural and (I)ADL), higher in case of poorer physical function and higher in case of more comorbidities. Although formal care came forward from the review as having an

**Table 8b:** Mean scores (cross-sectional models), mean change scores (longitudinal models) and ranges of the variables included in the regression models for the group as a whole (DOC-PG and usual care).

	Cross-sectional model Baseline	Longitudinal model 6 months – Baseline	Longitudinal model 12 months – Baseline	% Change
<b>Methods:</b>				
Proxy-good (hours)	8.5 (9.0) [0 – 64]	-0.1 (9.0) [-29 – 69]	2.0 (19.3) [-29 – 168]	+24%
Opp.cost A (hours)	-4.5 (7.7) [-48– 0]	-1.2 (8.0) [-33 – 35]	-2.7 (16.1) [-70 – 27]	-62%
Opp.cost B (hours)	5.3 (7.4) [0 – 48]	0.0 (7.2) [-27 – 23]	0.5 (7.5) [-39 – 20]	+ 9%
CarerQoL	7.2 (1.7) [2 – 10]	-0.2 (1.5) [-6 – 5]	-0.3 (1.7) [-9 – 5]	- 4%
<b>Patient:</b>				
Gender (% female)	68.2%	n.a.	n.a.	n.a
Age	77.3 (6.7) [39]	n.a.	n.a.	n.a
Health	59.7 (19.2) [10-100]	- 1.4 (17.8) [-50 – 60]	-2.5 (20.9) [-70 –60]	- 4%
Comorbidity	3.0 (2.0) [10]	0.03 (0.57) [-2 – 2]	0.05 (0.50) [-2 – 2]	n.a.
Dementia severity	4.2 (1.1) [1 – 6]	0.3 (0.8) [-2 – 3]	0.5 (0.8) [-2 – 3]	+ 12%
Cognitive impairment	20.2 (5.7) [4 – 29]	- 1.1 (4.4) [-18 – 9]	-2.1 (4.9) [-21 – 11]	- 10%
Behavioural disturbances	22.1 (14.9) [0 – 67]	2.7 (14.4) [-36 – 60]	5.9 (15.8) [-31 – 72]	+ 27%
(I)ADL impairments	16.7 (6.1) [6 – 31]	1.9 (3.9) [-13 – 16]	3.5 (5.1) [-15 – 16]	+ 21%
Physical function	50.9 (28.9) [0–100]	- 1.9 (18.9) [-80 – 50]	- 8.6 (22.6) [-75 – 55]	+ 17%
Formal care (%yes)	58.8%	0.3 (0.7) [-1 – 3]	0.4 (0.7) [-1 – 2]	n.a.
<b>Caregiver:</b>				
Gender	67.1%	n.a.	n.a.	n.a
Age	58.8 (14.0) [57]	n.a.	n.a.	n.a
Income	5.0 (1.3) [1 – 7]	n.a.	n.a.	n.a
Education (%high)	57.1%	n.a.	n.a	n.a
Number of tasks	6.0 (3.1) [0 – 15]	0.4 (2.4) [-7 – 7]	0.4 (2.6) [-7 – 7]	+ 7%
Duration of caregiving	45.3 (57.6) [1 - 300]	n.a.	n.a.	n.a
Relation: Spouse	32.4%	n.a.	n.a.	n.a
Child (in law)	58.8%	n.a.	n.a.	n.a
Living together (%yes)	37.1%	- 0.01 (0.13) [-1 – 0]	- 0.04 (0.19) [-1 – 0]	n.a
Caregiver	1.3 (1.5) [0 – 8]	- 0.01 (0.5) [-2 – 2]	- 0.01 (0.5) [-2 – 2]	n.a

association with informal care, the direction of the association differed between studies. Therefore, no hypothesis was made regarding this factor. Furthermore, it was hypothesized based on the identified factors relating to burden that burden is: higher in females, higher in older caregivers, higher in caregivers with lower education, higher in caregivers with a lower income, higher as the duration of caregiving increases, higher when the number of caregiving tasks increases, and higher when the number of other caregivers is smaller.

Table 8b shows the mean scores at baseline (used for examination of the construct validity) and the mean difference scores for the group as a whole (used for examination of responsiveness). With respect to the monetary methods, the amount of informal care increased over time; after one year, the mean amount of informal care increased with 2 hours (24%) in the proxy-good method, with 2.8 hours (62%) in the opportunity-cost-A-method, and with 0.5 hours (9%) in the opportunity-cost-B-method. The CarerQol remained stable and slightly decreased with a mean of 0.3 points (4%). Overall, the patients showed a (slight) decrease in clinical characteristics. The number of comorbidities, number of caregivers and living situation remained relatively stable over time. More people made use of formal services during the course of the study and the number of caregiving tasks also increased slightly.

## **Research question 2: construct validity (Cross-sectional regression models)**

### *Explorative univariate analyses*

Tables 9a and 9b show the results of the regression models (based on baseline data). Because number of hours and costs of informal care (in all monetary methods) were skewed to the left (in the proxy-good and opportunity-cost B methods), and to the right (in the opportunity-cost A method), log transformed hours and costs were used as dependent variables. In the univariate analyses, the three monetary methods to measure and value informal care performed equally well. Of the 19 variables, 10, 9 and 10 variables were found to be significant in the proxy-good-hours method, the opportunity-cost A-hours method and the opportunity-cost B-hours method respectively, and 12, 10 and 12 significant variables were found when the costs were used as dependent variables. With respect to the CarerQol method, 5 variables out of 19 were significant. The factors health, dementia severity, behavioural disturbances and physical function were associated with informal care in all methods. Comorbidities, (I)ADL impairments and formal care were associated with all monetary methods, but not with the CarerQol. All associations found were as hypothesized. When significant, formal care was positively associated with informal care.

### *Multivariate analyses*

Regarding the construct validity, the proxy-good method performed best as is shown by R-Squares of 63% and 60% for number of hours and costs as the dependent variables respectively. Gender (hours only), dementia severity (hours only), number of tasks and living situation were the significant factors; males receive more informal care than females, the amount of informal care increases with dementia severity and number of tasks, and more informal care is provided when the caregiver is a co-resident of the patient (spouse).

R-Squares of the proxy-good method were about twice as high as the R-Squares of the other monetary methods. The CarerQol method had the lowest R-Square. Again, all associations found, were as hypothesized.

**Table 9a:** Univariate cross-sectional regression models (baseline measurement)

	Proxy-good		Proxy-good		Opp. Cost A		Opp. Cost B		Opp. Cost B		CarerQoL	
	Total hours	Total costs	Total hours	VAS								
<b>Patient</b>												
Gender	ns	ns										
Age	ns	+	ns	ns								
Health	-	-	-	-	-	-	-	-	-	-	-	+
Comorbidity	+	+	+	+	+	+	+	+	+	+	+	ns
Dementia severity	+	+	+	+	+	+	+	+	+	+	+	-
Cognitive impairment	-	-	ns	ns								
Behavioral disturbances	+	+	+	+	+	+	+	+	+	+	+	-
(I)ADL impairments	+	+	+	+	+	+	+	+	+	+	+	ns
Physical function	-	-	-	-	-	-	-	-	-	-	-	+
Formal care	+	+	+	+	+	+	+	+	+	+	+	ns
<b>Caregiver:</b>												
Gender	ns	ns										
Age	ns	ns										
Income	-(0.07)	ns	ns									
Education	ns	ns										
Number of tasks	+	+	+	+	+	+	+	+	+	+	+	ns
Duration of caregiving	ns	+(0.06)	ns	ns								
Relation	ns	ns										
Living together	ns	ns										
Caregiver network	ns	+	ns	ns								

+ = positively associated (p<0.05); - = negatively associated (p<0.05); ns = not significant; p-values between 0.05 – 0.10 are shown between brackets



**Research question 3: responsiveness (longitudinal regression models)***Explorative univariate analyses*

Tables 10a and 10b show the results of the univariate regression analyses. In the univariate analyses, none of the factors were commonly associated with changes in informal care in all methods, neither at 6 nor at 12 months. Regarding the change scores at 6 months (table 10a), the proxy-good method had the highest number of factors being associated with changes in informal care (i.e. 6 and 5 out of 11), with significant factors being (changes in) health, behavioural disturbances, IADL impairments, physical function, amount of formal care and number of tasks. Regarding the change scores after 12 months (table 10b) the opportunity-cost B method and the proxy-good method both had 7 out of 11 significant factors, with hours as the dependent variable. Factors common to both methods were (changes in) behavioural disturbances, IADL impairments, formal care, and number of tasks. When costs was used as dependent variable, the proxy-good method had the highest number of significant factors, i.e. 6 out of 11, with (changes in) cognitive impairment, behavioural disturbances, physical function, formal care, number of tasks and caregiver network as significant variables. The opportunity-cost-A method and the CarerQol method had the lowest number of significant factors both at 6 months and 12 months. All associations found, were in the expected direction. When significant, an increase in formal care was associated with an increase in informal care.

*Multivariate analyses*

Tables 10c and 10d show the results of the multivariate regression analyses.

Regarding the change scores after 6 months (table 10c), the proxy-good method was most responsive to changes in the factors with R-Squares of 20% and 14% for respectively number of hours and costs as dependant variables. The change in physical function and number of tasks were acting as independent factors when the change in number of hours was the dependent variable, whereas change in number of tasks was the only significant factor when change in costs was the dependant variable. The opportunity-cost-A method and the CarerQol method were least responsive as indicated by the lowest R-Squares.

Regarding the change scores after 12 months (table 10d), the opportunity-cost B method was most responsive with R-Squares of 21% and 23% for (change in) number of hours and costs respectively as dependant variables, followed by the proxy-good method with R-Squares of 13% and 10% for (change in) number of hours and costs. In the opportunity-cost B method, change in informal care was associated with changes in behavioural disturbances, formal care and living situation. The only factor common to both the opportunity-cost B and proxy-good regression models was the change in formal care. Again, the opportunity-cost-A method and the CarerQol method were least responsive. All associations found, were in the expected direction. When significant, an increase in formal care was associated with an increase in informal care.

**Table 10a:** Univariate longitudinal regression models (change scores: 6 months - baseline)

	Proxy-good		Proxy-good		Opp. Cost A		Opp. Cost B		Opp. Cost B		CarerQoI	
	Change total hours	Change total costs	Change VAS	Change VAS								
<b>Patient</b>												
Health	-	-	ns	ns	ns	ns	-	-	ns	ns	+	+
Comorbidity	ns	ns	ns									
Dementia severity	ns	ns	ns									
Cognitive impairment	ns	ns	ns									
Behavioral disturbances	+	+	ns	ns	ns	ns	+	+	ns	ns	ns	ns
(I)ADL impairments	+	+	ns	ns	ns	ns	+	+	ns	ns	ns	ns
Physical function	-	-	ns	ns	ns							
Formal care	+(0.08)	ns	ns	ns								
<b>Caregiver:</b>												
Number of tasks	+	+	-	ns	-	-						
Living together	ns	ns	ns	+	ns	+	ns	ns	ns	ns	ns	ns
Caregiver network	ns	ns	ns									

+ = positively associated ( $p < 0.05$ ); - = negatively associated ( $p < 0.05$ ); ns = not significant; p-values between 0.05 - 0.10 are shown between brackets

**Table 10b:** Univariate longitudinal regression models (change scores: 12 months - baseline)

	Proxy-good		Proxy-good		Opp. Cost A		Opp. Cost B		Opp. Cost B		CarerQoI	
	Change total hours	Change total costs	Change VAS	Change VAS								
<b>Patient</b>												
Health	ns	ns	+	ns	ns	ns	- (0.08)	ns	ns	ns	+	ns
Comorbidity	ns	ns	ns									
Dementia severity	ns	ns	ns	ns	ns	ns	+	ns	ns	ns	ns	ns
Cognitive impairment	-	-	ns	ns	ns							
Behavioral disturbances	+	+ (0.07)	ns	ns	ns	ns	+	ns	ns	ns	ns	ns
(I)ADL impairments	+	ns	ns	ns	ns	ns	+	ns	ns	ns	ns	ns
Physical function	-	- (0.10)	+	ns	+	ns						
Formal care	+	+	ns	ns	ns	ns	+	ns	ns	ns	ns	ns
<b>Caregiver:</b>												
Number of tasks	+	+	ns	ns	ns	ns	+ (0.06)	ns	ns	ns	ns	ns
Living together	ns	ns	+	ns	ns	ns	+	ns	ns	ns	ns	ns
Caregiver network	+	+ (0.09)	ns	ns	ns							

+ = positively associated (p<0.05); - = negatively associated (p<0.05); ns = not significant; p-values between 0.05 – 0.10 are shown between brackets

**Table 10c:** Multivariate longitudinal regression models (change scores: 6 months - baseline)

	Proxy-good		Proxy-good		Opp. Cost A		Opp. Cost B		CarerQoI	
	Change total hours	Change total costs								
<b>Patient</b>										
Health										
Comorbidity										
Dementia severity										
Cognitive impairment										
Behavioral disturbances										
(I)ADL impairments										
Physical function	-									
Formal care										
<b>Caregiver:</b>										
Number of tasks	+	+								
Living together						+				
Caregiver network										
<b>R-Square final model</b>	<b>0.197</b>	<b>0.141</b>	<b>ns</b>	<b>0.046</b>	<b>0.057</b>	<b>0.080</b>	<b>0.046</b>	<b>0.080</b>	<b>0.046</b>	<b>0.046</b>

+ = positively associated (p<0.05); - = negatively associated (p<0.05) in the final models; ns: no significant variables found

**Table 10d:** Multivariate longitudinal regression models (change scores: 12 months - baseline)

	Proxy-good		Proxy-good		Opp. Cost A		Opp. Cost B		Opp. Cost B		CarerQoI	
	Change total hours	Change total costs										
<b>Patient</b>												
Health			+									
Comorbidity												
Dementia severity												
Cognitive impairment												
Behavioral disturbances												
(I)ADL impairments												
Physical function												
Formal care	+			+								
<b>Caregiver:</b>												
Number of tasks	+			+								
Living together												
Caregiver network												
<b>R-Square final model</b>	<b>0.131</b>	<b>0.104</b>	<b>0.044</b>	<b>0.104</b>	<b>ns</b>	<b>ns</b>	<b>0.206</b>	<b>0.233</b>	<b>0.206</b>	<b>0.233</b>	<b>0.032</b>	<b>0.032</b>

+ = positively associated (p<0.05); - = negatively associated (p<0.05) in the final models; ns: no significant variables found

Overall, R-Squares of the multivariate longitudinal models were substantially lower compared to R-Squares of the multivariate cross-sectional models.

### **Research question 4**

#### *Most suitable method to measure and value informal care in our study population*

Based on both the cross-sectional and longitudinal regression analyses to compare construct validity and responsiveness of each method respectively, it can be concluded that the proxy-good method is the best method to measure and value informal care in patients with dementia or another cognitive impairment and their caregivers.

## **Discussion**

This chapter describes and discusses different methods to measure informal care (by means of a questionnaire, developed by the iMTA in Rotterdam, the Netherlands) in patients with dementia or another cognitive disorder. These methods include the proxy-good-method, the opportunity-cost-method (time forgone and preferred time) and the CarerQol-method.

The literature search revealed various factors (mainly patient characteristics), which were associated with (the amount of) informal care. It should be noted that very little caregiver characteristics were used in these studies and that for the greater part cross-sectional analyses were conducted. The variety of the factors found (i.e. clinical characteristics of the patients, characteristics of the caregiver and social circumstances) indicate the complexity of informal care. Apparently, informal care is interrelated with many different factors. Factors emerging from the literature search as being associated with informal care, were included in our regression analyses.

Regarding construct validity, the proxy-good method performed best in the multivariate cross-sectional model as indicated by the highest R-Squares explaining more than 60% of the variance. Regarding responsiveness, the proxy-good method was most responsive to changes occurring after six months, as indicated by the highest R-Squares in the multivariate longitudinal models. After 12 months, the opportunity-cost B method was most responsive, followed by the opportunity cost method. Overall, based on the regression analyses, it can be concluded that the proxy-good method is most suitable to measure and value informal care, followed by the opportunity-cost B method (preferred time) when applied in patients with dementia or another cognitive impairment and their caregivers. The opportunity-cost A method and the CarerQol method performed worst with respect to both construct validity and responsiveness.

It should be noted that judgments regarding construct validity and responsiveness were merely based on comparison of the R-Squares, which only provides a relative outcome. Therefore, based on results from this study it cannot be concluded that the proxy-good method has a 'good' construct validity and is actually 'responsive' to changes, as no absolute criterion was established to determine this. Nevertheless, the R-Squares of the

cross-sectional models were fairly high for the proxy-good method (over 60%), which may indicate a good construct validity. On the other hand, the R-Squares of the longitudinal models were considerably lower. Therefore, it cannot be concluded that the responsiveness of proxy-good method is satisfactory, but only that – based on the change scores at 6 months- the proxy-good method performed best compared to the other methods, and – based on change scores at 12 months - the opportunity cost B method was the most responsive. Why the opportunity costs B method performed best at 12 months, is yet unclear.

Furthermore, it should be noted that lack of clarity exists regarding the definition and most appropriate approach for assessing responsiveness (Husted et al., 2000; Terwee et al., 2003). The methodology of assessing responsiveness is thus not so straightforward. We studied the so called external responsiveness which reflects the extent to which changes in an instrument (informal care) relate to corresponding changes in the external standard or anchor. Therefore, this method is also referred to as an anchor-based approach. For example, with respect to a study we performed examining the construct validity and responsiveness of the EQ-5D(+C), the MMSE was used as an anchor (Wolfs et al., 2007a). However, in the current study, we did not use a single anchor or standard. All factors emerging from the review that were included in our univariate regression analyses can in fact be considered possible anchors. However, some factors are highly correlated with each other. Also, the influence of one factor on informal care may be offset by the influence of another factor. Therefore, it is clear that informal care is a multi-component construct which should be evaluated by means of multivariate regression analyses. The use of regression analyses to study external responsiveness is recommended (Husted et al., 2000).

Regarding changes over time, the amount of informal care increased over time (with a mean increase of 1.7 hours) in all monetary methods, which is in line with the progressive decline of the patients. In addition, more people made use of formal care as the disease progressed. As results of the longitudinal regression models pointed out that formal care was positively associated with informal care, this indicates that formal care complements rather than substitutes informal care. It is plausible that informal care becomes a way of life, integrated in the daily routine of the caregivers. Once a person starts caring for a loved one, he or she may not be willing to give up these activities and most likely will take on other tasks or increase the intensity of the tasks. The CarerQol remained relatively stable over time and was least responsive to changes. This finding is confirmed in the literature on caregivers burden, in which many intervention studies have shown that caregivers burden is remarkably insensitive to change (Verhey et al., 2007b). Overall, changes in the factors – in general- were small, which may in part explain the relatively low responsiveness in this study. Moreover, as mentioned before, the influence of one factor on informal care may be offset by the influence of another factor.

Based on the results from this study, it is impossible to indicate which combination of factors has an undisputable association with informal care. Although the univariate cross-sectional regression analyses identified a number of factors common to each method, the factors emerging from the multivariate cross-sectional analyses were rather

diverse. However, regarding the monetary methods, number of tasks was a significant factor in each regression model. Inclusion of this number of tasks in the regression models was based on the finding that it emerged from the literature search as being associated with burden, which we used as a proxy for the CarerQoL. A striking result was that, in this study, number of tasks was not associated with the CarerQoL. Moreover, the positive association of number of tasks with the monetary methods seems rather obvious, as they constitute almost the same concept.

In the longitudinal analyses, none of the factors had an association with each of the methods. However, formal care, number of tasks and living together seem important in explaining changes in informal care.

Nevertheless, it is difficult to indicate which variables were most important in general.

A remarkable result of this study deserves some attention. The greater part of the caregivers were children (in-law) of the patients instead of their spouses which is generally the case in dementia (de Vugt et al., 2005b). It is therefore possible that we did not capture the primary caregiver for all patients (although this was asked explicitly) and the amount of informal care presented in this study could therefore be an underestimation of the actual amount. However, we did account for additional caregivers and we included this caregiver network in our regression analyses. Furthermore, we expect that this does not affect the comparison of the different methods to measure and value informal care regarding their construct validity and responsiveness, as these methods were applied to the same caregivers.

Another point of interest concerns the valuation part of the opportunity cost methods, especially the valuation of leisure time. In this study, leisure time was valued against the wage rate of a paid housekeeper. Some argue that the valuation of leisure time should be based on the wage rate of the caregiver, on which a correction is applied (Torgerson et al., 1994). The costs presented in this study would then have been substantially higher for both opportunity costs methods (2A and 2B), since most caregivers gave up leisure time (method A) to provide informal care and preferred more leisure time (method B). Whether this would have changed our results (regarding the comparison of the different methods) is yet unclear but will be investigated. In this study, the valuation method of van den Berg (van den Berg et al., 2006) was used for comparability with published results on this issue.

A notable finding is that, with regard to the monetary methods, most hours were measured by means of the proxy-good method whereas both opportunity-cost methods revealed (almost) similar and less hours. The higher amount of hours in the proxy-good method may be due to an overestimation of the provided care. This has also been suggested by van den Berg et al (van den Berg et al., 2006). On the other hand, it can be argued that it is less difficult to indicate time spent on specific tasks than to indicate the amount of time forgone or preferred time, especially since the greater part of the caregivers was not involved with paid work anymore. It is therefore also plausible that an underestimation occurred when the opportunity costs methods are applied. It is however

not clear how this would have effected the comparison of construct validity and responsiveness of the methods.

There are some possible limitations that need to be discussed

In our study, informal care was divided in ADL activities, HDL activities and IADL activities. However, there is great variability in the literature regarding the definition and quantification of informal care time especially concerning the inclusion of supervision as a separate dimension of informal care (McDaid, 2001; Wimo et al., 2002). Caregivers may be able to perform activities while providing care such as supervision at the same time. Since supervision is not included explicitly in our survey, this could be a limitation. More uniformity about the concept of informal care is warranted. Obviously, this limitation does not apply to the opportunity-cost methods. With regard to construct validity, it should also be noted that only the baseline results were used for the regression models. Although not expected, regression analyses on the 6 months and 12 months results may lead to different conclusions regarding the construct validity of each of the three methods.

A final limitation could be that a selection bias took place because we merely included those patients who completed all three measurements. Data from the non-completers did however not significantly differ from the data included in this study (data not shown). Although we do not expect that this will change our conclusions, we plan to perform additional analyses using multiple imputation to account for missing data.

Overall, it can be concluded that, for now, the proxy-good method seems a worth-while method to measure and value informal care. Future research into the opportunity-cost B method (preferred time) might be useful, as responsiveness of this method was better than the responsiveness of the proxy-good method when 12-months data were used. Although the opportunity-cost method theoretically is preferred since it reproduces an actual amount of time forgone (or preferred time in this case), we contend that the proxy-good method is the most suitable method in economic evaluations to measure and value informal care in patients with dementia or another cognitive impairment and their caregivers. Future research is however essential to confirm the generalizability of our results in other study populations. Moreover, more research is necessary to decide which tasks should be incorporated (i.e. the definition of informal care) and how leisure time should be valued in using the opportunity cost methods.

The CarerQol method and the opportunity-cost A method do not seem useful to measure and value informal care in this patient population and their caregivers. Regarding the CarerQol, the practical application and integration in economic evaluation is yet unclear. In this respect, we argue that a monetary measure is used in economic evaluations since this can easily be implemented in the numerator of the incremental cost-effectiveness ratio as costs.



# **PART THREE**

Findings of the MEDICIE-study



# Chapter 6

Effectiveness of an integrated multidisciplinary diagnostic approach for psychogeriatric patients: a randomized controlled trial



**British Journal of Psychiatry, Accepted**

Claire AG Wolfs, Alfons Kessels, Carmen D Dirksen, Johan L Severens and Frans RJ Verhey

## Abstract

**Declaration of interest:** None

**Background:** An integrated multidisciplinary approach to dementia is often recommended but has rarely been evaluated.

**Aim:** To evaluate the clinical effects of an integrated multidisciplinary diagnostic facility for psychogeriatric patients (DOC-PG).

**Method:** Patients suspected of having complex psychogeriatric problems were randomly allocated to DOC-PG (n=137) or to treatment as usual (n=93). They were assessed at baseline, and at 6 and 12 months later by means of personal interviews with the patient's proxy. The primary outcome was health-related quality of life (HRQoL), assessed using the visual analogue scale (VAS) of the EuroQoL-5D.

**Results:** HRQoL was improved at 6 months in the intervention group, whereas HRQoL in the control group decreased. Furthermore, more patients in the intervention group experienced a clinically relevant improvement of 10 points or more on the VAS at both follow-up measurements.

**Conclusion:** An integrated multidisciplinary approach improves dementia care.

## Introduction

An integrated multidisciplinary approach to diagnosing and managing complex disorders such as dementia is generally recommended (APA, 1997; CBO, 2005; NICE, 2006) because no single medical specialty has the expertise to deal with the complex range of mental, physical, and social problems that accompany dementia (Collighan et al., 1993; Verhey et al., 1993). However, to date no randomized clinical trials have investigated the value of such an approach to dementia care (Wolfs et al., 2006). Recently, an outpatient diagnostic facility, the Diagnostic Observation Centre for PsychoGeriatric patients (DOC-PG), was established in Maastricht, the Netherlands. This facility combines the hospital-based approach of a memory clinic with the care-oriented approach of a regional community mental health team and aims to provide general practitioners (GPs) with detailed diagnostic and therapeutic advice for patients with cognitive disorders.

The MEDICIE study (Maastricht Evaluation of a Diagnostic Intervention for Cognitively Impaired Elderly) is a randomized controlled trial comparing the efficacy and efficiency of DOC-PG and usual care. We hypothesized that DOC-PG would have beneficial effects on health-related quality of life (HRQoL) compared with usual care, based on the assumption that both diagnosis of the cognitive disorder according to specialist guidelines and appropriate assessment of the patient's social circumstances are prerequisites for the best possible care for the patient and his/her family. We here report on our findings.

## Methods

We used a cluster randomized study design. The sample size was determined using a power calculation that ensures the detection of at least 80% of the differences in the mean score on the visual analogue scale (VAS) of the EQ-5D at 5% significance. Assuming an intra-class correlation coefficient of 0.10, 108 patients per group were required.

### Study participants

The MEDICIE study was approved by the Medical Ethics Committee of Maastricht University Hospital. Patients were recruited from July 2002 to August 2004 from 60 general practices in the Maastricht region, 7 practices in the Sittard region, and 3 practices in the east Heerlen region (all three regions are in the province of Limburg, in the south of the Netherlands). GPs in these practices were asked to refer all patients with possible dementia or a cognitive disorder. The inclusion and exclusion criteria are given in the Table.

**Trial inclusion and exclusion criteria****Inclusion criteria:**

Aged 55 years or older;  
 Suspected of having dementia or a cognitive disorder;  
 No referral to other local/regional services in the last two years; and  
 Availability of a proxy (visiting the patient at least once a week).

**Exclusion criteria:**

Acute disorders that need a prompt therapeutic intervention; and  
 Living in a nursing home.

**Randomization**

Randomization took place at the general practice level to prevent contamination at a patient and GP level. In order to control for effects related to differences in general practices, all practices were asked to supply information about the practitioner's experience, demographic characteristics of the practice population, and the practitioner's affinity with geriatric problems. On the basis of these data, two groups of practices were formed, and the patients from these practices were randomly assigned (by means of a computer program, randols) to either the intervention group (DOC-PG) or the control group (usual care). GPs were initially blinded to this procedure and the random allocation sequence was concealed for most of the participants.

**Interventions***DOC-PG*

The DOC-PG has expertise in the fields of old age psychiatry, geriatric medicine, neuropsychology, physiotherapy, occupational therapy, geriatric nursing, and mental health nursing and hence enables multidisciplinary assessment of patients, covering aspects such as somatic screening, psychogeriatric assessment, and evaluation of the required levels of care for the patient and his/her carer. GPs can refer patients to the DOC-PG if a cognitive disorder is suspected. During a 2-week diagnostic screening procedure, patients are visited once at their homes and are asked to visit the University Hospital departments of geriatric medicine and geriatric psychiatry. A CT scan and various blood tests are performed. The results are then discussed at a weekly interdisciplinary meeting in which a definite diagnosis is made and a treatment plan is formulated. The patient's GP is sent a summary of the assessments, the multi-axis diagnosis, and recommendations for treatment and management. Thereafter the GP is responsible for his/her patients even though further investigations may have been recommended.

*Usual care*

In the control group, GPs provided usual care. This means that the patients were not referred to the DOC-PG and that either the diagnosis was made by the GP or the patient was referred to one of the separate regional services, e.g. the Maastricht Memory Clinic, geriatric medicine, or the department of mental health for the elderly at the mental health community service (Verhey et al., 1993).

**Outcome measures**

Interviewers, who for practical reasons could not be kept totally blinded to the treatment assignment, assessed participants at baseline (within 2 weeks of the DOC-PG or usual care intervention) and at 6 and 12 months after the baseline measurement. All outcome measures, except the Mini Mental Scale Examination (MMSE), were collected through personal interviews with the patient's proxy (i.e., we measured the proxy's perception of the patient's health). The health-related quality of life (HRQoL) of the patient and the carer was the primary outcome of this study. Because we expected that the patients would show a complex range of mental, physical, and social problems, we chose to use the EQ 5D to measure HRQoL. This instrument has been validated in a number of European countries including the Netherlands (Brooks, 1996) and provides a simple descriptive profile and a single index value for health status. It is widely used in cost-utility analyses (Dolan, 1997; Kind, 2003). The EQ-5D consists of a visual analogue scale (VAS) ranging from 0 (worst imaginable health state) to 100 (best imaginable health state). Changes in VAS scores over the course of one year was the primary outcome variable in this study. A difference of 10% or more between the intervention group and the control group on the VAS was a priori considered a clinically relevant difference, as described in the original protocol that preceded the start of the study. The number of patients experiencing this clinically relevant difference were compared between both groups.

As secondary outcome instrument, we used the SF-36, a 36-item generic questionnaire used to measure nine relevant aspects of the health-related functioning of patients. Higher scores reflect better functioning (Brazier et al., 1992; VanderZee et al., 1996; Ware et al., 1992). Additional secondary outcome measures were scores on the MMSE, the Global Deterioration Scale (GDS), NeuroPsychiatric Inventory (NPI), the Instrumental Activities of Daily Living Scale (IADL), and the Cornell Scale of Depression in Dementia (CSDD). The MMSE (Folstein et al., 1975b; Kok et al., 2002) assesses the severity of cognitive decline; the GDS evaluates seven stages of global functioning in patients with a primary degenerative dementia such as Alzheimer's disease (Reisberg et al., 1982); and the NPI (Cummings, 1997; Cummings et al., 2001) appraises patients' behavioural and psychological problems. The IADL scale measures seven areas of more complex activities required for optimal independent functioning, with scores reflecting whether patients are completely independent, in need of assistance, or are completely dependent on others for the performance of specific activities (Lawton et al., 1969). The CSDD (Alexopoulos et al., 1988; Kurlowicz et al., 2002) is a 19-item depression scale that was developed specifically to measure the severity of depressive symptoms in older adults with dementia. Higher scores on all instruments, except on the MMSE, are indicative of more severe problems.

**Statistical analyses***Missing data*

Missing items were imputed using a regression model, and missing data or data missing covariates were imputed using Rubin's multiple imputation procedure (Rubin, 1987). This

method generates ten different data sets for imputed data. All analyses were performed with each of these ten data sets and the results were pooled. Complete missing data were imputed if participants had completed the instrument on two occasions but not if they had completed only the baseline measurement. These patients were considered study dropouts. With a logistic regression analysis, the probability of being a study dropout was assessed and, with these probabilities, p-weights were calculated ( $1/(1-\text{predicted probability})$ ). This allowed for differential weighting of people in data analysis (Little et al., 1987). The data of patients who had died after the baseline measurement and before the follow-up investigations were not analyzed.

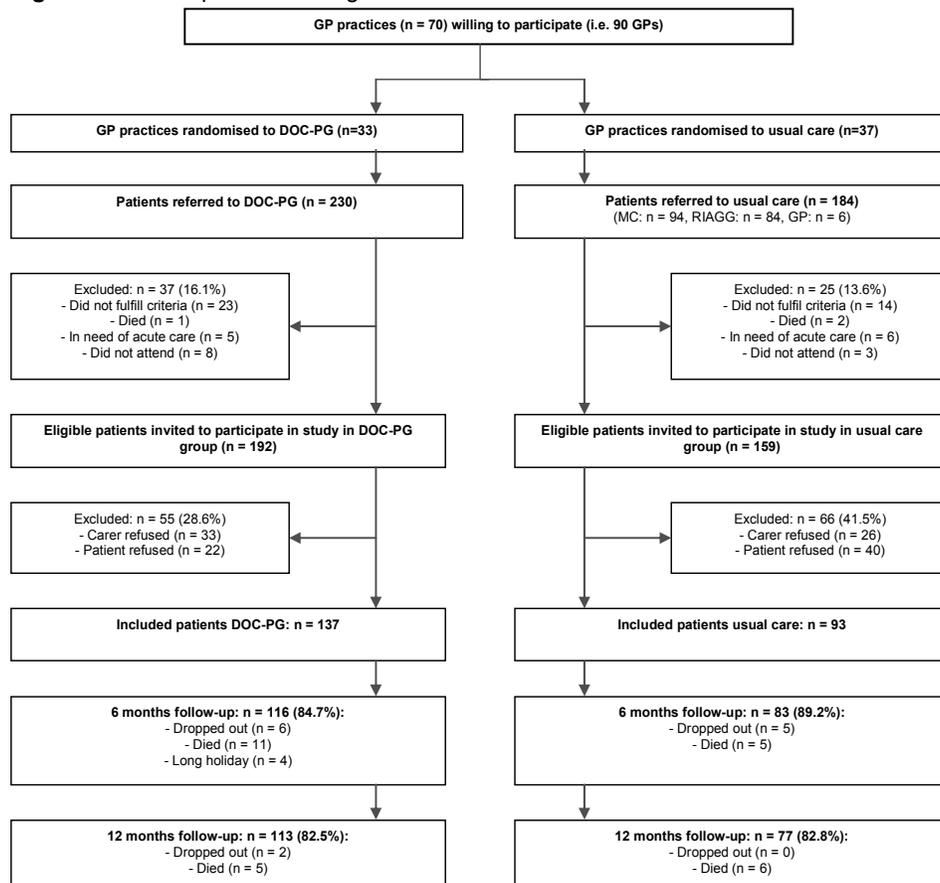
### *Data analysis*

Weighted regression models, clustered on general practice level, were used to examine the influence of group (intervention or usual care) on outcome on each of the instruments. The cluster option was used to account for the correlated data within general practices. The dependent variables in the models were the scores and the change, over time, of the participants' scores on the instruments, with (baseline) characteristics (i.e. group, gender, age, diagnosis, and MMSE score) as independent variables. SPSS version 12.0.1 was used to calculate the p-weights, to examine group differences, and to impute the missing items by means of a regression model (missing value analysis). Rubin's multiple imputation procedure and our main regression analyses were performed using STATA version 8.2. The background characteristics of the participants (both the patients and their proxies) were summarized using descriptive statistics. Response distributions of the instruments are provided.

## **Results**

### **Participants**

Of the general practices included in this study, 33 were randomized to the DOC-PG intervention and 37 to usual care. Between July 2002 and August 2004, 414 patients were referred for further treatment. Of these patients, 351 were eligible for this study and 230 (65%) agreed to participate. Non-participants were comparable to the participants with respect to age (77.8 (SD 6.4) and 77.8 (SD 6.7), respectively) and gender (59.5% women and 66.2%, respectively). The main reason for not participating was that participation would be too much of a burden for either the patient or the proxy. We followed up 94.3% of the patients. Eight patients (5.8%) in the intervention group and five patients (5.4%) in the control group withdrew from the study because "the burden is too high" (intervention group, n=2; control group n=2), "the proxy has health problems" (intervention group, n=4; control group n=1), and "participation in the study has no beneficial effects for the patient" (intervention group, n=2; control group n=2). The flow of participants through the trial is shown in Figure 1. The demographic and clinical characteristics of the participants were similar at baseline in both groups (Tables 1, 2 and 3).

**Figure 1:** Flow of patients through trial

## Outcomes

The mean score on the social functioning component of the SF-36 was significantly higher ( $p=0.03$ ) in the intervention group than in the control group at 6 months (see Table 2); no other differences in mean scores were found between the groups. The mean difference scores for the EQ-5D over time were significantly different between the two groups ( $p=0.04$ ). HRQoL, measured with the VAS, improved slightly in the intervention group (1.5 points on the VAS) but decreased in the control group (4 points on the VAS).

We found a mean group difference of 9.6% on the VAS of the EQ-5D after 12 months, which is very close to our initial expectations. The proportion of patients who improved more than 10% (of the group difference) on the VAS was significantly greater ( $p=0.01$ ) in the intervention group (39.0%) than in the control group (22.1%). With an improvement of 0.03 on the population utility score of the EQ-5D being considered a clinically relevant improvement (Marra et al., 2005), significantly ( $p=0.04$ ) more patients in the intervention

**Table 1:** Baseline demographic characteristics of patients assigned to the intervention or control group

		Intervention group (n = 137)	Control group (n = 93)
<b>Sex patient</b>	Female (%)	89 (63.1%)	59 (63.4%)
<b>Age patient</b>	Mean (SD)	78.3 (6.5)	77.3 (6.8)
	Range	[55 – 93]	[60 – 94]
<b>Relationship proxy:</b>			
	Spouse	51 (37.2%)	37 (39.7%)
	Child (in law)	73 (53.3%)	46 (49.6%)
	Other	13 (9.5%)	10 (10.7%)
<b>Sex proxy</b>	Female (%)	90 (65.7%)	62 (66.7%)
<b>Age proxy</b>	Mean (SD)	60.4 (13.5)	59.8 (13.9)
	Range	[30 – 84]	[34 – 91]
<b>Dementia:</b>		97 (70.8%)	66 (71.0%)
	Alzheimer's Disease (AD)	66 (48.2%)	31 (33.3%)
	Vascular Dementia (VD)	12 (8.8%)	14 (15.1%)
	Mixed Dementia	15 (10.9%)	6 (6.5%)
	Other Dementia	4 (2.9%)	15 (16.1%)
<b>No Dementia:</b>		40 (29.2%)	27 (29.0%)
	Cognitive Impairment / MCI	24 (17.5%)	15 (16.1%)
	Other Cognitive Impairment	16 (11.7%)	12 (12.9%)

group than in the control group showed a clinically relevant improvement after 6 months (42.1% versus 37.7%, respectively). Furthermore, the groups differed significantly ( $p=0.02$ ) on the change score in the social functioning component of the SF-36, with patients in the intervention group showing a larger improvement than patients in the control group. After 12 months, more patients in the intervention group than in the control group showed a clinically relevant improvement in HRQoL measured as an improvement of more than 10% of the group difference on the VAS (32.6% versus 18.6% respectively  $p=0.01$ ) and on the utility score of the EQ-5D (40.6% versus 24.7%, respectively;  $p=0.00$ ). The groups did not differ in terms of clinical outcome measures (see Table 3). We investigated whether these differences in HRQoL between the groups were related to the use of cholinesterase inhibitors in a post-hoc analysis. In general, few patients received cholinesterase inhibitors (mean 14.6%), but significantly more patients in the intervention group than in the control group were treated with cholinesterase inhibitors (18.3% and 9.1%, respectively;  $p=0.01$ ). However, the use of cholinesterase inhibitors had no influence on the proportion of patients who showed a clinically relevant improvement in HRQoL (after 6 months  $p=0.15$ ; and after 12 months  $p=0.53$ ).

**Table 2:** Mean HRQoL outcomes at follow up and results of regression analyses (group differences). Values are the means, standard deviations and mean changes over time between the groups for each instrument.

DOC-PG: n=115 Usual care: n=77		Baseline (T1)	6 months (T2)	Mean difference between groups on change score (T2 - T1) [CI]	12 months (T3)	Mean difference between groups on change score (T3 - T1) [CI]
<b>EQ-5D:</b>						
VAS	DOC-PG Usual care	58.7 (20.7) 60.0 (19.3)	60.2 (18.3) 56.1 (18.8)	5.4 [0.29 - 10.45] *	58.3 (20.5) 54.4 (21.8)	5.2 [-0.58 - 10.94]
VAS>10	DOC-PG Usual care	- -	39.0% 22.1%	0.8 [0.17 - 1.40]**	32.6% 18.2%	0.8 [0.16 - 1.45]**
Utility score	DOC-PG Usual care	0.54 (0.33) 0.54 (0.30)	0.58 (0.33) 0.53 (0.33)	0.1 [-0.04 - 0.12]	0.49 (0.35) 0.43 (0.34)	0.1 [-0.04 - 0.14]
<b>SF36:</b>						
Physical function	DOC-PG Usual care	50.5 (29.3) 51.7 (30.1)	48.6 (29.6) 49.3 (31.8)	0.4 [-5.67 - 6.50]	41.9 (31.0) 44.1 (31.8)	-1.1 [-8.55 - 6.33]
Role physical	DOC-PG Usual care	50.3 (41.4) 47.0 (43.0)	54.8 (40.2) 46.8 (39.8)	3.4 [-7.78 - 14.48]	48.5 (39.7) 45.7 (39.6)	-1.1 [-14.49 - 12.30]
Bodily pain	DOC-PG Usual care	68.2 (29.7) 68.1 (27.9)	71.7 (26.3) 69.2 (25.0)	1.6 [-6.70 - 9.92]	70.3 (27.9) 74.4 (25.7)	-4.5 [-13.27 - 4.27]
General health	DOC-PG Usual care	44.3 (22.0) 40.7 (21.0)	46.6 (21.1) 43.9 (21.7)	-1.3 [-6.75 - 4.16]	45.4 (21.7) 40.5 (20.9)	1.2 [-4.03 - 6.40]
Vitality	DOC-PG Usual care	48.3 (23.6) 47.5 (22.9)	49.4 (20.4) 45.7 (22.2)	2.3 [-3.64 - 8.16]	44.8 (21.8) 44.7 (21.5)	-1.3 [-7.48 - 4.86]
Social function	DOC-PG Usual care	53.2 (33.2) 55.4 (29.2)	66.4 (28.1) 55.6 (33.6)	11.8 [1.96 - 21.54]***	55.7 (33.6) 52.5 (35.0)	4.7 [-4.97 - 14.46]
Mental health	DOC-PG Usual care	57.0 (22.9) 57.5 (21.5)	60.7 (21.5) 56.3 (22.8)	4.3 [-0.77 - 9.40]	57.9 (23.3) 58.9 (20.8)	-0.7 [-6.60 - 5.14]
Role emotional	DOC-PG Usual care	41.4 (40.6) 45.0 (40.5)	50.5 (40.2) 46.7 (42.8)	6.2 [-7.60 - 20.06]	48.3 (41.7) 54.7 (40.8)	-2.9 [-16.45 - 10.57]

[CI] = confidence interval; VAS>10 = Proportion of patients who improved 10% or more of the group difference on the VAS compared with baseline (0=<10, 1=>10); \* = significant (p=0.04); \*\* = significant (p=0.01); \*\*\* = significant (p=0.02)

**Table 3:** Mean clinical outcomes at follow up and results of regression analyses (group differences). Values are the means, standard deviations, and mean changes over time between groups for each instrument.

DOC-PG: n=115 Usual care: n=77		Baseline (T1)	6 months (T2)	Mean difference between groups on change score (T2 - T1) [CI]	12 months (T3)	Mean difference between groups on change score (T3 - T1) [CI]
MMSE	DOC-PG	20.5 (6.0)	18.8 (7.8)	-0.9 [-2.23 - 0.34]	18.0 (7.7)	0.0 [-1.43 - 1.48]
	Usual care	19.8 (6.6)	19.2 (17.5)		17.4 (8.8)	
GDS	DOC-PG	4.2 (0.9)	4.4 (1.2)	-0.1 [-0.44 - 0.14]	4.7 (1.0)	0.0 [-0.23 - 0.21]
	Usual care	4.2 (1.1)	4.6 (1.2)		4.7 (1.1)	
IADL	DOC-PG	17.1 (5.7)	18.7 (6.2)	-0.1 [-1.16 - 1.06]	20.2 (6.1)	-0.7 [-1.85 - 0.46]
	Usual care	16.5 (6.1)	18.12 (6.3)		20.4 (6.5)	
Cornell	DOC-PG	7.9 (4.3)	7.2 (4.7)	-1.3 [-2.62 - 0.07]	7.5 (5.0)	-0.8 [-2.24 - 0.69]
	Usual care	7.4 (3.8)	7.9 (5.0)		7.8 (4.9)	
NPI total	DOC-PG	23.4 (15.6)	24.3 (18.5)	-4.0 [-8.46 - 0.54]	28.4 (20.8)	-1.2 [-6.06 - 3.63]
	Usual care	22.6 (16.5)	27.3 (20.8)		29.0 (21.0)	

[CI] = [confidence interval]

## Discussion

In this study of the effects of an integrated multidisciplinary diagnostic observation centre for psychogeriatric patients, we found a modest but significant improvement in the proxy perception of HRQoL 6 months after the baseline measurement, confirming our initial hypothesis. Furthermore, more patients in the intervention group experienced a clinically relevant improvement of 10% or more of the group difference on the VAS and 0.03 or more on the utility score of the EQ-5D, after both 6 and 12 months. These differences were not attributable to the use of cholinesterase inhibitors. We found no differences in cognitive functioning, behavioural and psychological problems, ability to perform activities of daily living, and emotional functioning. In the absence of any significant effects on the secondary clinical outcome measures, it is difficult to establish what may have caused this favourable outcome in the intervention group. Because the DOC-PG provides GPs with different types of advice, such as adaptation of medication, improvement of sensory function by ear syringing or testing eyesight, further referral to other hospital departments and to paramedical disciplines, and advice to initiate extra care, such nursing home placement, respite care, or services like 'meals on wheels' (Wolfs et al., 2007b), the improved outcome probably reflects the sum of the different advice and recommendations given.

The similarity of outcomes other than HRQoL in the two groups may be because the intervention (DOC-PG) provides access to two healthcare facilities that are available to the control group. While medical centres tend to focus on medical diagnostics and pharmacotherapy, community mental health services focus on the provision of appropriate levels of care and support for patients and their carers. It is thus not surprising that the two approaches had comparable effects on psychological and behavioural problems, emotional functioning, and ability to manage daily life. In this context, usual care in our region is very good and is provided by an active university medical centre and a community mental health service that have collaborated in the past on several projects. Thus, the contrast between DOC-PG and usual care may have been smaller than would be the case in other regions.

To our knowledge, this is the first randomized controlled trial of a multidisciplinary diagnostic approach to dementia. Our results suggest that an integrated approach to dementia, as recommended by dementia guidelines (APA, 1997; CBO, 2005; NICE, 2006), improves patient outcomes. In the absence of a cure for dementia, the finding that (the proxy perception of) HRQoL can be improved with an integrated treatment plan formulated on the basis of a multidisciplinary diagnostic evaluation is important. It should however be noted that the results of this study can not be generalized to nursing home care.

The study had potential weaknesses. Firstly, the design of the study was not optimal because it was not feasible to blind the interviewers assessing the patients and their carers to the treatment assigned. However, all instruments were standardized and the participants received neutral instructions for every instrument. Another potential problem is linked to our inability to keep the random allocation sequence completely concealed, because the person responsible for the allocation of patients also recruited a small number of patients (5%). However, the people who recruited the majority of the patients

were blinded with respect to patient allocation. The blindness of the referring GPs could not be maintained until the end of the study. In order to investigate the potential effects of this on the study results, we compared post-hoc the characteristics of patients in the two groups who were recruited in the first year and in the second year. We did not find any differences within the intervention group (with respect to age, gender, diagnosis, MMSE score and Global Deterioration Scale score); however, there were non-significant differences in diagnosis and GDS score in the control group: in the second year of the inclusion period, more people with a cognitive disorder other than dementia and with a lower GDS score were included. The GPs probably wanted to refer patients to DOC-PG but this was only possible after recruitment was completed. The inclusion of slightly healthier patients (with consequently higher QOL and lower costs) in the latter half of the inclusion period, probably resulted in a less favourable outcome for DOC-PG. Another potential limitation is the use of proxies to complete the questionnaires. We chose to use proxies because of the longitudinal nature of the study, the complex health problems of the study population, and the anticipated progressive global deterioration of intellect and personality of the study population. In the later stages of dementia, proxy measures are generally considered necessary because patients are no longer able to evaluate their own health (Jonsson et al., 2006a; Selai, 1998). The proxy scores on the various instruments may have been biased because of a perceived caregiver burden (Logsdon et al., 2002), but this bias would apply to both groups. Furthermore, it should be emphasized that we measured the proxy's perception of the HRQoL of the patient and not a direct estimate of HRQoL. Another problem is the presence of missing data, which could have affected the statistical analyses. However, very few data (5%) were missing and multiple imputation procedures provide a useful strategy for dealing with data sets with missing values. Instead of filling in a single value for each missing value, the missing value is replaced by a set of ten plausible values that represent the uncertainty with respect to choosing the right value to impute. This results in statistically valid inferences that properly reflect the uncertainty brought on by missing values (Rubin, 1987; Rubin et al., 1991).

We chose the VAS of the EQ-5D as main outcome because it has good clinimetric properties, is reliable (Parkin et al., 2004), and is easy to administer. Unfortunately, the VAS is more subjective than the descriptive component of the EQ-5D and this could be considered a limitation. A person's state of mind, goals and expectations can influence VAS scores (Dolan, 1996; Ubel et al., 2003); however, we expected that these effects would be present in both groups. Moreover, the VAS enables a personal valuation of the patient's HRQoL, which is an important outcome in the absence of a cure.

There is a growing interest in studies on effectiveness and efficacy of multidisciplinary healthcare models. Currently, an economic evaluation is conducted, comparing the costs of DOC-PG and of usual care. Although a multidisciplinary model is more effective than a monodisciplinary model, it is also more complex, requiring a higher level of organization. It is therefore a challenge for clinicians to combine their professional expertise and share responsibility for a patient given their different, and sometimes opposing, approaches and views on patient care and management. For instance, the role of memory clinics is debated. While some claim that these clinics merely prescribe and monitor drug

treatment (Pelosi et al., 2006), such clinics are becoming increasingly more integrated in the standard care for dementia in the Netherlands (Verhey et al., 2007a). We recommend that all services involved with dementia care integrate (such as in the DOC-PG) rather than polarize, because greater integration will lead to greater continuity of care for patients with dementia. The value of DOC-PG has already been recognized by GPs, as evidenced by the high referral rate among GPs and with the high compliance with DOC-PG recommendations (Wolfs et al., 2007b).

We thank the patients, their carers and the GPs who participated in this study. Additionally, we thank Daniëlle Willems, Audrey Fiddeler, Heidi Lansdaal and Ilvy Mayen for collecting and entering the data. We also thank everyone working at the DOC-PG, the MMC and the CMHT for their help in recruiting patients.

Contributors: FV was the principal investigator and is guarantor. FV, AK and CD designed the trial. CW and JS helped to refine the study design and CW carried out the outcome measures. AK and CW were responsible for statistical analysis and interpretation of the study, and all authors contributed to interpretation. CW drafted the manuscript, and all authors critically revised it for scientific content and approved the final version.

Funding: The study is funded by the Dutch Research Institute for Care- Medical Sciences (ZorgOnderzoek Nederland-Medische wetenschappen, ZON-Mw), grant nr 945-02-055

Ethical approval: Local ethics committee approved the research.



# Chapter 7

Economic evaluation of an integrated diagnostic approach for psychogeriatric patients: Results of a randomized controlled trial



**Submitted**

Claire AG Wolfs, Carmen D Dirksen, Alfons Kessels, Johan L Severens and Frans RJ Verhey

## Abstract

**Declaration of interest:** None

**Background:** Due to the increasing number of elderly people with dementia, the costs of dementia and dementia care are expected to grow rapidly in the coming decades.

**Aim:** To evaluate the cost-effectiveness of an integrated multidisciplinary diagnostic facility for ambulant psychogeriatric patients (DOC-PG) compared to usual care.

**Method:** An economic evaluation was carried out alongside a randomised controlled trial (the MEDICIE-study), comparing the costs and consequences of an integrated multidisciplinary approach to diagnosing dementia with usual care. The economic evaluation was conducted from a societal perspective according to (inter)national guidelines, using a one year time horizon. An incremental cost effectiveness ratio (ICER) was calculated by dividing the difference in costs by the difference in Quality Adjusted Life Years (QALYs). Non-parametric bootstrapping and one-way sensitivity analyses were used to assess the uncertainty in the costs and effects.

**Results:** Compared to patients receiving usual care, patients who visited DOC-PG gained a mean 0.05 QALYs. The incremental costs per QALY amounted to € 1267. This point estimate lies beneath the commonly accepted thresholds and is surrounded by an acceptable range of uncertainty. Sensitivity analyses showed that cost-effectiveness results were robust. DOC-PG is therefore considered cost-effective.

**Conclusion:** Compared to usual care in the Netherlands, DOC-PG is not demonstrable more expensive, despite of the investment in the diagnostic intervention, but it showed a high probability of being more effective. Therefore, we contend that DOC-PG is indeed a cost-effective facility for the diagnosis and management of dementia in ambulant patients.

## Introduction

Dementia is an expensive condition and the costs of dementia contribute significantly to the total health care expenditures (Jonsson, 2004; Slobbe et al., 2006; Wimo et al., 1997). Due to the increasing number of individuals suffering from dementia, the costs of dementia care are expected to increase considerably in the coming years. People affected by dementia also pay a high price in terms of their quality of life. Studies focusing on dementia and dementia care are therefore of great importance. While many studies on this topic have been conducted, randomised studies are rare (Wimo et al., 1997). Economic evaluations of dementia and dementia care are often laden with methodological problems such as difficulties with the measurement and valuation of informal care (Jonsson, 2003). To our knowledge, no studies concerning an integrated approach to dementia and dementia care with a randomised design and a complete health economic evaluation have previously been performed (NICE, 2006; Wolfs et al., 2006).

This article reports on an economic evaluation of a multidisciplinary integrated approach to the diagnosis and management of dementia alongside a randomised controlled trial. A cost-effectiveness analysis was performed in which the Diagnostic Observation Centre for PsychoGeriatric Patients (DOC-PG) was compared to usual care in the Netherlands.

## Methods

### Study participants

The MEDICIE-study (Maastricht Evaluation of a Diagnostic Intervention for Cognitively Impaired Elderly) was approved by the medical ethics committee at Maastricht University Hospital. Patients were recruited from July 2002 to August 2004 from 67 general practices in the Maastricht region and surroundings. General practitioners (GPs) were asked to refer all patients suspected of having dementia or a cognitive disorder. Participants were included if they: 1) were aged 55 years or older; 2) were suspected of having dementia or a psychiatric disorder with co-morbid somatic problems; 3) had not been referred to other local/regional services in the last two years; and 4) had the availability of a proxy (someone who visits the patient at least once a week). Participants were excluded if they: 1) suffered from acute disorders that needed a prompt therapeutic intervention; or 2) lived in a nursing home.

Randomization took place at the level of the general practice to prevent contamination on both the patient level and the GP level. In order to control for effects related to differences in GP practices, an inquiry with GPs was conducted. In this inquiry, each GP's level of experience, demographic characteristics of the GP's patient population and affinity with geriatric problems were assessed. Based on these data, two groups of practices were formed to ensure that the GP practices were comparable on these relevant aspects. Following this, the two groups were randomly assigned (by means of a computer programme (randols)) to either the intervention group (DOC-PG) or the control group (usual care). GPs were initially kept blinded with respect to the group to which their patients were assigned (DOC-PG or usual care) and the random allocation

sequence was concealed for most of the participants. The clinical results of the MEDICIE-study are described elsewhere in more detail (Wolfs et al., 2007e).

### **Interventions**

#### *DOC-PG*

The function of the DOC-PG was to provide multidisciplinary assessment and advice through somatic screening, psychogeriatric assessment, and evaluation of the required levels of care for the patient and his/her carer. Participating disciplines in the DOC-PG included old age psychiatry, geriatric medicine, neuropsychology, physiotherapy, occupational therapy, geriatric nursing and mental health nursing. The diagnostic screening conducted by the DOC-PG took place over the course of two weeks, and consisted of a home visit and two visits to the University Hospital departments of geriatric medicine and geriatric psychiatry. In addition, a CT scan and various blood tests were performed. The results were then discussed at a weekly interdisciplinary meeting in which a definitive diagnosis was made and a treatment plan was formulated. Correspondence was sent to the referring GP in which a summary of the assessments, the multi-axis diagnosis and recommendations for treatment and management were described. After the assessment, the GP was responsible for the patients. However, the recommendations provided to the GP often indicated that the involvement of one or more of the disciplines represented in the DOC-PG was warranted in further treatment.

#### *Usual care*

In the usual care group, the GP provided care as usual. This means that the GP could not refer the patient to the DOC-PG and that the diagnosis was either made by the GP or the GP referred the patient to one of the existing separate regional services, such as the Maastricht Memory Clinic (MMC) (Verhey et al., 2007c), geriatric medicine or the department of mental health for the elderly at mental health community services (CMHT).

### **Cost-effectiveness analysis**

An economic evaluation was carried out alongside the randomised controlled trial (the MEDICIE-study). The economic evaluation compared the costs and consequences of an integrated multidisciplinary approach to diagnosing dementia with usual care. The economic evaluation was conducted from a societal perspective according to (inter)national guidelines (Drummond et al., 1997; Gold et al., 1996; Oostenbrink et al., 2004). An incremental cost effectiveness ratio (ICER) was calculated by dividing the difference in costs by the difference in Quality Adjusted Life Years (QALYs) to reflect a possible investment needed to achieve a higher patient outcome. The maximum amount of QALYs that can be reached within one year is 1. The time horizon was one year.

### **Effectiveness**

The EQ-5D was used to measure patient's Health Related Quality of Life (HRQoL) at baseline, at 6 months follow-up and at 12 months follow-up. The EQ-5D was filled out by the patient's proxy. The EQ-5D instrument was developed and validated in a number of European countries including the Netherlands (Brooks, 1996; Lamers et al., (in press); Lamers et al., 2005) and it has been validated in patients with dementia (Jonsson et al., 2006b). The EQ-5D describes health status according to five dimensions. Each

dimension has three levels, which yields 243 potential combinations of health states across the five dimensions. Dolan et al. (Dolan, 1997) presented 42 of these health states to members of a representative sample of the UK general population, which were then valued using the Time Trade-Off (TTO) method. Based on these valuations, utility scores can be deduced by means of an additive function (Kind, 2003). The utility scores were used to calculate QALYs using the following formula:  $((\text{utility score}_{\text{baseline}} + \text{utility score}_{6 \text{ months}}) / 2) * 0.5 + ((\text{utility score}_{6 \text{ months}} + \text{utility score}_{12 \text{ months}}) / 2) * 0.5$ . Patients who had past away during the year covered by the evaluation were given a utility score of zero from the exact time of death.

### **Costs**

The cost-analysis was performed according to the Dutch guidelines for cost calculations in health care (Oostenbrink et al., 2004). Costs were calculated by multiplying volumes of resource use by the cost price per resource unit and included the period from the first visit (baseline measurement) up to four weeks after the last visit (12 months following randomization). The cost-analysis included health care costs and costs outside the health care sector. Health care costs included intervention costs (DOC-PG or usual care) and costs during follow up such as hospital costs, contacts with the CMHT, contacts with the GP, contacts with other health care professionals, home care, day care, admissions to a nursing home or elderly home, medication and acquisition of goods/aids. Costs outside the health care sector consisted of out-of-pocket costs, travelling costs and informal care costs. All costs were expressed in Euros (1.00 Euro is 1.32 US Dollars and 0.68 British Pounds) of the year 2005. Whenever necessary, cost prices were converted to this reference year (2005) by means of price index numbers. All cost prices were adopted from Oostenbrink (Oostenbrink et al., 2004) unless stated otherwise. The costs of the patients who had past away during the year covered by the evaluation were set at zero from the exact time of death.

#### *Volumes of resource use and costs*

The hospital information system (ZIS), the Electronic Patient Files of the local CMHTs, the registries of local pharmacies, an informal care survey and cost diaries were used to determine the volumes of resource use for each participating patient. Table 1 summarizes the sources of the volumes and the sources of the cost prices.

**Table 1:** Costs per unit. source of the costs and source of the volumes

	Costs per unit (2005)	Source (costs)	Source (volumes)
<b>Hospital:</b> Outpatient	€ 102.91 /visit	1	A
Inpatient	€ 489.86 /day	1	A
Intervention (various)#	Various	2	A
Parking	€ 2.50 /visit	1	A
<b>CMHT:</b> All contacts	€ 127.60 /contact	1	B
<b>Medication:</b> Various#	Various /DDD*	3	C & E
<b>Admission<sup>X</sup>:</b> Elderly home	€ 87.47 /day	1	D
Nursing home	€ 212.00 /day	1	D
<b>Home care<sup>X</sup>:</b> Domestic help	€ 22.33 /hour	1	D
Nursing	€ 41.58 /hour	1	D
<b>Day care:</b> Community centre	€ 5.13 /day	4	D
Elderly home	€ 49.89 /day	4	D
Nursing home	€ 124.52 /day	1	D
<b>General Practitioner:</b> Practice	€ 20.79 /contact	1	E
Home	€ 41.58 /contact	1	E
Telephone	€ 10.39 /contact	1	E
<b>Health care professionals</b>			
Psychiatrist	€ 78.21 /contact	1	E
Psychologist	€ 75.00 /contact	4	E
Physical therapist	€ 23.41 /contact	1	E
Occupational therapist	€ 23.67 /contact	1	E
Speech therapist	€ 25.73 /contact	1	E
Dietician	€ 59.60 /contact	4	E
Podotherapist	€ 38.00 /contact	4	E
Audiologist	€ 53.40 /contact	4	E
Pedicurist	€ 15.00 /contact	4	E
Dentist	€ 18.40 /contact	4	E
Alternative professionals			
Reflexologist	€ 32.00 /contact	4	E
Magnetizer	€ 40.00 /contact	4	E
<b>Informal care<sup>X</sup>:</b> Various	€ 8.54 /hour	1	D
Paid house keeper	€ 8.54 /hour	1	E
<b>Resources / aids bought:</b>			
Consumable goods**	€ 9.85 /4 weeks	5	E
Durable goods***	€ 480.00 /year	5	E
Travelling expenses****	€ 0.17 /km	1	A & E
Patient (meals on wheels etc.)	various	6	E
Proxy (telephone, gasoline, etc.)	various	6	D

Source volumes: A Hospital Information System; B Electronic Patient File CMHT; C Registries of local pharmacies; D Informal Care Questionnaire; E Cost diaries.

\* DDD = Daily Defined Dose; \*\* Consumable goods = Incontinence material, bandages, adhesive dressings, cleansing tissues and plastic gloves; \*\*\* Durable goods = Mobility aids (wheelchair tri-

walker, walking frame, walking cane), hearing aids, bathroom aids (bath elevator, grab rail, shower seat), stair lift, wheelchair pillow, therapeutic foot gear, special undergarments, bed elevator, toilet elevator, glasses, special bed, stair gate, girdle, alarm, medication box, special clock, therapeutic stockings. \*\*\*\* Mean distance to GP or health care professional is 1.8km, mean distance to hospital is 7.0km (Oostenbrink); # The interventions and medication are highly varied: a full listing can be made available upon request. X: Sensitivity analyses were performed on admission to a nursing home (€ 106.00 - € 318.00), nursing home care (range € 20.79 - € 62.37) and informal care (range € 8.54 - € 32.67).

A): The Hospital Information System (ZIS) was used to determine, per patient, the number of days in the hospital, visits to the outpatient departments and all diagnostic and therapeutic interventions. Four patients were recruited from the CMHT in Sittard (Prins Claus Centre) and their hospital costs were collected from the hospital in Sittard.

B): The CMHT costs were collected from the Electronic Patient Files maintained by the RIAGG Maastricht and the Prins Claus Centre in Sittard.

C): With the help of the Society of Pharmacists in Maastricht (AVM), the pharmacists were asked to provide us with medication overviews of each patient during the time patients participated in the MEDICIE study. The cost diaries were used to calculate medication costs when no overview was available.

D): With respect to informal care activities, we used a survey that had been developed for the measurement and valuation of informal care (van den Berg et al., 2005a; van den Berg et al., 2005b; van den Berg et al., 2006). In the survey, informal caregivers were asked to indicate the average time spent on 16 different informal care tasks per week. This survey was adapted and extended to make it suitable for use in the MEDICIE study. It was used to determine the costs of informal care and out-of-pocket costs for the informal caregiver, home care, day care, institutionalization and housekeeping. The time devoted to informal care was valued against the wage rate of a paid housekeeper for all 16 tasks (Oostenbrink et al., 2004). Out-of-pocket costs made by the informal caregiver comprised mostly telephone costs and travelling expenses. With regard to home care, we distinguished between domestic help and physical help. Day care was specified into day care in a community centre, day care in an elderly home and day care in a nursing home. We furthermore distinguished admissions to an elderly home and to a nursing home. The informal care survey was administered by means of an interview with the primary caregiver of the patient at baseline and at 6 and 12 months follow-up.

E): Cost diaries completed by the proxy of the patient were used to determine the costs made outside the hospital that could not be gathered from hospital or pharmacist's registrations. In these diaries, we asked the proxies to document visits to the GP (we distinguished between visits to the GP, home visits and phone contacts), visits to health care professionals (i.e. physical therapist, occupational therapist, psychologist, psychiatrist, speech therapist, dietician, podotherapist, pedicurist, audiologist and dentist) and visits to alternative health care professionals (i.e. reflexologist and magnetizer). Furthermore, proxies were asked to list resources or aids that had been bought. These resources were categorized either as consumable goods (i.e. incontinence materials, bandages, plasters, plastic gloves and cleansing tissues) or durable goods (i.e. mobility aids, hearing aids, visual aids, bathroom aids). Lastly, out-of-

pocket costs for the patients were reported. These included, for the greater part, 'meals on wheels'. The proxies were asked to fill in these diaries prospectively on five occasions during the study (baseline, 3 months, 6 months, 9 months and 12 months) for four successive weeks.

### **Statistical analyses**

Incomplete data on the questionnaires (EQ-5D, cost diary and informal care survey) were imputed using a regression model. Complete missing data or data missing covariates were imputed using Rubin's multiple imputation (MI) procedure (Rubin, 1987). This method generates ten different data sets for imputed data. All analyses were performed with each of these ten data sets and these results were pooled. The complete missing data (on all questionnaires) of patients was imputed when a participant had completed the EQ-5D on baseline and at least one other occasion. Missing data of patients who had deceased, but completed every EQ-5D (once or twice according to their time of death) were also imputed. Patients who had completed the EQ-5D on less than two occasions or patients with a missing baseline measurement, were considered study drop-outs. With a logistic regression analysis, the probability of being a study drop-out was assessed and, with these probabilities, p-weights were calculated ( $1/(1-\text{predicted probability})$ ). This allowed for differential weighting of patients (according to their gender and severity of the dementia which were the variables significantly associated with the probability of being a study drop-out) during the analysis and description of the data (Little et al., 1987). P-weights were calculated based on the drop out-patients who had not passed away and were not institutionalized in a nursing home. It should be noted that the p-weight corrected data (with regard to the costs, effects and cost-effectiveness) turned out to be somewhat less favourable for DOC-PG than the uncorrected data. In order to maintain clarity and readability, the uncorrected data are not shown (but can be made available upon request).

SPSS version 12.0.1 was used to calculate the p-weights, to examine group differences and to impute the missing items by means of a regression model (missing value analysis). Rubin's multiple imputation procedure was performed using STATA version 8.2. Background characteristics of the participants (both the patients and their proxies) were summarized using descriptive statistics. Response distributions of the instruments are provided.

To represent the uncertainty in the costs and effects (STATA and MS-Excel), we undertook non-parametric bootstrapping on the incremental costs and effectiveness with 1000 replications (Briggs et al., 1997). The bootstrap simulations were performed with a p-weight correction. These p-weights were multiplied by 100 and each patient was added N times in the data file according to this p-weight.

The incremental costs and effects can be represented visually by using the incremental cost-effectiveness plane (Black, 1990). The horizontal axis divides the plane according to incremental effects while the vertical axis divides the plane according to incremental costs. This results in four quadrants: 1) southeast quadrant (SE): DOC-PG is less costly and more effective than usual care: DOC-PG is dominant 2) northwest quadrant (NW): DOC-PG is more costly and less effective than usual care: DOC-PG is inferior; 3) southwest quadrant (SW): DOC-PG is less costly and less effective than usual care: the

cost-effectiveness depends on the minimum amount of money society is willing to accept for a QALY loss and 4) northeast quadrant (NE): DOC-PG is more costly and more effective than usual care: the cost-effectiveness depends on the maximum amount of money society is willing to pay for a QALY gain. The choice of an intervention depends on what society is prepared to pay for a gain in effectiveness or willing to accept for a loss of effectiveness. This is termed the cost-effectiveness ceiling ratio. Decision makers are likely to be interested in the probability that a new intervention is cost-effective when compared to the existing alternative. In other words, the probability that a treatment is cost-effective varies according to the ceiling ratio. This probability can be shown in a cost-effectiveness acceptability curve (CEAC), which is based on the non-parametric bootstrapping (van Hout et al., 1994). One-way sensitivity analyses were performed to study the robustness of the cost-effectiveness results related to deterministic variables in the calculations, such as unit prices for home care and informal care. The cost prices of the resources with the largest cost differences between both groups were varied (-50%, +50%). Furthermore, the 'proxy good method' was applied as an alternative for valuing informal care, whereby housekeeping activities were valued against the wage rate of a housekeeper and personal care was valued against the wage rate of a nurse (van den Berg et al., 2006).

## Results

### Participants

Table 2 summarizes sample characteristics. Of the 230 patients, 64.3% were female. Of the proxies, 66.1% were female and most often a child (-in-law) or spouse of the patient (90.0%). In most cases, dementia (present in 70.9% of the patients) was associated with Alzheimer's disease (42.2%). Patients whose etiological diagnoses could not be determined were assigned to the "other" groups (i.e. other dementia or other cognitive impairment). Six patients (4.4%) in the intervention group and five patients (5.4%) in the control group withdrew from the study after the baseline measurement and these were considered study drop-outs.

The 219 patients that could be analyzed in the cost-effectiveness analysis (CEA) were comparable to the total study population. Of these 219 patients, 16 patients (11.7%) in the intervention group and 11 patients (11.8%) in the control group passed away during the course of the study.

### Effectiveness

Table 3 summarizes the percentage of patients reporting some or severe problems, their utility scores and QALYs of the EQ-5D with p-weight correction at the three measurements. After one year, most severe problems were encountered by the participants in both groups regarding self-care and usual activities. A larger proportion of patients in the control group suffered from severe problems on these domains. This was also reflected in the higher mean utility scores for the intervention group (0.43 versus 0.37) although these two domains have the lowest weight in the additive function (in case of severe problems) to deduce the utility score.

**Table 2:** Baseline demographic characteristics of the total study population and those who were analyzed in the cost-effectiveness analysis (CEA)

		Total		CEA	
		Intervention group (n = 137)	Control group (n = 93)	Intervention group (n = 131)	Control group (n = 88)
<b>Sex patient</b>	Female (%)	89 (64.9%)	59 (63.4%)	85 (64.9%)	58 (65.9%)
<b>Age patient</b>	Mean (SD)	78.3 (6.5)	77.3 (6.8)	78.3 (6.6)	77.5 (6.9)
	Range	[55 – 93]	[60 – 94]	[55 – 93]	[60 – 94]
<b>Relationship proxy:</b>					
	Spouse	51 (37.2%)	37 (39.7%)	46 (35.1%)	34 (38.6%)
	Child (in law)	73 (53.3%)	46 (49.6%)	72 (55.0%)	45 (51.1%)
	Other	13 (9.5%)	10 (10.7%)	13 (9.9%)	9 (10.2%)
<b>Sex proxy</b>	Female (%)	90 (65.7%)	62 (66.7%)	87 (66.4%)	58 (65.9%)
<b>Age proxy</b>	Mean (SD)	60.4 (13.5)	59.8 (13.9)	59.9 (13.4)	59.4 (14.6)
	Range	[30 – 84]	[34 – 91]	[30 – 84]	[34 – 91]
<b>Dementia:</b>		97 (70.8%)	66 (71.0%)	93 (71.0%)	64 (72.7%)
	Alzheimer's Disease (AD)	66 (48.2%)	31 (33.3%)	62 (47.3%)	31 (35.2%)
	Vascular Dementia (VD)	12 (8.8%)	14 (15.1%)	12 (9.2%)	13 (14.8%)
	Mixed Dementia	15 (10.9%)	6 (6.5%)	15 (11.4%)	5 (5.7%)
	Other Dementia	4 (2.9%)	15 (16.1%)	4 (3.1%)	15 (17.0%)
<b>No Dementia:</b>		40 (29.2%)	27 (29.0%)	38 (29.0%)	24 (27.3%)
	Cognitive Impairment / MCI	24 (17.5%)	15 (16.1%)	22 (16.8%)	12 (13.6%)
	Other Cognitive Impairment	16 (11.7%)	12 (12.9%)	16 (12.2%)	12 (13.6%)
<b>MMSE</b>		20.4 (5.6)	20.1 (6.0)	20.5 (5.5)	19.9 (6.1)

Furthermore, patients in the control group showed a greater decline in utility score after one year when compared to patients in the intervention group (0.09 versus 0.16). The average incremental effects (QALYs) were 0.05 (0.50 versus 0.45). The utility scores at the six and 12 months follow-up measurement, as well as the QALYs, differed significantly between the groups ( $p=0.00$ ).

### Costs

The use of health care resources is summarized in table 4. We calculated all p-weight corrected costs for 219 patients (table 4). It should be noted that the cost diaries were poorly completed (60%). We therefore had to impute GP costs, the costs of other professionals and patient's expenses in more cases than the other cost categories. The total mean costs over a period of one year amounted to € 38,396 in the DOC-PG group and € 38,331 in the control group. Admissions (to a nursing home and an elderly

**Table 3:** Percentage of patients reporting any problems, utility scores and QALYs of the EQ-5D with p-weight correction (for the utility scores and QALYs) at the three measurements. The utility scores and QALYs include the data of those patients who had deceased during the course of the study.

DOC-PG	Baseline (n=131)		6 months (n=120)		12 months (n=115)	
	Some	Severe	Some	Severe	Some	Severe
<b>Mobility:</b>						
Number of patients (%)	66.4%	0%	51.7%	6.7%	62.6%	1.7%
<b>Self-Care:</b>						
Number of patients (%)	38.2%	14.5%	35.3%	17.2%	34.8%	27.8%
<b>Usual Activities:</b>						
Number of patients (%)	50.8%	22.3%	44.0%	27.6%	40.9%	33.9%
<b>Pain/Discomfort:</b>						
Number of patients (%)	42.7%	9.9%	37.1 %	5.2%	38.9%	7.1%
<b>Anxiety/Depression:</b>						
Number of patients (%)	45.8%	11.5%	40.5%	2.6%	27.8%	10.4%
<b>Utility score:</b>						
Mean [CI]	(n=131) 0.52 [0.50 0.54]		(n=131) 0.53 [0.51 0.55]		(n=131) 0.43 [0.41 0.45]	
<b>QALYs Mean [CI]</b>						
	0.50 [0.49 0.52]					

Usual care	Baseline (n=88)		6 months (n=83)		12 months (n=77)	
	Some	Severe	Some	Severe	Some	Severe
<b>Mobility:</b>						
Number of patients (%)	72.7%	0%	65.1%	7.2%	63.6%	1.3%
<b>Self-Care:</b>						
Number of patients (%)	46.6%	11.4%	36.1%	27.7%	37.7%	33.8%
<b>Usual Activities:</b>						
Number of patients (%)	45.5%	25.0%	43.4%	33.7%	28.6%	50.6%
<b>Pain/Discomfort:</b>						
Number of patients (%)	44.3%	6.8%	47.0%	6.0%	41.6%	6.5%
<b>Anxiety/Depression:</b>						
Number of patients (%)	48.9%	6.8%	43.4%	9.6%	41.6%	10.4%
<b>Utility score:</b>						
Mean [CI]	(n=88) 0.53 [0.51 0.55]		(n=88) 0.46 [0.44 0.49]		(n=88) 0.37 [0.35 0.40]	
<b>QALYs Mean [CI]</b>						
	0.45 [0.43 0.47]					

**Table 4:** Number of patients making use of health care resources and mean usage (i.e. number of visits, days in nursing home etc.) during 12 months follow-up, mean costs per patient, mean effects (QALYs) and the incremental cost-effectiveness ratios (ICER) with p-weight correction.

	No of patients		Mean visits per patient [CI]		Mean costs per patient [CI]		Incremental costs/effects
	DOC-PG	USC-PG	DOC-PG	USC-PG	DOC-PG	USC-PG	
<b>Intervention costs<sup>A, B</sup></b>	131 (100%)	/	/	/	€ 857 [ 854 860]	€ 492 [ 473 511]	€ 365
<b>Hospital costs<sup>A</sup></b>	100 (76%)	/	/	/	€ 3457 [3125 3787]	€ 2033 [1750 2316]	€ 1424
<b>CMHT costs<sup>B</sup></b>	93 (71%)	/	8.7 [ 8.0 9.4]	7.8 [ 7.1 8.5]	€ 1112 [1018 1204]	€ 1001 [ 917 1084]	€ 111
<b>Medication<sup>C, E</sup></b>	127 (97%)	/	/	/	€ 1457 [1375 1539]	€ 1586 [1483 1688]	-€ 129
<b>Admission<sup>D</sup></b>							
Nursing home	18 (14%)	15 (17%)	23.8 [ 19.9 27.6]	34.8 [ 29.1 40.4]	€ 5039 [4224 5853]	€ 7370 [6177 8563]	-€ 2331
Elderly home	20 (15%)	14 (16%)	41.4 [ 35.3 47.5]	35.3 [ 28.6 41.9]	€ 3620 [3086 4153]	€ 3084 [2504 3665]	€ 536
<b>Home care<sup>D</sup></b>							
Domestic	71 (54%)	51 (58%)	90.8 [ 84.2 97.4]	91.5 [ 83.4 99.6]	€ 2027 [1879 2174]	€ 2043 [1862 2224]	-€ 16
Nursing	66 (50%)	46 (52%)	106.2 [ 94.4 118.0]	88.5 [ 76.3 100.7]	€ 4415 [3926 4904]	€ 3677 [3170 4185]	€ 738
<b>Day care in<sup>D</sup></b>							
Nursing home	18 (14%)	16 (18%)	10.3 [ 8.5 12.0]	13.0 [ 10.6 15.4]	€ 1277 [1056 1498]	€ 1616 [1318 1913]	-€ 339
Elderly home	34 (26%)	30 (34%)	31.2 [ 27.4 35.0]	24.7 [ 21.6 27.8]	€ 1557 [1366 1747]	€ 1232 [1075 1388]	€ 325
Community centre	2 (2%)	8 (9%)	0.4 [ 0.2 0.6]	4.7 [ 3.5 5.9]	€ 2 [ 1 3]	€ 24 [ 18 30]	-€ 22
<b>General Practitioner<sup>E</sup></b>	117 (89%)	78 (89%)	22.6 [ 21.3 23.9]	25.9 [ 24.3 27.6]	€ 491 [ 462 520]	€ 543 [ 505 581]	-€ 52
<b>Other professionals<sup>E</sup></b>	88 (67%)	62 (70%)	51.6 [ 48.4 54.8]	53.5 [ 49.5 57.4]	€ 1447 [1361 1533]	€ 1470 [1364 1576]	-€ 23
<b>Informal care<sup>D</sup></b>	131 (100%)	88 (100%)	807.2 [739.2 875.1]	888.0 [792.2 983.7]	€ 6311 [5793 6829]	€ 6934 [6207 7661]	-€ 623
Housekeeper <sup>D</sup>	83 (63%)	66 (75%)	83.4 [ 77.9 88.9]	92.3 [ 85.5 99.0]	€ 631 [ 588 673]	€ 696 [ 642 749]	-€ 65
Durable goods <sup>E</sup>	60 (46%)	40 (46%)	/	/	€ 453 [ 414 493]	€ 531 [ 476 585]	-€ 78
Consumable goods <sup>E</sup>	86 (66%)	65 (74%)	/	/	€ 145 [ 137 153]	€ 180 [ 169 191]	-€ 35
Out-of-pocket patient <sup>E</sup>	89 (68%)	66 (75%)	/	/	€ 1042 [ 988 1095]	€ 1170 [1098 1241]	-€ 128
Out-of-pocket caregiver <sup>D</sup>	119 (91%)	77 (88%)	/	/	€ 2399 [2324 2475]	€ 2027 [1932 2123]	€ 372
Travelling costs GP & profs <sup>E</sup>	/	/	/	/	€ 35 [ 32 37]	€ 37 [ 34 39]	-€ 2
*Travelling costs hospital <sup>A</sup>	/	/	/	/	€ 9 [ 8 10]	€ 10 [ 9 11]	-€ 1
Travelling costs day care <sup>D</sup>	/	/	/	/	€ 598 [ 538 658]	€ 558 [ 501 614]	€ 40
*Parking costs <sup>A</sup>	/	/	/	/	€ 9 [ 8 10]	€ 11 [ 10 12]	-€ 2
<b>Total costs<sup>A, B, C, D, E</sup></b>	/	/	/	/	€ 38,396 [37053 39738]	€ 38,331 [36428 40234]	€ 65
<b>Effects (QALY)</b>					<b>0.503 [0.487 0.519]</b>	<b>0.452 [0.432 0.472]</b>	<b>0.05 QALY</b>
<b>ICER</b>							<b>€ 1267 / QALY</b>

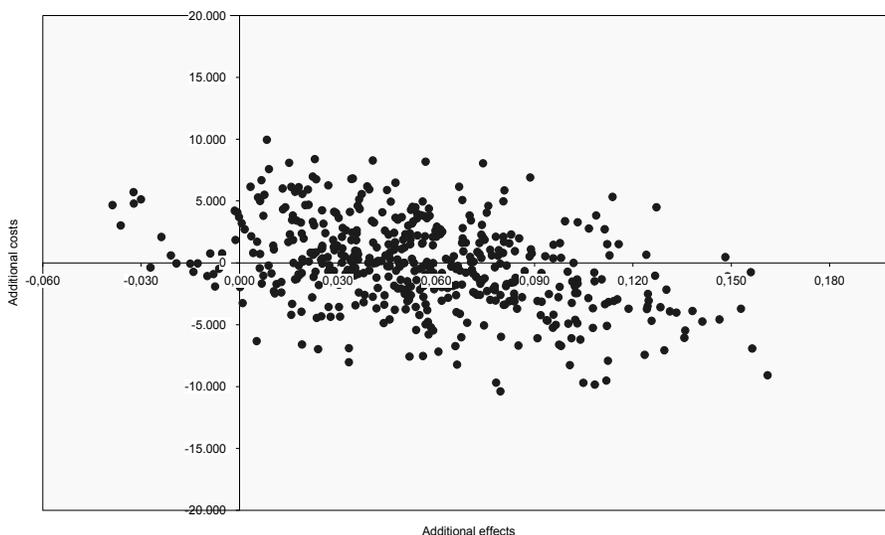
A Hospital Information System; B Electronic Patient File CMHT; C Registries of local pharmacies; D Informal Care Questionnaire; E Cost diaries.

home) and informal care constituted the largest portions of the costs (25% and 17% respectively) in both groups. Additionally, 16% of the costs were spent on contacts with various health care professionals (i.e. intervention costs, hospital costs, CMHT costs, GP costs and the costs of other professionals) and on home care, 9% on out-of-pocket costs, 7% on day care, 4% on medication and the remainder on durables, consumables and travelling expenses. The largest cost differences between both groups were most evident for admissions to a nursing home, hospital costs, home care (nursing home), and informal care. Of these, the costs of admissions and informal care were lower for the intervention group. The average incremental costs were € 65.

### Cost-effectiveness

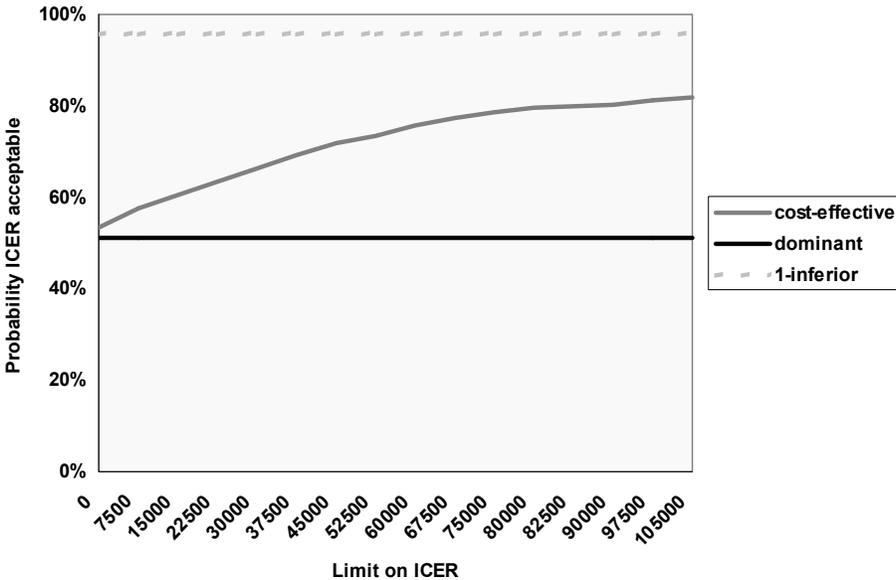
The mean Incremental Cost-Effectiveness Ratio (ICER) in the bootstrap simulation was € 1267 / QALY. The incremental costs in the bootstrap simulation ranged from € -7,435 (2.5 percentile) to € 6,750 (97.5 percentile). The incremental effectiveness ranged from -0.01 (2.5 percentile) to 0.13 (97.5 percentile). The scatter plot in figure 1 provides a visual representation of the results of the bootstrap simulation. The majority of the incremental cost-effectiveness pairs are situated to the right of the vertical axis (y-axis), meaning that 94% of the pairs indicate that the DOC-PG is more effective than usual care. Most of these incremental cost-effectiveness pairs (51%) are situated in the quadrant indicating dominance for DOC-PG (less costs and more effectiveness for DOC-PG), whereas 43% is situated in the northeast quadrant (more costs and more effectiveness for DOC-PG). The cost-effectiveness acceptability curve (CEAC) is shown in figure 2. When the ceiling ratio is € 45,000 (corresponding with the threshold put forth

**Figure 1:** Scatter plot of the estimated incremental costs and incremental effects of DOC-PG versus usual care obtained by bootstrap simulations.



by the NICE guidelines: + £30000 (Raftery, 2001)), the probability that the DOC-PG is cost-effective is 72%. When the ceiling ratio is € 80,000 (the threshold determined by the Dutch Council for Public Health and Health Care (RVZ, 2006)), this probability increases to 80%. Applying a more conservative ceiling ratio of € 20,000 results in a 63% probability that DOC-PG is cost-effective.

**Figure 2:** Cost-effectiveness acceptability curves showing the probability that DOC-PG is cost-effective when compared to usual care over a range of values for the maximum acceptable ceiling ratio (limit on ICER)



The results of the one-way sensitivity analyses are shown in table 5. Since the unit prices of hospital interventions varied considerably due to the numerous types of interventions, these unit prices were excluded from the sensitivity analyses. Results of the sensitivity analyses showed that cost-effectiveness results were robust and even demonstrated dominance of DOC-PG in three of the five cases, except when the costs for admission to a nursing home were half the genuine price and when the costs for home care in a nursing home were double the genuine price.

**Table 5:** Results of the p-weighted sensitivity analyses, with and without regression correction

		Mean total costs per patient			
	Cost price	DOC-PG (n=131)	Usual care (n=88)	ICER	
<b>Admission:</b>					
Nursing	-50%	€ 106.00	€ 2,847	€ 3,950	€ 25,420 / QALY
	+50%	€ 318.00	€ 8,443	€ 11,851	dominance
<b>Home care:</b>					
Nursing	-50%	€ 20.79	€ 2,448	€ 2,016	dominance
	+50%	€ 62.37	€ 7,346	€ 6,048	€ 11,698 / QALY
<b>Informal care:</b>					
Proxy good method*:					
HDL		€ 8.54	€ 3,366	€ 3,181	
ADL		€ 32.67	€ 2,821	€ 4,083	
IADL		€ 32.67	€ 16,858	€ 20,023	
Total			€ 23,045	€ 27,287	dominance

\* Method used by B. van den Berg et al. (2006).

## Discussion

This article reports on the first economic evaluation of an integrated multidisciplinary facility for diagnosing dementia conducted alongside a randomised trial. The results show that, when compared to usual care, patients who visited the DOC-PG, gained a mean of 0.05 QALYs. The incremental costs per QALY amounted to € 1267. The probability that DOC-PG is cost-effective ranges from 63% when applying a conservative ceiling ratio of € 20,000 to 80% with a ceiling ratio of € 80,000. The NICE guidelines (Raftery, 2001) state that the reimbursement of interventions costing less than £30.000 (approximately € 45.000) are generally never questioned. This threshold has recently been set at € 80.000 by the Dutch Council for Public Health and Health Care (RVZ, 2006) for diseases with a high burden such as Alzheimer's Disease. Clearly, our point estimate lies beneath this threshold and is surrounded by an acceptable amount of uncertainty. Therefore, we consider an integrated approach to dementia as provided by the DOC-PG to be cost-effective. Good economic and social arguments exist for increasing expenditures on dementia services when improved or preserved health status or quality of life are substantial or when deterioration is delayed (Kavanagh et al., 2002; Knapp et al., 1998a; Knapp et al., 1998b).

The significant difference in effectiveness (i.e. HRQoL) between both groups in favour of DOC-PG, mainly pertained to self care and usual activities. However, these domains (in case of severe problems) have the smallest impact on the utility score. Although the results showed a gain in QALYs for the intervention group when compared with the control group, the mean HRQoL (measured by the utility score of the EQ-5D) for both groups deteriorated over the course of one year. Cognitive functioning, neuropsychiatric problems, ability to perform activities of daily living and emotional functioning also

worsened after a year (Wolfs et al., 2007e), which is not surprising given that the majority of patients suffered from dementia. However, the deterioration, especially with regard to HRQoL, was more substantial for the control group.

In our study, the mean total costs per year per patient amounted to € 38,396 for the intervention group and € 38,331 for the control group. In the intervention group, more costs were spent on the hospital costs. A more complete diagnostic assessment will likely reveal previously undiagnosed somatic or psychiatric disorders. Alternatively, it may make the unmet needs of patients and (or) caregivers more apparent, which then would lead to more interventions and thus to more costs. It is very important that these coexistent medical problems in patients with dementia are carefully managed as treatment of these problems can increase a patient's wellbeing (NICE, 2006).

The largest difference in costs between DOC-PG and usual care, in favour of DOC-PG, relates to admissions to a nursing home. Patients in the control group resided in a nursing home for a longer period of time than patients in the intervention group ( $p=0.01$ ), which obviously results in a large cost difference. It remains unclear whether this difference in costs diminishes in the course of time. Due to the lack of evidence, no extrapolation beyond our follow-up period could be performed to explore this. Another notable difference in costs between the intervention group and the control group is related to informal care. Informal care costs were considerably higher in the latter group. This is possibly due to the fact that the DOC-PG team advised the (more extensive) involvement of professional care such as home care (Wolfs et al., 2007b). These services, which are more expensive than informal care, were indeed utilised more extensively in the intervention group. An integrated approach may therefore impact the process (i.e. accessibility and swiftness) of service delivery. This possibility has also been considered by Brown et al. (Brown et al., 2003) who examined the impact of integrated health and social care teams in (healthy) older adults. It should be noted that informal care is rarely included in economic evaluation studies, since the collection of these data is often complex (Evers et al., 2000; Evers et al., 1997). The measurement and valuation of informal care is still a matter of debate and the costs are highly dependent on the valuation method used. In this study, we valued the time caregivers spent on informal care using the hourly wage for housekeeping services (Oostenbrink et al., 2004). This rather conservative approach probably resulted in an underestimation of the costs of informal care. Given that more time was spent on informal care in the control group, it is likely that the cost-difference between the groups would become smaller when other methods for valuing informal care are applied (van den Berg et al., 2006). This would, in turn, result in a more attractive ICER for DOC-PG. In fact, applying a different valuation method (i.e. the proxy good method) in the sensitivity analysis resulted in dominance for DOC-PG.

The costs of dementia care in this study (+ € 38,500 per year) are comparable to those in other studies using similar cost categories (31,000 – 46,700 US dollars per year) (Wimo et al., 1997). To date, no studies have estimated the costs and/or cost-effectiveness of an integrated approach to diagnosing dementia nor have they estimated the costs of other diagnostic services (NICE, 2006; Wolfs et al., 2006). Few studies exist on the costs and cost-effectiveness of psychosocial therapies and scanning and imaging in diagnosing dementia (Knapp et al., 2006; McMahon et al., 2000; McMahon et al.,

2003; Wimo et al., 1997). There have been a fair amount of pharmaco-economic studies that focus on dementia (Fillit et al., 2005; Jonsson, 2003). However, cost-effectiveness studies often include different costs and different effects (i.e. MMSE or QoL-AD instead of QALYs). Furthermore, there are also international differences in the organization of health care. Therefore, it is difficult to draw comparisons between the different studies that focus on the cost-effectiveness of dementia care.

It should be noted that usual care in our region consisted mainly of an active university MC and a CMHT that have collaborated in the past on several projects. Thus, the contrast between DOC-PG and usual care may have been lower as may have been the case in other regions.

Some limitations of this study need to be considered. A first weakness of this study relates to the poor completion of the cost diaries which were used to calculate part of the cost volumes. These diaries appeared to be very time-consuming and difficult to complete. Retrospective questionnaires would probably have resulted in a higher completion rate. Indeed, it may be more beneficial to use a cost questionnaire with structured closed questions for the assessment of health care utilization in economic evaluations alongside clinical trials (van den Brink et al., 2005). As there were no other possibilities, other than the diary, for determining each patient's number of visits to the GP, visits to other health care professionals and number of medical aids purchased elsewhere, these costs were imputed in a larger sample than the rest of the costs. However, these costs constitute only a small proportion of the total costs (7%).

An additional potential limitation was our use of proxies to complete the questionnaires. The method of proxy rating was chosen because of the longitudinal nature of the study, the complex health problems of the study population, and the progressive global deterioration of intellect and personality characteristics in members of the study population. It is generally acknowledged that, in the later stages of dementia, proxy measures are required since patients are often incapable of adequately evaluating their own health (Jonsson et al., 2006a; Selai, 1998). Nonetheless, it is possible that the scores on the instruments were biased because of a perceived caregiver burden (Logsdon et al., 2002). This potential bias, however, applies to both groups.

A final potential limitation concerns our follow-up period of one year. It can be argued that a longer duration of follow-up is important when examining treatment benefits for dementia, a disorder that results in a gradual progressive decline over several years (NICE, 2006). Due to the lack of evidence, no extrapolation beyond our follow-up period could be performed.

In conclusion, this economic evaluation, conducted from a societal perspective using a one year time horizon, shows that an integrated approach to dementia by means of DOC-PG is not demonstrable more expensive, despite of the investment in the diagnostic intervention, but that it has a high probability of being more effective. Therefore, we contend that DOC-PG is indeed a cost-effective facility for the diagnosis and management of dementia. In light of our ageing population with increasingly more patients suffering from dementia, appropriate multidisciplinary organizational models are necessary to ensure the quality of care while controlling the costs.

*Acknowledgement*

We thank the patients, their caregivers and the GPs who participated in this study. Additionally, we thank Daniëlle Willems for collecting and inputting a great deal of the data, and Audrey Fiddeler, Heidi Lansdaal, Ilvy Mayen and Andrea Wolf for collecting and inputting the remainder of the (cost)data. We also thank everyone working at the DOC-PG, the MMC and the CMHT for their help with respect to recruiting patients.

Contributors: FV was the principal investigator and is guarantor. FV, CD and AK designed the trial. CW and JS helped to refine the study design and carried out the outcome measures. CW drafted the manuscript, and all authors critically revised it for scientific content and approved the final version.

Funding: The study was funded by the Dutch Research Institute for Care and Medical Sciences (ZorgOnderzoek Nederland-Medische wetenschappen, ZON-Mw), grant nr 945-02-055

Ethics committee approval: The local ethics committee approved the research.

# Chapter 8

Informal caregiving in dementia: results of a survey



Claire AG Wolfs, Carmen D Dirksen, Alfons Kessels, Frans RJ Verhey and Johan L Severens

## Abstract

**Declaration of interest:** None

**Background:** Informal care plays a substantial role in the provision of total care. The number is expected to increase in the coming decades due to demographic and social-economic developments. More attention to informal care, the position of informal caregivers, the consequences of providing informal care and the inclusion of informal care in economic evaluations is therefore essential.

**Aim:** The aim of the current study is to evaluate possible effects of an integrated multidisciplinary approach with regard to dementia on caregiver outcomes.

**Method:** The study is part of a randomized controlled trial into the effects of an integrated approach to dementia by means of a Diagnostic Observation Centre for PsychoGeriatric patients (DOC-PG). Several caregiver outcomes (patient characteristics, caregiver characteristics, social context and amount and costs of informal care) were compared between the intervention group (DOC-PG) and the control group.

**Results:** We found that significantly more patients in the intervention group made use of informal day care (i.e. day care provided by friends or family) at the baseline measurement. Another difference related to the costs of informal care, measured and valued with the proxy-good method, which were significantly higher in the control group at the 12 months follow-up. Finally, significantly more caregivers in the intervention group gave up unpaid work in order to care for the patients at the baseline measurement. No other differences between the groups were found.

**Conclusion:** Overall, few significant differences were found between the groups. However, our results suggest that DOC-PG may have a positive effect on the amount of informal care since this amount increased more in the control group than in the intervention group after one year.

## Introduction

Informal care plays a substantial role in health care. In the Netherlands, it was estimated in 2001 that 3.75 million people provided informal care. This number is expected to increase even further due to demographic and socio-economic developments (Boer de et al., 2003). As the population is ageing, more people will suffer from dementia in the coming decades. Furthermore, the number of single elderly is increasing and there is an increasing shortage of personnel in health care. These developments cause tension between the availability of, and demand for informal care. Due to the nature of dementia, where the majority of service provision falls on care rather than treatment, there is a large demand on informal care (NICE, 2006). Informal caregivers may encounter positive aspects in the process of caring for a loved one, but more often they experience a number of negative outcomes such as emotional strain, financial losses, feelings of isolation and health declines (Schulz et al., 1999; van den Berg et al., 2004). However, caregiver's burden is a multi-component construct, influenced by characteristics of the patient (i.e. severity of the disease, behavioural problems, independent functioning and depressive symptoms), characteristics of the caregiver (i.e. perceived stress, psychological wellbeing, lack of knowledge, immature coping), and the social context (i.e. finances, social support, family help, number and nature of the caregiving tasks). This construct is complex and to date only partially understood (Verhey et al., 2007b). The current study is part of a randomized controlled trial into the effects of an integrated approach to dementia (Wolfs et al., 2007e). The aim of this study is to evaluate possible effects of such an integrated multidisciplinary approach on caregiver outcomes.

## Methods

The proposed study is part of a prospective controlled randomized trial into the effects of the multidisciplinary psychogeriatric diagnostic observation centre (DOC-PG) on somatic and mental health, and the costs and use of health care facilities of patients with psychogeriatric problems (dementia, somatic problems with comorbid psychiatric disorders). The results of the MEDICIE-study (Maastricht Evaluation of a Diagnostic Intervention for Cognitively Impaired Elderly) are described elsewhere in more detail (Wolfs et al., 2007c; Wolfs et al., 2007d; Wolfs et al., 2007e). In the current study, the intervention group (DOC-PG) and the control group (usual care) were compared regarding the results of an informal care questionnaire, which was developed by the iMTA in Rotterdam, to measure and value informal care. This questionnaire was adapted and extended to make it suitable for use in the MEDICIE study. It was administered by means of an interview with the primary caregiver of the patient at baseline and at 6 and 12 months follow-up. The results are categorized in patient characteristics, characteristics of the caregiver and social context. Please note that the method section in this chapter is similar to the method section in chapter 5.

### **Study population**

The MEDICIE-study was approved by the medical ethical committee at Maastricht University Hospital. Patients were recruited from July 2002 to August 2004 from 70 GP practices in the Maastricht region and surroundings. GPs were asked to refer all patients that may have dementia or a cognitive disorder.

Patients were included who 1) were older than 55 years; 2) were suspected by the GP of having dementia, or were suspected of having both a somatic and a psychological disorder; 3) had not been referred in the two years prior to this study to one of the participating services; 4) had the availability of a proxy, i.e. a person visiting the patient at least once a week).

Patients were excluded when 1) they suffered from acute disorders that needed a prompt therapeutic intervention; or 2) they lived in a nursing home, or received care that is comparable to that of a nursing home (e.g. substitution of care in an elderly home);

### **Interventions**

#### *DOC-PG*

The function of the DOC-PG was to provide multidisciplinary assessment and advice through somatic screening, psychogeriatric assessment, and evaluation of the required levels of care for the patient and his/her carer. Participating disciplines in the DOC-PG included old age psychiatry, geriatric medicine, neuropsychology, physiotherapy, occupational therapy, geriatric nursing and mental health nursing. The diagnostic screening conducted by the DOC-PG took place over the course of two weeks, and consisted of a home visit and two visits to the University Hospital departments of geriatric medicine and geriatric psychiatry. In addition, a CT scan and various blood tests were performed. The results were then discussed at a weekly interdisciplinary meeting in which a definitive diagnosis was made and a treatment plan was formulated. Correspondence was sent to the referring GP in which a summary of the assessments, the multi-axis diagnosis and recommendations for treatment and management were described. After the assessment, the GP was responsible for the patients. However, the recommendations provided to the GP often indicated that the involvement of one or more of the disciplines represented in the DOC-PG was warranted in further treatment.

#### *Usual care*

In the control group, the GP provided usual care. This means that the GP could not refer the patient to the DOC-PG and that the diagnosis was either made by the GP or the GP referred the patient to one of the existing separate regional services, such as the Maastricht Memory Clinic (MMC) (Verhey et al., 2007c), geriatric medicine or the department of mental health for the elderly at mental health community services (CMHT).

### **Patient characteristics:**

Patient characteristics described in this study are gender, age, diagnosis, severity of the dementia, independent functioning and behavioural problems (which have been collected in the MEDICIE-study), number of (self-reported) comorbidities, health,

physical function and health related quality of life (HRQoL). The severity of the dementia was assessed by means of the Mini Mental State Examination (MMSE) and the Global Deterioration Scale (GDS). The MMSE (Folstein et al., 1975b; Kok et al., 2002) assessed the severity of cognitive decline. The GDS (Reisberg et al., 1982) evaluated patients on seven stages of global functioning for those suffering from a primary degenerative dementia such as Alzheimer's disease. The Instrumental Activities of Daily Living Scale (IADL) (Lawton et al., 1969) measured seven areas of more complex activities required for optimal independent functioning. The scoring for this scale indicated whether the patient was completely independent, in need of assistance or was completely dependent on others for the performance of specific activities. Finally, the NeuroPsychiatric Inventory (NPI) (Cummings, 1997; Cummings et al., 2001; Cummings et al., 1994) appraised patients' behavioural and psychological problems. Higher scores on all instruments, except on the MMSE, are indicative of more severe problems. Health and physical function were assessed by means of the Visual Analogue Scale (VAS) of the EQ-5D (Health Related Quality of Life) ("EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group", 1990) and physical functioning aspect of the SF36 (Brazier et al., 1992; VanderZee et al., 1996; Ware et al., 1992) (also measuring Health Related Quality of Life) respectively.

### **Caregiver characteristics**

Caregiver characteristics described in this study are gender, age, relationship with the patient, education, income, self-rated health, health-related quality of life (HRQoL), and burden.

Self-rated health and HRQoL were assessed by means of the EQ-5D (Health Related Quality of Life) ("EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group", 1990).

Regarding burden, a distinction is made between objective burden and subjective burden. Objective burden concerns the number of caregiving tasks, and whether the caregiver shares a house with the care recipient. The subjective burden of informal care indicates how the caregiver encounters the impact of objective burden (Brouwer et al., 2004). This subjective burden may change over time as a result of changes in caregiving situation, the deterioration of the patient's health or coping strategies. We used the self-rated burden scale (SRB) to measure subjective burden (Brouwer et al., 2004; van Exel et al., 2005; van Exel et al., 2004). The SRB consists of a single question: "How burdensome is the provision of informal care to you at this moment?" and caregivers are asked to indicate their burden on a horizontal visual analogue scale ranging from 0 ("not at all straining") to 10 ("much too straining"). Subjective burden was also assessed using the Perceived Stress caused by Informal Caregiving scale (Dutch translation: *Ervaren Druk door Informele Zorg (EDIZ)*) (Pot et al., 1995). The EDIZ is a 9-item measurement to assess self perceived pressure from informal care.

### **Social context**

The social context of both the patient and the caregiver are the caregiver network (i.e. the availability and number of other caregivers and the number of hours spent on informal care by these caregivers) and formal care (home care, day care and admission).

Furthermore, other daily activities of the patients (that support the caregiver), respite care and paid leave from work were considered.

### **Measurement and valuation of informal care**

In the informal care questionnaire applied in the MEDICIE study, three different methods (two monetary methods (1 and 2) and one non-monetary method (3)) for the measurement and valuation of informal care were described. Costs of informal care were expressed in 2005 Euros.

1) The proxy-good-method (van den Berg et al., 2006): This method (also called the market cost method or replacement cost method) values the measured time spent on providing informal care at the (labour) market price of a close substitute. This value can differ per task. Caregivers were asked to indicate the average time spent on 18 different informal care tasks per week. This time can be valued using different methods, resulting in a monetary outcome (Oostenbrink et al., 2004; van den Berg et al., 2004; van den Berg et al., 2006). We subdivided the 18 tasks into household activities of daily living (HDL), activities of daily living (ADL) and instrumental activities of daily living (IADL) according to the classification of van den Berg et al (van den Berg et al., 2006). Household (HDL) activities were valued against the wage rate of a housekeeper (€ 8,54). Personal care (ADL and IADL) was valued against the wage rate of a nurse (€ 32,67) (Oostenbrink et al., 2004; van den Berg et al., 2006). Measuring time spent on informal care is the essence of the proxy-good-method.

2) The opportunity-cost-method (van den Berg et al., 2006): In this method, we valued the lost time as the opportunity cost (i.e. hours of paid work, unpaid work or leisure time forgone), or the value of the time in its current best alternative use (hours of paid work, unpaid work or leisure time). We asked the caregivers:

a) what types and amount of time they actually gave up in order to provide informal care (i.e. number of hours of paid work, number of hours of unpaid work and/or number of hours of leisure time).

b) which activities (paid work, unpaid work or leisure time) they would rather perform, instead of giving informal care and to indicate the preferred amount of these activities.

Hours of paid work were valued against the hourly wage of € 49,21 for men and € 37,47 for women, based on a mean age of caregivers in this sample of 59 years. For unpaid work and leisure time, the hourly wage of a housekeeper was used as shadow price (€ 8,54) which was again in accordance with van den Berg et al (Oostenbrink et al., 2004; van den Berg et al., 2006) .

3) The CarerQol-method (Brouwer et al., 2006): This method provides us with a non-monetary outcome regarding the measurement and valuation of informal care. The CarerQol is aimed at measuring the care-related quality of life of informal caregivers (Brouwer et al., 2006). It combines seven important burden dimensions with a valuation component (a visual analogue scale (VAS) for happiness. The seven burden dimensions are 1) fulfilment; 2) relational problems; 3) mental problems; 4) problems with daily activities; 5) financial problems; 6) support; and 7) physical problems. The CarerQol-VAS

ranges from 0 (“completely unhappy”) to 10 (“completely happy”). The CarerQoI has been validated in a Dutch sample of heterogeneous caregivers.

### **Statistical analyses**

In this study, data of the informal care surveys were analyzed for those patients who completed the survey at all three measurements (i.e. at the baseline measurement, at the 6 months follow-up and at the 12 months follow-up) and who had the same primary caregiver during the course of the study. Missing items within the survey were imputed by means of a regression model (missing value analysis in SPSS version 12.0.1).

Background characteristics of the participants (both the patients and their caregivers) are summarized using descriptive statistics. Response distributions of the informal care survey are provided. Independent sample t-test were used to study differences between the DOG-PG group and the control group. Significant differences ( $p < 0.05$ ) are indicated by means of \* in the tables.

## **Results**

In the MEDICIE study, 230 patients were included. In the course of the study, 27 patients (11.7%) died (16 in the intervention group and 11 the control group ) and 13 patients (5.7%) dropped out (8 in the intervention group and 5 in the control group).

Of the study completers, 170 patients, 100 (73.0%) in the intervention group and 70 (75.3%) in the control group completed all three informal care questionnaires and were included in the analyses.

Table 1 summarizes the patient characteristics at the three measurements. Patients in the intervention group and the control group were highly comparable with regard to these characteristics. The clinical characteristics of the patients showed a progressive deterioration over time.

Table 2 summarizes the caregiver characteristics. Caregivers in the intervention group and the control group were also highly comparable. The greater part of the caregivers were female and most often a child (-in-law) or spouse of the patient. Objective and subjective burden remained stable over time in both groups.

Table 3 describes the social context of the patients and their caregivers. We found that significantly more patients in the intervention group made use of informal day care (i.e. day care provided by friends or family) at the baseline measurement. No other differences between the groups were found regarding social context. Over the course of one year, more people made use of formal care in both groups. Hardly any of the caregivers used services in order to support them (i.e. paid leave from work or respite care).

**Table 1:** Patient characteristics (intervention group and control group) at the three measurements.

	Intervention group (n=100)			Control group (n=70)		
	Baseline	6 months	12 months	Baseline	6 months	12 months
<b>Sex patient</b> Female (%)	68.0%	n.a.	n.a.	68.6%	n.a.	n.a.
<b>Age patient</b> Mean (SD)	77.8 (6.4)	n.a.	n.a.	76.6 (7.0)	n.a.	n.a.
Range	55 - 91	n.a.	n.a.	60 - 94	n.a.	n.a.
<b>Dementia:</b>	<b>69 (69.0%)</b>	n.a.	n.a.	<b>47 (67.1%)</b>	n.a.	n.a.
Alzheimer's Disease (AD)	46 (46.0%)	n.a.	n.a.	22 (31.4%)	n.a.	n.a.
Vascular Dementia (VD)	7 ( 7.0%)	n.a.	n.a.	10 (14.3%)	n.a.	n.a.
Mixed Dementia	13 (13.0%)	n.a.	n.a.	4 ( 5.7%)	n.a.	n.a.
Other Dementia	3 ( 3.0%)	n.a.	n.a.	11 (15.7%)	n.a.	n.a.
<b>No Dementia:</b>	<b>31 (31.0%)</b>	n.a.	n.a.	<b>23 (32.9%)</b>	n.a.	n.a.
Cognitive Impairment / MCI	15 (15.0%)	n.a.	n.a.	11 (15.7%)	n.a.	n.a.
Other Cognitive Impairment	16 (16.0%)	n.a.	n.a.	12 (17.2%)	n.a.	n.a.
<b>MMSE<sup>1</sup></b>	20.3 (5.5)	18.9 (7.5)	18.3 (6.9)	19.9 (6.1)	19.5 (6.8)	17.8 (8.4)
<b>GDS<sup>2</sup></b>	4.2 (1.1)	4.5 (1.2)	4.7 (1.1)	4.1 (1.2)	4.5 (1.3)	4.7 (1.3)
<b>IADL<sup>3</sup></b>	17.1 (5.9)	19.0 (6.3)	20.4 (6.3)	16.1 (6.4)	18.0 (6.8)	19.9 (6.9)
<b>NPI<sup>4</sup></b>	22.5 (14.5)	24.2 (16.9)	28.4 (20.5)	21.6 (15.4)	25.7 (19.8)	27.7 (20.3)
<b>Number of comorbidities</b>	2.9 (19.7)	2.9 (1.8)	3.1 (1.7)	3.1 (2.0)	3.2 (1.9)	3.1 (1.8)
<b>Health<sup>5</sup> Mean (SD)</b>	59.5 (19.7)	60.0 (17.6)	59.2 (19.2)	60.1 (18.5)	56.0 (17.8)	54.5 (20.7)
<b>Physical function<sup>6</sup> Mean (SD)</b>	49.8 (28.4)	48.3 (28.3)	41.3 (29.0)	52.6 (29.6)	50.2 (30.8)	43.9 (30.7)

<sup>1</sup> Mini Mental State Examination; <sup>2</sup> Global Deterioration Scale; <sup>3</sup> Instrumental Activities of Daily Living Scale; <sup>4</sup> NeuroPsychiatric Inventory; <sup>5</sup> Visual Analogue Scale EQ-5D; <sup>6</sup> Physical functioning part of the SF36; n.a.: characteristics which do not change over time

**Table 2:** Caregiver characteristics (intervention group and control group) at the three measurements

	Intervention group (n=100)			Control group (n=70)		
	Baseline	6 months	12 months	Baseline	6 months	12 months
<b>Sex caregiver</b> Female (%)	66.0%	n.a.	n.a.	68.6%	n.a.	n.a.
<b>Age caregiver</b> Mean (SD)	59.3 (13.5)	n.a.	n.a.	58.2 (14.7)	n.a.	n.a.
Range	30 – 83	n.a.	n.a.	34 - 87	n.a.	n.a.
<b>Relationship patient:</b>						
Spouse	31.0%	n.a.	n.a.	34.3%	n.a.	n.a.
Child (in law)	61.0%	n.a.	n.a.	55.8%	n.a.	n.a.
Other	8.0%	n.a.	n.a.	9.9%	n.a.	n.a.
<b>Education:</b>						
Low	49 (49.0%)	n.a.	n.a.	24 (34.3)	n.a.	n.a.
High	51 (51.0%)	n.a.	n.a.	46 (65.7)	n.a.	n.a.
<b>Income (net):</b>						
€ 545 - € 726	3 ( 3.0%)	n.a.	n.a.	2 ( 2.9%)	n.a.	n.a.
€ 726 - € 908	3 ( 3.0%)	n.a.	n.a.	2 ( 2.9%)	n.a.	n.a.
€ 908 - € 1135	6 ( 6.0%)	n.a.	n.a.	4 ( 5.7%)	n.a.	n.a.
€ 1135 - € 1589	21 (21.0%)	n.a.	n.a.	10 (14.3%)	n.a.	n.a.
€ 1589 - € 2269	15 (15.0%)	n.a.	n.a.	14 (20.0%)	n.a.	n.a.
> € 2269	52 (52.0%)	n.a.	n.a.	38 (54.3%)	n.a.	n.a.
<b>EQ-5D</b>						
VAS Mean (SD)	76.7 (16.7)	73.1 (17.7)	73.8 (17.1)	77.9 (12.9)	73.1 (13.5)	73.3 (13.9)
Utility score Mean (SD)	0.86 (0.21)	0.87 (0.18)	0.85 (0.22)	0.89 (0.16)	0.87 (0.17)	0.88 (0.16)
<b>Objective burden:</b>						
Number of care giving tasks	5.9 ( 3.0)	6.2 ( 2.8)	6.2 ( 2.8)	6.0 ( 3.1)	6.5 ( 3.6)	6.6 ( 3.0)
Share house with care recipient (%yes)	36.0%	34.0%	34.0%	38.6%	37.1%	34.3%
<b>Subjective burden:</b>						
SRB <sup>1</sup>	3.4 (2.8)	3.9 (2.6)	4.1 (2.6)	3.6 (2.9)	4.1 (2.8)	4.2 (2.4)
EDIZ <sup>2</sup>	4.6 (2.7)	4.5 (2.7)	4.3 (2.7)	4.4 (2.7)	4.4 (3.0)	4.4 (2.9)

<sup>1</sup> Self-Rated Burden scale; <sup>2</sup> Perceived Stress caused by Informal Caregiving scale;

n.a.: characteristics which do not change over time

**Table 3:** Social context (intervention group and control group) at the three measurements

	Intervention group (n=100)			Control group (n=70)		
	Baseline	6 months	12 months	Baseline	6 months	12 months
<b>Caregiver network:</b>						
Other caregivers? (%yes)	62.0%	56.0%	55.0%	60.0%	67.1%	68.6%
Number of other care givers	1.4 (1.5)	1.4 (1.6)	1.4 (1.7)	1.3 (1.6)	1.7 (2.5)	1.6 (2.4)
Number of hours per week	4.1 (6.6)	3.7 (5.6)	4.3 (7.4)	3.2 (4.6)	3.7 (5.9)	6.1 (9.7)
<b>Other day care</b>						
(e.g. family, % yes)	9.0%*	6.0%	10.0%	0.0%	4.3%	7.1%
Hours per month (mean, SD)	0.7 (2.7)	0.7 (3.2)	1.2 (5.0)	n.a	1.1 (6.7)	1.7 (8.4)
<b>Monthly activities** (% yes)</b>						
Hours per month (mean, SD)	3.6 (12.5)	2.7 (9.6)	3.2 (8.0)	6.4 (16.1)	8.1 (26.1)	4.4 (10.3)
<b>Formal care: (%yes)</b>						
<b>Waiting list home care (%yes)</b>	2.0%	1.0%	1.0%	1.4%	0.0%	0.0%
<b>Home care (% yes)</b>	47.0%	45.0%	39.0%	47.1%	45.7%	42.9%
Hours per week (mean, SD)	2.9 (5.9)	2.9 (4.4)	2.9 (5.7)	2.4 (3.5)	3.3 (6.0)	2.6 (4.1)
<b>Waiting list day care (%yes)</b>	3.0%	2.0%	5.0%	7.1%	4.3%	5.7%
<b>Formal day care (%yes)</b>	17.0%	37.0%	41.0%	22.9%	34.3%	40.0%
Hours per week (mean, SD)	0.6 (1.6)	1.3 (2.2)	1.7 (2.5)	0.5 (1.2)	0.9 (1.5)	1.4 (2.1)
<b>Elderly home:</b>						
Waiting list (%yes)	8.0%	9.0%	6.0%	14.3%	10.0%	7.1%
Live (% yes)	11.0%	12.0%	17.0%	5.7%	10.0%	15.7%
<b>Nursing home:</b>						
Waiting list (% yes)	1.0%	1.0%	5.0%	1.4%	7.1%	2.9%
Live (% yes)	n.a.	9.0%	12.0%	n.a.	10.0%	15.7%
<b>Support (caregiver):</b>						
Paid leave from work (% yes)	0.0%	0.0%	1.0%	0.0%	0.0%	0.0%
Respite care (% yes)	0.0%	1.0%	0.0%	1.4%	1.4%	5.7%

\* significant difference between groups ( $p=0.01$ ); \*\* Monthly activities of the patients such as playing cards, play pool etc.

Table 4 presents the results of the proxy-good method to measure informal care (number of hours per week per activity and the percentages of caregivers performing the activities). Most caregivers spent time on IADL and HDL tasks (>90% in both groups). The percentage of caregivers performing these tasks remained stable over time. At baseline, 37.0% and 40.0% of the caregivers in the intervention group and the control group respectively spent time on ADL tasks and these percentages increased to 46.0% and 55.7% at the 12 months follow-up measurement. Overall, most time was spent on the HDL activities. The total amount of time spent on informal care decreased with 0.4 hours (which is a decrease of 4.8%) in the intervention group over the course of one year. The total amount of time spent on informal care in the control group increased with 5.4 hours (which is an increase of 60.6%) over the course of one year. In the intervention group the total costs slightly decreased with 5.3% whereas the total costs in the control group increased with 100.3% after a year. The total costs differed significantly between both groups after one year.

Table 5a presents the results of the opportunity-cost-A method regarding types and time forgone. Most caregivers indicated that they used leisure time in order to care for the patients in both groups. Significantly more caregivers in the intervention group gave up unpaid work in order to care for the patients at the baseline measurement. In total, a mean of 5.8 hours and 6.0 hours per week were forgone in order to provide caregiving in the intervention group and the control group respectively.

The total amount of time forgone increased with 1.8 hours (in the intervention group) and 4.3 hours (in the control group) over the course of one year. This is an increase of 36% in the intervention group and 110% in the control group. The total costs increase with 60.7% and 77.4% in the intervention group and control group respectively after one year.

Table 5b presents the results of the opportunity-cost-B method regarding the preferred usage of time in case no more care needed to be given to the patient. Most patients preferred to spend freed time on leisure at the three measurements in both groups. The mean number of hours amounted to 5.5 hours (in the intervention group) and 5.4 hours (in the control group), which was similar to the number of time forgone (opportunity-cost method A). The total amount of preferred time increased with 0.4 hours (in the intervention group) and 0.6 hours (in the control group) over the course of one year, which is an increase of 7.1% and 12.2% respectively. However, the costs decreased in both groups over the course of one year (15.3% in the intervention group and 5.1% in the control group), due to a change in preferences i.e. after 12 months the number of caregivers who would like to spend freed time on paid work had decreased whereas the number of caregivers who would like to spend freed time on leisure (which is cheaper) had increased.

**Table 4:** Proxy-good method (1); informal care time in mean hours per week at baseline, 6 months follow-up and 12 months follow-up (mean, SD, percentage of people performing the task)

	Intervention group (n=100)			Control group (n=70)		
	Baseline	6 months	12 months	Baseline	6 months	12 months
<b>HDL activities (total):</b>	<b>4.7 (6.1): 88.0%</b>	<b>4.1 (4.0): 93.0%</b>	<b>4.5 (5.3): 90.0%</b>	<b>4.9 (6.5): 90.0%</b>	<b>4.7 (7.3): 91.4%</b>	<b>5.1 (6.6): 90.0%</b>
Preparation of food and drinks	1.4 (3.3): (38.0%)	1.2 (2.5): (42.0%)	1.4 (3.1): (41.0%)	1.5 (2.8): (47.1%)	1.4 (3.5): (37.1%)	1.9 (3.5): (47.0%)
Cleaning the house	0.8 (2.6): (28.0%)	0.6 (1.1): (40.0%)	0.5 (1.2): (36.0%)	0.9 (2.2): (30.0%)	0.9 (2.0): (30.6%)	0.8 (1.9): (40.0%)
Washing, ironing or sewing	0.6 (1.1): (35.0%)	0.6 (1.2): (39.0%)	0.7 (1.1): (41.0%)	0.4 (1.2): (25.7%)	0.4 (0.8): (31.4%)	0.9 (1.3): (44.3%)
Shopping, groceries	1.1 (1.4): (71.0%)	0.8 (1.0): (70.0%)	0.8 (1.2): (65.0%)	0.9 (1.4): (61.4%)	0.9 (1.6): (54.3%)	0.9 (1.3): (62.9%)
Chores, gardening, maintenance	0.4 (1.2): (27.0%)	0.3 (1.5): (22.0%)	0.3 (0.9): (25.0%)	0.3 (1.1): (32.9%)	0.3 (0.6): (34.3%)	0.1 (0.4): (25.7%)
Other (inspection fridge, garbage etc.)	0.5 (1.0): (46.0%)	0.6 (1.2): (60.0%)	0.7 (2.6): (52.0%)	0.8 (1.9): (50.0%)	0.9 (2.0): (57.1%)	0.5 (1.5): (45.7%)
<b>ADL activities (total):</b>	<b>0.7 (2.2): 37.0%</b>	<b>0.8 (2.1): 44.0%</b>	<b>0.8 (1.9): 46.0%</b>	<b>0.7 (2.4): 40.0%</b>	<b>1.8 (4.5): 48.6%</b>	<b>1.9 (3.9): 55.7%</b>
Personal care (un)dressing, wash etc.)	0.5 (1.4): (28.0%)	0.6 (2.0): (33.0%)	0.5 (1.4): (37.0%)	0.4 (1.3): (32.9%)	1.1 (2.8): (41.4%)	1.1 (2.0): (42.9%)
Visiting the toilet	0.1 (0.7): (3.0%)	0.1 (0.4): (8.0%)	0.1 (0.5): (11.0%)	0.1 (0.4): (10.0%)	0.4 (1.9): (15.7%)	0.4 (1.6): (18.6%)
Moving around within the house	0.1 (0.7): (9.0%)	0.1 (0.2): (12.0%)	0.1 (0.5): (14.0%)	0.1 (0.5): (10.0%)	0.3 (1.8): (17.1%)	0.3 (0.7): (17.1%)
Eating and drinking	0.0 (0.1): (8.0%)	0.0 (0.1): (8.0%)	0.1 (0.2): (12.0%)	0.1 (0.4): (8.6%)	0.1 (1.9): (12.9%)	0.2 (0.9): (18.6%)
<b>IADL activities (total):</b>	<b>2.9 (3.0): 96.0%</b>	<b>2.7 (3.7): 96.0%</b>	<b>2.6 (2.9): 94.0%</b>	<b>3.3 (3.6): 95.7%</b>	<b>3.2 (5.7): 92.9%</b>	<b>7.4 (28.0): 98.6%</b>
Moving around outside the house	0.4 (0.8): (31.0%)	0.4 (0.8): (32.0%)	0.5 (1.0): (35.0%)	0.5 (1.5): (28.6%)	0.5 (1.3): (32.9%)	0.4 (0.7): (31.4%)
Excursions, visits to friends/family	0.9 (1.5): (51.0%)	0.8 (1.7): (42.0%)	0.6 (1.0): (41.0%)	0.7 (1.2): (52.9%)	0.6 (0.9): (47.1%)	0.6 (1.5): (44.3%)
Contacts with health care services	0.6 (0.8): (81.0%)	0.6 (1.0): (77.0%)	0.5 (0.8): (71.0%)	0.9 (1.5): (77.1%)	0.6 (0.8): (80.0%)	0.4 (0.6): (71.4%)
Organizing home adaptations	0.2 (0.5): (40.0%)	0.2 (0.3): (36.0%)	0.2 (0.4): (31.0%)	0.1 (0.3): (28.6%)	0.2 (0.3): (47.1%)	0.2 (0.4): (40.0%)
Financial matters (insurance, rent)	0.4 (0.5): (56.0%)	0.3 (0.4): (56.0%)	0.5 (0.9): (65.0%)	0.5 (0.7): (54.3%)	0.4 (0.6): (58.6%)	0.4 (0.7): (60.0%)
Medication	0.1 (0.4): (26.0%)	0.2 (0.5): (35.0%)	0.1 (0.3): (30.0%)	0.3 (1.3): (31.4%)	0.1 (0.4): (27.1%)	0.2 (0.7): (27.1%)
Other (mail, remembering)	0.3 (1.0): (16.0%)	0.2 (2.0): (11.0%)	0.3 (1.5): (14.0%)	0.3 (1.1): (18.6%)	0.8 (4.8): (18.6%)	5.1 (27.5): (24.3%)
<b>Total activities:</b>	<b>8.3 (8.2): 100%</b>	<b>7.5 (6.7): 100%</b>	<b>7.9 (7.3): 100%</b>	<b>8.9 (10.0): 100%</b>	<b>9.7 (14.8): 100%</b>	<b>14.3 (32.5): 100%</b>
<b>Total costs:</b>	<b>€ 156.49 (154.9)</b>	<b>€ 148.24 (156.8)</b>	<b>€ 148.21 (136.1)</b>	<b>€ 172.92 (198.4)</b>	<b>€ 203.28 (323.6)</b>	<b>€ 346.41* (984.6)</b>

\* significant difference between groups (p=0.05);

**Table 5a:** Opportunity cost method (2A): Changes in paid work, unpaid work and leisure time (%), mean number of hours, (SD) caused by providing care

	Intervention group (n=100)			Control group (n=70)		
	Baseline	6 months	12 months	Baseline	6 months	12 months
<b>Paid work forgone (% yes)</b>						
Hours forgone (mean, SD)	7.0%	5.0%	1.0%	1.4%	1.4%	4.3%
No change	0.6 (3.9)	1.2 (3.6)	1.3 (1.0)	0.2 (1.9)	0.0 (1.7)	0.0 (2.8)
Not applicable (no paid work)	37.0%	37.0%	45.0%	47.2%	48.6%	45.7%
<b>Unpaid work forgone (% yes)</b>						
Hours forgone (mean, SD)	56.0%	57.0%	54.0%	51.4%	47.1%	48.6%
No change	19.0%*	12.0%	8.0%	7.1%	12.9%	8.5%
Not applicable (no unpaid work)	1.2 (3.4)	1.4 (2.5)	1.8 (2.1)	0.5 (2.6)	1.2 (2.8)	1.7 (7.2)
Hours less (mean, SD)	81.0%	81.0%	89.0%	90.0%	85.7%	90.1%
Not applicable (no leisure time work)	0.0%	3.0%	0.0%	2.9%	0.0%	0.0%
<b>Leisure time forgone (% yes)</b>						
Hours less (mean, SD)	43.0%	32.0%	24.0%	52.9%	30.0%	28.6%
No change	3.2 (5.8)	3.0 (7.7)	3.7 (5.0)	3.2 (4.4)	4.8 (3.3)	6.5 (9.2)
Not applicable (no leisure time work)	57.0%	57.0%	66.0%	47.1%	65.7%	62.9%
Total hours forgone (mean, SD)	5.0 (8.8)	5.6 (11.9)	6.8 (14.1)	3.9 (5.7)	6.0 (9.1)	8.2 (14.7)
Total costs	€ 60,09 (157.7)	€ 82,47 (263.7)	€ 96,54 (273.4)	€ 42,53 (104.9)	€ 54,59 (143.8)	€ 75,46 (216.2)

\* significant difference between groups (p=0.01);

**Table 5b:** Opportunity cost method (2B): Preferences regarding paid work, unpaid work and leisure time (%), mean number of hours, (SD)

	Intervention group (n=100)			Control group (n=70)		
	Baseline	6 months	12 months	Baseline	6 months	12 months
<b>When I no longer have to provide informal care, I would spend this freed time on:</b>						
More paid work (%yes)	8.0%	4.0%	4.0%	2.9%	7.1%	4.3%
Hours more (mean, SD)	0.9 (3.5)	0.4 (2.1)	0.4 (2.2)	0.7 (4.4)	1.1 (5.7)	0.4 (2.6)
More unpaid work (%yes)	16.0%	15.0%	14.0%	15.7%	17.1%	21.4%
Hours more (mean, SD)	0.9 (2.6)	0.9 (2.4)	0.8 (2.3)	0.7 (2.0)	0.8 (2.3)	0.8 (1.9)
More leisure time (%yes)	59.0%	67.0%	69.0%	61.4%	77.1%	71.4%
Hours more (mean, SD)	3.8 (5.3)	3.6 (4.2)	4.8 (5.0)	3.5 (4.8)	4.0 (3.9)	4.3 (4.8)
Total number of hours (mean, SD)	5.6 (7.3)	4.9 (4.7)	6.0 (6.2)	4.9 (7.7)	5.9 (3.3)	5.5 (5.5)
Total costs	€ 75,33 (159.9)	€ 53,93 (82.5)	€ 63,82 (114.5)	€ 62,40 (183.5)	€ 86,43 (246.4)	€ 59,22 (100.1)

Table 6 summarizes the results of the CarerQoL, used as a non-monetary outcome for the measurement of informal care. Physical problems (of the caregiver) were most often reported in the intervention group (46.7%) at the three measurements (43.3% in the control group), whereas most problems (48.6%) in the control group were reported regarding relational problems with the patient (42.0% in the intervention group). Furthermore, 43.4% in the intervention group and 41.0% in the control group reported mental problems and 39.0% and 38.2% of the caregivers in the intervention group and control group respectively indicated having problems with carrying out their daily activities. Most caregivers found fulfilment in caring for their loved one. The mean baseline score on the CarerQoL VAS was 7.2 in both group. These scores slightly decreased to 7.0 in the intervention group and 6.8 in the control group at the 12 months follow up measurement. In total, there was a decrease of 2.7% (intervention group) and 5.5% (control group) on the CarerQoL. No differences between both groups were significant with regard to the CarerQoL.

Table 6: CarerQoL method (3): Subjective care related quality of life questionnaire (non-monetary outcome regarding the measurement and valuation of informal care).

		Intervention group (n=100)			Control group (n=70)		
		Baseline	6 months	12 months	Baseline	6months	12 months
<b>Fulfilment (%)</b>	No	9.0%	7.0%	16.0%	8.6%	8.6%	10.0%
	Some	32.0%	29.0%	31.0%	42.9%	40.0%	41.4%
	A lot of	59.0%	64.0%	53.0%	48.6%	51.4%	48.6%
<b>Relational problems (%)</b>	No	55.0%	66.0%	53.0%	52.9%	51.4%	50.0%
	Some	36.0%	27.0%	37.0%	35.7%	37.2%	40.0%
	A lot of	9.0%	7.0%	10.0%	11.4%	11.4%	10.0%
<b>Mental problems (%)</b>	No	59.0%	57.0%	54.0%	57.1%	62.8%	57.1%
	Some	30.0%	29.0%	35.0%	35.8%	28.6%	30.0%
	A lot of	11.0%	14.0%	11.0%	7.1%	8.6%	12.9%
<b>Problems with daily activities (%)</b>							
	No	59.0%	61.0%	63.0%	64.3%	61.4%	60.0%
	Some	36.0%	34.0%	29.0%	24.3%	34.3%	35.7%
	A lot of	5.0%	5.0%	8.0%	11.4%	4.3%	4.3%
<b>Financial problems (%)</b>	No	99.0%	95.0%	95.0%	97.1%	94.3%	91.4%
	Some	1.0%	5.0%	5.0%	2.9%	4.3%	7.1%
	A lot of	0%	0%	0%	0%	1.4%	1.4%
<b>Support (%)</b>	No	25.0%	23.0%	22.0%	17.1%	18.6%	15.7%
	Some	31.0%	27.0%	37.0%	37.1%	31.4%	42.9%
	A lot of	44.0%	50.0%	41.0%	45.8%	50.0%	41.4%
<b>Physical problems (%)</b>	No	52.0%	53.0%	55.0%	55.7%	58.6%	55.7%
	Some	35.0%	30.0%	24.0%	31.4%	25.7%	31.4%
	A lot of	13.0%	17.0%	21.0%	12.9%	15.7%	12.9%
<b>VAS (mean, SD)</b>		<b>7.2 (1.7)</b>	<b>7.1 (1.7)</b>	<b>7.0 (1.7)</b>	<b>7.2 (1.6)</b>	<b>6.8 (1.7)</b>	<b>6.8 (1.6)</b>

## Discussion

This study describes the effects of an integrated multidisciplinary approach to dementia on caregiver outcomes. The intervention group (DOC-PG) and the control group (usual care) were compared regarding the results of an informal care questionnaire, which was developed by the iMTA in Rotterdam, to measure and value informal care.

We found that significantly more patients in the intervention group made use of informal day care (i.e. day care provided by friends or family) at the baseline measurement. Another difference at the baseline measurement related to opportunity cost method (time forgone) to measure informal care. Significantly more caregivers in the intervention group gave up unpaid work in order to care for the patients. These baseline differences disappeared in the course of the follow-up measurements and we therefore did not expect that they affected the outcomes after 6 and 12 months.

Another difference related to the costs of informal care, measured and valued with the proxy-good method which were significantly higher in the control group at the 12 months follow-up.

No other differences between the groups were found.

It should be noted that the increase of the amount of informal care over time was higher in the control group than in the intervention group in all monetary methods. In fact, the amount of informal care measured with the proxy-good method increased with 60.6% in the control group after one year whereas this amount decreased with 4.8% in the intervention group. The opportunity-cost-methods indicated increases of the amount of informal care in both groups after a year, with increases of 110% (time forgone) and 12.2% (time preferred) in the control group as opposed to 36.0% (time forgone) and 7.1% (time preferred) in the intervention group. It is therefore plausible that DOC-PG had a positive effect on the amount of informal care provided and consequently also on the costs. The CarerQol remained relatively stable over time in both groups, although the control group showed a larger mean decrease on the happiness scale than the intervention group (5.5% in the control group and 2.7% in the intervention group). This finding is confirmed in the literature on caregivers burden, in which many intervention studies have shown that caregivers burden is remarkably stable over time (Verhey et al., 2007b).

A result of this study deserves some attention. The greater part of the caregivers were children (in-law) of the patients instead of their spouses which is generally the case in dementia (de Vugt et al., 2005b). It is therefore possible that we did not capture the primary caregiver for all patients (although this was asked explicitly) and the amount of informal care presented in this study could therefore be an underestimation of the actual amount. However, since the number of caregivers that were spouses and children (in-law) were similar in both groups, we expect that this did not affect the comparison between the two groups.

There are some possible limitations that need to be discussed.

In our study, informal care was divided in ADL activities, HDL activities and IADL activities. However, there is great variability in the literature regarding the definition and quantification of informal care time especially concerning the inclusion of supervision as a separate dimension of informal care (McDaid, 2001; Wimo et al., 2002). Caregivers may be able to perform activities while providing care such as supervision at the same time. Since supervision is not included explicitly in our survey, this could be a limitation. More uniformity about the concept of informal care is warranted.

Another limitation could be that a selection bias took place because we merely included those patients who completed all three measurements. Data from the non-completers did however not significantly differ from the data included in this study (data not shown). Although we do not expect that this will change our conclusions, we plan to perform additional analyses using multiple imputation to account for missing data.

Overall, few significant differences were found between the groups. The similarity of the groups on most outcomes may be due to the fact that our intervention (DOC-PG) was a combination of the two health care facilities that are available to the control group. The strengths of both facilities were combined in DOC-PG and thus result in comparable effects. In this context, it should be noted that usual care in our region consisted mainly of an active university MC and a CMHT that have collaborated in the past on several projects. Thus, the contrast between DOC-PG and usual care may have been lower as may have been the case in other regions.

However, our results indicate that DOC-PG did have a positive effect on the amount of informal care since this amount increased more in the control group than in the intervention group after one year. As the number of people receiving formal care and the amount of formal care were similar in both groups, it is not clear what causes this difference in the amount of informal care. A more efficient allocation of services and care according to the specific needs of the caregivers and the patients after the integrated multidisciplinary assessment by DOC-PG are possible explanations. Future research should however explore this topic further.

# Chapter 9

GP concordance with advice for treatment following a multidisciplinary psychogeriatric assessment



Published in *International Journal of Geriatric Psychiatry*, 2006; 22 (3): 233-240

Claire A.G. Wolfs, Frans R.J. Verhey, Alfons Kessels, Ron A.G. Winkens, Johan L. Severens and Carmen D. Dirksen

## Abstract

**Objective:** To evaluate the concordance of General Practitioners (GPs) with advice for treatment after a multidisciplinary psychogeriatric assessment by the Diagnostic Observation Centre for PsychoGeriatric patients (DOC-PG).

**Method:** Concordance checklists, listing the recommendations from the multidisciplinary team, were sent to the GPs in order to establish GP concordance. Regression models were used to study the associations between various patient and GP characteristics and level of concordance. Furthermore, results of a questionnaire (to identify the level of satisfaction regarding the services provided by the DOC-PG) were compared with the level of GP concordance.

**Results:** Based on 530 recommendations, the overall GP concordance rate amounted to 71%. The most common types of advice pertained to medication, GP follow-up/advice and referral. GP concordance with advice regarding admissions was the highest, followed by advice concerning the arrangement of day care, home care and the adaptation of medication. GP concordance was lowest for referral recommendations to other specialties and recommendations regarding psychoeducation. Concordance was higher for patients who lived alone, for patients with fewer cognitive problems, when the number of recommendations did not exceed six and in group practices. Concordance was dependent on the type of advice. Satisfaction with DOC-PG did not correlate with the level of concordance.

**Conclusions:** In general, GPs showed a high level of concordance with advice from the DOC-PG. Enhancement of GP concordance can be achieved by limiting the number of recommendations, giving detailed explanations about the purpose of recommendations and by doing this continuously educating GPs.

## Introduction

General Practitioners (GPs) are key figures in identifying, diagnosing as well as managing dementia (Waldorff et al., 2001). As in many other countries, Dutch GPs have developed their own guidelines for the diagnosis of dementia (Stoppe, 2005; Wind et al., 2003), in which early identification and adequate diagnosis of cognitive disorders are warranted. Achieving this goal can be facilitated by further referral to a specialized service such as multidisciplinary memory clinics (MMCs), which may provide the GP with specific advice or recommendations. The extent to which these recommendations have been complied with is important for the functioning of disease management models, and has been determined in various patient groups (Allen et al., 1986; Mainprize et al., 1987a; Podgorski et al., 2002a; Reed et al., 1990; Reuben et al., 1996; Ruskin, 1985; Sears et al., 1983; Wrigley et al., 1991), but to the best of our knowledge only one study (Teitelbaum et al., 1996) has been published about the level of concordance by referrers in psychogeriatric outpatients. Furthermore, no studies have been conducted in the health care systems comparable with the Dutch model, in which the GP acts as gatekeeper.

In many countries, the GPs have several specialized services for patients with psychogeriatric problems at their disposal, such as MMCs or mental health community centres. There is a new diagnostic facility in Maastricht, the Netherlands. It is the Diagnostic Observation Centre (DOC-PG), an outpatient facility that combines the medical approach of the MMC, the Department of Geriatric Medicine at Maastricht University Hospital and the regional community mental health centre, using a programmed investigative package. The objective of the present study was to examine GPs level of concordance following a multidisciplinary psychogeriatric assessment by the DOC-PG. In particular we were interested in factors influencing concordance.

## Materials & Methods

The current study was part of the larger MEDICIE study (Maastricht Evaluation of a Diagnostic Intervention for Cognitively Impaired Elderly), a randomized clinical trial (RCT) comparing clinical effects and health care economic effects of the multidisciplinary DOC-PG with usual care (Wolfs et al., 2005b). The study was approved by the medical ethical committee of Maastricht University Hospital.

### DOC-PG

The function of the DOC-PG is to provide multidisciplinary assessment by somatic screening, psychogeriatric assessment, and evaluation of the required levels of care for the patient and his (her) carer. Participating disciplines in the DOC-PG include old age psychiatry, geriatric medicine, neuropsychology, physiotherapy, occupational therapy, geriatric nursing and mental health nursing. The GPs can refer to the DOC-PG, when a cognitive disorder or a combined somatic/psychiatric disorder is suspected. The diagnostic screening of DOC-PG takes place in two weeks and consists of a visit at home and two visits to the University Hospital (to the internal medicine outpatient clinic

and the psychiatric outpatient clinic). In addition, a CT or MRI scan is performed. Thereafter, the results are discussed at a weekly interdisciplinary meeting, in which a definite diagnosis is made, and a treatment plan is formulated. A letter is sent to the referring GP in which a summary of the assessments, the multi-axis diagnosis and recommendations for treatment and management are described. After the assessment the GP is responsible for the patients, although the advice may be that further treatment by one or more of the disciplines involved in the DOC-PG is warranted.

### **Data collection**

Data were collected by means of a concordance checklist (a) and a satisfaction questionnaire (b), which were sent to each GP who had referred a patient to DOC-PG during the index period. The GPs who did not respond repeatedly received reminders.

(a) The concordance checklist listed all advice or recommendations at patient level from the multidisciplinary team in the period from July 2002 until August 2004. An inventory of the recommendations for each patient (i.e. a concordance checklist) was sent to the GPs, on which the GP was requested to indicate whether an advice was complied with or not (yes/no/other). The “other” option was added in case difficulties were encountered or when the GPs wanted to add some remarks. Concordance was defined as a “yes” score on the checklists as reported by the GPs. When patients refused to comply with a recommendation for treatment (such as day care), or when advice had already been carried out, (i.e. patients already received the care that had been recommended), it was scored as concordant.

The recommendations were then classified into 12 categories i.e.: 1) Refer to MC for Alzheimer medication; 2) Adapt other medication; 3) Perform diagnostics (i.e. check blood pressure, diabetes, vitamin deficiencies etc.); 4) Arrange follow up (i.e. regularly check weight, personal hygiene, skin problems etc.); 5) Give general advice (i.e. stop smoking, wear glasses or hearing aid, increase fluid intake etc.); 6) Provide psychoeducation; 7) Refer to hospital specialties other than DOC-PG; 8) Refer to paramedical disciplines; 9) Refer to care authorities (home care); 10) Arrange nursing home admission; 11) Arrange care home admission; 12) Arrange daily activities (day care centre).

(b) A short satisfaction questionnaire (see appendix A) was developed to identify the level of satisfaction regarding the services provided by the DOC-PG. The questionnaire was sent to the GPs three months before sending the checklists in order not to influence either outcome. The questionnaire consisted of six statements, with five possible reactions to each statement, namely “completely agree” (score +2), “agree” (score +1), “no opinion” (score 0), “disagree” (score -1) and “completely disagree” (score -2).

The Mini Mental State Examination (MMSE) was used as a global indicator of the severity of cognitive impairment. (Folstein et al., 1975a).

### **Statistical analyses**

The software used for the analyses was SPSS version 12.0.1 and STATA version 8.0. Background characteristics (means, standard deviations, ranges) of patients and GPs

were summarized using SPSS. Furthermore, frequencies and concordance rates for the 12 groups of recommendations are given.

A logistic stepwise regression model was developed in STATA, in order to examine the influence of various covariables on concordance. As the concordance per advice was expected to be correlated within a GP, the logistic regression analysis was adjusted for intragroup correlation by using the cluster option in STATA. In order to be able to incorporate the categorical variable "type of advice" into the regression equations, the 12 categories of advice were combined to three groups (see table 3) and two dummy variables were used. Each category was compared with a reference category (i.e. "carry out").

In the model, three types of independent variables can be distinguished: 1) patient characteristics (age, gender, living situation (alone or not alone), diagnosis (dementia yes/no), and MMSE), 2) advice characteristics (the number (< 6 or > 6) and type of recommendations) and 3) GP characteristics (age, group practice/solo practice, city/rural and percentage elderly (>age 65) per practice). First, it was determined using a stepwise logistic regression for each of these three groups which covariables were most influential on concordance. Thereafter, the selected variables from the three analyses significantly associated with concordance were entered into one model.

Spearman rank-order correlations (two-tailed) were calculated between the six statements of the satisfaction questionnaire and the mean concordance rate of each GP, just using data of GPs who returned the checklists as well as the questionnaires.

## Results

### Practitioner and patient characteristics

In total, 137 patients, who were referred by 49 GPs, gave their informed consent to participate in the MEDICIE study and their data were analyzed. The DOC-PG team formulated a total of 783 recommendations, of which 640 (81.7%) were to be concorded with by the GP, and the others by disciplines of the DOC-PG. Characteristics of the GPs, the GPs' practices and the patients are summarized in table 1 and table 2.

### Response

After two reminders by telephone, 44 GPs (89.8%) returned the concordance checklists for a total of 110 patients. In four cases the advice could not be concorded with because of the premature death of the patient and they were therefore omitted from the analyses. Eventually 530 recommendations were to be concorded with by the GP, with a mean of 4.8 recommendations per patient (SD = 2.2, range 1-11).

The satisfaction questionnaires were returned by 42 GPs (85.7%). Seven had not been filled in and one was returned anonymously. Of the remaining 34 questionnaires, five GPs did not return their concordance checklists. Therefore, 29 questionnaires were used to study the influence of GP satisfaction on the level of GP concordance. Characteristics

**Table 1:** Characteristics of GPs (practices) referring to DOC-PG

Number of GPs	Total (n=49)	Participating (n=44)
<b>Male/female</b>	43 / 6	38 / 6
<b>Practices (type)</b>	32	29
Solo practice	20	18
Group practice	12	11
<b>Urbanization</b>		
City	33	31
Rural areas	16	13
<b>Proportion of patients older than 65 years</b>		
<b>Total (mean)</b>	19.2%	18.7%
Minimum	5.0%	5.0%
Maximum	34.6%	34.6%

**Table 2:** Characteristics of patients

	Total	Participating
<b>No. of patients</b>	137	110
<b>Age</b>		
Mean (SD)	78.3 (6.5)	78.6 (6.4)
Range	55 - 93	61 - 93
<b>Mini Mental State Examination (MMSE)</b>		
Mean (SD)	20.4 (5.6)	20.4 (5.9)
<b>Diagnosis</b>		
Dementia	96 (70.1%)	76 (69.1%)
No Dementia	41 (29.9%)	34 (30.9%)
<b>Living situation</b>		
Living alone	52 (38.0%)	44 (40.0%)
Not living alone	85 (62.0%)	66 (66.0%)

of the responders and non- responders among the GPs and the patients, as well as the number and type of recommendations, were comparable.

### Advice or recommendations

The types of advice to be concurred with by the GP are summarized in table 3. Generally speaking the most common types of advice related to medication, diagnostics by GP and referrals. The other medication recommendations (112) included advice to start new medication (45), change/reorganize existing medication (31), discontinue medication (26) or continue medication as it was (10). A variety of further assessments was also suggested, which comprised mainly of referrals to hospital departments not involved with DOC-PG (80) and to paramedical disciplines (50 to physiotherapists, 29 to occupational therapists and 25 to others such as speech therapists or dieticians).

### Concordance

The results concerning GP concordance are also summarized in table 3. The GPs reported a total concordance rate of 71%. The highest concordance occurred for recommendations regarding admissions, either to a nursing home or a care home (93% and 100% respectively). Concordance was also high for Alzheimer medication recommendations (87%) and other medication recommendations (total 83%, start new medication 84%, change/reorganize existing medication 93%, discontinue medication 60% and continue medication as it was 88%), for advice concerning daily activities (85%) and for home care recommendations (84%). The lowest concordance occurred for advice concerning psychoeducation (64%), referral to hospital departments not involved with DOC-PG (54%) and for referral to paramedical disciplines (48%). Referral to occupational therapists had a particularly low concordance rate (29%). Concordance rates for recommendations concerning diagnostics, follow-up and general advice by the GPs were in between (77%, 75% and 72% respectively).

Non-concordance was stated for 152 recommendations. With regard to nine recommendations the “other” option was chosen, since the GP did not know whether the advice had been adhered to (four had already been admitted to a nursing home).

### Satisfaction questionnaire

The results of the questionnaire are summarized in appendix A. While scores could vary between -2 (completely disagree) and +2 (completely agree), a mean positive score was found for each statement, which suggests a high level of satisfaction with the services provided by the DOC-PG. GPs agreed most (a mean score of +1.1) with the statement about the added value of DOC-PG (question 6) and least (a mean score of +0.2) with the statement concerning the time period between referral and diagnosis (question 3). No significant Spearman rank-order correlations were found between any of the statements of the satisfaction questionnaire and the level of GP concordance.

### Logistic stepwise regression analysis

The results of the logistic regression analyses are outlined in table 4. Several characteristics of the patients, the GPs and the advice were significantly associated with GP concordance as demonstrated by the first three analyses. Concordance was higher for patients living alone (than for patients not living alone) and for patients with fewer cognitive problems, i.e. a higher MMSE score (than for patients with more cognitive problems i.e. a lower MMSE score). Concordance was also higher in group practices (than in solo practices). Concordance significantly decreased when more than six recommendations were given. The cut-off score for this number of recommendations was changed until statistical significance was reached. Furthermore, the type of advice influenced the level of concordance, i.e. the level of concordance regarding further referral was significantly lower than the level of concordance regarding recommendations to be performed by GPs themselves. Finally, when these variables were consolidated into one model, all variables except living situation significantly ( $p < .05$ ) influenced concordance.

**Table 3:** Types of advice and GP concordance rates

Type of advice	Type of advice	Total (137 patients)	Checklists (110 patients)	Concordance with	GP concordance
1. Adapt medication	(A)	112	88	73	73/88: 83.0%
2. Perform diagnostics	(A)	101	84	65	65/84: 77.4%
3. Arrange follow up	(A)	26	24	18	18/24: 75.0%
4. Give general advice	(A)	71	57	41	41/57: 71.9%
5. Provide psycho education	(B)	15	14	9	9/14: 64.3%
6. Arrange home care	(B)	41	32	27	27/32: 84.4%
7. Arrange admission to nursing home	(B)	15	15	14	14/15: 93.3%
8. Arrange admission to care home	(B)	2	2	2	2/2: 100.0%
9. Arrange daily activities (daycare)	(B)	37	27	23	23/27: 85.2%
10. Refer to MMC for AD medication	(B)	36	30	26	26/30: 86.7%
11. Refer to hospital (not DOC-PG)	(C)	80	70	38	38/70: 54.3%
12. Refer to paramedical disciplines	(C)	104	87	42	42/87: 48.3%
<b>Total:</b>		<b>640</b>	<b>530</b>	<b>378</b>	<b>378/530: 71.3%</b>

MMC = Multidisciplinary Memory Clinic

AD = Alzheimer's disease

Type of advice:

A = To be performed by GPs themselves

B = The arrangement of particular types of service (e.g. daycare)

C = The referral to hospital departments (not DOC-PG) and paramedical disciplines

Table 4: Results of the 4 logistic stepwise regression analyses

Concordant with yes/no	Coef.	Robust Std. Err.	P> z	95% Confidence Interval
<b>1: MMSE</b>				
Living alone (1) / not living alone (2)	0.044	0.016	0.006	[0.013, 0.074]
	-0.456	0.190	0.016	[-0.829, -0.083]
<b>2: Practice (group (1) / solo (2))</b>	-0.469	0.208	0.024	[-0.878, -0.060]
<b>3: Number of recommendations (&lt; 6 (1) or &gt; 6 (2))</b>	-.0499	0.175	0.004	[-0.842, -0.156]
<b>4: MMSE</b>				
Living alone/not living alone	0.042	0.018	0.019	[0.007, 0.077]
Number of recommendations	-0.391	0.212	0.065	[-0.806, 0.245]
Practice (group/solo)	-0.469	0.178	0.009	[-0.818, -0.120]
	-0.506	0.241	0.036	[-0.979, -0.034]

530 observations, clustering at GP level

1-4: Dependent variable compiled with yes/no (p<0.05)

1: Independent variables patient characteristics

2: Independent variables GP characteristics

3: Independent variables type and number of recommendations

4: Independent variables significant variables of analyses 1, 2 and 3

NB The dummy variables (type of advice) were left out of the table.

## Discussion

In general, the results of this study showed good GP concordance with specified advice arising from a multidisciplinary psychogeriatric assessment by the DOC-PG. Furthermore, GPs were on average satisfied with the services provided by the DOC-PG. They generally recognized the added value of the DOC-PG compared with existing diagnostic facilities. The relatively high response rate of the GPs possibly reflects this satisfaction.

Most recommendations from the DOC-PG concerned medical recommendations such as adapting medication, performing diagnostics and referring to hospital departments (other than DOC-PG), which is comparable to other studies (Mainprize et al., 1987b; Reed et al., 1990; Ruskin, 1985; Sears et al., 1983; Teitelbaum et al., 1996). The wide range of recommendations reflects the complexity of our study population's health problems.

In total, the GPs concurred with 71% of the recommendations. In line with other studies (Mainprize et al., 1987b; Podgorski et al., 2002a; Podgorski et al., 2002b; Reed et al., 1990; Ruskin, 1985; Sears et al., 1983; Teitelbaum et al., 1996), recommendations of a medical nature were concurred with best. Interestingly, with regard to medication recommendations, GPs were more likely to start new medication or to adapt existing regimens than to discontinue medication. This result, also reported by Reed et al. (1990), could be due to a reluctance to change longstanding drug regimens common in older patients. Recommendations concerning referral and psychoeducation were the least concurred with. It is possible that referrals, especially to physiotherapists and occupational therapists, are less common for GPs or are considered unnecessary. This may be related to the fact that the role of these paramedical disciplines in managing psychogeriatric patients is not well established yet, at least not in the Netherlands. More knowledge about the added value of paramedical interventions in psychogeriatric patients is advisable. With regard to psychoeducation, this is only a small part of the professional responsibilities of Dutch GPs. According to the dementia guidelines (Wind et al., 2003), GPs merely have to guide patients in need of psychosocial assistance to the appropriate services. It is therefore possible that GPs in fact did concord, but were under the impression that they did not carry out the recommendations completely since they only helped patients to get started finding help. It is advisable to provide more knowledge about the professional responsibilities of GPs regarding psychosocial issues. It remains speculative how this key finding, i.e. more concordance with medical recommendations compared with psychosocial recommendations, should be interpreted. It is possible that GPs carried out the easy tasks first and left harder tasks until later, potentially 'loosing' them. It is possibly due to resource limitations (e.g. the availability of occupational therapists). Finally, it is also possible that GPs disagreed with recommendations regarding patients whom they knew much better than did the specialist team.

It should be noted that during the study period the DOC-PG procedure with regard to the initiation of Alzheimer medication was changed. In the first period GPs were responsible for referral to the MC, whereas in the latter phase the MC directly called up patients. This explains the relatively low proportion of patients with the advice to start Alzheimer medication.

Other important findings arising from this study relate to the various predictors of GP concordance, examined by means of logistic stepwise regression analyses. Living situation and the score on the MMSE were the patient characteristics that influenced concordance. GPs concurred better with recommendations when these pertained to patients who are living alone. Recommendations concerning patients not living alone might be seen as less urgent since there is a carer present to help the patient in case of need. However, in the final model (combining the patient characteristics, the GP characteristics and number and type of recommendations), living situation lost its significant influence, which was caused by type of advice. This is possibly due to the fact that fewer referral recommendations were given to patients living alone, which in turn is possibly due to a poorer quality of reported health by patients and proxies living alone. No other differences (i.e. in cognitive functioning, age, number and type of advice) were found between patients living alone and patients not living alone. Inconsistent with previous literature (Sears et al., 1983; Teitelbaum et al., 1996), results on the MMSE score were positively associated with the concordance rate, meaning that better concordance occurred in cases of less cognitive impairment. This may indicate that GPs recognize the need for an early diagnosis of dementia. Given the necessity of an aetiological diagnosis of dementia for current pharmacotherapy (e.g. cholinesterase inhibitors), this finding is relevant to clinical practice.

A variable negatively associated with concordance is the number of recommendations formulated, a result which is also reported in other studies (Reuben et al., 1996; Sears et al., 1983). Although comprehensive evaluation is essential, exhaustive lists of recommendations may be counterproductive. Our study indicated better concordance when six or fewer recommendations are given, which compares well with Sears and Charlson (1983), who found a cut off number of five recommendations. It is advisable to minimize the number of recommendations, focusing on central issues in patient care. The type of advice given also influenced concordance. As mentioned above, GPs were more likely to concur with advice they actually had to carry out themselves (such as adapt medication and perform diagnostics) and with advice concerning arrangements (such as the arrangement of home care and day care), than with advice concerning referrals. It is advisable to give detailed explanations about the purpose of the recommendations made on an individual basis.

Finally, our study showed that GPs working in group practices displayed higher concordance rates than GPs who work in solo practices. This may be related to an exchange of experience, resulting in a continuous learning process. This and other advantages of working in a group practice (such as improved quality of life for the GP and continuity of care for the patient) have been reported by Feron and co-workers (Feron et al., 2003). It should be noted that the reasons that are provided to explain the various associations remain unclear and more research is necessary to clarify these issues. No relationship between the level of satisfaction with DOC-PG and GP concordance was found in this study.

There are limitations to this study that need to be considered. First, our methodology may be limited since we relied on self-reports by the GPs to determine GP concordance. This method raises several concerns since these reports are totally dependent on the level of accessibility and detail (i.e. probably more elaborate reporting on medication

than on psychoeducation) of the data recorded by the GPs. Since the period between DOC-PG assessment and the examination of concordance was relatively long (between one and two years), GP reports may be inaccurate in the case of incomplete patient records. As mentioned previously, this study is a side study of a (larger) RCT. Since GP concordance was not a major outcome but part of the process evaluation of DOC-PG, we did not use a more targeted method to estimate concordance. In future, a more elaborate method will be used to study GP concordance. A second limitation relates to the reasons for non-concordance. GPs were not explicitly requested to systematically provide reasons for non-concordance with the recommendations given. Therefore, the underlying motives for non-concordance remain unclear. Future studies should address this issue in order to improve the interface between GPs and specialized services.

In conclusion, the relation between the GPs and DOC-PG is satisfactory as indicated by the high level of concordance. Enhancement of GP concordance can be achieved by limiting the number of recommendations, giving detailed explanations about the purpose of recommendations and by doing so continuously educating GPs. The positive findings of this study indicate the acceptance of the DOC-PG by the direct care providers, which was the first barrier to overcome. Results of the MEDICIE study, available in 2006, will reveal the possible benefits of DOC-PG for the patients.

**Key points:**

GPs displayed a high level of concordance with advice (or recommendations) from the Diagnostic Observation Centre for Psycho Geriatric patients (DOC-PG).  
Enhancement of GP concordance can be achieved by limiting the number of recommendations, giving detailed explanations about the purpose of recommendations and by doing so continuously educating GPs.

# Chapter 10

General discussion



## Introduction

The aim of this thesis was to evaluate the effectiveness and cost-effectiveness of an integrated multidisciplinary approach by the Diagnostic Observation Centre for PsychoGeriatric patients (DOC-PG) to the diagnosis and management of dementia. In the first part of the thesis, we systematically searched the literature for empirical studies on the added value of a multidisciplinary approach to dementia (chapter 3). Part two addressed the construct validity and responsiveness of the EQ-5D and the EQ-5D+C (used for measuring Health-Related Quality of Life) and of a number of different methods for the measurement and valuation of informal care (chapters 4 and 5). The final part described the findings of the MEDICIE-study. First, we studied the effectiveness and cost-effectiveness of DOC-PG compared with usual care (chapters 6 and 7). Subsequently the differences between DOC-PG and usual care regarding caregiver outcomes are described (Chapter 8). Finally, chapter 9 focused on the concordance with advice for treatment and level of satisfaction of the General Practitioners (GPs), after an assessment by the DOC-PG. In this last chapter of the thesis we will briefly summarize the main findings and discuss methodological issues regarding care for dementia and economic evaluations within this field, implications for (clinical) practice and future directions.

## Summary of findings

### Part one

Although an integrated policy regarding dementia is generally recommended in several practice parameters (APA, 1997; CBO, 2005), empirical evidence is very scarce. A review of the literature in chapter 3 identified merely five studies on this topic. This underlines the importance of the MEDICIE-study, as this is the first randomized controlled trial on the effectiveness and cost-effectiveness of such an integrated multidisciplinary approach to dementia.

### Part two: Methodological issues

The increasing number of older adults who are diagnosed with dementia has far-reaching implications for health service delivery and expenditures (Holland et al., 2004a). Economic evaluations are performed more often to assist decision-makers in setting priorities, especially with regard to resource allocation (van Velden et al., 2005). A central component of economic evaluations in health care is the use of preference-based instruments to measure changes in Health-Related Quality of Life (HRQoL). To date, the EQ-5D (Brooks, 1996; , "EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group", 1990) and the SF-6D (Brazier et al., 2002b) are the only two utility-generating instruments recommended by NICE (NICE, 2006), although concern has been raised that these are insufficiently sensitive to pick up (clinical) changes in dementia patients (Knapp, 2007). The finding in chapter 4 that the EQ-5D performs well (regarding construct validity and sensitivity to change) in our population

with cognitive impairments using proxy ratings is therefore a relevant outcome as it suggests that this concern is ungrounded.

Another important issue in economic evaluations in dementia relates to informal care. It is generally recommended (Drummond et al., 1997; Gold et al., 1996; Oostenbrink et al., 2004) that economic evaluation studies should be conducted from a broad societal perspective, including all relevant costs and health effects. Yet, systematic reviews (Evers et al., 2000; Evers et al., 1997; Goossens et al., 1997, 2000) show that informal care is rarely included in economic evaluation studies, since the collection and valuation of these data is often complex. In chapter 5 we addressed this issue by applying different methods to measure and value informal care. These methods included the proxy-good-method (van den Berg et al., 2006), two opportunity-cost-methods (time forgone and preferred time) (van den Berg et al., 2006) and the CarerQol-method (Brouwer et al., 2006). The proxy-good-method (also called the market cost method or replacement cost method) values the measured time spent on providing informal care at the (labour) market price of a close substitute. In the opportunity-cost-methods, we valued the lost time as the opportunity cost (i.e. hours of paid work, unpaid work or leisure time forgone), or the value of the time in its current best alternative use (hours of paid work, unpaid work or leisure time). The CarerQol-method provides us with a non-monetary outcome regarding the measurement and valuation of informal care. It is aimed at measuring the care-related quality of life of informal caregivers. Based on the results of chapter 5, the proxy-good method seems as yet the most valid method to measure and value informal care in patients suffering from dementia (or another cognitive disorder) and their caregivers.

### **Part three: Findings of the MEDICIE-study**

The outcomes of the MEDICIE-study were evaluated at multiple levels. We studied the outcomes on patient level (chapters 6 and 7), on caregiver level (chapter 8) and on the level of the General Practitioner (chapter 9).

Health-related quality of life (HRQoL) of the patient was the primary outcome to evaluate the effectiveness of DOC-PG (chapter 6). The EQ-5D was used to measure HRQoL. Changes in VAS scores of the EQ-5D over the course of one year was the primary outcome variable. A difference of 10% or more between the intervention group and the control group on the VAS was a priori considered a clinically relevant difference. The number of patients experiencing this clinically relevant difference were compared between both groups. We found a mean group difference of 9.6% on the VAS of the EQ-5D after 12 months, which is very close to our initial expectations. The proportion of patients who improved more than 10% (of the group difference) on the VAS was significantly greater ( $p=0.01$ ) in the intervention group (39.0%) than in the control group (22.1%). The groups did not differ in terms of clinical outcome measures.

Alongside the randomized controlled trial, an economic evaluation was carried out, comparing the costs and consequences of DOC-PG with usual care (chapter 7). The economic evaluation was conducted from a societal perspective according to (inter)national guidelines, using a one year time horizon. An incremental cost

effectiveness ratio (ICER) was calculated by dividing the difference in costs by the difference in Quality Adjusted Life Years (QALYs). Non-parametric bootstrapping and one-way sensitivity analyses were used to assess the uncertainty in the costs and effects. The EQ-5D instrument was used to measure HRQoL and the proxy-good method was used to measure and value informal care, which corresponds well with the recommendations made in the methodological chapters (4 and 5) of this thesis. Compared to patients receiving usual care, patients who visited DOC-PG gained a mean 0.05 QALYs. The incremental costs per QALY amounted to € 1267. DOC-PG is not demonstrable more expensive than usual care in the Netherlands, despite of the investment in the diagnostic intervention, but it showed a high probability of being more effective. Therefore, we contend that DOC-PG is indeed a cost-effective facility for the diagnosis and management of dementia in ambulant patients.

Chapter 8 describes differences in caregiver outcome between the intervention group and the control group after an integrated assessment by DOC-PG. A questionnaire, developed by the iMTA in Rotterdam, the Netherlands was used to measure and value informal care, to assess caregiver characteristics (relationship with the patient, education, income, self-rated health, HRQoL, and burden) and to assess social context (caregiver network, formal care, other daily activities of the patients (that support the caregiver), respite care and paid leave from work. We found that significantly more patients in the intervention group made use of informal day care (i.e. day care provided by friends or family) at the baseline measurement. Another difference related to the costs of informal care, measured and valued with the proxy-good method which were significantly higher in the control group at the 12 months follow-up. Finally, significantly more caregivers in the intervention group gave up unpaid work in order to care for the patients as indicated at the baseline measurement. No other significant differences between the groups were found.

Overall, although few differences were found between both groups, the results suggest that DOC-PG may have a positive effect on the amount of informal care since this amount increased more in the control group (mean increase of 61%) than in the intervention group (mean increase of 12%) after one year.

The value of DOC-PG has already been recognized by GPs as evidenced by the high referral rate among GPs and by the high concordance with DOC-PG recommendations (chapter 9). Acceptance of the DOC-PG by the direct care providers was the first barrier to overcome (Wolfs et al., 2007b).

Which perspective should be considered most important, when evaluating an integrated model, such as DOC-PG, using a multitude of perspectives? Isn't it true that a positive effect is a gain, no matter who incurs it? In fact, eventually, positive effects on GPs will most likely have positive effects on patients and caregivers. Likewise, positive effects for patients will likely improve caregiver's well-being, and vice versa. It is therefore plausible that the positive results presented in this thesis will have beneficial consequences regarding these mutual relationships.

## Methodological considerations

### Strengths of the study

The MEDICIE-study is the first study using a randomized design and a complete health economic evaluation regarding an integrated approach to dementia.

Furthermore, the evaluation of the effects of DOC-PG took place on multiple levels (i.e. the level of the patients and their caregivers and on the level of the GP's) from a broad societal perspective, which aids to establish a more complete picture. Another strength is that valid and reliable research instruments were used to assess the patients. In addition, the MEDICIE-study used a highly systematic approach with respect to the assessment. Each participant was interviewed at home and followed up by the same researcher.

A final strength is the longitudinal nature of the study (i.e. a follow-up period of one year) to evaluate the effects of an integrated approach, although some argue that a longer duration of follow-up is necessary when examining treatment benefits for dementia, a disorder that results in a gradual progressive decline over several years (NICE, 2006; Wimo, 2007). The follow-up period in dementia studies should be at least 12 months to capture all effects and costs, as major cost-driving events such as institutionalization will rarely occur in periods shorter than 12 months. It is therefore possible that we missed certain outcomes that occurred in the year following our follow-up period. Due to the lack of evidence, no extrapolation beyond our follow-up period could be performed. It should be noted that longitudinal designs usually result in a large proportion of study drop-outs. However, in the MEDICIE-study less than 5% were loss to follow-up after 1 year.

### Limitations and research challenges

The MEDICIE-study also has methodological limitations and research challenges that need to be discussed. These issues are addressed below, as well as how they may have influenced the results presented in this thesis.

#### *Design of the study*

It was not feasible to blind the interviewers assessing the patients and their carers to the treatment assigned. However, all instruments were standardized and the participants received neutral instructions for every instrument.

Another potential problem is linked to our inability to keep the random allocation sequence completely concealed, because the person responsible for the allocation of patients also recruited a small number of patients (5%). However, the people who recruited the majority of the patients were blinded with respect to patient allocation. The blindness of the referring GPs could not be maintained until the end of the study. In order to investigate the potential effects of this on the study results, we compared post-hoc the characteristics of patients in the two groups who were recruited in the first year (GPs were blinded) and in the second year (GPs were not blinded). We did not find any differences within the intervention group (with respect to age, gender, diagnosis, MMSE score and Global Deterioration Scale (GDS) score); however, there were differences in diagnosis and GDS score in the control group: in the second year of the inclusion period, more people with a cognitive disorder other than dementia and with a lower GDS score were included. The GPs probably wanted to refer patients to DOC-PG but this was only

possible after recruitment was completed. Although these patients fitted the inclusion criteria, this resulted in a somewhat more heterogeneous patient population. The inclusion of slightly healthier patients (with consequently higher QOL and lower costs) in the latter half of the inclusion period, probably resulted in a less favourable outcome for DOC-PG.

#### *Data collection*

All data were collected by means of personal interviews with a patient's proxy. The method of proxy rating was chosen because of the longitudinal nature of the study, the complex health problems of the study population, and the progressive global deterioration of intellect and personality characteristic to members of the study population. It is generally acknowledged that in the later stages of dementia, proxy measures are indeed required since patients are no longer capable of adequately evaluating their own health (Jonsson et al., 2006a; Selai, 1998). It is therefore possible that the scores on the instruments were biased because of a perceived caregiver burden (Logsdon et al., 2002). This bias, however, applies to both groups. Furthermore, it should be emphasized that we measured the proxy's perception of the HRQoL of the patient and not a direct estimate of HRQoL. This issue should be considered as a research challenge rather than an actual limitation of the MEDICIE-study.

Proxies were also asked to complete cost diaries which were used to calculate part of the cost volumes. These diaries appeared to be very time-consuming and difficult to complete. Retrospective questionnaires with a limited recall period would probably have resulted in a higher completion rate. Indeed, it may be more beneficial to use a cost questionnaire with structured closed questions for the assessment of health care utilisation in economic evaluations alongside clinical trials (van den Brink et al., 2005). As there were no other possibilities in our study, other than the diary, for determining each patient's number of visits to the GP, visits to other health care professionals and number of medical aids purchased elsewhere (since too much time had passed to ask retrospectively), more missing values appeared in these costs and therefore imputation had to be conducted in a larger proportion than with the other cost items. However, these costs constitute only a small proportion of the total costs (7%).

A final research challenge concerns the measurement of informal care. In our study, informal care was divided in ADL activities, HDL activities and IADL activities. However, there is great variability in the literature regarding the definition and quantification of informal care time, especially concerning the inclusion of supervision as a separate dimension of informal care (McDaid, 2001; Wimo et al., 2002). Caregivers may be able to perform activities while providing care such as supervision at the same time. Since supervision is not included explicitly in our survey, this could be a limitation. More uniformity about the concept of informal care is warranted.

## Implications

In daily practice (in health care) there often exists a division between a disease-oriented approach and a care-oriented approach (CBO, 2005). The disease-oriented diagnostic approach of the secondary health care addresses itself to the nature, cause, treatment and prognosis of the dementia syndrome, leading to specific advice to the referrer. In contrast, the community care-oriented approach of social services is aimed at the level of care needed, and the strength and weakness of the patient's carers. The DOC-PG model combines these complementary approaches as care-oriented diagnostics without adequate diagnosis of the underlying disease is undesirable. Integration rather than polarization is important, because greater integration will lead to greater continuity of care for patients with dementia. To increase this continuity even further, DOC-PG uses a tailor based approach. The DOC-PG provides the patients with individual advice for treatment and management, customized to their needs. Furthermore, the DOC-PG is a functional service which is indicated by the involvement of disciplines such as occupational therapy and physical therapy. DOC-PG is aimed at maximizing the empowerment of the patients and herewith enabling them to live independently as long as possible.

### Multidisciplinary organisation models

There is a growing interest in multidisciplinary organisation models in health care. Furthermore, current policy promotes the integration of health and social care. These recommendations are mainly practice based instead of evidence based. The MEDICIE-study provides the evidence based scientific proof that the subjective and objective Health Related Quality of Life can be improved through a comprehensive assessment by the DOC-PG. This could be an important impulse to use this DOC-PG model in other places in the Netherlands. It is likely that the findings of this study are cited in national or international clinical guidelines.

Currently, approximately half of the Memory Clinics in the Netherlands structurally collaborate with community mental health teams such as the CMHT (Verhey et al., 2005). We expect this number to increase over the next years. This expectation is line with the policy applied by the government i.e. guideline development (and its implementation activities) and the "National Dementia Program" (LDP). The LDP provides for the development of integrated care from a client-focussed principle.

Although a multidisciplinary model is more effective than single disciplines, it is also more complex. Additionally, it requires a higher level of organisation. It is therefore a challenge for clinicians to combine their professional expertise and share responsibility for a patient given their different, and sometimes opposing, approaches and views on patient care and management. However, combining professional expertise will result in a continuous learning process for the participating parties.

The DOC-PG already has blossomed into a firm and close team as indicated by the personal development of the single disciplines as well as the satisfactory mutual teamwork.

**Health economics in dementia**

The far-reaching implications for health service delivery and expenditures due to the increasing number of older adults who are diagnosed with dementia has led to a growing interest in analyzing the cost-effectiveness of new interventions. Clinical outcomes must therefore be combined with economic outcomes. Health utility measures such as QALYs are commonly used in economic evaluations. However, more research on these measures in dementia is necessary to ensure their validity (Katona et al., 2007). While HRQoL measures have certain strengths such as the recognition of the differing values placed on health states by different individuals and groups of individuals, the encompassment of a person's needs and allowing for the comparison of different outcomes in different patient groups, they also have serious challenges (Rabins et al., 2007). An important challenge to using HRQoL as an outcome in dementia, is its subjective nature. It remains unclear who is most competent to rate a patient's HRQoL: the patient, the caregiver or the physician. Furthermore, the stability of an individual's construct of HRQoL might change over the course of a progressive illness such as dementia (Rabins et al., 2007). More research is necessary to shed light on these issues.

As there is no cure or adequate treatment for dementia, it is also possible to adapt a cost-consequence approach. A cost-consequence analysis is an analysis in which costs and effects are calculated but not aggregated into QALYs or cost-effectiveness ratios (Mauskopf et al., 1998). It is a listing of all relevant costs (health care costs and costs outside the health care sector), outcomes (such as clinical outcomes, patient satisfaction) and consequences (quality of life impacts and utility impacts). Such a listing makes the impact of a new intervention as comprehensive and transparent as possible and enables decision-makers to make adequate resource allocation decisions (Mauskopf et al., 1998). However, in a cost-consequence analysis it is not possible to compare different interventions and efficiency can not be demonstrated. Taking this into account, when studying new interventions in dementia care, the QALY is an appropriate measure. In the MEDICIE-study, we evaluated DOC-PG by using various clinical outcomes, caregiver outcomes, GP satisfaction and concordance as well as QALYs and costs. Herewith, a complete picture of the new intervention is available and, by including a cost-effectiveness analysis, a better understanding of the economics of dementia is provided.

Another difficulty in economic evaluations in dementia concerns the inclusion of informal care. As there is no cure for dementia, and the majority of service provision falls on care rather than treatment, there is a large demand on informal care. It is however unclear what constitutes informal care and how to measure the input of informal caregivers. How many hours of informal care are actually delivered and can informal care be distinguished from the usual interaction of a marriage or family life? There is a lack of consensus and uniformity of the methodology used in evaluating informal care. Obviously, this can lead to problems in the interpretation of results and in comparing different studies (van den Berg et al., 2004).

Currently, the caregiver activities time survey (CATS) (Clipp et al., 1995), the caregiver activity survey (CAS) (Davis et al., 1997), and the Resource Utilization in Dementia (RUD) tool (Wimo et al., 2000) are widely used instruments to measure informal care.

However, the validation of these instruments is unsatisfactory both regarding the measurement (time-factor) in general and the supervision-factor in particular since supervision is rather difficult to quantify (Wimo et al., 2007). It can be argued whether time should be used as the relevant aspect to value informal care. Other approaches, such as willingness-to-pay or cost-benefit analyses (see chapter one) might be interesting alternatives. To date, these approaches have been rarely studied in dementia (Knapp, 2007; Wimo et al., 2007).

Although dementia care is without doubt of the utmost importance, money issues are always present in policy and practice discussions in health care (Knapp, 2007). Scarcity is a permanent and persistent part of health care and policy makers have to make difficult choices. However, they should not base their resource allocation decisions solely on economics, but they need to be concerned with all other relevant outcomes regarding dementia care. This underlines the need for broad assessment and evaluation.

Overall, the positive results of the MEDICIE-study will be relevant for health care professionals, GPs, policy makers and of course the patients. Implementing a cost-effective integrated diagnostic facility will lead to benefits for the entire health care organisation by improving quality of life of the patients, facilitating their management and controlling the costs. This is important considering the broad impact of dementia in the next decades. The results of this study support the trend towards multidisciplinary organisation models in health care.

## **Future directions**

### **Directions for future clinical research**

Although an integrated approach to dementia by means of a DOC-PG can be considered effective, more research is necessary to determine which aspect or which combination of aspects involved with the integrated facility is likely to result in the best patient outcome. This can be achieved by comparing different multidisciplinary teams who are active in the Netherlands. An observational study using standardized measures seems most appropriate for this purpose and follow-up of dementia patients by means of a standardized protocol (including an economic evaluation) is desirable.

### **Directions for future HTA research**

More research is necessary to study the relationship between patient rating of HRQoL and proxy rating of HRQoL. It needs to be (further) verified who is most competent to rate a patient's HRQoL: the patient, the caregiver or the physician. International guidelines should be developed for the measurement and valuation of informal care. First a clear definition of the concept should be formulated. Then uniformity on the measurement and valuation needs to be established for the comparability of international economic evaluations.

### **Directions for clinical practice**

In addition to the broad evaluation we applied in this study, i.e. studying the effects on multiple levels, it might also be worthwhile to study process parameters such as patient and caregiver satisfaction, efficiency, accessibility and swiftness of service delivery. More research regarding a collaboration with (local) GPs is important to gain more insight in their needs and imperfections. Recently, the AD-Euro study in Nijmegen, the Netherlands has commenced in close collaboration with Maastricht, concerning a comparison between the care provided by the GP and care provided by the Memory Clinic after a diagnosis has been made. This study is an interesting continuation of the MEDICIE-study which will give us more insight in the behaviour of GPs regarding the management of patients with dementia. This insight is important since in the Netherlands, the GP fulfils the role of gatekeeper in the health care system. Implementing a new service is therefore largely dependent on the co-operation with these GPs. A first challenge in order to implement a new diagnostic facility such as the DOC-PG will be to change the view on patient care and management of the GPs.

### **Conclusions**

Although an integrated approach to the diagnosis and management is generally recommended, this has been scarcely empirically investigated. To our knowledge, the MEDICIE-study is the first study concerning an integrated approach to dementia and dementia care with a randomised design and a comprehensive health economic evaluation. GPs already recognize the added value of such an approach to dementia. DOC-PG can be considered an effective as well as a cost-effective facility for the diagnosis and management of dementia. In light of our ageing population with increasingly more patients suffering from dementia, appropriate multidisciplinary organisational models are necessary to ensure the quality of care while controlling the costs.

More research is essential to determine which aspect or which combination of aspects involved with the integrated facility is likely to result in the best Quality of Life. We further recommend to assess patients diagnostically by means of a tailor based approach. It is possible to adjust an integrated multidisciplinary model for each patient. DOC-PG is already successful regarding this tailor based approach as indicated by the versatility of the recommendations formulated by the team. In order to implement an integrated model for dementia, the cooperation between GPs and health and social care services involved with dementia should be (further) promoted. Especially, a close collaboration with the local GPs is a prerequisite.

# Samenvatting



Het doel van dit onderzoek is het evalueren van een integrale multidisciplinaire aanpak voor het diagnosticeren en managen van dementie door het Diagnostisch OnderzoeksCentrum voor PsychoGeriatrische patiënten (DOC-PG). Het DOC-PG combineert de ziektegerichte diagnostische aanpak van een aantal ziekenhuisafdelingen met de zorggerichte aanpak van de RIAGG in Maastricht. Een dergelijke benadering bij dementie werd nog nooit op wetenschappelijk verantwoorde wijze geëvalueerd.

In de inleiding (hoofdstuk 1) van dit proefschrift wordt de relevantie en aanleiding van het onderzoek beschreven. Het aantal ouderen in de populatie groeit snel en zal de komende tijd ook sterk blijven toenemen. Daarbij neemt het aantal ouderen met gecombineerde somatische en psychologische aandoeningen toe en hiermee ook de noodzaak om geïntegreerde multidisciplinaire gezondheidszorgvoorzieningen te creëren en evalueren. Verder wordt in dit hoofdstuk aangegeven dat nieuwe gezondheidszorgvoorzieningen uitgebreid geëvalueerd dienen te worden, zowel klinisch als economisch.

### **Deel 1**

De twee hoofdvraagstellingen van dit proefschrift zijn dan ook:

- 1) *Is het DOC-PG een effectieve voorziening vergeleken met gebruikelijke zorg?*
- 2) *Is het DOC-PG een kosteneffectieve voorziening vergeleken met gebruikelijke zorg?*

Deze vraagstellingen werden onderzocht in de MEDICIE-studie (Maastricht Evaluation of a Diagnostic Intervention for Cognitively Impaired Elderly). Een beschrijving van de studie wordt gegeven in hoofdstuk 2.

In hoofdstuk 3 wordt een overzicht gegeven van alle gepubliceerde empirische studies wat betreft een integrale multidisciplinaire aanpak bij dementie. Hoewel een dergelijke aanpak wordt geadviseerd, is wetenschappelijk bewijs schaars. Dit benadrukt het belang van de MEDICIE-studie: het is de 1e gerandomiseerde studie naar de effectiviteit en kosteneffectiviteit van een geïntegreerde benadering bij dementie.

### **Deel 2: Methodologische aspecten**

Het doel van het exploratieve hoofdstuk 4 is om de construct validiteit en de responsiviteit van de EQ-5D en de EQ-5D+C (dit is de EQ-5D uitgebreid met een cognitieve dimensie) te vergelijken. Wat betreft construct validiteit vonden we vergelijkbare correlaties tussen de EQ-5D en de Mini Mental State Examination (MMSE) en tussen de EQ-5D+C en de MMSE. Verder waren zowel de EQ-5D als de EQ-5D+C gevoelig voor veranderingen in de MMSE. De EQ-5D presteerde echter iets beter. Concluderend kunnen we zeggen dat de EQ-5D prima presteert m.b.t. het evalueren van gezondheidgerelateerde kwaliteit van leven in een populatie met cognitieve beperkingen

Hoofdstuk 5 beschrijft de resultaten van 3 methodes voor het meten van informele zorg namelijk de proxy-good-methode, de opportuniteitskosten-methode (beide monetaire methodes) en de CarerQoI-methode (niet monetaire methode). De opportuniteitskosten methode werd verder opgesplitst in de opportuniteitskosten-A-methode (namelijk verloren uren betaald werk, onbetaald werk of vrije tijd), en de opportuniteitskosten-B-methode (namelijk geprefereerde uren betaald werk, onbetaald werk of vrije tijd). Een

literatuuronderzoek werd gedaan om na te gaan welke factoren de mate van informele zorg bij dementia beïnvloeden. Verder werd voor elk van deze methodes zowel de construct validiteit als de responsiviteit onderzocht. Gebaseerd op de resultaten ten aanzien van zowel de construct validiteit en responsiviteit is de proxy-good methode de beste methode voor het meten en waarderen van informele zorg in patiënten met dementie of een andere cognitieve stoornis en hun mantelzorgers.

### **Deel 3: Bevindingen van de MEDICIE-studie**

De effectiviteit en kosteneffectiviteit van DOC-PG vergeleken met gebruikelijke zorg worden beschreven in de hoofdstukken 6 en 7.

Gezondheidsgerelateerde kwaliteit van leven van de patiënt, was de primaire uitkomstmaat om de effectiviteit van DOC-PG te evalueren (hoofdstuk 6). De EQ-5D werd gebruikt om gezondheidsgerelateerde kwaliteit van leven te meten. Veranderingen in de Visual Analogue Scale (VAS) scores van de EQ-5D gedurende 1 jaar was de primaire uitkomst. Een verschil van 10% of meer tussen de interventiegroep en de gebruikelijke zorg groep op de VAS werd beschouwd als een klinisch relevant verschil. Het aantal patiënten die dit klinisch relevante verschil ervaarden werd vergeleken tussen beide groepen. We vonden een gemiddeld groepsverschil van 9.6% op de VAS van de EQ-5D na 12 maanden, en dit komt zeer dicht in de buurt van onze initiële verwachtingen. De proportie patiënten die meer dan 10% (van het groepsverschil) verbeterden op de VAS was significant groter in de interventiegroep (39.0%) dan in de gebruikelijke zorg groep (22.1%). De groepen verschilden niet van elkaar wat betreft de klinische uitkomstmaten.

Naast de gerandomiseerde klinische trial (RCT) werd een economische evaluatie uitgevoerd. Hierin werden de kosten en effecten van DOC-PG vergeleken met die van de gebruikelijke zorg (hoofdstuk 7). De economische evaluatie werd uitgevoerd vanuit een maatschappelijk perspectief volgens (inter)nationale richtlijnen met een 1-jaar durende tijdshorizon. Een incrementele kosteneffectiviteitsratio (IKER) werd berekend door het verschil in kosten te delen door het verschil in Quality Adjusted Life Years (voor kwaliteit gecorrigeerde levensjaren of QALYs). Een non-parametrische bootstrap-analyse en een-weg gevoeligheidsanalyses werden uitgevoerd om de onzekerheid in kosten en effecten te bepalen. De EQ-5D werd gebruikt om gezondheidsgerelateerde kwaliteit van leven te meten en de proxy-good methode werd gebruikt om informele zorg te meten en te waarderen. Vergeleken met de patiënten die gebruikelijke zorg ontvingen, wonnen patiënten die DOC-PG bezochten gemiddeld 0.05 QALYs. De incrementele kosten per QALY bedroegen € 1267. DOC-PG is niet aantoonbaar duurder dan gebruikelijke zorg in Nederland, ondanks de investering in de diagnostische interventie, maar heeft een grote waarschijnlijkheid effectiever te zijn. DOC-PG kan dan ook beschouwd worden als een kosteneffectieve voorziening voor de diagnose en management van dementie in ambulante patiënten.

Hoofdstuk 8 beschrijft de verschillen tussen de interventiegroep en de gebruikelijke zorg groep na een uitgebreid DOC-PG onderzoek. Een vragenlijst, ontwikkeld door het iMTA in Rotterdam in Nederland, werd gebruikt om informele zorg te meten en te waarderen,

om een aantal karakteristieken van de mantelzorger te beschrijven (relatie met de patiënt, educatie, inkomen, gezondheid, gezondheidgerelateerde kwaliteit van leven en belasting) en om de sociale context te beschrijven (mantelzorg netwerk, formele zorg, dagelijkse activiteiten van de patiënt (bijv. kaarten) ter ontlasting van de mantelzorger, respijtzorg en betaald zorgverlof). We vonden dat significant meer mensen in de interventiegroep gebruik maakten van informele dagopvang (dagopvang door familie of vrienden) op de baseline meting. Een ander verschil had betrekking op de kosten van informele zorg, gemeten en gewaardeerd met behulp van de proxy-good methode. Deze waren hoger in de gebruikelijke zorg groep dan in de interventiegroep na 12 maanden. Tenslotte gaven meer mantelzorgers in de interventiegroep onbetaald werk op om voor de patiënten te kunnen zorgen tijdens de baseline meting. Geen andere significante verschillen werden gevonden tussen beide groepen.

Hoewel maar weinig verschillen tussen beide groepen werden gevonden, suggereren de resultaten dat DOC-PG wellicht een positief effect op de hoeveelheid informele zorg heeft aangezien deze hoeveelheid na 12 maanden veel meer gestegen was in de gebruikelijke zorg groep (gemiddelde stijging van 61%) dan in de interventiegroep (gemiddelde stijging van 12%).

De waarde van het DOC-PG wordt ook al herkend door de huisartsen wat geïndiceerd wordt door het grote aantal verwijzingen naar het DOC-PG en de hoge mate van concordantie met DOC-PG aanbevelingen (hoofdstuk 9). Acceptatie van DOC-PG door de directe zorgverstrekkers was de eerste barrière die overwonnen moest worden.

Tenslotte wordt in hoofdstuk 10 een algehele discussie over de resultaten van de MEDICIE-studie gegeven. Methodologische punten (m.b.t. sterktes van de studie, design van de studie en dataverzameling), implicaties voor de praktijk (multidisciplinaire modellen en gezondheidseconomie bij dementie) en aanbevelingen voor toekomstig onderzoek (bij klinisch onderzoek, bij HTA onderzoek en in de klinische praktijk) komen in dit hoofdstuk aan bod.

Concluderend kan gesteld worden dat een geïntegreerde multidisciplinaire aanpak bij dementie effectiever is wat betreft kwaliteit van leven van de patiënten dan de gebruikelijke zorg. DOC-PG kan verder beschouwd worden als een kosteneffectieve voorziening voor de diagnose en management van dementie in ambulante patiënten.

## List of abbreviations





<b>AD</b>	Alzheimer's Disease
<b>ADL</b>	Activities of Daily Living
<b>CEA</b>	Cost-effectiveness analysis
<b>CEAC</b>	Cost-effectiveness acceptability curve
<b>CMHT</b>	Community Mental Health Team
<b>CPN</b>	Community Psychiatric Nurse
<b>CSDD</b>	Cornell Scale of Depression in Dementia
<b>CUA</b>	Cost-utility analysis
<b>DOC-PG</b>	Diagnostic Observation Centre for PsychoGeriatric patients
<b>DSM</b>	Diagnostic and Statistical Manual
<b>EDIZ</b>	Ervaren Druk door Informele Zorg
<b>EQ-5D</b>	EuroQoL-5D
<b>EQ-5D+C</b>	EuroQoL-5D + Cognition
<b>GDS</b>	Global Deterioration Scale
<b>GP</b>	General Practitioner
<b>HDL</b>	Household Activities of Daily Living
<b>HRQoL</b>	Health Related Quality of Life
<b>HTA</b>	Health Technology Assessment
<b>IADL</b>	Instrumental Activities of Daily Living
<b>ICER</b>	Incremental Cost-Effectiveness Ratio
<b>MC</b>	Memory Clinic
<b>MEDICIE</b>	Maastricht Evaluation of a Diagnostic Intervention for Cognitively Impaired Elderly
<b>MMSE</b>	Mini Mental State Examination
<b>MT</b>	Multidisciplinary Team
<b>MTA</b>	Medical Technology Assessment
<b>NPI</b>	NeuroPsychiatric Inventory
<b>PTO</b>	Person Trade-Off
<b>QALY</b>	Quality Adjusted Life Year
<b>QoL</b>	Quality of Life
<b>SD</b>	Standard Deviation
<b>SF-36</b>	Short Form 36
<b>SF-6D</b>	Short Form 6 dimensions
<b>SG</b>	Standard Gamble
<b>SPSS</b>	Statistical Package for the Social Sciences
<b>SRB</b>	Self-Rated Burden
<b>TTO</b>	Time Trade-Off
<b>VAS</b>	Visual Analogue Scale
<b>VD</b>	Vascular Dementia



# References



## REFERENCES

- Aalten, P., de Vugt, M. E., Jaspers, N., Jolles, J., & Verhey, F. R. (2005). The course of neuropsychiatric symptoms in dementia. Part I: findings from the two-year longitudinal Maasbed study. *Int J Geriatr Psychiatry*, 20(6), 523-530.
- AAN. (1994). Practice parameter for diagnosis and evaluation of dementia (summary statement). Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology*, 44(11), 2203-2206.
- Albert, S. M., Del Castillo-Castaneda, C., Sano, M., Jacobs, D. M., Marder, K., Bell, K., Bylsma, F., Lafleche, G., Brandt, J., Albert, M., & Stern, Y. (1996). Quality of life in patients with Alzheimer's disease as reported by patient proxies. *Journal of the American Geriatrics Society*, 44(11), 1342-1347.
- Albert, S. M., Sano, M., Bell, K., Merchant, C., Small, S., & Stern, Y. (1998). Hourly care received by people with Alzheimer's disease: results from an urban, community survey. *Gerontologist*, 38(6), 704-714.
- Alexopoulos, G. S., Abrams, R. C., Young, R. C., & Shamoian, C. A. (1988). Cornell Scale for Depression in Dementia. *Biological Psychiatry*, 23(3), 271-284.
- Allen, C. M., Becker, P. M., McVey, L. J., Saltz, C., Feussner, J. R., & Cohen, H. J. (1986). A randomized, controlled clinical trial of a geriatric consultation team. Compliance with recommendations. *Jama*, 255(19), 2617-2621.
- Anderson, C., Mhurchu, C. N., Rubenach, S., Clark, M., Spencer, C., & Winsor, A. (2000). Home or hospital for stroke Rehabilitation? Results of a randomized controlled trial : II: cost minimization analysis at 6 months. *Stroke*, 31(5), 1032-1037.
- Ankri, J., Beaufils, B., Novella, J. L., Morrone, I., Guillemin, F., Jolly, D., Ploton, L., & Blanchard, F. (2003). Use of the EQ-5D among patients suffering from dementia. *Journal of Clinical Epidemiology*, 56(11), 1055-1063.
- APA. (1980). *Diagnostic and Statistical Manual of Mental Disorders (3rd edition)*. Washington DC: American Psychiatric Association.
- APA. (1987). *Diagnostic and Statistical Manual of Mental Disorders (3rd edition revised)*. Washington DC: American Psychiatric Association.
- APA. (1994). *Diagnostic and Statistical Manual of Mental Disorders (fourth edition)*. Washington DC: American Psychiatric Association.
- APA. (1997). Practice guideline for the treatment of patients with Alzheimer's disease and other dementias of late life. American Psychiatric Association. *American Journal of Psychiatry*, 154(5 Suppl), 1-39.
- Applegate, W. B., Miller, S. T., Graney, M. J., Elam, J. T., Burns, R., & Akins, D. E. (1990). A randomized controlled trial of a geriatric assessment unit in a community rehabilitation hospital. *New England Journal of Medicine* 322(22), 1572-1578.
- Baillon, S., Muhommad, S., Marudkar, M., Suribhatla, S., Dennis, M., Spreadbury, C., Munro, D., & Lindesay, J. (2003). Neuropsychological performance in Alzheimer's disease and vascular dementia: comparisons in a memory clinic population. *Int J Geriatr Psychiatry*, 18(7), 602-608.
- Banerjee, S., Smith, S. C., Lamping, D. L., Harwood, R. H., Foley, B., Smith, P., Murray, J., Prince, M., Levin, E., Mann, A., & Knapp, M. (2006). Quality of life in dementia: more than just cognition. An analysis of associations with quality of life in dementia. *Journal of Neurology, Neurosurgery and Psychiatry*, 77(2), 146-148.
- Bayer, A. J., Pathy, M. S., & Twining, C. (1987). The memory clinic. A new approach to the detection of early dementia. *Drugs*, 33 Suppl 2, 84-89.
- Black, W. (1990). The CE plane: A graphic representation of cost-effectiveness. *Medical Decision Making*, 10, 212-214.
- Blazer, D. G., & Williams, C. D. (1980). Epidemiology of dysphoria and depression in an elderly population. *American Journal of Psychiatry*, 137, 439-444.
- Boer de, A., Schellingerhout, R., & Timmermans, J. M. (2003). *Mantelzorg in getallen (Informal care in numbers)*. The Hague: Sociaal en Cultureel Planbureau (Social and Cultural Planning Office of the Netherlands)
- Brazier, J., Deverill, M., Green, C., Harper, R., & Booth, A. (1999). A review of the use of health status measures in economic evaluation. *Health Technology Assessment*, 3(9), i-iv, 1-164.
- Brazier, J., Roberts, J., & Deverill, M. (2002a). The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics*, 21(2), 271-292.
- Brazier, J., Roberts, J., & Deverill, M. (2002b). The estimation of a preference-based measure of health from the SF-36. *J Health Econ*, 21(2), 271-292.
- Brazier, J., Roberts, J., Tsuchiya, A., & Busschbach, J. (2004). A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Economics*, 13(9), 873-884.
- Brazier, J. E., Harper, R., Jones, N. M., O'Cathain, A., Thomas, K. J., Usherwood, T., & Westlake, L. (1992). Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *British Medical Journal*, 305(6846), 160-164.

- Briggs, A. H., Wonderling, D. E., & Mooney, C. Z. (1997). Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. *Health Economics*, 6, 327-340.
- Brodsky, H. (2007). Meaning and measurement of caregiver outcomes. *Int Psychogeriatr*, 19(3), 363-381.
- Brooks, R. (1996). EuroQol: the current state of play. *Health Policy*, 37(1), 53-72.
- Brouwer, W. B., van Exel, N. J., Koopmanschap, M. A., & Rutten, F. F. (1999). The valuation of informal care in economic appraisal. A consideration of individual choice and societal costs of time. *Int J Technol Assess Health Care*, 15(1), 147-160.
- Brouwer, W. B., van Exel, N. J., van de Berg, B., Dinant, H. J., Koopmanschap, M. A., & van den Bos, G. A. (2004). Burden of caregiving: evidence of objective burden, subjective burden, and quality of life impacts on informal caregivers of patients with rheumatoid arthritis. *Arthritis & Rheumatism*, 51(4), 570-577.
- Brouwer, W. B., van Exel, N. J., van Gorp, B., & Redekop, W. K. (2006). The CarerQol instrument: a new instrument to measure care-related quality of life of informal caregivers for use in economic evaluations. *Qual Life Res*, 15(6), 1005-1021.
- Brown, L., Tucker, C., & Domokos, T. (2003). Evaluating the impact of integrated health and social care teams on older people living in the community. *Health & Social Care in the Community*, 11(2), 85-94.
- Bryan, S., Hardyman, W., Bentham, P., Buckley, A., & Laight, A. (2005). Proxy completion of EQ-5D in patients with dementia. *Quality of Life Research*, 14(1), 107-118.
- Busschbach, J. J., Brouwer, W. B., van der Donk, A., Passchier, J., & Rutten, F. F. (1998). An outline for a cost-effectiveness analysis of a drug for patients with Alzheimer's disease. *Pharmacoeconomics*, 13(1 Pt 1), 21-34.
- CBO. (2005). *Richtlijn diagnostiek en medicamenteuze behandeling van dementie (Guideline diagnostics and medicinal treatment of dementia)*. Utrecht: CBO/MWR.
- Centraal Bureau voor Statistiek: *Statistisch jaarboek*. (2000). Voorburg/Heerlen: CBS.
- Chiu, L., Tang, K. Y., Liu, Y. H., Shyu, W. C., & Chang, T. P. (1999). Cost comparisons between family-based care and nursing home care for dementia. *J Adv Nurs*, 29(4), 1005-1012.
- Clipp, E. C., & Moore, M. J. (1995). Caregiver time use: an outcome measure in clinical trial research on Alzheimer's disease. *Clin Pharmacol Ther*, 58(2), 228-236.
- Cohen, R. A., Moser, D. J., Clark, M. M., Aloia, M. S., Cargill, B. R., Stefanik, S., Albrecht, A., Tilkemeier, P., & Forman, D. E. (1999). Neurocognitive functioning and improvement in quality of life following participation in cardiac rehabilitation. *American Journal of Cardiology*, 83(9), 1374-1378.
- Cole, M. G., Fenton, F. R., Engelsmann, F., & Mansouri, I. (1991). Effectiveness of geriatric psychiatry consultation in an acute care hospital: a randomized clinical trial. *Journal of the American Geriatrics Society* 39(12), 1183-1188.
- Collighan, G., Macdonald, A., Herzberg, J., Philpot, M., & Lindesay, J. (1993). An evaluation of the multidisciplinary approach to psychiatric diagnosis in elderly people. *British Medical Journal*, 306(6881), 821-824.
- Cummings, J. L. (1997). The Neuropsychiatric Inventory: assessing psychopathology in dementia patients. *Neurology*, 48(5 Suppl 6), S10-16.
- Cummings, J. L., & McPherson, S. (2001). Neuropsychiatric assessment of Alzheimer's disease and related dementias. *Aging (Milano)*, 13(3), 240-246.
- Cummings, J. L., Mega, M., Gray, K., Rosenberg-Thompson, S., Carusi, D. A., & Gornbein, J. (1994). The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology*, 44(12), 2308-2314.
- Davis, K. L., Marin, D. B., Kane, R., Patrick, D., Peskind, E. R., Raskind, M. A., & Puder, K. L. (1997). The Caregiver Activity Survey (CAS): development and validation of a new measure for caregivers of persons with Alzheimer's disease. *Int J Geriatr Psychiatry*, 12(10), 978-988.
- De Bruyne, G. M., Meyboom-de Jong, B., Muskens, J. B., Veltman, M. T. M., Weijtens, J. T. N. M., & Wind, A. W. (1991). De NHG-standaard dementiesyndroom. (Guideline on dementia of the Dutch College of General Practitioners.). *Huisarts & Wetenschap*, 34 (13), 598-617.
- de Vugt, M. E., Nicolson, N. A., Aalten, P., Lousberg, R., Jolle, J., & Verhey, F. R. (2005a). Behavioral problems in dementia patients and salivary cortisol patterns in caregivers. *J Neuropsychiatry Clin Neurosci*, 17(2), 201-207.
- de Vugt, M. E., Stevens, F., Aalten, P., Lousberg, R., Jaspers, N., & Verhey, F. R. (2005b). A prospective study of the effects of behavioral symptoms on the institutionalization of patients with dementia. *Int Psychogeriatr*, 17(4), 577-589.
- de Vugt, M. E., Stevens, F., Aalten, P., Lousberg, R., Jaspers, N., Winkens, I., Jolles, J., & Verhey, F. R. (2004). Do caregiver management strategies influence patient behaviour in dementia? *Int J Geriatr Psychiatry*, 19(1), 85-92.

## REFERENCES

- Degl'Innocenti, A., Elmfeldt, D., Hansson, L., Breteler, M., James, O., Lithell, H., Olofsson, B., Skoog, I., Trenkwalder, P., Zanchetti, A., & Wiklund, I. (2002). Cognitive function and health-related quality of life in elderly patients with hypertension—baseline data from the study on cognition and prognosis in the elderly (SCOPE). *Blood Press*, *11*(3), 157-165.
- Dennis, M., Furness, L., Lindsay, J., & Wright, N. (1998). Assessment of patients with memory problems using a nurse-administered instrument to detect early dementia and dementia subtypes. *Int J Geriatr Psychiatry*, *13*(6), 405-409.
- Dolan, P. (1996). The effect of experience of illness on health state valuations. *Journal of Clinical Epidemiology*, *49*(5), 551-564.
- Dolan, P. (1997). Modeling valuations for EuroQol health states. *Medical Care*, *35*(11), 1095-1108.
- Drummond, M., O'Brien, B., Stoddart, G., & Torrance, G. (1997). *Methods for the economic evaluation of health care programmes* (2nd ed.). New York: Oxford University Press.
- Dunkin, J. J., & Anderson-Hanley, C. (1998). Dementia caregiver burden: a review of the literature and guidelines for assessment and intervention. *Neurology*, *51*(1 Suppl 1), S53-60; discussion S65-57.
- EuroQol—a new facility for the measurement of health-related quality of life. The EuroQol Group. (1990). *Health Policy*, *16*(3), 199-208.
- Evers, S. M. A. A., Ament, A. J. A. H., & Blaauw, G. (2000). Economic evaluation in stroke research: a systematic review. *Stroke*, *31*, 1046-1053.
- Evers, S. M. A. A., van Wijk, A. S., & Ament, A. J. A. H. (1997). Economic evaluation of mental health care interventions: a review. *Health Economics*, *6*, 161-177.
- Feeny, D., Furlong, W., Boyle, M., & Torrance, G. W. (1995). Multi-attribute health status classification systems. Health Utilities Index. *Pharmacoeconomics*, *7*(6), 490-502.
- Feldman, H. H., Van Baelen, B., Kavanagh, S. M., & Torfs, K. E. (2005). Cognition, function, and caregiving time patterns in patients with mild-to-moderate Alzheimer disease: a 12-month analysis. *Alzheimer Dis Assoc Disord*, *19*(1), 29-36.
- Feron, J. M., Cerexhe, F., Pestiaux, D., Roland, M., Giet, D., Montrieux, C., & Paulus, D. (2003). GPs working in solo practice: obstacles and motivations for working in a group? A qualitative study. *Family Practice*, *20*(2), 167-172.
- Ferran, J., Wilson, K., Doran, M., Ghadiali, E., Johnson, F., Cooper, P., & McCracken, C. (1996). The early onset dementias: a study of clinical characteristics and service use. *International Journal of Geriatric Psychiatry*, *11*, 863-869.
- Fillit, H., & Hill, J. (2005). Economics of dementia and pharmacoeconomics of dementia therapy. *The American Journal of Geriatric Pharmacotherapy*, *3*(1), 39-49.
- Folstein, M., Folstein, S., & McHugh, P. (1975a). "Mini Mental State" a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 196-198.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975b). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189-198.
- Girling, D. M., Huppert, F. A., Brayne, C., Paykel, E. S., Gill, C., & Mathewson, D. (1995). Depressive symptoms in the very elderly—their prevalence and significance. *International Journal of Geriatric Psychiatry*, *10*(6), 497-504.
- Gold, M. R., Siegel, J. E., Russell, L. B., & Weinstein, M. C. (1996). *Cost-effectiveness in health and medicine*. New York: Oxford University Press.
- Goossens, M. E. J. B., & Evers, S. M. A. A. (1997). Economic evaluation of back pain interventions. *Journal of Occupational Rehabilitation* *7*, 15-32.
- Goossens, M. E. J. B., & Evers, S. M. A. A. (2000). Cost-effectiveness of treatment for neck and low back pain. In A. L. Nachemson & E. Jonsson (Eds.), *Neck and Back Pain: The Scientific Evidence of Causes, Diagnosis and Treatment*. Philadelphia: Williams & Wilkins.
- Green, C. (2001). On the societal value of health care: what do we know about the person trade-off technique? . *Health Economics* *10*, 233-243.
- Harrow, B. S., Mahoney, D. F., Mendelsohn, A. B., Ory, M. G., Coon, D. W., Belle, S. H., & Nichols, L. O. (2004). Variation in cost of informal caregiving and formal-service use for people with Alzheimer's disease. *Am J Alzheimers Dis Other Dement*, *19*(5), 299-308.
- Haywood, K. L., Garratt, A. M., Dziedzic, K., & Dawes, P. T. (2002). Generic measures of health-related quality of life in ankylosing spondylitis: reliability, validity and responsiveness. *Rheumatology (Oxford)*, *41*(12), 1380-1387.
- Hejl, A., Høgh, P., & Waldemar, G. (2002). Potentially reversible conditions in 1000 consecutive memory clinic patients. *J Neurol Neurosurg Psychiatry*, *73*(4), 390-394.

- Hoeymans, N., van Lindert, H., & Westert, G. P. (2005). The health status of the Dutch population as assessed by the EQ-6D. *Quality of Life Research*, 14(3), 655-663.
- Hogh, P., Waldemar, G., Knudsen, G. M., Bruhn, P., Mortensen, H., Wildschiodtz, G., Bech, R. A., Juhler, M., & Paulson, O. B. (1999). A multidisciplinary memory clinic in a neurological setting: diagnostic evaluation of 400 consecutive patients. *European Journal of Neurology*, 6(3), 279-288.
- Holland, R., Smith, R. D., Harvey, I., Swift, L., & Lenaghan, E. (2004a). Assessing quality of life in the elderly: a direct comparison of the EQ-5D and AQoL. *Health Economics*, 13(8), 793-805.
- Holland, R., Smith, R. D., Harvey, I., Swift, L., & Lenaghan, E. (2004b). Assessing quality of life in the elderly: a direct comparison of the EQ-5D and AQoL. *Health Econ*, 13(8), 793-805.
- Hout, H. v. (1999). *Diagnosis dementia. Evaluation of the clinical practice of general practitioners and a memory clinic*. Catholic University Nijmegen, Nijmegen.
- Husted, J. A., Cook, R. J., Farewell, V. T., & Gladman, D. D. (2000). Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol*, 53(5), 459-468.
- Jonsson, L. (2003). Pharmacoeconomics of cholinesterase inhibitors in the treatment of Alzheimer's disease. *Pharmacoeconomics*, 21(14), 1025-1037.
- Jonsson, L. (2004). Economic evidence in dementia: a review. *European Journal of Health Economics*, 5 Suppl 1, S30-35.
- Jonsson, L., Andreassen, N., Kilander, L., Soininen, H., Waldemar, G., Nygaard, H., Winblad, B., Jonhagen, M. E., Hallikainen, M., & Wimo, A. (2006a). Patient- and proxy-reported utility in Alzheimer disease using the EuroQoL. *Alzheimer Disease and Associated Disorders*, 20(1), 49-55.
- Jonsson, L., Eriksdotter Jonhagen, M., Kilander, L., Soininen, H., Hallikainen, M., Waldemar, G., Nygaard, H., Andreassen, N., Winblad, B., & Wimo, A. (2006b). Determinants of costs of care for patients with Alzheimer's disease. *International Journal of Geriatric Psychiatry*, 21(5), 449-459.
- Kat, M. G., de Jonghe, J. F. M., Aalten, P., Kalisvaart, C. J., Droës, R. M., & Verhey, F. R. J. (2002). Neuropsychiatric symptoms of dementia: psychometric aspects of the Dutch version of the Neuropsychiatric Inventory. *Tijdschrift voor Gerontologie en Geriatrie*, 33(4), 150-155.
- Katona, C., Livingston, G., Cooper, C., Ames, D., Brodaty, H., & Chiu, E. (2007). International Psychogeriatric Association consensus statement on defining and measuring treatment benefits in dementia. *Int Psychogeriatr*, 19(3), 345-354.
- Kavanagh, S., & Knapp, M. (2002). Costs and cognitive disability: modelling the underlying associations. *British Journal of Psychiatry*, 180, 120-125.
- Kemper, P. (1992). The use of formal and informal home care by the disabled elderly. *Health Serv Res*, 27(4), 421-451.
- Kim, S. W., Kim, J. M., Stewart, R., Bae, K. L., Yang, S. J., Shin, I. S., Shin, H. Y., & Yoon, J. S. (2006). Correlates of caregiver burden for Korean elders according to cognitive and functional status. *Int J Geriatr Psychiatry*, 21(9), 853-861.
- Kind, P. (2003). Guidelines for value sets in economic and on-economic studies using EQ-5D. In R. Brooks, R. Rabin & F. d. Charro (Eds.), *The measurement and valuation of health status using EQ-5D: A European perspective*. The Netherlands: Kluwer Academic Publishers.
- Knapp, M. (2007). Economic outcomes and levers: impacts for individuals and society. *Int Psychogeriatr*, 19(3), 483-495.
- Knapp, M., Thorgrimsen, L., Patel, A., Spector, A., Hallam, A., Woods, B., & Orrell, M. (2006). Cognitive stimulation therapy for people with dementia: cost-effectiveness analysis. *British Journal of Psychiatry*, 188, 574-580.
- Knapp, M., & Wigglesworth, R. (1998a). Costing community care for older people. In A. Wimo, B. Jönssen, G. Karlsson & B. Winblad (Eds.), *Health Economics of Dementia* (pp. 231-244). Chichester: Wiley.
- Knapp, M., Wilkinson, D., & Wigglesworth, R. (1998b). The economic consequences of Alzheimer's disease in the context of new drug developments. *International Journal of Geriatric Psychiatry*, 13(8), 531-543.
- Kok, R., & Verhey, F. (2002). Dutch translation of the Mini Mental State Examination (Folstein et al., 1975).
- Konig, H. H., Ulshofer, A., Gregor, M., von Tirpitz, C., Reinshagen, M., Adler, G., & Leidl, R. (2002). Validation of the EuroQol questionnaire in patients with inflammatory bowel disease. *European Journal of Gastroenterology & Hepatology* 14(11), 1205-1215.
- Krabbe, P. F., Stouthard, M. E., Essink-Bot, M. L., & Bonsel, G. J. (1999a). The effect of adding a cognitive dimension to the EuroQol multiattribute health-status classification system. *J Clin Epidemiol*, 52(4), 293-301.
- Krabbe, P. F., Stouthard, M. E., Essink-Bot, M. L., & Bonsel, G. J. (1999b). The effect of adding a cognitive dimension to the EuroQol multiattribute health-status classification system. *Journal of Clinical Epidemiology*, 52(4), 293-301.

## REFERENCES

- Kurlowicz, L. H., Evans, L. K., Strumpf, N. E., & Maislin, G. (2002). A psychometric evaluation of the Cornell Scale for Depression in Dementia in a frail, nursing home population. *American Journal of Geriatric Psychiatry*, 10(5), 600-608.
- Lamers, L., McDonnell, J., Stalmeier, P., PFM Krabbe, & van, J. B. ((in press)). The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. *Health Economics*
- Lamers, L. M., Stalmeier, P. F., McDonnell, J., Krabbe, P. F., & van Busschbach, J. J. (2005). Kwaliteit van leven meten in economische evaluaties: het Nederlands EQ-5D-tarief. *Nederlands Tijdschrift voor Geneeskunde*, 149(28), 1574-1578.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.
- Langa, K. M., Chernew, M. E., Kabeto, M. U., Herzog, A. R., Ofstedal, M. B., Willis, R. J., Wallace, R. B., Mucha, L. M., Straus, W. L., & Fendrick, A. M. (2001). National estimates of the quantity and cost of informal caregiving for the elderly with dementia. *J Gen Intern Med*, 16(11), 770-778.
- Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist*, 9(3), 179-186.
- Lindeboom, R., Sprangers, M. A., & Zwinderman, A. H. (2005). Responsiveness: a reinvention of the wheel? *Health & Quality of Life Outcomes*, 3(1), 8.
- Little, R. J. A., & Rubin, D. B. (1987). Nonresponse in Sample Surveys. In *Statistical Analysis with Missing Data* (pp. 50-72). Canada: John Wiley & Sons, Inc.
- Logsdon, R. G., Gibbons, L. E., McCurry, S. M., & Teri, L. (2002). Assessing quality of life in older adults with cognitive impairment. *Psychosomatic Medicine*, 64(3), 510-519.
- Luce, B. R., Manning, W. G., Siegel, J. E., & Lipscomb, J. (1996). Estimating costs in cost-effectiveness analysis. In M. R. Gold, J. E. Siegel, L. B. Russel & M. C. Weinstein (Eds.), *Cost effectiveness in health and medicine*. Oxford: Oxford University Press.
- Magaziner, J. (1997). Use of proxies to measure health and functional outcomes in effectiveness research in persons with Alzheimer disease and related disorders. *Alzheimer Disease and Associated Disorders*, 11 Suppl 6, 168-174.
- Mainprize, E., & Rodin, G. (1987a). Geriatric referrals to a psychiatric consultation-liaison service. *Can J Psychiatry*, 32(1), 5-9.
- Mainprize, E., & Rodin, G. (1987b). Geriatric referrals to a psychiatric consultation-liaison service. *Canadian Journal of Psychiatry*, 32, 5-9.
- Maniadakis, N., & Gray, A. (2000). The economic burden of back pain in the UK. *Pain*, 84(1), 95-103.
- Marra, C. A., Woolcott, J. C., Kopec, J. A., Shojania, K., Offer, R., Brazier, J. E., Esdaile, J. M., & Anis, A. H. (2005). A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. *Social Science & Medicine*, 60(7), 1571-1582.
- Mathers, C. D., Vos, E. T., & Stevenson, C. E. (1999). *The burden of disease and injury in Australia*. Canberra: Australian Institute of Health and Welfare.
- Mathers, C. D., Vos, E. T., Stevenson, C. E., & Begg, S. J. (2001). The burden of disease and injury in Australia. *Bull World Health Organ*, 79(11), 1076-1084.
- Mauskopf, J. A., Paul, J. E., Grant, D. M., & Stergachis, A. (1998). The role of cost-consequence analysis in healthcare decision-making. *Pharmacoeconomics*, 13(3), 277-288.
- McDaid, D. (2001). Estimating the costs of informal care for people with Alzheimer's disease: methodological and practical challenges. *Int J Geriatr Psychiatry*, 16(4), 400-405.
- McKeith, I., Fairbairn, A., Perry, R., Thompson, P., & Perry, E. (1992). Neuroleptic sensitivity in patients with senile dementia of Lewy body type. *Brnj*, 305(6855), 673-678.
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, 34(7), 939-944.
- McMahon, P. M., Araki, S. S., Neumann, P. J., Harris, G. J., & Gazelle, G. S. (2000). Cost-effectiveness of functional imaging tests in the diagnosis of Alzheimer disease. *Radiology*, 217(1), 58-68.
- McMahon, P. M., Araki, S. S., Sandberg, E. A., Neumann, P. J., & Gazelle, G. S. (2003). Cost-effectiveness of PET in the diagnosis of Alzheimer disease. *Radiology*, 228(2), 515-522.
- Murray, C. J. L., & Lopez, L. D. (1996). *The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020* (Vol. 1). Geneva: Harvard School of Public Health / World Health Organization.
- NICE. (2006). *Clinical Guideline: Dementia. Supporting people with dementia and their carers in health and social care*. London: National Collaborating Centre for Mental Health.

- Nordberg, G., von Strauss, E., Kareholt, I., Johansson, L., & Wimo, A. (2005). The amount of informal and formal care among non-demented and demented elderly persons-results from a Swedish population-based study. *Int J Geriatr Psychiatry*, 20(9), 862-871.
- Norman, G. R., Stratford, P., & Regehr, G. (1997). Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. *Journal of Clinical Epidemiology*, 50(8), 869-879.
- Nowels, D., McGloin, J., Westfall, J. M., & Holcomb, S. (2005). Validation of the EQ-5D quality of life instrument in patients after myocardial infarction. *Quality of Life Research*, 14(1), 95-105.
- Olde Rikkert, M. G., van der Flier, W. M., de Leeuw, F. E., Verbeek, M., Jansen, R. W., Verhey, F., & Scheltens, P. (2006). Multiple diagnostic tests are needed to assess multiple causes of dementia. *Arch Neurol*, 63(1), 144-146.
- Oortwijn, W. (2000). *First things first. Priority setting for Health Technology Assessment*. Leiden University, Leiden.
- Oostenbrink, J. B., Koopmanschap, M. A., & Rutten, F. F. H. (2004). *Handleiding voor kostenonderzoek: methoden en richtlijnrijzen voor economische evaluaties in de gezondheidszorg*. Amstelveen: College voor zorgverzekeringen.
- Parkin, D., Rice, N., Jacoby, A., & Doughty, J. (2004). Use of a visual analogue scale in a daily patient diary: modelling cross-sectional time-series data on health-related quality of life. *Social Science & Medicine*, 59(2), 351-360.
- Patterson, M. B., & Whitehouse, P. J. (1990). The diagnostic assessment of patients with dementia. In N. L. Mace (Ed.), *Dementia care: Patient, family and community* (pp. 3-21). Baltimore: Johns Hopkins University Press.
- Pelosi, A. J., McNulty, S. V., & Jackson, G. A. (2006). Role of cholinesterase inhibitors in dementia care needs rethinking.[see comment]. *British Medical Journal*, 333(7566), 491-493.
- Pickard, A. S., Johnson, J. A., & Feeny, D. H. (2005). Responsiveness of generic health-related quality of life measures in stroke. *Quality of Life Research*, 14(1), 207-219.
- Pinto-Prades, J. (1997). Is the person trade-off a valid method for allocating health care resources? . *Health Economics* 6, 71-81.
- Plugge, L. A., Verhey, F. R., van Everdingen, J. J., & Jolles, J. (1991). Differential diagnosis of dementia: intra- and interdisciplinary agreement. *J Geriatr Psychiatry Neurol*, 4(2), 90-97.
- Podgorski, C. A., Lanning, B. D., Casaceli, C. J., Nardi, A. L., & Cox, C. (2002a). Dementia consults: predictors of compliance by primary care physicians. *Am J Alzheimers Dis Other Demen*, 17(1), 44-50.
- Podgorski, C. A., Lanning, B. D., Casaceli, C. J., Nardi, A. L., & Cox, C. (2002b). Dementia consults: predictors of compliance by primary care physicians. *American journal of Alzheimer's disease and other dementias*, 17(1), 44-50.
- Posnett, J., & Jan, S. (1996). Indirect cost in economic evaluation: the opportunity cost of unpaid inputs. *Health Econ*, 5(1), 13-23.
- Pot, A. M., van Dyck, R., & Deeg, D. J. (1995). [Perceived stress caused by informal caregiving. Construction of a scale]. *Tijdschrift voor Gerontologie en Geriatrie*, 26(5), 214-219.
- Rabins, P. V., & Black, B. S. (2007). Measuring quality of life in dementia: purposes, goals, challenges and progress. *Int Psychogeriatr*, 19(3), 401-407.
- Raftery, J. (2001). NICE: faster access to modern treatments? Analysis of guidance on health technologies. *British Medical Journal*, 323(7324), 1300-1303.
- Ready, R. E., & Ott, B. R. (2003). Quality of Life measures for dementia. *Health Qual Life Outcomes*, 1(1), 11.
- Reed, R. L., Kligman, E. W., & Weiss, B. D. (1990). Comprehensive geriatric assessment recommendations: adherence of family practice residents. *J Fam Pract*, 31(4), 389-392.
- Reisberg, B. (2007). Global measures: utility in defining and measuring treatment response in dementia. *Int Psychogeriatr*, 19(3), 421-456.
- Reisberg, B., Ferris, S. H., de Leon, M. J., & Crook, T. (1982). The Global Deterioration Scale for assessment of primary degenerative dementia. *American Journal of Psychiatry*, 139(9), 1136-1139.
- Reuben, D. B., Maly, R. C., Hirsch, S. H., Frank, J. C., Oakes, A. M., Siu, A. L., & Hays, R. D. (1996). Physician implementation of and patient adherence to recommendations from comprehensive geriatric assessment. *Am J Med*, 100(4), 444-451.
- Robinson, K. M. (1997). Family caregiving: who provides the care, and at what cost? *Nurs Econ*, 15(5), 243-247.
- Rockwood, K. (2007). The measuring, meaning and importance of activities of daily living (ADLs) as an outcome. *Int Psychogeriatr*, 19(3), 467-482.
- Rothman, K. J. (1986). *Modern Epidemiology*. Boston: Little, Brown & Co.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons.
- Rubin, D. B., & Schenker, N. (1991). Multiple imputation in health-care databases: an overview and some applications. *Statistics in Medicine*, 10(4), 585-598.

## REFERENCES

- Ruskin, P. E. (1985). Geropsychiatric consultation in a university hospital: a report on 67 referrals. *Am J Psychiatry*, 142(3), 333-336.
- RVZ. (2006). *Zinnige en duurzame zorg*. Zoetermeer: Raad voor de Volksgezondheid en Zorg.
- Schneider, J., Hallam, A., Murray, J., Foley, B., Atkin, L., Banerjee, S., Islam, M. K., & Mann, A. (2002). Formal and informal care for people with dementia: factors associated with service receipt. *Aging Ment Health*, 6(3), 255-265.
- Schrag, A., Selai, C., Davis, J., Lees, A. J., Jahanshahi, M., & Quinn, N. (2003a). Health-related quality of life in patients with progressive supranuclear palsy. *Movement Disorders*, 18(12), 1464-1469.
- Schrag, A., Selai, C., Davis, J., Lees, A. J., Jahanshahi, M., & Quinn, N. (2003b). Health-related quality of life in patients with progressive supranuclear palsy. *Mov Disord*, 18(12), 1464-1469.
- Schramm, U., Berger, G., Muller, R., Kratzsch, T., Peters, J., & Frolich, L. (2002). Psychometric properties of Clock Drawing Test and MMSE or Short Performance Test (SKT) in dementia screening in a memory clinic population. *Int J Geriatr Psychiatry*, 17(3), 254-260.
- Schrijnemakers, V. J. J., & Haveman, M. J. (1995a). Effects of preventive outpatient geriatric assessment: short term results of a randomized controlled study. *Home health care services quarterly* 15(2), 81-97.
- Schrijnemakers, V. J. J., Haveman, M. J., & Kessels, A. (1995b). Effects of preventive outpatient geriatric assessment for frail elderly persons: results of a randomized controlled study. *International Journal of Health Sciences* 6, 177-186.
- Schulz, R., & Beach, S. R. (1999). Caregiving as a risk factor for mortality: the Caregiver Health Effects Study. *Jama*, 282(23), 2215-2219.
- Schulz, R., O'Brien, A. T., Bookwala, J., & Fleissner, K. (1995). Psychiatric and physical morbidity effects of dementia caregiving: prevalence, correlates, and causes. *Gerontologist*, 35(6), 771-791.
- Sears, C. L., & Charlson, M. E. (1983). The effectiveness of a consultation. Compliance with initial recommendations. *Am J Med*, 74(5), 870-876.
- Selai, C. (1998). Using the EuroQol (EQ-5D) in Dementia. In R. Rabin, J. Busschbach, F. Charro, M. Essink-Bot & G. Bonsel (Eds.), *Proceedings of the EuroQol Plenary Meeting, 2-3 October 1997* (pp. 157-168). Rotterdam: Erasmus University.
- Selai, C. (2001). Assessing quality of life in dementia. *Medical Care*, 39(8), 753-755.
- Slobbe, I. C. J., Kommer, G. J., Smit, J. M., Groen, J., Meering, W. J., & Polder, J. J. (2006). *Kosten van Ziekten in Nederland (Cost of Illness in the Netherlands)* (No. 270751010). Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu (RIVM), Centrum voor Volksgezondheid Toekomst Verkenningen; Centraal Bureau voor de Statistiek (CBS); Erasmus MC, afdeling Maatschappelijke Gezondheidszorg.
- Smith, K., & Wright, K. (1994). Informal care and economic appraisal: a discussion of possible methodological approaches. *Health Econ*, 3(3), 137-148.
- Smith, S. C., Lamping, D. L., Banerjee, S., Harwood, R., Foley, B., Smith, P., Cook, J. C., Murray, J., Prince, M., Levin, E., Mann, A., & Knapp, M. (2005). Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology. *Health Technology Assessment*, 9(10), 1-93, iii-iv.
- Solberg, T. K., Olsen, J. A., Ingebrigtsen, T., Hofoss, D., & Nygaard, O. P. (2005). Health-related quality of life assessment by the EuroQol-5D can provide cost-utility data in the field of low-back surgery. *European Spine Journal*, 14(10), 1000-1007.
- Stoppe, G. (2005). Primary care: Fit for the dementia challenge? In A. Burns (Ed.), *Standards in Dementia Care* (1 ed., pp. 185-189). Manchester U.K.: Taylor & Francis.
- Stouthard, M., Essink-Bot, M., Bonsel, G., Barendregt, J., Kramer, P., van de Water, H., Gunning-Schepers, L., & van der Maas, P. (1997). *Disability Weights for Diseases in the Netherlands*. Rotterdam: Erasmus University.
- Stouthard, M., Essink-Bot, M., & GJ, B. (2000). (on behalf of the Dutch Disability Weights (DDW) Group). Disability weights for diseases. A modified protocol and results for a Western European region. *European Journal of Public Health* 10, 24-30.
- Streiner, D. L. (2003). Clinimetrics vs. psychometrics: an unnecessary distinction.[comment]. *Journal of Clinical Epidemiology*, 56(12), 1142-1145; discussion 1146-1149.
- Stuck, A. E., Siu, A. L., Wieland, G. D., Adams, J., & Rubenstein, L. Z. (1993). Comprehensive geriatric assessment: a meta-analysis of controlled trials. *Lancet*, 342(8878), 1032-1036.
- Teitelbaum, L., Cotton, D., Ginsburg, M. L., & Nashed, Y. H. (1996). Psychogeriatric Consultation Services: Effect and Effectiveness. *Canadian Journal of Psychiatry*, 41(December), 638-644.
- Terwee, C. B., Dekker, F. W., Wiersinga, W. M., Prummel, M. F., & Bossuyt, P. M. (2003). On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. *Quality of Life Research*, 12(4), 349-362.

- Torgerson, D. J., Donaldson, C., & Reid, D. M. (1994). Private versus social opportunity cost of time: valuing time in the demand for health care. *Health Econ*, 3(3), 149-155.
- Torti, F. M., Jr., Gwyther, L. P., Reed, S. D., Friedman, J. Y., & Schulman, K. A. (2004). A multinational review of recent trends and reports in dementia caregiver burden. *Alzheimer Dis Assoc Disord*, 18(2), 99-109.
- Toseland, R. J., O'Donnell, J. C., Engelhardt, J. B., Richie, J., Jue, D., & Banks, S. M. (1997). Outpatient geriatric evaluation and management: is there an investment effect? . *Gerontologist* 37(3), 324-332.
- Turpie, I. D. (2000). The geriatric day hospital. In J. Grimley Evans, T. F. Williams, B. L. Beattie, J. P. Michel & G. K. Wilcock (Eds.), *Oxford textbook of geriatric medicine* Oxford: Oxford University Press
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453-458.
- Ubel, P. A., Loewenstein, G., & Jepson, C. (2003). Whose quality of life? A commentary exploring discrepancies between health state evaluations of patients and the general public.[see comment]. *Quality of Life Research*, 12(6), 599-607.
- van den Berg, B., Al, M., Brouwer, W., van Exel, J., & Koopmanschap, M. (2005a). Economic valuation of informal care: the conjoint measurement method applied to informal caregiving. *Social Science & Medicine*, 61(6), 1342-1355.
- van den Berg, B., Brouwer, W., & Koopmanschap, M. (2004). Economic valuation of informal care. An overview of methods and applications. *European Journal of Health Economics*, 5, 36-45.
- van den Berg, B., Brouwer, W., van Exel, J., & Koopmanschap, M. (2005b). Economic valuation of informal care: the contingent valuation method applied to informal caregiving. *Health Economics*, 14(2), 169-183.
- van den Berg, B., Brouwer, W., van Exel, J., Koopmanschap, M., van den Bos, G. A., & Rutten, F. (2006). Economic valuation of informal care: lessons from the application of the opportunity costs and proxy good methods. *Social Science & Medicine*, 62(4), 835-845.
- van den Brink, M., van den Hout, W. B., Stiggelbout, A. M., Putter, H., van de Velde, C. J., & Kievit, J. (2005). Self-reports of health-care utilization: diary or questionnaire? *International Journal of Technology Assessment in Health Care*, 21(3), 298-304.
- Van der Cammen, T. J., Simpson, J. M., Fraser, R. M., Preker, A. S., & Exton-Smith, A. N. (1987). The Memory Clinic. A new approach to the detection of dementia. *Br J Psychiatry*, 150, 359-364.
- van der Roer, N., Goes, E. S., Blom, M., & Busschbach, J. J. (2000). [Model of costs of care for dementia: community-dwelling vs. institutionalization]. *Tijdschr Gerontol Geriatr*, 31(2), 55-61.
- van Exel, N. J., Koopmanschap, M. A., van den Berg, B., Brouwer, W. B., & van den Bos, G. A. (2005). Burden of informal caregiving for stroke patients. Identification of caregivers at risk of adverse health effects. *Cerebrovascular Diseases*, 19(1), 11-17.
- van Exel, N. J., Scholte op Reimer, W. J., Brouwer, W. B., van den Berg, B., Koopmanschap, M. A., & van den Bos, G. A. (2004). Instruments for assessing the burden of informal caregiving for stroke patients in clinical practice: a comparison of CSI, CRA, SCQ and self-rated burden. *Clinical Rehabilitation*, 18(2), 203-214.
- van Hout, B., Al, M., Gordon, G., & Rutten, F. F. H. (1994). Costs, effects and c/e-ratios alongside a clinical trial. *Health Economics*, 3, 309-319.
- van Hout, H., Teunisse, S., Derix, M., Poels, P., Kuin, Y., Vernooij-Dassen, M., Grol, R., & Hoefnagels, W. (2001). CAMDEX, can it be more efficient? Observational study on the contribution of four screening measures to the diagnosis of dementia by a memory clinic team. *Int J Geriatr Psychiatry*, 16(1), 64-69.
- van Hout, H., Vernooij-Dassen, M., Poels, P., Hoefnagels, W., & Grol, R. (2000). Are general practitioners able to accurately diagnose dementia and identify Alzheimer's disease? A comparison with an outpatient memory clinic. *Br J Gen Pract*, 50(453), 311-312.
- van Velden, M. E., Severens, J. L., & Novak, A. (2005). Economic evaluations of healthcare programmes and decision making: the influence of economic evaluations on different healthcare decision-making levels. *Pharmacoeconomics*, 23(11), 1075-1082.
- VanderZee, K. I., Sanderman, R., Heyink, J. W., & de Haes, H. (1996). Psychometric qualities of the RAND 36-Item Health Survey 1.0: a multidimensional measure of general health status. *International Journal of Behavioral Medicine*, 3(2), 104-122.
- Verhey, F. R., Jolles, J., Ponds, R. W., Rozendaal, N., Plugge, L. A., de Vet, R. C., Vreeling, F. W., & van der Lugt, P. J. (1993). Diagnosing dementia: a comparison between a monodisciplinary and a multidisciplinary approach. *Journal of Neuropsychiatry & Clinical Neurosciences*, 5(1), 78-85.
- Verhey, F. R., Scheltens, P., & Olde Rikkert, M. G. (2007a). [The development of memory clinics in The Netherlands]. *Nederlands Tijdschrift voor Geneeskunde*, 151(10), 578-580.
- Verhey, F. R. J., de Vugt, M. E., Aalten, P., Vernooij-Dassen, M. J., Byrne, E. J., & Robert, P. (2007b). Caregiver outcomes in disease modifying trials. *Submitted*.

## REFERENCES

- Verhey, F. R. J., & Honig, A. (1997). Depression in the elderly. In A. Honig & H. M. v. Praag (Eds.), *Recent advances in affective disorders* New York: Wiley and sons.
- Verhey, F. R. J., Ramakers, I. H. G. B., Jolles, J., Scheltens, P. H., Vernooij-Dassen, M., Olde Rikkert, M. G. M., & Blom, M. (2005). *Geheugenpoliklinieken in Nederland: inventarisatie 2004 (Memory Clinics in the Netherlands: inventarisatie 2004)* (No. ISBN: 9075579004). Maastricht: University of Maastricht.
- Verhey, F. R. J., Scheltens, P. H., & Olde Rikkert, M. G. M. (2007c). De Geheugenpolikliniek: recente ontwikkelingen in Nederland (The Memory Clinic: recent developments in the Netherlands). *Nederlands Tijdschrift voor Geneeskunde*, (Accepted).
- Vraamark Elberling, T., Stokholm, J., Hogh, P., & Waldemar, G. (2002). Diagnostic profile of young and middle-aged memory clinic patients. *Neurology*, 59(8), 1259-1262.
- W.H.O. (1993). *The ICD-10 Classification of Mental and Behavioural Disorders. Diagnostic Criteria for Research*. Geneva.: World Health Organization.
- Wagner, M. T., Nayak, M., & Fink, C. (1995). Bedside screening of neurocognitive function. In L. A. Cushman & M. J. Scherer (Eds.), *Psychological Assessment in Medical Rehabilitation* (pp. 145-198). Washington, DC: American Psychological Association.
- Waldorff, F. B., Bulow, L. B., Malterud, K., & Waldemar, G. (2001). Management of dementia in primary health care: the experience of collaboration between the GP and the district nurse. *Family practice*, 18(5), 549-552.
- Walstra, G. J., Teunisse, S., van Gool, W. A., & van Crevel, H. (1997). Symptomatic treatment of elderly patients with early Alzheimer's disease at a memory clinic. *J Geriatr Psychiatry Neurol*, 10(1), 33-38.
- Ware, J. E., Jr., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care*, 30(6), 473-483.
- Wimo, A. (2007). Clinical and economic outcomes--friend or foe? *Int Psychogeriatr*, 19(3), 497-507.
- Wimo, A., Ljunggren, G., & Winblad, B. (1997). Costs of dementia and dementia care: a review. *International Journal of Geriatric Psychiatry*, 12(8), 841-856.
- Wimo, A., & Nordberg, G. (2007). Validity and reliability of assessments of time. Comparisons of direct observations and estimates of time by the use of the resource utilization in dementia (RUD)-instrument. *Arch Gerontol Geriatr*, 44(1), 71-81.
- Wimo, A., Nordberg, G., Jansson, W., & Grafstrom, M. (2000). Assessment of informal services to demented people with the RUD instrument. *Int J Geriatr Psychiatry*, 15(10), 969-971.
- Wimo, A., von Strauss, E., Nordberg, G., Sassi, F., & Johansson, L. (2002). Time spent on informal and formal care giving for persons with dementia in Sweden. *Health Policy*, 61(3), 255-268.
- Wind, A., Gusselkloo, J., Vernooij-Dassen, M., Bouma, M., Boomsma, L., & Boukes, F. (2003). NHG-Standaard Dementie (tweede herziening). *Huisarts Wet*, 46(13), 754-767.
- Wolfs, C., Severens, J., Dirksen, C., Kessels, A., Jaspers, N., & Verhey, F. (2005a). A Randomized Clinical Trial on the Effects of a Diagnostic Observation Centre for PsychoGeriatric Patients (DOC-PG). *International Psychogeriatrics*, 17, 370-371.
- Wolfs, C. A., Dirksen, C. D., Kessels, A., Willems, D. C., Verhey, F. R., & Severens, J. L. (2007a). Performance of the EQ-5D and the EQ-5D+C in elderly patients with cognitive impairments. *Health Qual Life Outcomes*, 5(1), 33.
- Wolfs, C. A., Dirksen, C. D., Severens, J. L., & Verhey, F. R. (2006). The added value of a multidisciplinary approach in diagnosing dementia: a review. *International Journal of Geriatric Psychiatry*, 21(3), 223-232.
- Wolfs, C. A., Verhey, F. R., Kessels, A., Winkens, R. A., Severens, J. L., & Dirksen, C. D. (2007b). GP concordance with advice for treatment following a multidisciplinary psychogeriatric assessment. *International Journal of Geriatric Psychiatry*, 22(3), 233-240.
- Wolfs, C. A., Verhey, F. R., Severens, J. L., Dirksen, C. D., Kessels, A., & Winkens, R. A. (2007c). *Eindrappingage Doelmatigheidsonderzoek (nr. 945-02-055)*. Maastricht: ZonMw.
- Wolfs, C. A. G., Dirksen, C. D., Kessels, A., Severens, J. L., & Verhey, F. R. J. (2007d). Economic evaluation of an integrated diagnostic approach for psychogeriatric patients: Results of a randomised controlled trial. *Submitted*.
- Wolfs, C. A. G., Kessels, A., Dirksen, C. D., Severens, J. L., & Verhey, F. R. (2007e). Effectiveness of an integrated multidisciplinary diagnostic approach for psychogeriatric patients: a randomized controlled trial. *British Journal of Psychiatry (Accepted)*.
- Wolfs, C. A. G., Severens, J. L., Dirksen, C. D., Kessels, A., Jaspers, N., & Verhey, F. R. J. (2005b). A Randomized Clinical Trial on the Effects of a Diagnostic Observation Centre for PsychoGeriatric patients (DOC-PG). *International Psychogeriatrics*, 17(Supplement 2), 370-371.
- Wrigley, M., & Loane, R. (1991). Consultation-liaison referrals to the north Dublin old age psychiatry service. *Ir Med J*, 84(3), 89-91.

- Wu, A. W., Jacobson, K. L., Frick, K. D., Clark, R., Revicki, D. A., Freedberg, K. A., Scott-Lennox, J., & Feinberg, J. (2002). Validity and responsiveness of the euroqol as a measure of health-related quality of life in people enrolled in an AIDS clinical trial. *Quality of Life Research, 11*(3), 273-282.
- Zee vd, K. I., Sanderman, R., & Heyink, J. W. (1993). De psychometrische kwaliteiten van de MOS 36-item Short Form Health Survey (SF-36) in een Nederlandse populatie. . *Tijdschrift voor sociale gezondheidszorg 4*, 183-191.
- Zhu, C. W., Scarmeas, N., Torgan, R., Albert, M., Brandt, J., Blacker, D., Sano, M., & Stern, Y. (2006a). Clinical characteristics and longitudinal changes of informal cost of Alzheimer's disease in the community. *J Am Geriatr Soc, 54*(10), 1596-1602.
- Zhu, C. W., Scarmeas, N., Torgan, R., Albert, M., Brandt, J., Blacker, D., Sano, M., & Stern, Y. (2006b). Clinical features associated with costs in early AD: baseline data from the Predictors Study. *Neurology, 66*(7), 1021-1028.



Dankwoord



## DANKWOORD

Mijn proefschrift is klaar! Alleen nog het dankwoord.....

In de afgelopen vijf jaren heb ik zóveel mensen ontmoet die op een of andere manier een bijdrage hebben geleverd aan dit proefschrift, dat ik bij voorbaat mijn excuses aanbied, mocht ik iemand vergeten zijn.

Allereerst wil ik alle patiënten en hun partners, kinderen en andere naaste verzorgers hartelijk bedanken. Heel erg bedankt voor jullie inzet, gastvrijheid en openhartige gesprekken die vaak tijdovrend en confronterend waren.

Vervolgens wil ik mijn promotoren Frans Verhey, Hans Severens en mijn co-promotor Carmen Dirksen bedanken voor de plezierige samenwerking de afgelopen jaren. Frans, ik ben je dankbaar voor je enthousiasme, de grote mate van vrijheid en vertrouwen die je me schonk en voor al hetgeen je me geleerd hebt.

Hans en Carmen: bedankt voor mijn inwijding in de wereld van de MTA! Hans, ik ben je dankbaar voor je wetenschappelijke expertise, je creatieve inzichten, je belangstelling en vertrouwen. Carmen, ik heb je heldere uitleg, kritische kijk op analyses en artikelen altijd zeer gewaardeerd. Bedankt dat je deur altijd voor me open stond!

Beste Frans, Hans en Carmen, ik hoop dat deze nuttige samenwerking tussen “kliniek “ en “ economie” ook in de toekomst vruchtbaar zal blijven!

Fons Kessels, methodoloog en statisticus: Beste Fons, bedankt voor je hulp bij het analyseren en structureren van mijn data en je betrokkenheid bij de opzet en uitwerking van de artikelen. Ondanks de vele “multiple imputation” pogingen, die soms resulteerden in wederzijdse “multiple irritation”, heb ik onze samenwerking erg gewaardeerd en er veel van geleerd!

Daniëlle Willems, lieve Daan, we zijn samen begonnen aan het MEDICIE-avontuur en inmiddels ben je een hele goede vriendin. Bedankt voor jouw hulp en betrokkenheid bij het includeren en bezoeken van de 230 patiënten. Zonder jou was dit project nooit zo'n succes geworden (en niet zo gezellig)! Ik ben blij dat jij mijn paranimf wilt zijn en dat ik dit 31 januari voor jou terug mag doen.

Audrey Fiddelers, Ilvy Mayen, Heidi Lansdaal en Andrea Wolf, ik bedank ook jullie voor de hulp bij het verzamelen en verwerken van alle onderzoeksgegevens.

Het (oude en nieuwe) DOC-PG team: Frank Amory, Yvonne van Zutphen, Birgit Senden, Ingrid van Laar, Ron Warnier, Marian Freuling, Evelien Pijpers, Hilde Heijnen, Marianne van de Boogerd, Wubbo Mulder, Albert Leentjens, Tanja Hamelers, Debbie van Menxel, Marianne Maas, de arts-assistenten psychiatrie, Dick Nagelhout en Annemie Warnier, bedankt voor jullie belangstelling en medewerking aan het onderzoek.

Alle medewerkers van de geheugenpoli (ook de secretaresses voor de hulp en uitleg bij het zoeken naar dossiers), bedankt voor jullie medewerking!

Niek Jaspers en Mike Verkaaik, bedankt voor de goede samenwerking met de RIAGG en de mogelijkheid het onderzoek ook op de RIAGG te mogen uitvoeren. Ook de overige medewerkers van de afdeling Ouderenzorg, met name Debbie Peeters, Paul Wiltling,

Manon Hanssen, Magda Reumkens en de dames van het secretariaat wil ik bedanken voor alle hulp.

Alle huisartsen (en hun assistentes), bedankt voor jullie aanmeldingen, het invullen van de tevredenheidsvragenlijst en de checklijsten. Jan van Rooij en José Bindels, bedankt voor jullie hulp met het aanschrijven van de huisartsen. Ron Winkens, beste Ron, bedankt voor je waardevolle bijdrage aan het “concordance” artikel en je inzet tijdens het MEDICIE-overleg.

De medewerkers van de Apothekers Vereniging Maastricht, bedankt voor jullie hulp bij het verzamelen van alle medicatie-overzichten.

De leden van de leescommissie, prof. dr. J. Metsemakers (voorzitter), prof. dr. J. Schols, dr. L. Jönssen, dr. A. Leentjens en dr. W. Brouwer, bedankt voor het snelle lezen van het manuscript tijdens de zomervakantie.

Mijn (oud)collega's van de afdeling Klinische Epidemiologie en Medical Technology Assessment (KEMTA), met wie ik bijna vijf jaar met veel plezier heb samengewerkt:

Allereerst mijn kamergenootje, lieve Audrey, bedankt voor de gezelligheid en het delen van de lekkere theetjes en soepjes! Ik wens je heel veel succes met jouw promotie! Bert, Brigitte, Carmen, Daniëlle B, Daniëlle W, Debby, Fons, Fred, Gemma, Ghislaine, Hans, Irene, Isabel, Janneke, Jody, Linda, Margriet, Martin en Thea, bedankt voor jullie interesse, jullie collegialiteit en de gezellige KEMTA-borrels.

Mijn collega's van het instituut Hersenen en Gedrag, bedankt voor de fijne werksfeer tijdens de afronding van mijn proefschrift (waardoor ik niet de gezelligste collega ben geweest, hè Esther en Deliane!).

Mijn collega's van het Alzheimer Centrum Limburg; Geert Leenders, bedankt voor jouw enthousiasme, interesse en zangkunsten (schön ist es auf der welt zu sein...). Marjolein de Vugt, bedankt voor je begrip wat betreft de vertraging in het “Kiezen in de Zorg” project, je bereidheid in de corona deel te nemen en onze gezellige samenwerking! Els Ketelslegers, bedankt voor al je hulp, zeker in deze laatste fase van de afronding van mijn proefschrift!

Mien familie en sjoenfamilie, danke veur eur belangstelling en steun.

Pap en mam, danke veur alles... veur eur leefde, vertrouwe en intresse. Sjoenpap en sjoenmam, danke veur eur steun en veur de vitamienesjes es Tom weer 's in 't boeteland waar. Jeroen, danke tot ste miene paranimf wils zien. Vief, bedaank veur dien intresse en luusterend oer (es iech mien ei weer 's kwiet mos).

Opa en oma, allewel iech 't uuch neet hub kinne vraoge, vind iech 't 'n ier tot geer op miene umslaag stoat.

Mien vrun en vrundinnekes, danke veur de broednudege oontspanning, de stomme kal en de gezèllege (dèks lange) weekendsjes.

Miene leve,.....Tom.....danke veur al.....danke veur de hõlp mit de computer elend (wat regelmaoteg veur kaom), danke veur dien gedõld, danke veur 't mit miech oet te hawwe (veural 't lètste jaor), en veur mit miech te trouwe..... Iech hauw vaan diech.



# Curriculum Vitae





## Curriculum Vitae

Claire Wolfs werd geboren op 20 mei 1977 in Eijsden, Nederland. In 1995 behaalde zij haar VWO diploma aan het Sint Maartenscollege in Maastricht. Van 1995 tot 1996 studeerde ze communicatiewetenschappen in Amsterdam en Nijmegen. In 1996 startte ze met de studie Psychologie, afstudeerrichting Biologische Psychologie in Maastricht. Na een onderzoeksstage (die betrekking had op hervonden herinneringen) behaalde ze in 2001 haar diploma. In 2002 werd ze als assistent in opleiding aangesteld bij de afdeling Klinische Epidemiologie en Medical Technology Assessment (KEMTA) van het academisch ziekenhuis in Maastricht en werkte ze nauw samen met de afdeling Psychiatrie en Neuropsychologie. Het doel van dit onderzoek was het evalueren van een integrale multidisciplinaire aanpak voor het diagnostiseren en managen van dementie door het Diagnostisch OnderzoeksCentrum voor PsychoGeriatrische patiënten (DOC-PG). Het onderzoek resulteerde in dit proefschrift. Momenteel is ze werkzaam als onderzoeker binnen de vakgroep Psychiatrie en Neuropsychologie van de Universiteit Maastricht.