Prevention of Depression

Filip Smit
We never know self-realisation.
We are two abysses – a well staring at the sky.

(from: *The Book of Disquiet*, Fernando Pessoa)
Doctoral Committee:

Prof. dr. A. van Balkom
Prof. dr. L. Bouter
Prof. dr. C. Hosman
Prof. dr. I. Marks
Prof. dr. F. Rutten
Prof. dr. Ph. Spinhoven

Faculty of Medicine

This thesis was prepared at the Departments of Psychiatry and Clinical Psychology of the Vrije Universiteit in Amsterdam, both of which are embedded in the Institute for Research in Extramural Medicine (EMGO), and in close cooperation with the Trimbos Institute, which is the Netherlands Institute of Mental Health and Addiction in Utrecht.

Financial support by The Netherlands Ministry of Health (VWS, The Hague), The Netherlands Health Research Council (ZonMw, The Hague), and the Netherlands Foundation of Mental Health (NFG, Utrecht) is gratefully acknowledged. The Trimbos Institute and the Department of Clinical Psychology of the Vrije Universiteit Amsterdam are acknowledged, with many thanks, for granting the author sabbatical leave. The Museum Boijmans van Beuningen in Rotterdam is acknowledged for kindly giving permission to make use of Albrecht Dürer’s copperplate “Melencolia I”.

Artikelnummer: AF0697
ISBN-10: 90-5253-562-0

Printing: Ladenius Communicatie BV, Houten.

Please, cite as:


© 2006, Filip Smit, Cilento (South of Naples) and Utrecht (The Netherlands).

All rights reserved. No parts of this book may be reproduced, in any form, without prior written permission of the author.
Prevention of Depression

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Geneeskunde
op dinsdag 16 januari 2007 om 13.45 uur
in het auditorium van de universiteit,
De Boelelaan 1105

door

Hans-Filip Engelbert Smit

geboren te Den Helder
promotoren: prof. dr. A.T.F. Beekman
prof. dr. W.J.M.J. Cuijpers
The plates at the beginning of each chapter represent St George – no doubt a historical figure who died in Lydda in 303, where his alleged tomb can be found. He became a figurehead of chivalry among the crusaders. This gave great impetus to the veneration of St George across Europe. He was adopted patron saint to England and Portugal, and to the Knights of the Garter under King Edward III. His well-known battle with the dragon is shrouded by fable, but might be traced to the myth of Perseus, and could, in the context of this thesis, be interpreted as an emblem of courageous determination, elsewhere called “mastery”, in the face of adversity.
Contents

Antithesis: A note to Albrecht Dürer’s Melancholia 9

General introduction 11

Part 1: Do we need it? 19
  1.1 Prevention of depression, rationale 21
  1.2 Costs of common mental disorders 27

Part 2: Do we know where to begin? 41
  2.1 Selecting key-variables for depression prevention 43
  2.2 Opportunities for cost-effective prevention of late-life depression 53
  2.3 Prevention of late-life depression in primary care, where to begin? 69

Part 3: Is it effective? 85
  3.1 A guided self-help intervention to prevent depression, randomised trial 87
  3.2 Psychological treatment of subthreshold depression, meta-analysis 99

Part 4: Is it affordable? 111
  4.1 Cost-effectiveness of preventing depression, randomised trial 113
  4.2 A risk-factor epidemiology of costs in mental health 125

Summary and general discussion 139
  5.1 Summary of preceding chapters 141
  5.2 General discussion 151
  5.3 Summary in Dutch 163

References 185

Acknowledgement 205

Curriculum vitae 207

Publications 209

Propositions 219

Jacket blurb 220
Prevention of Depression

“Melancholia” by Albrecht Dürer (1514)
Antithesis: a note to Albrecht Dürer’s Melancholia

As a whole, Dürer’s “Melancholia” (see title plate) is a somewhat disjointed composition; a mêlée of lavish allegory and harsh geometry. Nothing quite dominates the plate, and nothing lends it compositional unity. In fact, it is rather chaotic. However, the disharmony suggests hidden meaning and does so in a forlorn and forbidding way. It seems something is under construction, something will be revealed in due time, and yet, in the here and now, it resists decryption, both as a whole and in detail.

What do we see? The robed and winged figure of Melancholia sits at the corner of a building. We cannot see the face clearly. It is partly covered by a supporting hand, and for the remainder it is cast in shadows. Only two piercing eyes catch a flicker of light. The concentrated stare is unhappy in an unsentimental way. Observe this face. It is neither male nor female, and it looks strangely androgynous. It gazes at something we cannot see. The other hand of this figure rests on top of a closed book, and holds a drawing compass. Several keys dangle from a girdle. This suggests that something is closed, and can be opened, but we do not know what.

What else do we see? A greyhound lies on the ground, curled, starved and miserable. In fact, it looks rather less than a dog, and we find it difficult to recognise man’s most loyal companion. Next to Melancholia sits a putto on a millstone. The little cherub sits there as the pendant of a sinister bat, which hovers in the sky, and has “Melencolia I" written across its wings. These figures are surrounded by a number of geometrical forms, some tools, and cosmic phenomena. There is a globe, and the multifaceted shape of a rhomboid, both cut in solid stone. At the front we see tools for construction: a plane, a saw, a measuring rod, and some nails. Further to the back: pincers, a hammer, and rising against the building, a ladder. Above the head of the putto are scales. There is also an hourglass, a bell, and a magic square with numbers. In the bottom row we can read 1514, presumably the year when the engraving was made. The numbers are arranged as follows:

<table>
<thead>
<tr>
<th></th>
<th>16</th>
<th>3</th>
<th>2</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>10</td>
<td>11</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>7</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>14</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

and show, upon closer inspection, a pattern of almost maddeningly internal consistency. In the backdrop is a shimmering sea, a moonlit coastline, a nocturnal sky with the arc of a rainbow, and a meteor falling. The whole is illuminated by a dull light. The light, perhaps of the moon, falls from the upper right, and picks out the unconnected items one by one, and leaves them with barely a shadow. And there is the blinding, frontal, light of the ominous meteor. Caught in that black light is the pale arc of a rainbow.
Now, put this together, and think! It is neither day nor night: the meteor is a sign of the night, the rainbow of the day, and the two will not co-occur. This is an unworldly world, where the fluidity and animation of the allegorical figures stand out in sharp contrast to the stern shapes of geometry. This is life thrown out in an unconcerned, mathematical universe. This is *tedium vitae* pure and cold. This is a universe under partial construction – bare of harmony, without any connectedness – where even the distorted shape of the dog is hardly recognizable. But there is more. This is also the androgynous twin-cosmos of intellect and affect, of day and night, of male and female, of genius and foreboding, of the angelic *putto* and the demonic bat, the two pendants, the two aspects of the same world. It may look accidental and hapless, but there is method in this madness. Both the disparity, and the throwing together of the unconnected parts, create a world of new meaning. The numbers in the magic square add up to the same figure across the columns, across the rows, the diagonals, the four corners, and in the four corners. There is a pattern. There is encrypted meaning. The plate, as many have pointed out, may represent a spiritual self-portrait of Albrecht Dürer, one which links intellectual and artistic creation with melancholy, and makes them inseparable, as if the genius cannot, and will not, rise without depression.

This may be mere conjecture, but I fear the prevention of depression (never expected to be a complete success) may also invade this world of artistic and intellectual genius, this half-made world under permanent construction. Prevention of depression could perhaps save one precious thing, but at the expense of another. Let that serve as a warning for the remainder of this thesis.

**Literature**


Prevention of depression
General introduction
General introduction

Research questions
This thesis is about preventing depression. It is divided in four parts. Each part addresses one question. The questions are:

1. Do we need it?
2. Do we know where to begin?
3. Is it effective?
4. Is it cost-effective?

Clearly, these are common-sense questions. One may even find them unsurprising and simple. The amazing thing, however, is that the questions are rarely addressed jointly. However, the primary question, “Do we need prevention of depression?” cannot be answered without addressing the other questions. Assume for a moment that we have compelling reasons why we need depression prevention. However, if it is not effective, it remains a futile wish; and if it is not affordable, we have to replace this wish with another alternative that is economically more attainable. Accordingly, nothing comes of the original idea. In the same vein: we need to know where to begin with preventing depression, or else the whole venture will be reduced to a blindfold game of trial-and-error, which will be slow, painful and costly, and carries the risk that a possibly sound idea may be cast aside as impractical. Thus, all four questions are interrelated, like the numbers in Dürer’s Magic Square – and answering one question will require that the others also be addressed.

In the following sections we will first clarify the key-concepts, depression and prevention, then we will present the outline of this thesis and describe some of its limitations.

Depression
Depressive disorder is characterised by an abnormal depressed mood (dysphoria) and loss of pleasure (anhedonia). This blunted affect is present most of the day, nearly all days, for at least two weeks. The resulting lack of motivation can be quite crippling. In addition there are a number of other symptoms causing marked functional impairment, such as sleep disturbance (insomnia or hypersomnia), lack of energy (anemia), poor concentration, a lack or increase in appetite, inappropriate feelings of self-reproach, recurrent morbid thoughts about death and suicidal ideation (APA, 1987).

Depression often carries an unfavourable prognosis. A depressive episode lasts for six months on average (Kruijshaar et al, 2005) and in 20% of the cases it lasts longer than two years (Spijker et al, 2002). According to one estimate, there is a 85% probability that after recovery a new episode occurs in the five years following the index episode (Mueller et al, 1999), but other studies usually put recurrence rates somewhat lower (Fombonne et al, 2001; Keller et al, 1992;
Piccinelli and Wilkinson, 1994). Compulsively contemplating death and suicide is often symptomatic of this disorder, and makes suicide a real risk. About 60% of all suicides are committed by people who were depressed (Marquet et al, 2005), and mortality rates are higher by a factor of 1.65 in people with depression (Cuijpers and Smit, 2002).

It is not only a disabling condition, it is also highly prevalent. Annually 737,000 people suffer from depression in the Netherlands (Meijer et al, 2006). The impairment caused by this disorder and its high prevalence combine to create a substantial disease burden at national level, which is equal to an annual loss of 157,700 quality adjusted life years in the Dutch population alone (ibid). According to the World Health Organisation depression is one of the major causes of non-fatal disease burden world wide (WHO, 2000). Finally, it is well known that people with depression make more frequent use of health services and stay absent from their work more often, which has substantial economic ramifications (see Chapter 1.2).

**Why prevention of depression?**

Treatment of depressive disorder is as successful as any other medical treatment (Beekman et al, 2006). Nevertheless, there are two compelling reasons why we should like to also have preventive interventions for depressive disorder.

First, the annual influx of new (that is, first-ever) cases of depression is substantial. On an annual basis there are 358,500 new cases in the Netherlands, which is 49% of the prevalent cases. This suggests that it would be a sound idea to reduce this influx, which, in turn, requires primary prevention.

Second, curative interventions can only reduce the disease burden of depression (expressed as years lived with disability) to something in between 10% - 20% at population level, because not all cases are recognised as such, and when recognised not all will receive appropriate treatment or comply with the given treatment (Andrews et al, 2004; Chisholm et al, 2004).

From a public health point of view it seems therefore appropriate to reduce the disease duration by cure and to reduce the influx of new cases by primary prevention (see Chapter 1.1).

**Types of prevention**

At this point it might be well to introduce some terminology to describe prevention. Conventionally, prevention is divided into primary, secondary and tertiary prevention. Primary prevention is directed at avoiding new onsets, secondary prevention at early recognition and early intervention in people who have some preliminary symptoms of the disorder, while tertiary prevention aims to avoid impairments that may stem from a disease. It should further be noted that relapse prevention is aimed at avoiding relapse in remitting patients. Both primary and secondary prevention can also be described by a terminology introduced by Mrazek and Haggerty (1994). They distinguish between three types of prevention:
1. **Universal** prevention consists of interventions, often of a psycho-educational nature, directed at the whole population, regardless of risk status. The aim of universal prevention is to inform the general public about depressive disorder, how to recognise it, what people can do to avoid it, and how to improve the prognosis of its milder forms – and should the latter fail – to whom they can turn for help. Other forms of universal prevention may assume the form of screening programs, for example in schools or in the primary care setting. Universal prevention is important because it helps to create awareness in a population about the disorder, what can be done to prevent it, and what treatments are available. Sometimes it is thought that this awareness can act as a necessary catalyst to create a sense of readiness in a population to take a further step towards selective and indicated prevention, if so required.

2. **Selective** prevention is directed at population segments that are placed at a higher risk for depression because they have been exposed to risk factors known to be predictive of the onset of the disorder.

3. Finally, **indicated** prevention is directed at people who have some of the depressive symptoms, but do not meet the diagnostic criteria for the full-blown disorder. This group may be in an early disease stage, and is known to be at much higher risk of depressive disorder. It is these latter forms of – selective and indicated – prevention that will be discussed in this thesis. The idea is that people can make a transition from a relatively good to a worse health state. Presence or absence of risk factors and protective factors may promote or inhibit these transitions. Hence, conducting prevention needs to be based on knowledge of the relevant risk and protective factors. The aim of prevention is to contain the adverse effects of risk factors and to strengthen the effects of protective factors. In other words, the task at hand is to reduce risks for unfavourable transitions and improve prognosis in the face of adversity.

**Outline of the thesis**

The overall outline of the thesis can be summarised as follows. Chapter 1.1 introduces the rationale of preventing depression alongside some key statistics that underscore the importance, if not to say the necessity, of a much firmer role for prevention in public health. In Chapter 1.2 the economic costs of mental disorders are presented, providing, we hope, yet another incentive to attach greater value to the prevention of mental disorders, in particular that of depression. Both chapters help to answer the first question “Do we need it depression prevention?”

The next three chapters, 2.1 through 2.3, address the second question: “Do we know where to begin?” Here we employ epidemiological techniques and make use of large population-based psychiatric cohort studies: the Nemesis study, the Longitudinal Aging Study of Amsterdam (LASA) and the Amsterdam Study of the Elderly (Amstel study). The epidemiological studies are used to identify groups of people where prevention is most likely to yield substantial health gains with the
least effort and hence for the lowest costs. To put it in another way, the epidemiological data sets are used to generate hypotheses about the possible cost-effectiveness of prevention of depression.

As always, the proof of the pudding is in the eating, and this brings us to the questions: “Is prevention of depression effective and cost-effective?” To answer these questions we present the outcomes of a randomised prevention trial in Chapter 3.1 and a meta-analysis of similar trials in Chapter 3.2. These chapters explore and synthesise the first evidence that prevention of depression is possible and effective. Chapter 4.1 is based on the same trial, but presents data on the cost-effectiveness of preventing depression. Again, this study is one of the first of its kind.

In the preceding chapters we made a shift from population-based epidemiological studies (to generate hypotheses) to a clinical trial (to test one of the hypotheses). A parallel approach was adopted for the economic studies. We started with a population-based cost-of-illness study and concluded this line of investigation with a trial-based cost-effectiveness analysis. Thus, the shifts are from the general to the specific, from hypothesis generating to hypothesis testing, and follow two parallel courses, one epidemiological, the other economic.

In the final chapter, Chapter 4.2, the population-based epidemiological perspective is revisited in the form of a “risk-factor epidemiology of costs in mental health”. Here, former notions are brought together in a slightly new way, and we investigate the costs associated with risk factors for mental disorder. This chapter leads up to the General Discussion, which contains a summary of the main findings, followed by discussion, and directions for the future. I should like to mention that the latter section is partly borrowed from yet another study (Meijer et al, 2006), commissioned by the Netherlands Ministry of Health, to provide a scientific underpinning for a Policy White paper on the prevention of common mental disorders.

**Limitations**

This thesis has several limitations and it is only fair to point them out. First, it is restricted to the adult Dutch population, hence prevention of depression in children and adolescents will not be discussed.

Second, this thesis has a focus on selective prevention and to a larger extent an even more narrow focus on indicated prevention. In fact, our primary interest was to address the question how indicated prevention (in people with some early symptoms) could be made more restrictive by directing it towards ultra-high risk groups that do not only have some early symptoms, but are in addition exposed to a number of risk factors known to be predictive of the onset of depressive disorder. We took this angle for two reasons. Population segments having some depressive symptoms can be large, depending, of course, on the exact definition of “having some depressive symptoms”. It is hard, then, to see how all these people can be offered preventive interventions as limited economic resources will impose restrictions. Although we do not explicitly address the ethical question when
prevention is acceptable, ethic considerations played also a role in our decision to be restrictive. Selective prevention targets people who have been exposed to risk factors and have an increased risk of becoming depressed. Now, these people may have no knowledge about their risk status. Under those circumstances it might be inappropriate to offer them a (medical) intervention. After all, we cannot be sure that the disorder will occur without the intervention, and we can not be sure that the disorder will be avoided when the intervention is offered. In other words, some people in the target group will not benefit from the intervention, but will be confronted with a perhaps upsetting knowledge about their risk status. It is worth noting that these ethical considerations may take a different turn when it comes to indicated prevention. Now people have complaints and are likely to be aware about their condition. In fact, they may be looking for help or willing to accept help when it is offered.

Our approach to these logistic and ethical issues is to require that prevention is preferably directed at groups that have already some symptoms and are in addition exposed to risk factors indicative of an ultra-high calculable risk of depression. With the risk to create some terminological confusion: we wanted indicated prevention to become more selective. Hence the narrow focus.

Third, we should mention that this thesis represents work in progress. At one point prevention of depression in older people is discussed, at another we refer to younger adults. In other words, there are gaps, disconnected quantum leaps, and the answers to the research questions cannot be considered complete. More work remains to be done, and more work is being carried out. This thesis, then, is only a snapshot of the approaches we took to answer questions 1 through 4. Its relevance is that the approaches described in this thesis can be used for addressing similar questions about other mental disorders, in other age groups. Indeed, it gives me great personal satisfaction to see that some of the research lines are being extended to other fields, such as anxiety disorder, problem drinking, and to younger age groups.

One final issue needs to be addressed here. There is no thesis without an antithesis. That is why an extra section has been inserted. This section is formed by the title plate, consisting of Albrecht Dürer’s Melancholia, and a brief reflection on its possible meaning. The role of the antithesis is to make explicit that this thesis takes only one angle towards preventing depression. This is to say, that other views may be adopted, and must be adopted, in order to well and truly open up the debate.
Part I
Do we need prevention of depression?

Chapter 1.1
Prevention of depression, rationale

This chapter is based on:


Of related interest:


Chapter 1.1

Prevention of depression, rationale

1.1.1 Introduction

In this chapter some basic, and therefore fundamental, concepts, will be introduced about depression and its prevention. First, depression will be briefly described, as will its prognosis, and its adverse health consequences at individual level. Then, the individual level concepts will be generalised to population level. At this level, epidemiologic and public health perspectives come into play. In tandem with the introduction of the key-concepts some facts and figures about depression will be presented. Against this background notions about prevention of depression and its rationale are presented.

1.1.2 Depression at individual level

*Depression*

Depression is a disabling disorder, characterised by depressed mood (dysphoria) and loss of interest (anhedonia) during the larger part of the day, most days, during at least two weeks. In addition there are several symptoms, such as lack of energy, insomnia, loss or gain of body weight, preoccupation with guilt, poor concentration, and recurrent thoughts about death and suicide. Each of these symptoms brings about significant suffering and has an adverse impact on functioning. In order to meet the diagnostic criteria for depressive disorder, one must have one or both core symptoms, dysphoria and anhedonia, plus at least four other symptoms (APA, 1987). Throughout this thesis, attention will be restricted to unipolar depressive disorder (major depression) and its subthreshold manifestation (minor depression). We define the latter only loosely when some of the symptoms of depressive disorder are present, but do not (yet) meet the diagnostic criteria for depressive disorder. More rigorous criteria for minor depression can be found in the Research Diagnostic Criteria in the Appendix of the DSM-IV. The focus on depression is taken to the exclusion of other depressive conditions such as dysthymia and bipolar disorder.

*Course and prognosis*

Figure 1.1.1 depicts, in a schematic fashion, how a person can make transitions between levels of caseness. In brief, a symptom-free person can acquire symptoms, and become symptomatic. The condition may further deteriorate to such an extent that the diagnostic criteria for depressive disorder are met. After an episode of depressive disorder a person may progress towards remission, recover, or relapse.
before recovery. It should be noted that recovery can be complete or partial. After recovery a new episode of depression may occur, which is then called recurrence.

At this point it is worth noting that prevention will have to occur before disease onset, when a person is either symptom free, or when there are some symptoms that do not yet meet the diagnostic criteria of depressive disorder. The aim, then, is to try to avoid disease onset. By contrast, treatment will be directed toward people who already have the disease, and its objective is to promote recovery.

**Figure 1.1.1** Course of depression across symptom levels

<table>
<thead>
<tr>
<th>Symptom free</th>
<th>Symptomatic</th>
<th>Syndromatic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>onset</td>
<td>acute episode</td>
</tr>
<tr>
<td></td>
<td>remission</td>
<td></td>
</tr>
<tr>
<td></td>
<td>relapse</td>
<td></td>
</tr>
<tr>
<td></td>
<td>recovery</td>
<td></td>
</tr>
</tbody>
</table>

It has already been mentioned, but Figure 1.1.1 is a simplification of how people move between symptom levels. When taking the life-course perspective, it is often seen that depression has a remittent course, with recurrences, partial recoveries, and a certain amount of time spent in the twilight zone of having some depressive symptoms but not the disorder. All in all, depression carries a poor prognosis. According to estimates of the World Health Organisation, the mean episode duration is six months (Üstün et al, 2004). According to Dutch data about people aged 18 – 65 years, the average time spend in this condition is also six months (Kruijshaar et al, 2005), and 20% will be depressed for longer than two years (Spijkers et al, 2002). After recovery, the risk of recurrence is high: 85% will experience a new episode within five years after the index episode (Mueller et al, 1999), but this might be an over-estimate (cf. Fombonne et al, 2001).

**Quality of life in depression**
Depression compromises quality of life (Üstün et al, 2004). Depression has a disease burden comparable with, for example, near blindness (Van Oers et al, 2002). In the Netherlands, one year spent in depression is valued as six-and-a-half months in good health. The remaining five-and-a-half months (46%) are regarded as
“lost”, that is, as if the person had not lived at all, and was dead (Stouthard et al, 1997; Meijer et al, 2006). This gives depression, compared with other illness, a large disability weights.

The disability weight is a component of the well-know disability adjusted life year (DALY), to be discussed later. Disability weights are placed on a scale of 0 to 1, where 0 (no disability) is set equal to good health, and 1 (extreme disability) is set equal to death, and value the years lived with disability. In the calculation of disability weights of depression, three severity levels of a depressive episode are considered: mild, moderate and severe, having disability weights of 0.14, 0.35 and 0.76, respectively (Stouthard et al, 1997). The mean disability weight is 0.46, taking into account the relative prevalence rates of each of the severity levels the Dutch population (Meijer et al, 2006). According to a study of the World Health Organisation the mean disability weight of depressive disorder is 0.48 (Üstün et al, 2004), only marginally larger than the weight currently in use in the Netherlands.

The DALY is a metric that combines both morbidity (time spent in ill-health), and mortality (life years lost through case fatality), because life years lost to premature death can also be included. In the calculation of the DALY of depression the average time spent in a depressive episode assumes the dominant role, and outweighs the much smaller effect of premature death. Only the elevated risk of suicide, estimated at 6% in depressed people on a life-time basis, is incorporated in the DALY (cf. Üstün et al, 2004; Chisholm et al, 2004).

Elsewhere, we have criticised the way in which the calculation of DALY of depressive disorder fails to more comprehensively accommodate excess mortality due to depression. In a meta-analysis of community-based studies, we found that people who suffer from depression are placed at a high risk of premature death. Their risk of dying is a factor of 1.65 higher than in non-depressed people (Cuijpers and Smit, 2002). Although 60% of the people who complete suicide meet the diagnostic criteria for depressive disorder (Lönnqvist, 2000; Marquet et al, 2005), the excess mortality associated with depression is not solely related to suicide. Other causes of premature death in depressed people are probably attributable to their less healthy life styles, their larger susceptibility to diseases, and poorer recovery rates once they are ill. In an early study, Tsuang and colleagues (1980) estimated that depression-related fatality shortens life expectancy by seven years. These aspects of case fatality are not well incorporated in the current calculations of DALYS for depression, where only suicide is seen as a factor contributing to mortality.

1.1.3 Depression at population level

Incidence, prevalence and recurrence

Figure 1.1.2 depicts the main flows at population level that were first described at individual level. The influx, or incidence, of new cases into the “reservoir” of
prevalent cases is a function of the birth rate in a population, which places new cohorts of people at risk of becoming a case of depression (first-ever incidence). Incidence can also be a function of people who have been depressed, recovered, and then experience a new episode of depression (repeat-incidence). Prevalence is a function of the influx of all new cases multiplied by their time spent in the depressive condition. Once people are recovered they are at risk of experiencing a new episode of depressive disorder, and may thus become a case of recurrent depression. It should perhaps be noted that Figure 1.1.2 is a simplification in that we have not depicted another recursive loop, that of relapse before recovery.

**Figure 1.1.2** Incidence, prevalence, recovery, and recurrence at population level

![Diagram of incidence, prevalence, recovery, and recurrence at population level](image)

From the public health perspective it is important to recognise the fact that depression is characterised by a substantial influx of new cases. In the Netherlands, the annual influx of new cases (first-ever incidence) of depression is 289,000 people in the age group 18-65 years (Bijl et al, 2002). For comparison: the annual number of prevalent cases in the same age group is 589,000 (Bijl et al, 1998). Therefore, a formidable 49% of the prevalent cases are, in fact, new cases.

Conventionally, treatment is directed at prevalent cases, and does not influence the number of first-ever cases, but it does influence the time spent in a depressive condition. The objective of treatment is to shorten episode duration and thus to accelerate recovery. In other words, treatment helps to compress the “reservoir” of prevalent cases in Figure 1.1.2 across the time-dimension, but has no impact on first-ever incidence. Given the substantial influx of first-ever cases, it is questionable if solely relying on treatment is the most appropriate organisation of public mental health. It may be better to tackle both the width and the breadth of the problem, that is, not only to shorten episode duration, but also to reduce the influx of new cases. For the latter, prevention is needed.

**Disease burden**

Depression is not only disabling at individual level, but also highly prevalent at population level. In the Netherlands alone it affects 738,000 people older than 16 years on an annual basis (Meijer et al, 2006). The high prevalence coupled with the
substantial disability weight and the long mean duration makes depression one of the illnesses ranking highest in terms of burden of disease at population level. Expressed in disability adjusted life years (DALYs) depression is responsible for an annual loss of 157,700 DALYs in the Netherlands (Meijer et al, 2006). This makes depression one of the leading causes of disability in the Netherlands.

The same picture emerges at global level. In the World Health Organisation’s roster of disorders responsible for the largest disease burden, depressive disorder ranks number one for women and two for men aged 15 to 45 years, and is placed at a position far above all injuries inflicted in road accidents and in war combined (WHO, 2000; Murray and Lopez, 1996; Üstün, 1999). When men and women are considered together and depression-related disability is studied across all age groups, then depression assumes the fourth position of illnesses leading with the largest disability (Üstün et al, 2004). In Europe, where infectious diseases are not as prevalent as in several other parts of the world, depression comes only after ischaemic heart disease and cerebrovascular disease (ibid). Again, the ranking of depression as a disabling condition may be biased downward because the disability weight of depressive disorder fails to take into account depression-related causes of mortality other than suicide, but this may not much affect its position which is already high up in the hierarchy of most disabling illnesses.

From the viewpoint of public health there is yet another compelling reason why we may have to look for prevention. In a seminal study Andrews and colleagues (2004) showed that treatment of depression can at best avert 34% of the years lived with disability (YLD) due to depression. However, this result would only be obtained under a hypothetical scenario where all people meeting the diagnostic criteria of depressive disorder are recognised as such, and receive optimal, that is empirically supported, treatment. Thus even in the best possible scenario consisting of 100% coverage and 100% adherence by therapists to evidenced-based guidelines, this would leave a formidable gap between the need of many and what curative interventions can offer. In fact, important fractions of those suffering from depression will go unrecognised, while many of the people recognised as cases of depression will receive inadequate therapies, or the therapies will just fail to produce the desired effects (Ormel and Tiemens, 1997). By implication, the gap is likely to be larger than 66%.

In the same study Andrews and colleagues arrive at a more realistic estimate, which is based on the current health care regimes in Australia. According to these calculations, only 22% of the depression-related YLDs are averted in actual practice. Chisholm et al (2004) reach a similar conclusion, and indicate that 10 – 30% of the depression-related YLDs can be averted by available treatments, assuming rates of 50% for coverage and 70% for adherence. In this more realistic scenario the gap of unmet needs falls in the range of 70 – 90%. Chisholm and colleagues conclude that it is evident that the current curative approaches toward depression have a limited impact at population level. They suggest that, when possible and where possible, improvements be made in the coverage and adherence
rates, and in the effectiveness of treatments, but also that preventive strategies be developed. One such preventive strategy, they say, could consist of proactive collaborative care with the aim not only to provide episodic treatment, but also to offer long-term maintenance therapy for those at risk of recurrences (cf. Katon et al, 2001). We second these recommendations, but like to go one step further (Smit et al, 2003; Beekman et al, 2006). Cure will forever be the core-business of psychiatry and related disciplines, but notions about the role of treatment in public health are shifting and expanding. Relapse prevention has been accepted as good clinical practice in the treatment of depression, and its place as an integral part of treatment is without much controversy. In recognition of the often remittent course of depressive disorder, a proactive collaborative form of maintenance therapy for recurrent cases can be viewed as a further and important improvement, because this approach yields a considerably larger population-level health gain than episodic treatments (cf. Chisholm et al, 2004; Katon et al, 2001). In the visualisation of Figure 1.1.2 this approach would help to reduce the flow that loops from recovery back to the incidence of yet another episode. However, in view of Figure 1.1.2 it may still be interesting to also develop strategies for curtailing the formidable annual influx of first-ever cases of depression. We have said it before, but this major public health problem should be tackled along its breadth and width: the reservoir of prevalent cases should be reduced, not only by promoting recovery, but also by reducing the risk to become a case of depressive disorder in the first place.

1.1.4 Conclusion

In this chapter it was shown that depressive disorder is associated with human suffering on a formidable scale. From a public health point of view, curative treatment is necessary, but it can not be expected to lessen the depression-related disease-burden in a substantial way. A formidable gap remains to be bridged. Moreover, the annual influx of new cases is very substantial, which casts a dubious light on the rationality of overly relying on curative treatments. That would be tantamount to the proverbial mopping of the floor under a running tap. The question, therefore, is whether preventive strategies can be developed in a way that makes them at once acceptable, effective, and affordable. When the answer is in the affirmative, then it may be worthwhile to give room to a much firmer role of primary prevention within the realm of public mental health.
Part I
Do we need prevention of depression?

Chapter 1.2
Costs of common mental disorders

This chapter is based on:


Of related interest:

Chapter 1.2

Costs of common mental disorders

1.2.1 Introduction

It is well known that mental disorders are highly prevalent (Kessler, 1994; Bijl et al, 1998b; ESEMed, 2004a), and are associated with a substantial disease burden (Murray and Lopez, 1994, WHO, 2000; Ustun, 2004), but their economic costs have been relatively less well researched (Berto et al, 2004; Greenberg and Birnbaum, 2005; Östergren and Olesen, 2004; Knerer et al, 2005; Marciniak et al, 2004). Yet, cost-of-illness studies can provide some guidance in debates on health service provision, resource allocation, and in particular in informing and setting research agendas for cost-effectiveness studies. These are reasons to conduct a comprehensive cost-of-illness study of common mental disorders in the general population. The studied disorders are depression, dysthymia, panic disorder, agoraphobia, generalized anxiety disorder, social phobia, simple phobias, alcohol abuse, and alcohol dependence. The aim of this study is to calculate the per capita costs of the disorders, and to compare the costs attributable to prevalent and incident cases at population level. Implications for curative and preventive psychiatry are discussed.

1.2.2 Methods

Sample

This cost-of-illness study was conducted alongside the Netherlands Mental Health Survey and Incidence Study (NEMESIS), a large-scale, population-based, psychiatric epidemiological study in which data were collected on DSM-III-R axis-I disorders, medical consumption and work-loss days (Bijl et al, 1998a).

At baseline (t₀), a random, stratified, multistage sample was obtained in three steps. First, municipalities were stratified by urbanization, and 90 municipalities were drawn randomly and proportionately from these strata. Second, within each municipality, households were randomly drawn from the postal register. Within each household the person with the most recent birthday was selected on condition that he or she was between 18 and 65 years old and sufficiently fluent in Dutch to be interviewed. Eligible persons, who were not immediately available, were contacted later in the year. The response rate was 69.7% resulting in a sample of 7,076 people at t₀. The baseline sample followed the same distribution over age, sex, civil status and urbanity as the general Dutch population. Only the male 18-24 age group was slightly underrepresented (11).
At first follow-up \((t_1)\), which occurred one year \((M=379 \text{ days}, \text{Sd}=35)\) after base-line, 5,618 persons (79.4%) continued their participation. We evaluated the effect of attrition from \(t_0\) to \(t_1\) and found that it was not related to having a disorder at \(t_0\) but again, the younger males were more likely to be lost to follow-up (De Graaf et al, 2000). The analyses were based on the \(t_1\) sample because medical consumption was measured at \(t_1\).

In this paper the focus is on the common mental disorders, defined here as disorders with a 1-year prevalence above 1%. The following disorders were included: major depressive disorder, dysthymia, panic disorder with and without agoraphobia, agoraphobia without a history of panic, generalised anxiety disorder, social phobia, simple phobia, alcohol abuse, and alcohol dependence. Mental disorders with a 1-year prevalence below 1% such as bipolar disorder, obsessive-compulsive disorder, eating disorders, and schizophrenia were excluded (114 cases). This resulted in an effective sample size of \(N=5,504\).

**Measures**

DSM-III-R axis-I disorders (American Psychiatric Association, 1994) were assessed in both waves with the Composite International Diagnostic Interview (CIDI, the Dutch 1.1 computerized version; Ter Smitten et al, 1998). The CIDI is a psychiatric interview and can be used by trained interviewers who are not clinicians. It is known to have an acceptable interrater and test-retest reliability (Wittchen, 1994).

A list of 30 chronic medical conditions was compiled, which was based on the Health Questionnaire of Statistics Netherlands (Central Bureau of Statistics, 1989). Respondents could reply to this list indicating the presence (or absence) of such a medical condition in the past 12 months.

Socio-demographic characteristics were assessed in structured face-to-face interviews. The following variables were used: gender, age (in years), education, living with a partner, born in the Netherlands or elsewhere, and employment status.

**Methodological outline**

There are two types of cost-of-illness studies: top-down and bottom-up studies.

In top-down studies, the national expenditure on health services is attributed to the recipients of the services, and when their diagnosis is known then each diagnosis can receive a price tag. In bottom-up studies the starting point is a population-based survey in which participants receive a psychiatric interview to ascertain their diagnostic status. In a next step, their medical consumption is measured as well as their work loss days. Both the medical consumption and productivity losses are then monetarily valued, and the costs can be linked to diagnoses.

The bottom-up approach has several advantages. First, in bottom-up studies the costs of production losses can be included and these costs typically represent the bulk of all costs, but are usually missed in top-down studies. Second, bottom-up
studies can provide information on people who meet the diagnostic criteria of a disorder, but are not recipients of professional help. In top-down studies these people, and they form the vast majority, are missed. These people generate fewer costs in health care, but may generate substantial costs due to absenteeism from work. Finally, bottom-up surveys use a uniform and standardized system of reliably ascertaining the diagnostic status of subjects. By contrast, in top-down studies one has to rely on the diagnostic accuracy of registers of general practitioners, hospitals and so on. The diagnostic accuracy of these registers can often be questioned and their reliability may vary considerably. These were reasons to conduct a comprehensive, bottom-up cost-of-illness study. The time frame of this study is restricted to this single year. Therefore, we did not correct for inflation and did not discount costs. All costs are expressed in euro (€) for the reference year 2003. For conversion to US$, one could make use of Purchasing Power Parities, which are currency conversion rates that both convert currency and equalize the purchasing power of different currencies. The Organization for Economic Co-operation and Development (OECD) equates US$ 1.00 to € 0.920 in The Netherlands in 2003 (downloadable from: www.oecd.org/std/ppp).

Resource use and costing

Conceptually, medical costs can be distinguished from direct non-medical costs and indirect non-medical costs. We will describe each type of the cost types in turn.

Direct medical costs are the costs of health service resource use. Information on the subjects’ use of health services was obtained with a prototype version of the Trimbos and Institute of Medical Technology Assessment Cost Questionnaire for Psychiatry (Hakkaart - Van Roijen et al, 2002). With this questionnaire subjects register the number of GP visits, sessions with psychiatrists, hospital days, etc. In a next step, medical resource use was costed by multiplying the number of health service units (consultations, hospital days, etc.) by their integral cost price (Oostenbrink et al, 2004), see Table 1.2.1. To these we added the costs of pharmacological interventions as the cost price per standard daily dose (obtained from the Pharmaceutical Compass at www.fk.cvz.nl), plus 6% VAT, multiplied by the number of prescription days, plus the pharmacist’s dispensing costs of € 6.45 per prescription.

Direct non-medical costs arise when patients travel to health service providers, and pay for parking. In the Netherlands, the average travel distance between an address and a general practitioner’s practice is 1.8 km. Similarly, travel distances to other health services are also known (ibid.). This information was combined with the information about actual health service uptake. Travel distances were valued at € 0.16 per km and 1 hour parking time was valued at € 2.50. To this we added the costs of the patients’ time spent in travel, waiting and in treatment at € 8.30 per hour, see Table 1.2.1.
### Table 1.2.1 Direct medical and direct non-medical costs by health service type

<table>
<thead>
<tr>
<th>Health service type</th>
<th>Direct Medical Costs, € in 2003</th>
<th>Direct Non-Medical Costs, € in 2003</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>unit</td>
<td>cost price a</td>
</tr>
<tr>
<td>Medical doctor</td>
<td>Consult</td>
<td>20.20</td>
</tr>
<tr>
<td>Medical specialist</td>
<td>Consult</td>
<td>98.00</td>
</tr>
<tr>
<td>Regional mental health service</td>
<td>Contact</td>
<td>124.00</td>
</tr>
<tr>
<td>Regional addiction service d</td>
<td>Contact</td>
<td>124.00</td>
</tr>
<tr>
<td>Mental Hospital – Outpatient</td>
<td>Consult</td>
<td>88.00</td>
</tr>
<tr>
<td>Mental Hospital – Day care</td>
<td>Contact</td>
<td>125.00</td>
</tr>
<tr>
<td>Mental Hospital – Inpatient</td>
<td>Day</td>
<td>250.00</td>
</tr>
<tr>
<td>General Hospital – Outpatient</td>
<td>Consult</td>
<td>56.00</td>
</tr>
<tr>
<td>General Hospital – Day care</td>
<td>Contact</td>
<td>229.00</td>
</tr>
<tr>
<td>General Hospital – Inpatient</td>
<td>Day</td>
<td>337.00</td>
</tr>
<tr>
<td>Teaching Hospital – Outpatient</td>
<td>Consult</td>
<td>100.00</td>
</tr>
<tr>
<td>Teaching Hospital – Day care</td>
<td>Contact</td>
<td>229.00</td>
</tr>
<tr>
<td>Teaching Hospital – Inpatient</td>
<td>Day</td>
<td>476.00</td>
</tr>
<tr>
<td>Private practice psychotherapist</td>
<td>Session</td>
<td>76.00</td>
</tr>
<tr>
<td>Social worker e</td>
<td>Contact</td>
<td>45.00</td>
</tr>
<tr>
<td>Physiotherapist</td>
<td>Contact</td>
<td>22.75</td>
</tr>
<tr>
<td>Home care</td>
<td>Hour</td>
<td>30.70</td>
</tr>
<tr>
<td>Informal care (family, friends) f</td>
<td>Hour</td>
<td>8.30</td>
</tr>
</tbody>
</table>

a Integral unit cost prices (Oostenbrink et al, 2004).
b Based on average distances (in km) and travel + waiting + treatment times (in hrs) for receiving treatment (cf. Oostenbrink et al, 2004)
c Costs of 1km = € 0.16, parking = € 2.50, 1h time = € 8.30 (Oostenbrink et al, 2004).
d Valued as outpatient mental health services.
e From DFL 77.00 in 1993, converted into Euro, indexed for 2003 (cf. www.cbs.nl) and rounded.

Indirect non-medical costs arise when production losses occur due to illness. Subjects were asked about the number of days spent in illness. These days were distributed proportionally over workdays (resulting in production losses due to work loss days) and days off work (resulting in production losses in the domestic sphere). To valuate a lost day in a paid job we used the age and gender specific monetary counter-value of production losses that occur during absence from work (ibid.), see Table 1.2.2. People may also be too ill to perform domestic tasks. These costs were valued at € 8.30 per hour, which corresponds with the price of domestic help.

The cost calculations were conducted in accordance with the latest Dutch guideline for health economic evaluations (ibid.), which closely resembles other international guidelines (Langley, 1995; Siegel et al, 1997; Torrance et al, 1996).
We had to depart from the Dutch guideline in one important respect. Instead of using the recommended friction cost method, we used the human capital method for calculating the costs of production losses. In the former the costs of production losses are restricted to a time period that would be necessary to replace a person with new labour, whereas in the human capital method the costs of productivity losses last as long as the time a person stays absent. Thus the estimated costs will be higher in the human capital method as compared with the friction cost method, especially in the presence of long-term absenteeism and when the time-horizon of the economic study is long. The fact that we made use of the human capital method was motivated by a lack of data-availability: we had no data on disease duration and therefore could not make the more complex calculations underlying the friction cost method.

<table>
<thead>
<tr>
<th>Age</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-24</td>
<td>20.49</td>
<td>20.07</td>
</tr>
<tr>
<td>25-34</td>
<td>32.74</td>
<td>29.88</td>
</tr>
<tr>
<td>35-44</td>
<td>40.86</td>
<td>33.60</td>
</tr>
<tr>
<td>45-54</td>
<td>45.37</td>
<td>34.21</td>
</tr>
<tr>
<td>55-64</td>
<td>47.82</td>
<td>36.41</td>
</tr>
<tr>
<td>65+</td>
<td>47.82</td>
<td>36.41</td>
</tr>
</tbody>
</table>

*Oostenbrink et al. (2004).*

**Analysis**

Multivariate regression analyses were performed to partial out the excess costs of each disorder while adjusting for the presence of other mental disorders and somatic disorders. In other words, the distinct types of costs (direct medical, direct non-medical, and indirect cost) and their total were regressed on the selected disorders, and on the presence of somatic conditions.

To account for initial non-response and drop-out, corrective post-stratification weights were used. After weighting, the sample followed exactly the same multivariate distribution over age, sex, civil status and urbanization as the population according to Statistics Netherlands (http://www.cbs.nl).

To account for the possible non-normality of the cost data, sample errors, 95% confidence intervals, and P-values were based on 1,000 bootstrap replications, while at each bootstrap step robust sample errors were obtained using the first-order Taylor-serieslinearization method. This was done to obtain correct 95% confidence intervals, and P-values.

For Table 1.2.3, the total per capita costs were regressed on the disorders, while adjusting for somatic illnesses. This was done following the procedure as outlined above.

For Table 1.2.4, the per capita cost estimates were used, and then projected on the population, taking into account the per capita costs and the prevalence rate of
each disorder. The prevalence rates of the distinct disorders were based on the same data-set (Bijl et al, 1998). In a third step the costs were calculated that are generated by the annual influx of new cases (incidence). The incidence rates were also based on the same dataset (Bijl et al, 2002). All analyses were conducted with Stata version 7.0/SE (StataCorp, 2001)

1.2.3 Results

Demographics
The sample (N= 5504) can be described as follows. 49% were women with a mean age of 40 years (range: 18 – 65) and 69% were living with a partner. The distribution over levels of education was as follows: elementary 5%, lower vocational 35%, secondary 30%, higher vocational and academic 30%. Of the sample, 94% was born in the Netherlands and 70% was employed. We refer to Table 1.2.4 for the prevalence and incidence rates of the common mental disorders.

Per capita costs
The annual per capita excess costs of having any common mental disorder are, on average, € 3,200 (SE= 563; 95% CI= 2117 – 4284). This is in the same order of magnitude as the somatic disorders included in this study, which average € 3,220 (SE= 245; 95% CI= 2740 – 3702). It is worth noting that the costs should be interpreted as excess costs, because every person generates, on average, an annual ‘base-rate’ of € 1,025 (SE= 138, 95% CI= 754 – 1297) irrespective of having a mental or somatic disorder. The ‘base-rate’ costs are generated by minor illnesses and injuries not related to the disorders under consideration. When a person experiences a mental disorder, then these costs are added to the base rate costs. Thus, the annual excess costs of a mental disorder are € 3,200, but a person suffering from it will generate € 1,025 + € 3,200 = € 4,225 in a given year.

Break-down of the per capita excess costs
The annual per capita excess costs of € 3,200 can be divided into three categories. First, there are direct medical costs, which are related to health service uptake. For any of the mental disorders, these amount on average to € 300 (SE= 79, 95% CI= 145 – 456), equivalent to 9.4% of the total costs. Second, there are the direct non-medical costs of € 179 (SE= 50, 95% CI= 80 – 278), representing 5.6% of the total costs. Finally, the bulk of the costs are due to production losses, the so called ‘indirect non-medical costs’ of € 2,725 (SE= 574, 95% CI= 1,601 – 3,849). Their share in the total costs is 85.2%.

It should perhaps be noted that the sum of different costs components is not exactly, but close to, the total costs. This discrepancy is due to the fact that a certain number of people make use of health services while another number is employed and can generate production losses in paid labour. Hence, the total costs are a sum
of the component costs weighted for the appropriate number of people, and not a simple addition.

**Per capita excess costs of the distinct disorders**

Table 1.2.3 further presents the annual per capita excess costs of the distinct disorders. The disorder which generates the highest costs is dysthymia (€ 10,332), followed by panic disorder (€ 8,390) and agoraphobia (€ 4,879). Of about the same magnitude are major depression (€ 2,278), simple phobia (€ 2,372) and alcohol dependence (€ 2,509). It is worth noting that some direct medical and non-medical excess costs, for example those of social phobia, are negative. Apparently, people with social phobia tend to make less use of medical services for their condition when controlling for other disorders – but this is all speculative, because these small amounts are not statistically significant.

**Table 1.2.3** Annual per capita excess costs (in €) by DSM-III-R disorder (weighted analysis, N= 5491)

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Direct medical</th>
<th>Direct non-medical</th>
<th>Indirect costs</th>
<th>Total costs</th>
<th>95% CI of total costs a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any disorder</td>
<td>300</td>
<td>179</td>
<td>2,725</td>
<td>3,200</td>
<td>2,117 - 4,284</td>
</tr>
<tr>
<td>Mood disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>. Depression</td>
<td>459</td>
<td>264</td>
<td>4,297</td>
<td>5,009</td>
<td>3,016 - 7,001</td>
</tr>
<tr>
<td>. Dysthymia</td>
<td>431</td>
<td>242</td>
<td>1,619</td>
<td>2,278</td>
<td>331 - 4,225</td>
</tr>
<tr>
<td>Anxiety disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>. Panic disorder</td>
<td>269</td>
<td>215</td>
<td>3,100</td>
<td>3,587</td>
<td>1,883 - 5,291</td>
</tr>
<tr>
<td>. Agoraphobia</td>
<td>270</td>
<td>629</td>
<td>7,486</td>
<td>8,390</td>
<td>2,772 - 14,008</td>
</tr>
<tr>
<td>. Social phobia</td>
<td>980</td>
<td>640</td>
<td>3,228</td>
<td>4,879</td>
<td>-514 - 10,272</td>
</tr>
<tr>
<td>. Simple phobia</td>
<td>-140</td>
<td>-222</td>
<td>2,210</td>
<td>1,848</td>
<td>-1,648 - 5,343</td>
</tr>
<tr>
<td>. Gen. anx. dis.</td>
<td>300</td>
<td>20</td>
<td>2,054</td>
<td>2,372</td>
<td>167 - 4,578</td>
</tr>
<tr>
<td>Alcohol-related</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>. Alcohol abuse</td>
<td>150</td>
<td>31</td>
<td>1,247</td>
<td>1,431</td>
<td>-194 - 3,057</td>
</tr>
<tr>
<td>. Alc. dependence</td>
<td>70</td>
<td>16</td>
<td>836</td>
<td>923</td>
<td>-887 - 2,733</td>
</tr>
<tr>
<td>Somatic illness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base rate b</td>
<td>6</td>
<td>11</td>
<td>1,033</td>
<td>1,049</td>
<td>784 - 1,315</td>
</tr>
</tbody>
</table>

*95% confidence intervals are based on robust sample errors and 1,000 non-parametric bootstraps
b Base rate for the regression model with all the distinct disorders. The base rate differs across models.

**Excess costs per one million prevalent cases**

At population level, the costs of a disorder are the annual per capita costs multiplied by the number of prevalent cases of that disorder in a year. These calculations were conducted per one million people in the age group of 18-65 years. As can be seen in
Table 1.2.4, dysthymia, panic disorder, simple phobia, depression and alcohol dependence generate the bulk of the costs at population level. In comparison with the per capita costs, the costs at population level have a somewhat different hierarchy: although the per capita costs of mood disorders are higher than those of the anxiety disorders, this ranking is reversed when taking into account the prevalence rates: being more prevalent, the anxiety disorders generate more costs.

Table 1.2.4  Annual excess costs per capita, and attributable to prevalent and incident cases per 1 million population, adjusted for somatic illnesses.

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Per capita costs</th>
<th>Prevalence (in %)</th>
<th>Prevalent costs (in mln. €)</th>
<th>Incidence (per 100 pyrs)</th>
<th>Incidence costs (in mln. €)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any disorder</td>
<td>3,200</td>
<td>20.9</td>
<td>669</td>
<td>8.20</td>
<td>262</td>
</tr>
<tr>
<td>Mood disorder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>5,009</td>
<td>6.2</td>
<td>311</td>
<td>2.09</td>
<td>105</td>
</tr>
<tr>
<td>Dysthymia</td>
<td>2,278</td>
<td>5.8</td>
<td>132</td>
<td>2.72</td>
<td>62</td>
</tr>
<tr>
<td>Anxiety disorder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panic disorder</td>
<td>10,322</td>
<td>2.3</td>
<td>237</td>
<td>0.38</td>
<td>39</td>
</tr>
<tr>
<td>Agoraphobia</td>
<td>3,587</td>
<td>11.3</td>
<td>405</td>
<td>2.65</td>
<td>95</td>
</tr>
<tr>
<td>Social phobia</td>
<td>8,390</td>
<td>2.2</td>
<td>185</td>
<td>0.78</td>
<td>65</td>
</tr>
<tr>
<td>Simple phobia</td>
<td>4,879</td>
<td>1.6</td>
<td>78</td>
<td>0.77</td>
<td>38</td>
</tr>
<tr>
<td>Gen. anx. dis.</td>
<td>2,372</td>
<td>7.1</td>
<td>168</td>
<td>2.20</td>
<td>52</td>
</tr>
<tr>
<td>Alcohol-related</td>
<td>917</td>
<td>1.2</td>
<td>11</td>
<td>0.72</td>
<td>7</td>
</tr>
<tr>
<td>Alc. abuse</td>
<td>1,431</td>
<td>7.6</td>
<td>109</td>
<td>1.56</td>
<td>22</td>
</tr>
<tr>
<td>Alc. depend.</td>
<td>923</td>
<td>11.7</td>
<td>108</td>
<td>2.38</td>
<td>22</td>
</tr>
<tr>
<td>Alc. depend.</td>
<td>2,509</td>
<td>5.5</td>
<td>138</td>
<td>0.49</td>
<td>12</td>
</tr>
</tbody>
</table>

2 Incidence rates (cf. Bijl et al. 2002).

Excess costs per one million incident cases

The ranking is different again when the focus is on the incidence. Then panic disorder comes first, followed by depression, simple phobia, and dysthymia, while alcohol dependence has lost its fifth position in the hierarchy to agoraphobia and social phobia.

Interestingly, the prevalent cases of common mental disorder generate € 669 million per 1 million population, while among them the new (incident) cases are responsible for € 262 million, which represents 39.2% of the costs. The contribution of the new cases to the total cost differs per type of disorder. The incident cases assume a share of one-fifth of the alcohol-related disorders, one-quarter of the anxiety disorders, and one-third of the mood disorders.
1.2.4 Discussion

Main findings
This study was conducted to quantify the excess costs of the common mental disorders in the general adult population. Three perspectives were taken: on the per capita costs, on the costs attributable to the prevalent cases at population level, and on the costs attributable to the incident cases at population level. The main findings can be summarized as follows.

The average per capita excess cost of a common mental disorder is € 3,200 (95% CI= 2,117 ~ 4,284), and are of the same magnitude as the costs of somatic illnesses at an average of € 3,220 (SE= 245; 95% CI= 2740 ~ 3702). The annual per capita excess costs of mood disorders (€ 5,009) are higher than those of anxiety disorders (€ 3,587) and alcohol-related disorders (€ 1,431). The disorder which generates the highest per capita costs is dysthymia (€ 10,332), followed by panic disorder (€ 8,849) and agoraphobia (€ 4,879). The per capita excess costs of major depression, simple phobia and alcohol dependence are of the same size and fall in the range of € 2,278 to € 2,509. The other common mental disorders are associated with lower per capita costs.

At population level, the per capita costs need to be multiplied by the number of prevalent cases in the population. Seen from this perspective, dysthymia and panic disorder rank among the disorders that generate the bulk of the costs. In addition, the simple phobias, alcohol dependence, alcohol abuse and depression, being highly prevalent disorders, find themselves closer to the top of the hierarchy.

Finally, a perspective can be adopted with the focus on the influx of new cases (incidence). It is then seen that panic disorder still ranks among the disorders that generate most costs, and depression has now risen to the second position in the hierarchy. It was further shown that the incident cases are responsible for 39.2% of the annual costs at population level.

Cost estimates are lower bounds
For several reasons the costs reported here should be interpreted as conservative estimates. First, the \( t_0 \) sample was drawn from the non-institutionalised population. This has resulted in an under-representation of those who were hospitalised, which, in turn, has resulted in an under-estimate of the direct medical costs, because in several cost-of-illness studies it was shown that hospitalization was one of the major cost drivers (cf. Berto et al, 2000). Second, the medical costs that were considered in this study were limited to those related to mental health services and prescription drugs for mental disorders. Yet, it is well known that mental disorders also generate costs in non-mental health services (Greenberg et al, 1999). Third, we only present the short-term costs that are incurred over a 1-year period. Excluded are the costs of morbidity and mortality that may occur later in the life-course. Fourth, the cost calculations were based on self-reported medical consumption, which is known to
result in an under-estimation of the actual medical consumption and the corresponding costs (Van den Brink et al, 2004; Brower et al, 1999). Fifth, apart from work loss days due to absenteeism, work cutback (while at work and not feeling well) will also result in production losses. The latter were not included in this study, and can be substantial (Kessler and Frank, 1997; Lim et al, 2000). Finally, it is known that excessive alcohol consumption is associated with crime-related costs (Andlin-Sobocki, 2004). In our study, these costs were not included. Therefore, the costs presented here are conservative estimates of the real costs associated with common mental disorders, and can only be generalized to the non-institutionalised population.

**Strengths and limitations**

This study has several strengths. DSM-III-R axis-I disorders were assessed with a reliable diagnostic instrument, the CIDI, in a large epidemiological study representative of the non-institutionalized population of 18 – 65 years in the Netherlands. In this way several disorders could be analysed simultaneously. The cost calculations encompassed direct medical, direct non-medical and indirect costs; and the costs of the mental disorders were adjusted for concurrent mental and somatic disorders. In this way it was possible to estimate the costs uniquely attributable to the distinct mental disorders, the so called ‘excess costs’. The ‘excess costs’ were obtained using robust statistical techniques. Finally, this is the first population-based cost-of-illness study in which the same data-set was used to arrive at cost estimates of both prevalence and incidence.

There are also a number of limitations. Despite the large sample, the 95% confidence intervals around the cost-estimates are broad. This is a common finding in health economic studies, because cost-data typically have large standard errors.

Two cost estimates seemed counter-intuitive: the costs of simple phobia appeared relatively large, and those of generalized anxiety disorder relatively small. Closer inspection of Table 1.2.3, however, reveals that the direct (treatment) costs of simple phobia are not excessively high, and that the high costs of simple phobia are mainly accounted for by the production losses associated with this disorder. This seems plausible, because simple phobia does not attract much medical attention, but does result in very substantial excess absenteeism from work as is corroborated in the European-wide ESEMed (2004b) study. The low costs of generalized anxiety disorder as found in this study may be the result of a selection process. It is well know that this disorder persists over a long time, and people suffering from it may drop out of paid employment. As a group, people meeting the diagnostic criteria of generalised anxiety disorder will then cease to generate production losses, because they are no longer seen as part of the labour force. Indeed, the European study did indicate that generalised anxiety disorder has only a limited impact on work loss days (ESEMed, 2004b). Nevertheless, the costs of generalised anxiety disorder as presented here seem low as compared to the cost-of-
illness studies reviewed by Löthgren (2004), but in line with the costs recently reported by Knerer et al. (2005).

**Implications**
The results show that the common mental disorders have a strong and independent economic impact, which appears comparable to that of somatic disorders, as is corroborated by other studies (cf. Berto et al, 2000) This finding underscores the importance of adequate resource allocation to health services for mental disorders vis-à-vis health services for physical illnesses.

With an average of € 2725 per person per year the production losses account for the vast majority (85.2%) of the total costs of mental disorders. This finding is supported by the European wide ESEDMed study (2004b) which demonstrated a strong impact of mental disorder on absenteeism from work in the general population. The € 2725 found in our data represent 9.6 workdays lost due to mental illness in a year. This seems plausible in light of the ESEDMed study and dovetails with the findings of a cost-of-illness study in depressive disorder in the UK (Thomas and Morris, 2003). In this context we must mention again that the costs of work cutback days could not be included in our analyses, but once included should have reinforced the notion that mental ill-health has a detrimental impact on production losses. By implication, employers must be seen as important economic beneficiaries of effective mental health services. This makes them pertinent stakeholders in mental health and improvements therein.

Within the spectrum of the common mental disorders, panic disorder has quite marked economic consequences, which warrant a much stronger interest in this disorder, as is corroborated by other research (Knerer et al, 2005; Greenberg et al, 1999; Salvador-Carulla et al, 1995). The same holds for dysthymia, which was found to be the disorder associated with the largest costs in this study, but not in the studies reviewed by Löthgren (2004a). It is therefore suggested that the agenda for research and development of preventive interventions place depression, dysthymia and panic disorder at the top of the priorities.

The distribution of national budgets for mental health over treatment and prevention should probably be reconsidered in many Western countries. This study shows that, at population level, 39.2% of the annual costs of mental disorders are accounted for by incident cases. However, typically only a small percentage (3% in the Netherlands) of the national budget for mental health is allocated to prevention. This is unlikely to be a rational distribution of investments. This issue has become more relevant now that recent studies have indicated that treatment of mental disorders can only be partially successful even under optimal, evidence-based, treatment regimes (Andrews et al, 2004; Chisholm et al, 2004); and that prevention of mental disorders, such as major depression, is not only a theoretical option, but can be offered effectively (Cuijpers et al, 2005; Willemse et al, 2004) and cost effective (Lynch et al, 2005; Smit et al, 2006a; 2006b). Nevertheless, the proof is in the eating of the pudding, and any decision to allocate larger funds to prevention is
best informed by cost-effectiveness studies. This cost-of-illness study only suggests that the time is right to conduct cost-effectiveness studies and to generate evidence in favour or against the idea that prevention of mental disorders is an economically viable idea.
Part II

Do we know where to begin?

Chapter 2.1

Selecting key-variables for depression prevention, an epidemiological approach

This chapter is based on:

Chapter 2.1
Selecting key-variables for depression prevention, an epidemiological approach

2.1.1 Introduction

In public mental health much effort is devoted to curing depression. The rationale is clear: depression is a disabling condition (Murray and Lopez, 1994; WHO, 2000), it carries a poor prognosis (Cuijpers and Smit, 2002) and affects many people (Bijl et al, 1998). However, a formidable 48.7% percent of the prevalent cases are in fact new, or incident, cases (see below in this chapter). It is therefore important to reduce the influx of new cases. Here, depression prevention could play a key role once several criteria are met such as feasibility, acceptability, effectiveness and affordability of the preventive interventions. However, testing interventions against these criteria is a necessary but slow and expensive stage in the development and evaluation cycle of preventive interventions. To inform and rationalise this process at its earliest possible stage, we propose using three criteria to identify possibly interesting preventive interventions: the candidate intervention is directed at one (or more) risk factors that (1) are strongly associated with the onset of depressive disorder, (2) would result in a substantial decrease in the incidence rate of the disorder when the adverse effect of the risk factors can be successfully contained, and (3) are associated with a low number needed to be treated. The benefit of these criteria is that they can be assessed before embarking on a costly and time-consuming prevention trial. The required data can be obtained from prospective, population-based, epidemiological studies of the putative risk factors and the incidence of depressive disorder. These studies yield the relevant statistics and help to inform and rationalise the debate about the most strategic targets for the primary prevention of depression.

2.1.2 Method

Subjects and procedures
The data were derived from the Netherlands Mental Health Survey and Incidence Study (Nemesis), a nationally representative cohort study based on a random sample of 7,076 people at baseline (t₀) of whom 5,618 were retained after one year at first follow-up (t₁). Elsewhere the subjects and the procedures are described in detail (Bijl et al, 1998a) as are the effects of loss-to-follow-up between t₀ and t₁ (De Graaf et al, 2000). To correct for the combined effect of initial non-response and dropout, post-stratification weights were calculated. After weighting, the multivariate distribution over the variables sex, age, marital status and degree of
urbanisation in the sample was exactly the same as in the population as described by the central bureau of statistics. From this data set we selected people who had never experienced a depression before and were therefore at risk of becoming first-ever incident cases of depression (N= 4,664). Of these, 143 developed a depressive disorder in the year between \( t_0 \) and \( t_1 \). The latter formed the index group of first-ever incident cases of depression. The reference group were people from the same cohort who did not meet the diagnostic criteria for depression in that year.

**Measures**

DSM-III-R axis-I diagnoses (APA, 1987) were assessed with the Composite International Diagnostic Interview, the CIDI, of the World Health Organisation (WHO, 1990). Use was made of the first Dutch version of the CIDI (Smeets and Dingemans, 1993). The CIDI assessments were carried out by trained lays in structured, computer-assisted, face-to-face, interviews both at \( t_0 \) and \( t_1 \). The CIDI is known to have excellent psychometric properties (Wittchen et al, 1991). In this study we restrict our focus on DSM-III-R major depression as the dependent variable.

All putative risk factors were measured at \( t_0 \). Their selection for this study was in part based on the vulnerability-stress theory of Brown and Harris (Brown and Harris, 1978; Harris, 2000) and previous risk factor epidemiological studies in depression (De Graaf et al, 2002a, 2002b).

Childhood trauma: i.e. emotional neglect, psychological abuse, physical abuse and sexual abuse before the age of sixteen. During the interviews the different types of abuse were explained to the respondents with examples. Two new variables were made: one indicating exposure to two or more forms of abuse, i.e. ‘multiple abuse’, and another indicating exposure to any form of neglect or abuse, i.e. ‘any abuse’.

Parental history concerned depressive problems, anxiety or phobic problems and problem drinking in one or both biological parents. The presence or absence of parental history was assessed by single questions. The reliability of this procedure was evaluated for the question about parental problem drinking and proved to be well in order (kappa= 0.83, cf. Cuijpers and Smit, 2001), but this is not known for the other questions about parental problems. Two new variables were calculated: one that indicated the presence of two or more parental problems, i.e. ‘multiple parental problems’, and another, which indicated the presence of ‘any parental problem’.

Vulnerability indicators were: Locus of control measured with the Mastery Scale (Pearlin and Schooler, 1978). We defined low mastery (low = 1, else = 0) as a score below the mean of the baseline sample (M= 3.9; Cronbach’s \( \alpha = 0.81 \)). Neuroticism was assessed with the Groningse Neuroticism Questionnaire (Ormel, 1980). We used a cut-off at the mean to obtain an indicator for above average levels of neuroticism (M= 2.7; \( \alpha = 0.80 \)). In a sensitivity analysis the 66th percentiles of both the Mastery and the Neuroticism scales were also used, because it was thought
that at these levels mastery and neuroticism become etiologically more relevant. Other indicators for vulnerability were: having a score equal to or larger than 2 on the 12-item version of General Health Questionnaire (GHQ, Goldberg and Hillier, 1997) in the Dutch version (Koeter and Ormel, 1991), and having ever felt depressed for two weeks or longer (i.e. presence of this core symptom of depression on a life time basis).

Antecedent mental disorders: any DSM-III-R axis-I anxiety disorder, and substance misuse or dependence disorder in the twelve months preceding \(t_0\).

Antecedent somatic illnesses: presence of somatic disorders (migraine, lower back-pain, complaints pointing towards cardiovascular, respiratory, and non-specific abdominal problems) in the twelve months preceding \(t_0\).

Demographic variables are: sex, age, living with a partner, ethnic descent (Dutch, other), education (elementary, lower vocational, secondary, higher vocational and academic) and degree of urbanisation (fewer or more than 500 addresses per square kilometre).

Key-statistics: IRR, AF, and NNT

The analyses (see below) had to produce indices that help to evaluate the utility of risk factors as seen from the prevention perspective. Here we describe these indices.

The incidence rate ratio, IRR, is a measure of strength between an exposure and outcome. It tells how much larger the incidence rate of depression is in the exposed group relative to the incidence rate among those not exposed to a risk factor. When the IRR has a value of 1, it indicates absence of any effect, values larger than 1 indicate an increased incidence rate and values below 1 indicate a reduced incidence rate in the exposed group relative to the unexposed group.

The attributable fraction, AF, is a measure of potential health gain in a population. It describes the percentage by which the incidence rate of depression in the population can be reduced when the risk factor is completely eliminated, or when its adverse effect is completely contained. Clearly, it is not realistic to expect that preventive interventions will be completely successful and the AF is best interpreted as a statistic that puts an upper limit to the health gain that could be achieved in a population. It is also possible to adjust the AF-statistic for partially successful interventions if so required (cf. Morgenstern and Bursie, 1982; Gunning-Scheper, 1988). For the AF-statistic to be valid it must further be assumed that the risk factor has an etiological role in the pathogenesis of depression.

The number needed to be treated, NNT, is a measure of the effort. It shows how many people in the population must receive the intervention in order to avoid one new case of depression. In the more realistic scenario that the intervention is successful for, say, 30%, then \((1/0.30) = 3.3\) NNT people must receive the intervention to avoid one case, but without this correction the NNT can serve as an indication of the effort that is at least required.

To summarise, IRR gives information on the strength of association between exposure and outcome, AF describes the maximum achievable health gain that
could be generated in a population when a preventive intervention is successful in blocking the adverse effect of an etiological factor, and NNT helps to quantify the minimum amount of effort that is needed to generate that health gain in a population. Together these statistics of impact and effort can help to inform and rationalise the debate on depression prevention.

Analysis
Stata (StataCorp, 1999) was used to carry out the analysis. The key-statistics were obtained in three steps. In the first step the incidence rate ratio, IRR, for each risk factor was obtained under a weighted Poisson regression model with the depression status at \( t_1 \) as the dependent and the exposure status at \( t_0 \) as the independent variable, while accounting for variable exposure times across subjects, adjusting for demographics, and weighting to correct for initial non-response and loss-to-follow-up (cf. Kleinbaum et al, 1998). In the second step, the population attributable fraction, AF, was obtained with help of the equation \( \text{AF} = \text{ER}(\text{IRR}-1) / (1+\text{ER}(\text{IRR}-1)) \), where \( \text{ER} \) is the exposure rate of the risk factor in the population at risk and where IRR was obtained as before (cf. Rothman and Greenland, 1998; see also Miettinen, 1974). In the third step, the number needed to be treated, NNT, was computed as the inverse of the absolute risk difference. The latter was obtained using the weighted linear probability model of the incidence of depression entered as a binominal outcome and the pertinent risk factor linked linearly in the equation (cf. McCullah and Nelder, 1989; Long, 1997). In all the equations we adjusted for demographics, because it was thought that the secular impacts of demographics are beyond the control of preventive interventions. In order to report correct 95% confidence intervals of the key-statistics under weighting, so called ‘robust’ variance-related estimates were obtained, using the first-order Taylor-series linearisation method, as implemented in Stata.

2.1.3 Results

Prevalence and incidence of depression in the Dutch population
On an annual basis 5.8% of the population aged 18-65 meet the DSM criteria of depressive disorder (Bijl et al, 1998). Projected to the 10.7 million Dutch people in this age range and using the gender and age specific prevalence rates, there are 589,000 depressed adults each year in the Netherlands. It is worth noting that the incidence of depression is high relative to its prevalence. The incidence occurs at a rate of 2.72 first-ever cases per 100 person-years (Bijl et al, 2002), which is equivalent to 287,000 new cases in a given year when using gender and age specific incidence rates. In other words, 48.7% of the prevalent cases are in fact new cases. Therefore, it is inappropriate to focus exclusively on the treatment of prevalent cases. Primary prevention has to play a supplementary role by reducing the annual influx of new cases.
**Description of the people at risk of becoming depressed**

The demographics of the people at risk of becoming first-ever cases of depressive disorder (i.e. all people except those who already experienced a depressive disorder) can be summarised as follows: female 51% (95% CI= 49.4−52.2), mean age 41 years (40.8−41.5), living with a partner 68.5% (67.0−70.0), immigrant 2.2% (1.8−2.7), employed 70% (68.3−70.9), living in an urban environment 81.9% (80.8−83.0), education: elementary 6% (5.5−6.8), lower vocational 36% (34.9−37.7), high-school 28% (26.9−29.3), higher vocational and academic 29% (28.2−30.8).

**Detailed results for childhood trauma**

As can be seen in Table 2.1.1, many people report neglect and childhood abuse. Emotional neglect is reported by 20.8% (ER). Neglect is associated with an elevated risk for depression. In people exposed to neglect it is a factor 2.85 larger than in unexposed people (IRR). Elimination of its adverse effect would help to reduce the incidence of depression by 27.8% (AF). Thus, the incidence rate which is now 2.72 per 100 person-years, would then, in the best possible scenario, decrease to 2.72 * (1 − 0.278) = 1.96. This is equivalent to avoiding the onset of depression in close to 70,000 people in the Dutch population aged 18-65: a very substantial health gain at population level.

This health gain also entails economic ramifications. In the Netherlands the annual incremental costs of depression, defined as the costs of excess health service use by depressed people, plus the costs of production losses due to absenteeism from work, are typically about € 2,300 per year per depressed person (Chapter 1.2; Simon et al, 2002 for other countries). Therefore, avoiding the onset of depression in 70,000 cases may save as much as 70,000 * 2,300= € 161,000,000 per year. If we were to spent the potential cost offsets of € 161 million euro on all 2,1 million people who were exposed to neglect, then the intervention may costs as much as € 77 per person before reaching a break-even point at which the costs of prevention would exceed its savings. Now, the aim of prevention is to create a health gain, not to save costs. In fact, the general public is most likely to show, up to some degree, a willingness to pay for a depression-free survival year. The break-even point of € 77 is thus best interpreted as an economically interesting starting point for prevention of depression, as it may suggest that prevention of depression may become cost-effective.

Finally, there is the NNT to consider in Table 2.1.1. It shows that the preventive intervention has to target 27.5 people who have been exposed to neglect during their childhood in order to prevent the onset of depression in one of them.

These indices of health gain, cost offsets, and efficiency represent, of course, the best possible scenario which is based on the unrealistic assumption of a completely effective intervention. Thus the real health gain and cost savings are less than indicated, but these ‘maximum’ statistics help to rank risk factors into a
hierarchy. By comparison, physical and multiple abuse have even better IRR, AF and NNT values overall. This makes them more important targeting points for prevention.

Table 2.1.1 Exposure rate (ER), incidence rate ratio (IRR), attributable fraction (AF), absolute risk difference (ARD) and numbers needed to be treated (NNT) for the risk factors. Weighted analysis, adjusted for demographics (N= 4,664).

<table>
<thead>
<tr>
<th>Risk indicator</th>
<th>ER, %</th>
<th>IRR</th>
<th>AF, %</th>
<th>ARD</th>
<th>NNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Childhood trauma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>. Emotional neglect</td>
<td>20.8</td>
<td>2.85**</td>
<td>27.8</td>
<td>0.036**</td>
<td>27.5</td>
</tr>
<tr>
<td>. Psychological abuse</td>
<td>10.3</td>
<td>3.20**</td>
<td>18.4</td>
<td>0.049**</td>
<td>20.4</td>
</tr>
<tr>
<td>. Physical abuse</td>
<td>7.5</td>
<td>3.84**</td>
<td>17.6</td>
<td>0.062**</td>
<td>16.1</td>
</tr>
<tr>
<td>. Sexual abuse</td>
<td>5.4</td>
<td>1.49 NS</td>
<td>2.6</td>
<td>0.015 NS</td>
<td>66.6</td>
</tr>
<tr>
<td>. ≥ 2 forms of abuse</td>
<td>10.7</td>
<td>3.38**</td>
<td>20.2</td>
<td>0.052**</td>
<td>19.4</td>
</tr>
<tr>
<td>. Any abuse</td>
<td>27.9</td>
<td>2.67**</td>
<td>31.8</td>
<td>0.031**</td>
<td>31.9</td>
</tr>
<tr>
<td>Parental history</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>. Depressive</td>
<td>19.7</td>
<td>1.46 NS</td>
<td>8.6</td>
<td>0.012 NS</td>
<td>86.3</td>
</tr>
<tr>
<td>. Anxious / phobic</td>
<td>9.4</td>
<td>1.52 NS</td>
<td>4.6</td>
<td>0.014 NS</td>
<td>71.0</td>
</tr>
<tr>
<td>. Alcoholic</td>
<td>7.3</td>
<td>1.14 NS</td>
<td>1.0</td>
<td>0.004 NS</td>
<td>261.4</td>
</tr>
<tr>
<td>. ≥ 2 problems</td>
<td>8.0</td>
<td>1.62 NS</td>
<td>4.7</td>
<td>0.017 NS</td>
<td>59.9</td>
</tr>
<tr>
<td>. Any problem</td>
<td>23.1</td>
<td>1.43 NS</td>
<td>9.0</td>
<td>0.011 NS</td>
<td>92.5</td>
</tr>
<tr>
<td>Vulnerability indicators</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>. Low mastery (&lt; M)</td>
<td>38.1</td>
<td>2.59**</td>
<td>37.7</td>
<td>0.028**</td>
<td>36.1</td>
</tr>
<tr>
<td>. High neuroticism (&gt; M)</td>
<td>33.5</td>
<td>4.96**</td>
<td>57.0</td>
<td>0.047**</td>
<td>21.4</td>
</tr>
<tr>
<td>. GHQ ≥ 2</td>
<td>20.3</td>
<td>3.60**</td>
<td>34.6</td>
<td>0.047**</td>
<td>21.3</td>
</tr>
<tr>
<td>. Felt depressed (≥ 2 weeks)</td>
<td>34.1</td>
<td>4.05**</td>
<td>51.0</td>
<td>0.041**</td>
<td>24.4</td>
</tr>
<tr>
<td>Antecedent mental disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>. Any anxiety disorder</td>
<td>9.0</td>
<td>3.95**</td>
<td>21.0</td>
<td>0.066**</td>
<td>22.7</td>
</tr>
<tr>
<td>. Any substance use disorder</td>
<td>8.6</td>
<td>2.53**</td>
<td>11.6</td>
<td>0.030**</td>
<td>33.7</td>
</tr>
<tr>
<td>Somatic complaints</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>. Migraine</td>
<td>8.1</td>
<td>1.96**</td>
<td>7.2</td>
<td>0.028 *</td>
<td>35.5</td>
</tr>
<tr>
<td>. Cardiovascular</td>
<td>1.1</td>
<td>1.36 NS</td>
<td>6.4</td>
<td>0.008 NS</td>
<td>127.3</td>
</tr>
<tr>
<td>. Respiratory</td>
<td>6.2</td>
<td>2.23**</td>
<td>7.1</td>
<td>0.031 *</td>
<td>31.8</td>
</tr>
<tr>
<td>. Abdominal</td>
<td>3.2</td>
<td>2.29 *</td>
<td>4.0</td>
<td>0.037 NS</td>
<td>27.3</td>
</tr>
<tr>
<td>. Lower back pain</td>
<td>10.0</td>
<td>1.63 *</td>
<td>5.9</td>
<td>0.017 NS</td>
<td>59.6</td>
</tr>
<tr>
<td>. ≥ 2 illnesses</td>
<td>4.3</td>
<td>2.85**</td>
<td>7.4</td>
<td>0.049**</td>
<td>20.4</td>
</tr>
<tr>
<td>. Any complaint</td>
<td>23.7</td>
<td>1.83**</td>
<td>16.3</td>
<td>0.020**</td>
<td>50.4</td>
</tr>
</tbody>
</table>

NS= not significant at α < 0.05, * significant at α < 0.05, ** significant at α < 0.01, M= mean

Taking childhood neglect and abuse as targeting points for depression prevention is not without drawbacks. The respective groups are rather large and it may be difficult to recognise them – although our own experience has told us that people are willing to volunteer information on a history of abuse when asked. Also,
information on a history of abuse may be available in medical records of a GP and could then be employed for selecting people eligible for a preventive intervention.

The example of childhood trauma illustrated a number of points: the incidence of depression depends on exposure to certain risk factors; it is possible to select risk factors that are especially interesting from the prevention perspective, and some tentative ante-hoc cost-benefit calculations can be made.

Other risk indicators
Now we turn to the remainder of Table 2.1.1. The general impression is that vulnerability indicators and childhood trauma are associated with relatively high IRR and AF values and some of them also have relatively low NNT values. These risk factors, and corresponding high-risk groups, warrant interest from the viewpoint of depression prevention in adult people. Parental history, by contrast, appears to be less interesting. More detailed inspection reveals that high IRR and AF values are found for neuroticism, presence of a depressed feeling, GHQ scores equal or higher than 2 and (physical) abuse. Moreover, (physical) abuse, GHQ ≥ 2 scores and high neuroticism are also associated with low NNT values.

Cumulative effects
The low NNT values for multiple exposures draw attention to another interesting feature of Table 2.1.1 (see childhood traumas and 2 somatic illnesses). This suggests that the risk to become a case of depressive disorder is much increased under multiple exposures. Further analysis shows that this is indeed the case, in particular for the presence of somatic illnesses.

Figure 2.1.1  The risk of becoming depressed (in %) by number of exposures (exp) to somatic complaints (migraine, respiratory and abdominal complaints).

Figure 2.1.1 shows the cumulative effect of having none, one, two or three somatic conditions on the risk of becoming depressive. In the presence of multiple
exposures the risk of becoming depressed increases sharply, reaching a risk of 46% of becoming depressed within a year. Similar risk curves can be produced for the joint exposure of other risk factors (results not shown). This indicates how high-risk groups, or even ultra-high risk groups, can be selected for primary prevention. Further, the risk of becoming depressed is 0.46 in the group with three illnesses against 0.03 in the group with no illness. The risk difference is therefore $0.46 - 0.03 = 0.43$, and the NNT is therefore $1/0.43 = 2.3$. The low NNT value for multiply exposed groups indicates that prevention in these groups may be more efficient than targeting singly exposed groups.

2.1.4 Discussion

Main findings

The outcomes of childhood trauma were already discussed in great detail and we shall not repeat that discussion here. Suffice it to recall that childhood adversities are strongly associated with the later onset of depressive disorder, as is also reported elsewhere (Brown et al, 1999; Kessler et al, 1997; Nelson et al. 2002; Portegijs et al, 1996). It was shown that the containment of their adverse effects would greatly help to reduce the incidence of depression in the population. It appears that this could be done in a cost-effective way.

It is generally acknowledged that parental history has a predictive value for the later onset of mental disorders in their children (cf. Mrazek and Haggerty, 1994; Kendler et al, 1997; Merikangas et al, 1998; Cuijpers et al, 1999). However, we found no support for this notion, although it should be noted that the statistics, as we found them, bordered on significance. Perhaps the fact that we modelled these risk factors prospectively and adjusted them for demographics has led to statistically insignificant results. By contrast, a cross-sectional analysis of our baseline data did show significant relationships between parental history and depression (Bijl et al, 2002). Nevertheless, we must conclude that parental history is not a particularly good starting point for depression prevention in adults. It should be emphasised, however, that a different conclusion might be drawn with respect to young children (cf. Beardslee and Podorefsky, 1988; Beardslee et al, 1997; Clarke et al, 1995, 1999) as opposed to the adults that were studied here. In addition, it should be born in mind that parental depression may not only result in depression in their offspring, but also elevates risks for a range of other disorders like anxiety disorders and substance-related disorders (Cuijpers et al, 2002).

Vulnerability indicators (cf. Harris, 2000) are associated with high IRRs and AFs and low NNTs and thus seem to offer an advantageous starting point for preventive interventions. Low mastery may be amenable to change through psychosocial interventions and be seen as an etiological factor on the causal pathway towards the onset of depression. Since a GHQ-score equal to or larger than two, and the presence of a depressive core symptom are conceptually overlapping
with depression, they cannot be seen as proper etiological factors – in a sense they are outcomes rather than predictors. Nevertheless, they may be valuable for depression prevention in terms of early detection, for example through screening.

Antecedent disorders, mental or somatic, may also be helpful in targeting selected groups, which is important for practical and financial reasons. It is well known that anxiety disorders and substance-use disorders may precede or accompany depression (Kessler et al, 1999; Preisig et al, 2001; De Graaf et al, 2002b) and the same can be said about somatic conditions like migraine, lower back-pain, cardiovascular disorders and (non-specific) respiratory and abdominal complaints (Jenkins et al, 1992; Geerlings et al, 2000). We found support for these notions, but found cardiovascular complaints to be unrelated to the onset of depression. However, it should be recalled that we studied adults under the age of 65, and for older people a different risk set may exist (cf. Geerlings et al, 2000).

**Strengths and limitations of this study**

The main findings and the conclusions (see below) must be placed in the context of the strengths and limitations of this study.

The strengths of this study are (-) the representativeness of the sample for the general Dutch population in the age range of 18 – 65, (-) the measurement of mental disorders with a reliable instrument, (-) the prospective design which enables the study of incidence and facilitates etiological inference, (-) the measurement of exposures which is not biased due to post-hoc rationalisation on the part of the respondents, because at \( t_0 \) they could not have any knowledge about their future health state at \( t_1 \). Furthermore, this study is among the first to show how a statistical technology could be applied to quantify potential health benefits and the required efforts to generate these health benefits in the field of prevention of mental disorders. It thus supplies the sort of technology which is of importance for setting a rational research and development agenda for depression prevention.

The limitations of this study consist in the not very detailed measurement of the exposures. We do not know at what age, for how long and how intensively subjects were exposed. Moreover, the number of studied risk factors is limited in that, for example, genetic and other biological risk factors were not included. We studied only dyads of a single exposure and a single disorder, but it would be interesting to extend this to the study of joint exposures. Figure 2.1.1 showed that this could be very interesting. It may be equally interesting not to restrict studies like these to a single disorder, but to extent the analyses to encompass a broader spectrum of disorders, because the studied risk factors rarely behave as specific risk factors for a distinct disorder, rather they behave as generic risk factors for the common mental disorders.

**Conclusions**

This study was set out to see whether, and how, epidemiology can make contributions to the formulation of a research and development agenda for
depression prevention. Our main conclusion is that epidemiology can be helpful in this respect. Statistics such as the incidence rate ratio IRR, population attributable fraction AF, and the number needed to be treated NNT, are helpful in selecting risk factors that have potential for prevention. These statistics can also be combined with information from cost-of-illness studies (see Chapter 1.2) and then used for antehoc cost-benefit analysis.

This claim is made on the understanding that, ultimately, the selection of risk factors and corresponding high-risk groups, cannot be decided by mechanical application of statistics alone. Other considerations are also important, such as the availability of preventive techniques, knowledge of their (assumed) success rates, and ethical and economic considerations. Once the research and development agenda for depression prevention is set, preventive interventions should be designed and tested in prevention trials to empirically ascertain their effectiveness and economic feasibility. Finally, projects with a proven effectiveness should be disseminated and properly implemented. Therefore, it is also understood that the risk-factor epidemiological contributions are only one element in a larger cycle of development and research. Nonetheless, the epidemiological contributions are crucial because they help at a very early stage to inform and rationalise the research and development agenda for depression prevention.

At the substantive level, the results suggest that depression prevention could be effective in primary care settings (cf. Jenkins et al, 1992). Patients with one or more somatic complaints will present themselves there and could be screened for the presence of a core symptom of depression, a GHQ-score above cut-off, or high levels of neuroticism. When these patients also report a history of childhood abuse or an antecedent anxiety disorder, then selective prevention appears to be indicated. The exact format of this intervention is, as yet, not clear, but presumably cost-effective interventions like bibliotherapy (Cuijpers, 1997) or brief counselling (Bower et al, 2003) are likely candidates.

In future research, thought must be given to the study of joint exposures of several risk-factors (see Chapter 2.2.), the cross-validation of this methodology (see Chapter 2.3), the quantification of impacts on the incidence not only of major depressive disorder but on the whole spectrum of mental disorders (see Chapter 4.2), and the accommodation of cost-benefit considerations in this sort of analysis (ibid). It should be hoped that these technologies will assist in targeting prevention there where it is likely to be most cost-effective.
Part II
Do we know where to begin?

Chapter 2.2
Opportunities for cost-effective prevention of late-life depression

This chapter is based on:

Chapter 2.2

Opportunities for cost-effective prevention of late-life depression: an epidemiological approach

2.2.1 Introduction

Late-life depression is characterized by high prevalence, unfavourable prognosis, reduced quality of life and excess mortality (Beekman et al, 2002; Cole et al, 1999; Cuijpers and Smit, 2002; Geerlings et al, 2000; 2001; 2002; Penninx et al, 1999; Stek et al, 2002). It is also associated with substantial societal costs (Beekman et al, 1997; Hunkeler et al, 2003; Unutzer et al, 1997; Von Korff et al, 1992; Katon et al, 2003). Late-life depression is further characterized by a large annual influx of new cases, because 1 in every 5 cases is a new case (this chapter). From the public health perspective, depression prevention may thus be an attractive, if not imperative, means to generate health gains in the population and to reduce future costs (Smit et al, 2004).

In this context it should be noted that depression is a treatable condition (Cuijpers, 1998; McCusker et al, 1998; Thase et al, 1997). However, according to a recent estimate the total disease burden associated with depression can only be reduced to about 34%, even under a hypothetical regimen of optimal evidence-based treatment (Andrews et al, 2000; 2004; Chisholm et al, 2004). This is another reason why prevention has to play an important role in public health. Recently, a meta-analysis of randomized trials of preventive interventions has shown that the incidence of depressive disorder can be reduced by 30% and this may indicate that prevention is a viable option (Cuijpers et al, 2005).

However, developing preventive interventions and testing their cost-effectiveness in randomised trials is time-consuming and expensive (Cuijpers, 2003). Therefore, one would like to be able to estimate the cost-effectiveness of future interventions at the earliest possible stage in the development and evaluation cycle, and target research efforts where they are likely to generate optimal yields. The aim of this paper is to describe a methodology that could help to identify cost-effective preventive interventions at the earliest possible stage and to apply this methodology to the case of late-life depression.

The methodology of identifying high-risk groups for prevention is not new (Miettinen, 1974; Morgenstern and Bursic, 1982), but in the field of psychiatric epidemiology and prevention research it has rarely been applied. The reason for this omission is that this methodology requires longitudinal data on the incidence of the disorder and its putative risk indicators in the general population. These data are not often available, but once there, they offer a wealth of information and can be
employed to set a rational research agenda in the field of preventive psychiatry (Smit et al, 2004).

2.2.2 Method

Subjects and procedures
The analyses were based on the data of the first two waves of the Longitudinal Aging Study Amsterdam (LASA). The sampling and procedures of this study have been described elsewhere in detail (Beekman et al, 2002). At baseline 3,056 community residents in the age group of 55 to 85 years were interviewed. Participating subjects had given their informed consent and underwent face-to-face interviews in their homes. The random sample was stratified by age and gender. The older age strata and men were over-sampled in anticipation of higher attrition rates among these groups during the course of the study. After three years (M= 1115, Sd= 59 days) 2,200 subjects (72%) were successfully re-interviewed. Loss-to-follow-up had occurred among 856 subjects, mainly because subjects were too ill or were no longer alive at the time of the first follow-up. Predictors of loss-to-follow-up were older age, male gender, lower education, functional limitations, chronic diseases and cognitive decline, but not depression status at baseline (ibid.). Corrective weights were used to account for the joint effect of intentional over-sampling and accidental attrition (see Analysis, below).

Depression
Depression was ascertained with the Center of Epidemiological Studies Depression Scale (CES-D; Radloff, 1977). The CES-D consists of 20 items and its total score has a range between 0 and 60. Scores ≥ 16 indicate clinically significant levels of depressive symptoms (Berkman et al, 1986). At this cutoff the sensitivity is 100% and the specificity is 88% for major depressive disorder in the elderly Dutch population (Beekman et al, 1997). In the remainder of this paper CES-D ≥ 16 will be referred to as “depression”. Measurements were taken at baseline ($t_0$) and at first follow-up ($t_1$). A person was deemed to be an incident case when three criteria were met: (1) absence of depression at $t_0$ (CES-D < 16), (2) presence of depression at $t_1$ (CES-D ≥ 16), and (3) significant change between $t_0$ and $t_1$ (change score on the CES-D ≥ 5).

Criterion 1 was used to ensure that the analysis was restricted to the group at risk, criterion 2 to ascertain depression status at $t_1$, and criterion 3 to prevent false-positive cases due to measurement error in the CES-D. In the latter criterion, a minimum change of 5 CES-D points was chosen because it represents, in clinical terms, a medium to large change (Lipsey and Wilson, 1993), and has the advantage that it has also been used in other studies (cf. Beekman et al, 2002). In addition, a change of 5 scale points on the CES-D is greater than 3.5 which, on this scale,
corresponds to the threshold for statistically reliable change (Jacobson and Truax, 1991). In short, a person was deemed to be an incident case when there was a change of 5 points or more, thereby crossing the cutoff of 16.

**Risk indicators**

Following the vulnerability-stress theory (Brown and Harris, 1978; Harris, 2000) and a more recently published review on risk indicators of late-life depression (Cole and Dendukuri, 2003) the following putative risk indicators were included.

Demographics: female gender (1 = female, 0 = male), age over 65 years, which is the age at which 30% of the sample makes a significant transition in their life due to retirement (1 = older than 65 years, 0 = younger), low education (dichotomized into 1 = elementary school and less, 0 = high school and more), living in an urban environment (1 = living in Amsterdam, 0 = living elsewhere).

Chronic illnesses, among them diabetes mellitus, chronic obstructive lung disease, cardiac disease, arthritis of knee or hip, and neoplasm were dichotomized at 1 = two or more, 0 = one or none (cf. Kriegsman et al, 1996). Cognitive impairment was assessed with help of the Mini Mental Health State (Folstein et al, 1975) and dichotomised at 1 = MMSE < 24, 0 = MMSE 24-30. Earlier studies have indicated that it is not so much the presence of chronic medical conditions that predicts the onset of depression, but rather the functional limitations that may stem from them, the subjective appraisal of one’s health, and finally the degree by which one’s sense of mastery (locus of control) is affected (Geerlings et al, 2000; Ormel et al, 1997; Zarit et a, 1999). Therefore, the following measures were also included: functional limitations (Van Sonsbeek, 1988, dichotomized as 1 = one or more, 0 = none), self-rated poor health (Central Bureau of Statistics, 1989; dichotomised as 1 = poor health, 0 = some times good / some times bad, fair, good, or, excellent health), and low mastery (Pearlin and Schooler, 1978), which was dichotomized as 1 = score below the 50th percentile on the scale, 0 = above 50th percentile.

Depressive symptoms (1 = CES-D scores between 5 and 15, 0 = CES-D < 5, i.e. below 50th percentile) at baseline were also relevant, because they can act as precursors of CES-D>16 depression. The distribution of the CES-D is as follows: 25% of the population falls in the range of 0 – 2, 50% in 0 – 5, 75% in 0–10 and 90% in the range of 0–16.

Finally, social vulnerability was assessed by two additional measures: Small social network (1 = below, 0 = above the median social network size of 13 persons), and widowhood (1 = ever widowed, 0 = other).

All risk indicators were measured at t₀ and were coded 1 as the index category for the (presumably) elevated risk status and 0 for the reference category. Dichotomization was carried out prior to the analysis.
**Analysis**

Analyses took into account that the data were generated by a sampling design with intentional over-sampling of the male and older age strata and loss-to-follow-up. This was done by weighting the data such that the multivariate distribution over gender and age in the sample was exactly the same as in the general Dutch population in the age range of 55 – 85 years as reported by Statistics Netherlands for the year 2002. In order to obtain correct 95% confidence intervals and p-values under weighting, all variance-related statistics were obtained with help of the first-order Taylor-serieslinearization method as implemented in Stata SE/7.0 (StataCorp, 2001). Weighted N’s are reported, rounded to the nearest integer, throughout the remainder of this paper. The (weighted) analyses were based on the 1925 people at risk to become cases of depression, i.e. the group without depression at the baseline. The subsequent analyses were carried out in several steps.

The Exposure Rate (ER) of each risk indicator was calculated on the basis of the weighted data. The ER gives the percentage of the elderly population exposed to the risk indicator.

For each risk indicator the Incidence Rate Ratio (IRR) was obtained by regressing the outcome (1 = incident case, 0= not an incident case) on the risk indicator in a weighted Poisson regression model. The IRRs were based on person-time data to account for the small differences in follow-up time between $t_0$ and $t_1$ across the subjects. The effect of each of the risk indicators was evaluated while adjusting for all other variables in the risk set. The IRR describes how much larger the incidence rate is in the exposed group relative to the incidence rate in the unexposed group, controlling for competing risks. IRR values larger than 1 signify an increased risk level in the exposed group and values smaller than 1 indicate a risk reduction.

A maximum-likelihood estimate of the population Attributable Fraction (AF) was obtained with the Aflogit-procedure in Stata for each of the risk indicators under a Poisson regression while adjusting for competing risks (Greenland and Drescher, 1993). When converted into a percentage, the AF denotes by how many percents the current incidence rate of depression in the population would be reduced if the adverse effect of the risk indicator is completely blocked (Miettinen, 1974; Rothman and Greenland, 1998). This equals the maximum possible impact of a completely successful preventive intervention. Since it cannot be realistically assumed that preventive interventions are completely successful in containing the adverse effect of a risk indicator, it follows that the AF-statistic represents the upper limit to the potential health gain in the population. Although it is possible to adjust the AF-statistic for interventions that are not completely effective (Morgenstern and Bursic, 1982), it is also understood that we need not correct the AF-statistic for the purpose of this paper: a measure of relative performance is good enough for ranking risk indicators by their utility for prevention. We will return to the interpretation of the AF later.
Finally, the number-needed-to-be-treated (NNT) of each risk indicator was calculated as the inverse of the absolute risk difference (ARD). The latter was obtained by regressing the incidence on a risk indicator in a linear probability model, while adjusting for all other competing risks in the model. The NNT denotes how many people should receive a preventive intervention in order to avoid one new case of late-life depression. Again we do not expect that preventive interventions are completely successful and it is thus understood that the NNT represents the lower limit of the effort that is required to generate a health gain in the population.

To summarize, we calculate the exposure rate (ER), the strength of association between risk indicator and outcome (IRR), maximum achievable health gain (AF), and minimum effort to generate that health gain (NNT).

Together these indices of impact and effort allow us to select high-risk groups for which depression prevention is likely to be associated with the highest health benefit in the population for the lowest cost. This selection process was carried out as follows. First, we computed ER, IRR, AF and NNT for all risk indicators simultaneously (Table 2.2.2). Then using conventional back-stepping procedures we selected the smallest set of risk indicators in which each risk indicator has a unique and significant contribution to the prediction of depression (Table 2.2.3). From this list the most promising risk indicator was then selected (with the highest IRR and AF, and lowest ER and NNT). This was followed by consecutively selecting and adding risk indicators in such a way that the values for the potential health benefit (IRR and AF) were kept as high as possible and the values for effort and cost (ER and NNT) as low as possible. This process of maximising – minimising is depicted in Figure 2.2.1.

Finally, when the economical costs of late-life depression are known, then the cost figures can be combined with the AF and the NNT. This gives an indication of the dollar value of both the costs and the savings of a future prevention. The method of this ante-hoc health-economical evaluation is straightforward, but best illustrated with real data (see below).

### 2.2.3 Results

**Incidence**

158 people (weighted N) became incident cases of depression in a sum total of 5,643 weighted person years. This translates into 2.8 incident cases per 100 person years. This weighted estimate is very close to the unweighted estimate of 170 cases per 5,861 person years, which is equal to 2.9 incident cases per 100 person years.

**Exposure status**

At $t_0$ the (weighted) study cohort consisted of 2,200 people. Of these 1,925 (87.5%) were ‘at risk’ of becoming depressed, of whom 158 (8.2%) became incident cases.
after three years at $t_f$. Table 2.2.1 describes the exposure status of both the group at risk and the incident group. As expected, the exposure rates are often significantly elevated in the incident group as compared to the group at risk.

**A model with all risk indicators**

Table 2.2.2 shows the exposure rate (ER), incidence rate ratio (IRR), the population attributable fraction (AF) and the numbers-needed-to-be-treated (NNT) for each of the risk indicators, after adjusting for the effects of all other risks in the model. The factors female gender, having two or more chronic diseases, experiencing functional limitations and having an above average number of depressive symptoms are associated with significant IRR, AF and NNT values (Table 2.2.2).

### Table 2.2.1 Exposure status (in %, 95% CI) in the group at risk and the incident group (weighted analyses).

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Group at risk (N= 1925)</th>
<th>Incident cases (N= 158)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% 95% CI</td>
<td>% 95% CI</td>
</tr>
<tr>
<td>Female gender</td>
<td>51.6 49.3–53.9</td>
<td>72.9 66.1–79.7*</td>
</tr>
<tr>
<td>Age $&gt;$ 65 yrs</td>
<td>52.2 49.8–54.5</td>
<td>66.7 58.9–74.6*</td>
</tr>
<tr>
<td>Low education</td>
<td>36.4 34.2–38.7</td>
<td>53.9 46.0–61.7*</td>
</tr>
<tr>
<td>Urban environment</td>
<td>24.7 22.7–26.7</td>
<td>29.4 22.3–36.5</td>
</tr>
<tr>
<td>$\geq$ 2 chronic diseases</td>
<td>32.8 30.6–34.9</td>
<td>51.7 43.8–59.5*</td>
</tr>
<tr>
<td>Functional limitations</td>
<td>28.5 26.4–30.5</td>
<td>52.1 44.2–60.0*</td>
</tr>
<tr>
<td>Self-rated poor health</td>
<td>0.7 0.3–1.0</td>
<td>2.6 0.0–5.3</td>
</tr>
<tr>
<td>Cognitive impairment</td>
<td>4.5 3.6–5.4</td>
<td>8.2 4.1–12.3</td>
</tr>
<tr>
<td>Depressive symptoms</td>
<td>40.3 38.0–42.6*</td>
<td>66.1 58.6–73.5*</td>
</tr>
<tr>
<td>Low mastery</td>
<td>53.8 51.5–56.2</td>
<td>69.2 61.8–76.5*</td>
</tr>
<tr>
<td>Small social network</td>
<td>45.5 43.1–47.8</td>
<td>58.5 50.6–66.3*</td>
</tr>
<tr>
<td>Ever widowed</td>
<td>19.9 18.2–21.7</td>
<td>33.4 26.2–40.6*</td>
</tr>
</tbody>
</table>

* Significant at $\alpha < 0.05$
### Table 2.2.2 Complete multivariate model of the risk indicators (N= 1,925, weighted analysis).

<table>
<thead>
<tr>
<th>Risk indicator (at t₀)</th>
<th>ER, %</th>
<th>IRR</th>
<th>95% CI</th>
<th>AF, %</th>
<th>95% CI</th>
<th>NNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female gender</td>
<td>51.6</td>
<td>1.79</td>
<td>1.24–2.60*</td>
<td>32.5</td>
<td>13.2–47.6*</td>
<td>26</td>
</tr>
<tr>
<td>Age &gt; 65 yrs</td>
<td>52.2</td>
<td>1.21</td>
<td>0.82–1.78</td>
<td>13.1</td>
<td>-14.0–33.7</td>
<td>83</td>
</tr>
<tr>
<td>Low education</td>
<td>36.4</td>
<td>1.38</td>
<td>0.99–1.92</td>
<td>15.8</td>
<td>-0.0–29.4</td>
<td>38</td>
</tr>
<tr>
<td>Urban environment</td>
<td>24.7</td>
<td>1.25</td>
<td>0.89–1.75</td>
<td>7.3</td>
<td>-3.8–17.2</td>
<td>67</td>
</tr>
<tr>
<td>≥ 2 chronic diseases</td>
<td>32.8</td>
<td>1.55</td>
<td>1.11–2.16*</td>
<td>20.6</td>
<td>5.2–33.4*</td>
<td>27</td>
</tr>
<tr>
<td>Functional limitations</td>
<td>28.5</td>
<td>1.52</td>
<td>1.06–2.18*</td>
<td>21.1</td>
<td>3.5–35.5*</td>
<td>25</td>
</tr>
<tr>
<td>Self-rated poor health</td>
<td>0.7</td>
<td>2.07</td>
<td>0.80–5.34</td>
<td>3.7</td>
<td>-1.6–8.7</td>
<td>6</td>
</tr>
<tr>
<td>Cognitive impairment</td>
<td>4.5</td>
<td>1.32</td>
<td>0.76–2.29</td>
<td>2.8</td>
<td>-1.1–7.0</td>
<td>24</td>
</tr>
<tr>
<td>Depressive symptoms</td>
<td>40.3</td>
<td>2.09</td>
<td>1.46–2.97*</td>
<td>39.4</td>
<td>22.2–52.8*</td>
<td>17</td>
</tr>
<tr>
<td>Low mastery</td>
<td>53.8</td>
<td>1.25</td>
<td>0.88–1.78</td>
<td>15.0</td>
<td>-8.6–33.4</td>
<td>70</td>
</tr>
<tr>
<td>Small social network</td>
<td>45.5</td>
<td>1.34</td>
<td>0.98–1.84</td>
<td>16.0</td>
<td>-1.0–30.0</td>
<td>42</td>
</tr>
<tr>
<td>Ever widowed</td>
<td>19.9</td>
<td>1.06</td>
<td>0.76–1.50</td>
<td>2.3</td>
<td>-10.1–13.3</td>
<td>80</td>
</tr>
<tr>
<td>Total AF</td>
<td></td>
<td></td>
<td></td>
<td>86.2</td>
<td>77.9–91.4*</td>
<td></td>
</tr>
</tbody>
</table>

ER= exposure rate (in %). IRR= incidence rate ratio. AF= attributable fraction. NNT= number needed to be treated. 95% CI= 95% confidence interval. * significant (p < 0.05)

### Selecting a smaller set of risk indicators

In a next step, we obtained a more parsimonious multivariate model with fewer risk indicators. This model is based on the smallest subset of statistically significant risk indicators (at α < 0.05) and was obtained using the backward stepping selection method in the respective regression equations. The rationale of this approach is that risk indicators are competing with each other and we need to select only the most competitive risk indicators. The results are presented in Table 2.2.3.
Table 2.2.3 Parsimonious model (all parameters significant at p < 0.05).

<table>
<thead>
<tr>
<th>Risk indicator</th>
<th>IRR</th>
<th>95% CI</th>
<th>AF, %</th>
<th>95% CI</th>
<th>NNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female gender</td>
<td>1.85</td>
<td>1.30~2.64</td>
<td>34.0</td>
<td>15.4~48.5</td>
<td>23</td>
</tr>
<tr>
<td>Low education</td>
<td>1.41</td>
<td>1.03~1.92</td>
<td>16.6</td>
<td>1.5~29.4</td>
<td>34</td>
</tr>
<tr>
<td>≥ 2 chronic diseases</td>
<td>1.57</td>
<td>1.14~2.16</td>
<td>20.7</td>
<td>6.2~33.0</td>
<td>26</td>
</tr>
<tr>
<td>Functional limitations</td>
<td>1.75</td>
<td>1.26~2.43</td>
<td>26.5</td>
<td>11.5~38.9</td>
<td>18</td>
</tr>
<tr>
<td>Depressive symptoms</td>
<td>2.18</td>
<td>1.56~3.04</td>
<td>40.3</td>
<td>24.5~52.8</td>
<td>16</td>
</tr>
<tr>
<td>Small social network</td>
<td>1.52</td>
<td>1.12~2.06</td>
<td>21.6</td>
<td>5.8~34.7</td>
<td>30</td>
</tr>
</tbody>
</table>

Total AF 82.8 74.3~88.5

IRR= incidence rate ratio, AF= population attributable fraction, 95% CI= 95% confidence interval, NNT= number needed to be treated.

In the parsimonious model a smaller number of risk indicators was retained (Table 2.2.3). These were: female gender, low education, having two or more chronic diseases, experience of functional limitations, having an above average number of depressive symptoms and having a small social network.

With this risk profile 82.8% of the future cases of clinically relevant depressive disorder can be predicted. In the complete model with all risk indicators (Table 2.2.2) the percentage was 86.2%. The implication is that the parsimonious risk-profile is nearly as good for predictive purposes as the one that contained all available risk indicators.

**Joint exposures**

In a next step, we assessed the potential health benefits when prevention is targeted at people who are exposed to combinations of risk indicators (Table 2.2.4).

We took depressive symptoms as a starting point, because this risk indicator is associated with the highest IRR and AF values and has the lowest NNT. From Table 2.2.4 it is clear that when these people with subsyndromatic depression also experience impairment, then the IRR and AF values rise and ER and NNT values drop even further (row 4 in Table 2.2.4). In other words, the joint exposure to both depressive symptoms and functional limitations is associated with better values overall.

The indices of impact and effort can be optimized further when these people also have a smaller than average social network. This group represents 11.7% of the population, has a risk of becoming depressed which is higher by a factor of 4.5; if the adverse effect of the exposure to all three risk indicators could be blocked completely, then the incidence rate of depression in the population would drop by 32.3% (row 9).
**Table 2.2.4 Cumulative effect of joint exposures**

<table>
<thead>
<tr>
<th>Joint exposure</th>
<th>ER, %</th>
<th>IRR</th>
<th>AF, %</th>
<th>NNT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Depressive symptoms</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. + female gender</td>
<td>29.3</td>
<td>3.00</td>
<td>37.2</td>
<td>9</td>
</tr>
<tr>
<td>2. + low education</td>
<td>20.6</td>
<td>3.06</td>
<td>31.8</td>
<td>8</td>
</tr>
<tr>
<td>3. + chronic diseases</td>
<td>20.4</td>
<td>2.86</td>
<td>28.7</td>
<td>8</td>
</tr>
<tr>
<td>4. + functional limitations</td>
<td>20.8</td>
<td>3.58</td>
<td>37.5</td>
<td>7</td>
</tr>
<tr>
<td>5. + small network</td>
<td>24.8</td>
<td>3.16</td>
<td>36.2</td>
<td>8</td>
</tr>
<tr>
<td><strong>Dep. Sympt. + functional limitations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. + female gender</td>
<td>14.3</td>
<td>3.60</td>
<td>28.7</td>
<td>6</td>
</tr>
<tr>
<td>7. + low education</td>
<td>11.1</td>
<td>3.79</td>
<td>25.9</td>
<td>5</td>
</tr>
<tr>
<td>8. + chronic diseases</td>
<td>12.4</td>
<td>3.16</td>
<td>23.0</td>
<td>7</td>
</tr>
<tr>
<td>9. + small social network</td>
<td>11.7</td>
<td>4.53</td>
<td>32.3</td>
<td>5</td>
</tr>
<tr>
<td><strong>Dep. + func. lim. + small network</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. + female gender</td>
<td>8.0</td>
<td>4.60</td>
<td>24.2</td>
<td>4</td>
</tr>
<tr>
<td>11. + low education</td>
<td>6.7</td>
<td>4.59</td>
<td>21.9</td>
<td>4</td>
</tr>
<tr>
<td>12. + chronic diseases</td>
<td>6.5</td>
<td>4.25</td>
<td>19.7</td>
<td>4</td>
</tr>
</tbody>
</table>

ER = exposure rate, IRR = incidence rate ratio, AF = population attributable fraction, 95% CI = 95% confidence interval, NNT = number needed to be treated

Preventive interventions are likely to become even more cost-effective when they are targeted at people who are in addition suffering from chronic diseases (row 12), or have attained only elementary education levels (row 11) or are women (row 10). It is of note that the latter risk-profiles correspond to relatively small groups (6.5 – 8.0% of the population). Smaller target groups will reduce the effort that has to be put into the logistics of the future interventions. The interventions are also likely to be more efficient judging by the low NNT value of only 4.
**Figure 2.2.1** Effect of adding extra risk indicators on the attributable fraction, AF (in %), exposure rate, ER (in %), and number-needed-to-be-treated, NNT.

This minimization - maximization process is depicted graphically in Figure 2.2.1. It shows how consecutively adding an additional risk indicator to the risk profile impacts on the AF, ER and NNT. In general, one would like to have the AF-curve as high as possible (in order to maximize the health benefit in the population), while, at the same time, it is better to have the ER and NNT-curves as low as possible (in order to target smaller population segments with greater efficiency). Thus, in the first step, the risk indicator ‘presence of depressive symptoms’ yielded an AF of 40.3%, indicating that the incidence rate of depression in the population would be reduced by as much as 40.3% when the occurrence of full-blown depressive disorder is prevented in all people with depressive symptoms. The ER indicates that this is a formidable task as 40.3% of the population should then receive the intervention. Furthermore, the efficiency of such an approach is not optimal, because the NNT indicates that 16 people of this target group must receive the intervention in order to avoid one new case of depressive disorder. This can be improved. In a next step there are several risk indicators to chose from, but ‘functional limitations’ keep the level of AF still high (now 37.5%), while a smaller segment of 20.8% of the population has to be targeted and the efficiency is better (NNT= 7). This process of adding new risk indicators continues until a target group has been selected that is associated with the best IRR, AF, ER and NNT values.
Quantifying the health gain

Consider the risk-profile in row 10 of Table 2.2.4: the risk of becoming depressed is a factor 4.6 higher in women who have some depressive symptoms, experience functional limitations and have a small social network. If the adverse effect of the joint exposure could be completely blocked, then the incidence rate in the population would be reduced by 24%. The current incidence rate of 2.8 cases per 100 person years would then become 2.8*(1-0.24) = 2.1 cases per 100 person years. In every one million older adults this would represent roughly 5,950 prevented cases per year. Of course, this number represents an upper limit of the achievable health gain, because preventive interventions are unlikely to be totally effective. Assuming that an intervention is 30% successful in avoiding new onsets (Cuijpers et al, 2005), then 1,785 onsets will be avoided. This still represents a substantial health gain in a source population of one million.

Cost-benefit considerations

Avoiding new onsets has economical ramifications. Katon and colleagues (2003) computed the excess costs of minor and major depression as at least $1045 per person per half year. They called this a conservative estimate. Preventing 1,785 onsets would thus result in a cost saving of at least 1.9 million dollars in every million elderly people in the population. It is also clear that the costs of the intervention will be balanced by its savings when the costs of the intervention will not exceed $1045 per avoided case. In the same vein, the NNT value can be used to calculate the costs per recipient of the intervention where the break-even point is reached of the costs and benefits of an intervention. This type of *ante-hoc* cost-benefit calculations can be used to select interesting preventive interventions for further cost-effectiveness research.

2.2.4 Discussion

We aimed to answer the question as to whether it would be possible to target depression prevention where it generates the best health benefits against the lowest cost. This would help to guide research efforts towards more promising research areas in preventive psychiatry.

Main findings

This study showed that it is possible to use longitudinal epidemiological data to select ‘cost-effective’ risk indicators. These are risk indicators that are associated with a substantial population attributable fraction (AF) and a low number-needed-to-be-treated (NNT). When the costs of the disorder are known from a cost-of-illness study, then it is also possible to combine the AF and NNT with the known costs into an ante-hoc cost-benefit analysis.
In short, we have the methodology at our disposal that could help to identify cost-effective preventive interventions at a very early stage of the costly and time-consuming cycle of development and evaluation of preventive interventions. Having said this, we need to add that ultimately the cost-effectiveness of preventive interventions has to be established in proper cost-effectiveness studies. The methodology of identifying interesting risk indicators for prevention is not new (Miettinen, 1974; Morgenstern and Bursic, 1982), but in the field of psychiatric epidemiology and prevention research it has rarely been applied (Smit et al, 2004). The reason for this omission is that this methodology requires unselected population-based longitudinal data on the incidence of the disorder and its putative risk indicators. These data are often not available, but once there, could be used to set a rational research and development agenda for preventive psychiatry. We applied this to late-life depression and came up with the following key-findings.

First, the incidence of clinically relevant late-life depression is 2.8 new cases per 100 person years. This must be seen as a conservative estimate, because rigorous criteria were used to avoid false-positive onsets. Therefore, the true incidence rate is likely to be higher.

Second, starting from a list of putative risk indicators only a few were identified as interesting from the prevention perspective when the effects of the risk indicators are adjusted for competing risks. These are (-) female gender, (-) low education, (-) having two or more chronic illnesses, (-) experience of functional limitations, (-) having above average symptom levels of depression not exceeding the threshold for clinically relevant depression, i.e., CES-D scores above 5 and below 16, and, finally, (-) having a small social network.

Third, the combined effect of being exposed to three or four selected risk indicators yield statistically significant and substantially interesting values on measures of potential health benefit (IRR, AF) and effort (ER, NNT). It is also worth noting that the joint exposure to more risk indicators is limited to a small fraction of the older population. The intervention thus has a narrow focus and the corresponding number of people is manageable from the prevention perspective at regional level. Even based on conservative assumptions, a preventive intervention could prevent 1,785 new onsets in every one million older adults.

Fourth, avoiding new onsets also has economical consequences. A study carried out in the United States shows that elderly people with minor or major depression generate on average $1045 in excess costs per person per half year. Again, this is a conservative estimate. Avoiding the aforementioned 1,785 onsets would thus create substantial savings. However, these savings are unlikely to completely offset the costs of a preventive intervention, but the cost savings form a good vantage point for cost-effective prevention of late-life depression.
**Strengths and limitations**

The findings have to be placed in the context of the strengths and limitations of this study. The strengths are (-) the use of population-based data, (-) the prospective design which enables the study of incidence and facilitates etiological inference, and (-) the measurement of exposures which is not biased due to *post-hoc* rationalization on the part of the respondents, because at $t_0$ they could not have any knowledge about their future health state at $t_1$. Furthermore, this study is among the first to show how a statistical technique could be applied to quantify potential health benefits and the effort required to generate these health benefits in the field of preventive psychiatry. It thus supplies the sort of methodology which is of importance for setting a rational Research and Development agenda for depression prevention.

The limitations of this study consist in the not very detailed measurement of the exposures. We do not know for how long and how intensely subjects were exposed. Moreover, the number of studied risk indicators is limited in that, for example, genetic and other biological risk indicators were not included. Another limitation is the measurement of depression with the CES-D, which is not a diagnostic instrument. However, it has good psychometrical properties. People who are exposed to several risk factors may well form a population segment that is not very responsive to health oriented interventions and this may affect the health gain that can be delivered by prevention. This is an important issue, which needs more research. Finally, it is not safe to generalize the economical findings of this study to countries other than the United States, because excess costs are likely to differ from one country to another.

Conceptually, it would be useful to distinguish between risk indicators that are amenable to change, such as depressive symptoms, from those that are not. It is also worth noting that some risk indicators are not modifiable, like chronic illnesses, but their adverse psychological effects might be contained. Finally, there are risk indicators, such as female gender, which are not modifiable, but are valuable from the perspective of identifying groups at risk – which was the principal aim of this paper.

**Directions for future research**

In future research, it would be valuable to cross-validate the methodology presented in this paper by comparing it with alternative methodologies, such as Classification and Regression Tree (CART) analysis (Lemon et al, 2003). The authors will present such a cross-validation study in Chapter 2.3. It is also recommended that in the future the methodology presented in this paper or related methodologies should be further developed and employed to targeting both research and prevention where they are likely to be most cost-effective.
Part II

Do we know where to begin?

Chapter 2.3

Prevention of late-life depression in primary care, where to begin?

This chapter is based on:


of related interest:

Chapter 2.3

Prevention of late-life depression in primary care, where to begin?

2.3.1 Introduction

The adverse consequences of depression are well established. Currently, of all illnesses, depressive disorder is causing the largest amount of non-fatal burden, accounting for almost 12% of all years lived with disability worldwide (Ustun et al, 2004). Depression is also associated with excess mortality (Cuijpers and Schoevers, 2004), higher demands on caregivers and higher service utilisation, and has substantial economical implications (Charney et al, 2003). In community dwelling elderly, the prevalence of depression requiring clinical attention is 13.5 % (Beekman et al, 1999), and more than 50% have a chronic course (Cole et al, 1999; Beekman et al, 2002). Although effective treatment is available, case finding and adequate treatment in primary care are generally poor (Tiemens et al, 1996; Sonnenberg et al, 2003). Still, even if all patients with depression were optimally treated with evidence-based interventions, only 34% of the depression-related disease burden could be averted (Andrews et al, 2004). From the public health perspective, it is therefore of great interest to consider the possibility of depression prevention (Smit et al, 2004; 2006a). However, evidence-based preventive strategies aimed at reducing the incidence of late life depression in community living elderly are sparse.

In prevention, different strategies can be chosen according to the stages or transitions in the development of a disorder (Haggerty and Mrazek, 1994). Universal prevention aims to influence the behaviour of the whole population in order to prevent the onset of disease. Examples of universal prevention are programs informing the population about the risk of alcohol intake or the benefits of physical exercise. Selective prevention is aimed at persons who are at risk because they have been exposed to certain risk factors. In the case of depression, risk indicators are for example loss of spouse and physical illness (Schoevers et al, 2000). The third form, indicated prevention, targets persons who already have early or subthreshold depression, in whom an intervention may reduce the likelihood of becoming a full-blown case (Haggerty and Mrazek, 1994; Munoz et al, 1996). An example of indicated prevention is cognitive therapy for the prevention of psychosis in people at ultra-high risk (McGorry et al, 2002a; 2002b), or pharmacotherapy for people with mild cognitive impairment to delay the onset of Alzheimer’s disease (Jelic and Winblad, 2003).

The choice for a specific type of prevention depends on the prognosis of the untreated disorder, in combination with the costs, benefits and possible adverse
consequences of different types of intervention. An ethical dilemma is that, by identifying a healthy person as a possible future case and starting some kind of intervention, this in itself carries certain negative consequences and should only be applied if the alternative of no intervention has a significantly higher probability of adverse consequences. In the case of depression, evidence-based universal prevention of depression is currently non-existent. Selective prevention may focus on persons with certain risk indicators such as exposure to loss events. Indicated prevention would be directed at persons with depressive symptoms below the diagnostic threshold for clinically relevant depression (Pincus et al, 1999; Cuijpers and Smit, 2004). For both ethical reasons and reasons of cost-effectiveness, preventive measures aimed at reducing incidence should target subjects with high a priori risk through exposure to multiple risk factors (Cuijpers, 2003). Furthermore, for practical reasons, persons at risk should be easily identifiable in primary care, and risk profiles have to be simple and unambiguous. From a public health perspective, prevention should lead to a substantial reduction of total disease burden and be economically affordable. This implies that the selected risk indicators should be indicative of a substantial proportion of new cases.

The current study explores selective and indicated prevention models to identify high-risk groups among older GP patients, bearing in mind the above qualifications. It seeks to identify those combinations of exposures that predict the largest health gains in the most cost-effective way when prevention is successful in blocking the adverse effects of these risk factors. These risk factors could then be used in primary care as an easy to use checklist or screener to identify patients at elevated risk of becoming a future case of depressive disorder. For indicated prevention, the use of a screening instrument to detect subthreshold depression would be required. As this is more time-consuming than recognising subjects on the basis of more straightforward risk factors needed for selective prevention, the benefits of both approaches are compared. The study is based on data from the Amsterdam Study of the Elderly (AMSTEL study), a large and prospective cohort study of depression in older people living in the community. The data-set encompasses a comprehensive set of risk factors associated with late-life depression (Schoevers et al, 2000; Schoevers et al, 2005).

2.3.2 Methods

Subjects and procedures
The AMSTEL study is a prospective study on the incidence of depression in a large and representative cohort of non-institutionalized persons over 65 years of age. Informed consent was obtained from each participant before the start of the study. The study was approved by the Medical Ethics Committee of the Vrije Universiteit Amsterdam. The population base for AMSTEL included all non-institutionalised individuals, in the 65-85 age bracket, who lived in the city of Amsterdam and were
registered with a general practitioner at baseline. The study was actively supported by the general practitioners, whose lists were used as the sampling frame. In the Netherlands, general practitioners are the gatekeepers to the health care system and almost every citizen is on the list of a general practitioner. In this role, general practitioners in the Netherlands provide social support and have a long-standing personal relationship with their patients. Thus, the source population consisted of nearly the whole non-institutionalised population. The sample was drawn from a list of 30 general practices spread throughout the city; practices were selected from all practices registered within the city of Amsterdam, 22 randomly and 8 by convenience from a network of general practitioners participating in medical research. The profile of the over-65 general practice-population in terms of age and gender, correspond to the non-institutionalised population in Amsterdam (Launer et al, 1993; 1994). Within each practice, respondents were randomly selected from four age strata spanning five years each (65-69 to 80-84). In order to obtain equally sized age-strata at follow-up, the older old were over-sampled. Out of a sampled total of 5,666, 4,051 (71.5%) responded and formed the baseline sample. Interviews with these subjects were conducted between May 1990 and November 1991. The sample corresponds to the non-institutionalised Amsterdam population in terms of age and gender. A flow chart is presented in Figure 2.3.1.

**Figure 2.3.1** Participant flow in the Amstel study

At follow-up three years later (median 38 months), 2,244 (55.4%) were re-interviewed, 656 (16.2%) people were deceased, 282 (7.0%) were too ill or too cognitively impaired to respond, 662 persons (16.3%) refused further co-operation, and 207 (5.1%) were not available for interview due to a variety of other reasons.
For this study, subjects with depression (523; 12.9%) or dementia (261; 6.4%) at baseline were excluded. The study cohort consisted of all respondents at follow-up who were neither depressed nor demented at baseline, for whom complete data were available at follow-up (N= 1,940).

**Measures**

An one-hour interview was developed to gather information on psychiatric symptoms, demographic and medical status, personal history of depression and family history. The interview consisted of the Dutch translation of the Mini-Mental State Examination (Folstein et al, 1975), all Geriatric Mental State Examination-items related to organic, affective and anxiety syndromes (Copeland et al, 1986; 1988), the Activities of Daily Living (ADL) scale (Katz et al, 1963), the Instrumental Activities of Daily Living (IADL) scale (Lawton and Brody, 1969), and part of the CAMDEX-interview (Roth et al, 1986). The interview was conducted during home visits by lay interviewers who were trained using video sessions and regularly supervised.

**Psychiatric diagnoses.** Diagnoses of dementia, depression and generalised anxiety disorder were made according to the GMS-AGECAT system. Diagnostic levels 3-5 correspond reliably to cases of depression requiring clinical attention in both the community (Copeland et al, 1990) and in elderly hospital patients (Ames and Tuckwell, 1994). GMS-AGECAT has a proven reliability in epidemiological studies, as became evident through replication studies (Copeland et al, 1988). GMS-AGECAT generates both a non-hierarchical syndrome level and a more narrowly defined diagnostic level. The diagnostic case level is calculated from the syndrome level using a hierarchy from organic to depression to anxiety disorder. In order to be able to also assess overlapping comorbidity influences, syndrome levels were used in the analyses as otherwise the diagnostic hierarchy would bias the results. Depression caseness at baseline and follow-up was defined as a GMS-AGECAT level 3 or higher (Copeland et al, 1988). Subjects not depressed at baseline who had become depressed at follow-up were considered to be incident cases. Subjects with depression case levels 1 or 2 at baseline were classified as having subthreshold depression. The same GMS-AGECAT package with an identical algorithm was used in the second wave of the study. When re-interviewing, raters were unaware of previous data and diagnoses.

**Risk indicators.** Socio-demographic variables included age, gender and educational level, the latter of which was dichotomised into lower (primary school or less) and higher (more than primary school) education. Living situation was assessed with the associated questions in GMS-AGECAT. A personal history of depression was ascertained by the relevant CAMDEX-question. The question was considered affirmative if treatment had previously been requested. The presence of chronic diseases was assessed with the pertinent CAMDEX questions on myocardial infarction, stroke, cancer, lung disease, diabetes, Parkinson's disease, arthritis and epilepsy. Cognitive status was assessed by MMSE-score. Cognitive
impairment was defined as a MMSE score below 24. Sleeping disorder was determined by two questions from the GMS on trouble falling asleep (yes, or, I would have problems if I didn’t use sleeping medication), and early wakening (at least two times per week at least two hours earlier than usual). Disability was measured using a combined scale consisting of all ADL and IADL items (Kempen and Stuurmeijer, 1990; Kempen et al, 1995). Subjects were considered disabled if the total score was two or more points below the maximum. This indicates that subjects were in need of help to perform at least two of the tasks mentioned, or were unable to perform at least one task.

**Analysis**

The analyses were conducted in several steps. First, standard epidemiological techniques were used to calculate the absolute (AR) and relative risk (RR) of becoming a case in subjects with or without exposure to a risk indicator.

In a second step, the importance of risk indicators for onset of disorder at population level was quantified by taking into account the exposure rate (ER), and the population attributable fraction (AF). The AF indicates by how many percents the incidence rate of depression in the population will be lowered when the adverse effect of the risk factor (or combination of risk factors) can be completely blocked. Similarly, a measure of the efficiency of a preventive intervention can be calculated: the number-needed-to-be-treated (NNT; Cook and Sackett, 1995). The NNT shows how many people must be protected from the adverse influence of a risk factor in order to avoid one new case of depression. The statistics, such as ER, RR, AF and NNT, were obtained using Stata SE / 7.0 (Stata Corporation, 2001).

In the third step, the statistics were used to find the best values for efficiency (low NNT values, preferably in small ER groups), and the best values for health gains (high RR and AF values) for *combinations* of exposures. To trace the effect of multiple exposures in all their combinations we used a method known as classification and regression tree (CART) analysis (Everitt, 2003; Lemon et al, 2003), and adapted it for our purposes. CART analysis can graphically be represented as tree-like figures (see Figure 2.3.2, later). First, a parent node is selected that optimises the NNT, RR and AF values in preferably the smallest (ER) groups. The parent node branches off in two directions, creating child nodes in which people are not only exposed to the risk factor in the parent node, but also to an additional risk factor in the child node. Therefore, the first generation of child nodes captures the effect of double exposures, the second generation of child nodes represents the effect of triple exposures, and so on. At each node, both branches (for subjects with and subjects without that risk indicator) are followed up to determine the ER, NNT, RR and AF values of the combined exposures. Thus, the sample is split into smaller subgroups at each step with the aim to obtain better values of ER, NNT, RR and AF in each step.
In order to determine the optimal child nodes, the following hierarchy of decision rules was used:

1. Select the child node with the lowest NNT to obtain the best efficiency.
2. Select the child node with the highest AF to obtain the greatest health gain.
3. When ties occur: select the child node with the smallest ER, because targeting smaller groups is logistically and economically less demanding when NNT and AF values are equal.

The following rules were employed for the termination of branches:

1. Terminate when no risk is involved: RR < 1.00 or RR not significant
2. Terminate when no health gain can be obtained: AF < 5% or AF not significant
3. Terminate when interventions are unlikely to be efficient: NNT > 50 or NNT not significant.

2.3.3 Results

Sample characteristics and response pattern
The baseline sample characteristics are shown in Table 2.3.1 and have been described in more detail elsewhere (Schoevers et al, 2000). In multivariate logistic regression analysis, non-response at the follow-up assessment was predicted by higher age, male gender, lower education, living alone, chronic disease(s), disability, organic caseness and depression (dichotomised as depression case level vs. no depression case level). When subjects who died between measurements were excluded from these analyses, only lower education, living alone and baseline organic syndrome predicted non-response. At follow-up 309 subjects (15.9% of the study sample) had become depressed. The following analyses investigate the associations of risk factors with incident depression.

Predictors of late-life depression
Table 2.3.1 shows the bivariate associations of the risk indicators (at $t_0$) with the onset of depression (at $t_1$). The NNT values for the risk factors in the model range from 242 (living alone, not statistically significant) to 4.1 (95% CI= 2.3 ~ 19.6) for subjects with generalised anxiety disorder (GAD). For each risk factor the associated AF was also calculated, ranging from 24.6 (subthreshold depression, 95% CI= 18.8 ~ 30.0) to 1.1 (living alone, not statistically significant). Statistically significant associations between risk indicators and depression incidence were found for the following risk factors: medical illness, disability, history of depression, recent loss of spouse, sleep disturbance, generalised anxiety disorder and subthreshold depression.
Table 2.3.1 Bivariate associations of risk indicators and onset of depression

<table>
<thead>
<tr>
<th>Risk indicators</th>
<th>ER, %</th>
<th>RR</th>
<th>AR, %</th>
<th>NNT</th>
<th>AF, %</th>
<th>P&lt;0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low education</td>
<td>35.6</td>
<td>1.12</td>
<td>17.1</td>
<td>54.9</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>Living alone</td>
<td>44.0</td>
<td>1.03</td>
<td>16.2</td>
<td>24.2</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>Medical illness</td>
<td>47.8</td>
<td>1.46</td>
<td>19.1</td>
<td>16.6</td>
<td>18.1</td>
<td>*</td>
</tr>
<tr>
<td>Functional impairment</td>
<td>17.7</td>
<td>1.74</td>
<td>24.5</td>
<td>9.6</td>
<td>11.5</td>
<td>*</td>
</tr>
<tr>
<td>Cognitive impairment</td>
<td>2.6</td>
<td>1.13</td>
<td>18.0</td>
<td>47.0</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>Depression history</td>
<td>13.0</td>
<td>1.61</td>
<td>23.7</td>
<td>11.2</td>
<td>7.3</td>
<td>*</td>
</tr>
<tr>
<td>Recent loss of spouse</td>
<td>6.9</td>
<td>2.29</td>
<td>33.6</td>
<td>5.3</td>
<td>8.2</td>
<td>*</td>
</tr>
<tr>
<td>Trouble sleeping</td>
<td>30.4</td>
<td>1.73</td>
<td>22.6</td>
<td>10.5</td>
<td>18.2</td>
<td>*</td>
</tr>
<tr>
<td>Generalised anxiety disorder</td>
<td>1.3</td>
<td>2.56</td>
<td>40.0</td>
<td>4.1</td>
<td>2.0</td>
<td>*</td>
</tr>
<tr>
<td>Subthreshold depression</td>
<td>22.7</td>
<td>2.43</td>
<td>29.3</td>
<td>5.8</td>
<td>24.6</td>
<td>*</td>
</tr>
</tbody>
</table>

ER= exposure rate, RR= relative risk, AR= absolute risk, NNT= number needed to be treated, AF= population attributable fraction, * risk indicator with significant RR, NNT and AF at P < 0.05. The 95% confidence intervals are omitted from this table, but reported in the original article (Schoevers et al, 2006).

Indicated prevention model

Following the algorithm described in the Methods section, subthreshold depression clearly was the preferred primary node (Figure 2.3.2). Although the NNT of 5.8 was higher than that of GAD (4.1, see Table 2.3.1), the associated AF of depression was far greater (24.6%), and that of GAD (2.0%) was below the threshold of 5% for the termination of branches. The prevention model was then continued for each branch, starting with subjects who had, and those who did not have subthreshold depression. In subjects with subthreshold depression, the optimum child node was having a disability. This step reduced the NNT to 3.9, but also limited the AF to 9.7%. The model was then further subdivided with living alone and female sex as subsequent risk indicators. As the AF fell below 5% in the next step (all subjects → subthreshold depression → disability → living alone → female sex → ...), the branch was terminated here. In subjects with subthreshold depression who were not disabled, having a chronic illness (AR 30.9%, NNT 6.0) and living alone (AR 32.9%, NNT 5.6, AF 4.9%) were chosen as subsequent child nodes. In subjects who did not have depressive symptoms, recent widowhood was the strongest predictor, with an AR of 31.4% of developing depression, an NNT of 6.1 and AF 5.4%. Further branching was not performed as the AF became too small (<5%).
Figure 2.3.2 Selecting high-risk groups for indicated prevention

AR = absolute risk, NNT = number needed to be treated, AF = population attributable fraction
The way in which AR, NNT and AF are interrelated is graphically illustrated in Figure 2.3.3. It shows how adding more risk factors to the model results in a higher AR, a lower NNT, but also in a lower AF.

**Figure 2.3.3** Increase in AR (%) and decrease in NNT (No) and AF (%) when risk factors are added to the model

![Graph showing the relationship between AR, NNT, and AF when risk factors are added to the model.](image)

**Selective prevention model**

Figure 2.3.4 shows the model without subthreshold depression. Now the focus is on selective prevention and easy-to-recognise cases. Following the algorithm described above, loss of spouse, with NNT 5.3 and AF 8.2%, was selected as the parental node. The selective prevention model then yielded medical illness as the optimum child node in subjects who lost their spouse. This step reduced the NNT to 3.7, but also limited the AF to 5.8%. Because of this, further branching was not performed. In subjects who had not recently lost their partner, being disabled was the most important risk indicator, with NNT 11.5 and AF 9.0%. Further risk indicators were the presence of a chronic medical illness, and female gender. The latter step reduced the AF, but only slightly affected the NNT (from 8.6 to 8.3).
2.3.4 Discussion

In this study, high-risk groups for incident depression were identified in a large cohort of elderly GP patients. As case finding is generally poor, and both prevalence and persistence of late-life depression are of major concern, the goal was to identify groups in primary care with a high vulnerability for depression according to easily identifiable criteria. Prevention is of interest, bearing in mind that even optimal care can only

AR= absolute risk, NNT= number needed to be treated, AF= population attributable fraction , Not sign. = not significant at P<0.05.
reduce the burden of depression by 34% (Adrews et al, 2004). Since, in primary care, the detection of subthreshold depression requires more effort than recognising easily identifiable risk indicators, two models were compared that either incorporated or excluded subthreshold depression as a predictor of the future onset of a full-blown depressive disorder. For both scenarios, the potential health benefits of preventive measures aimed at reducing the incidence of late-life depression could be determined. Although this approach, combining clinical, epidemiological and public health perspectives, has been strongly advocated (Munoz et al, 1996), its application is relatively new in the field of preventive psychiatry.

Our findings suggest that indicated prevention – which involves identifying subthreshold depression as the most important risk indicator – is the preferred option for identifying groups at high risk of developing depression. Subjects with depressive symptoms had a risk of almost 30%, and accounted for 24.6% of new cases at follow-up. With an NNT of 5.8 – which carries the promise of efficient interventions in this group – these subjects are of major interest for preventive interventions. Further subdivision of the sample according to combinations of risk factors yielded groups with an even higher risk of developing depression and lower NNT values, but these more narrowly focused approaches have a smaller impact on the incidence rate of depression in the population. For example, disabled women with subsyndromal symptoms who live alone had a 47% risk of becoming depressed, a NNT of only 3.2, but a relatively low AF of 5.9% indicating that the incidence rate would only fall by 5.9% at best.

In a recent meta-analysis on short-term psychotherapeutic interventions directed at persons with subthreshold depression, these were found to reduce the incidence of depression by 30% (Cuijpers et al, 2005). If this were applied to a sample such as ours, the NNT of intervention would be 5.8 / 0.3 = 19.3 for persons with subthreshold depression. For a relatively intensive form of treatment, this may be considered to be too great an effort, and it could be argued that one should start with those subjects who have the highest vulnerability. Further subdivision in the CART analysis then leads to smaller subgroups that were exposed to more risk factors. In women with disability who live alone, the adjusted NNT would be 10.7 (3.2 / 0.3). As an alternative to cognitive behavioural therapy, less costly forms of indicated prevention should also be taken into consideration. Both bibliotherapy (Cuijpers, 1997) and the Coping With Depression course, a manualised form of self-help with instructions on mood management, were found to be effective therapies for unipolar depression, with effect sizes that are comparable to those of other treatment modalities for depression (Cuijpers, 1998). More recently, minimal contact psychotherapy has been proven effective (Willemse et al, 2004) and cost-effective (Smit et al, 2006b) in reducing the onset of depression in primary care patients with subthreshold depression. Such interventions have also proved effective when offered as low-cost computerised self-help interventions (McCrone et al, 2004; Christensen et al, 2004).
When one considers selective prevention of late-life depression, our analyses show that, in line with earlier findings (Cole and Dendukuri, 2003), elderly persons who recently lost their spouses are at great risk of developing depression, and even more so when they have, in addition, a chronic medical condition (AR 41.8%, NNT 3.7). Overall, the AF values in this model were somewhat smaller, which means that fewer cases can be prevented. Examples of preventive programs directed at persons who lost their spouse include self-help groups that convene for emotional exchange and support, specific courses on competences needed to cope with single life, and widow-to-widow programs in which women who had lost their husband earlier visit recently widowed persons for emotional and practical support. Although these programs have shown promising results in terms of post bereavement adaptation (Vachon et al, 1980) and social functioning (Van Lammeren and Geelen, 1995), the evidence of efficacy in reducing depression onset is limited. In a meta-analysis of eight types of such interventions, Cuijpers and colleagues (2000) found a standardised effect size of 0.34 in comparison with controls, but the number of published studies is still limited and this difference was not statistically significant. Considering the fact that these are low-cost interventions, randomised controlled trials are urgently needed.

Indicated prevention thus has the best chances of detecting large groups of subjects at high risk of developing depression, with NNT values that could make preventive actions cost-effective, using available evidence-based interventions. Still, in comparison with selective prevention, indicated prevention requires the extra effort of screening subjects for subthreshold depression. A number of screening instruments have been validated for case finding in community living elderly. These include the CES-D, a 30-item questionnaire on depressive symptoms (Radloff, 1977; Beekman et al, 1999; Prince et al, 1999), but more recently also the 15-item version of the Geriatric Depression Scale has demonstrated its effectiveness for detecting elderly subjects with depressive symptoms in the community (Almeida and Almeida, 1999; De Cran et al, 2003; Osborn et al, 2002). Screening of large numbers of patients in primary care has been shown to be both feasible and effective (Sheikh and Yesavage, 1986; Bijl et al, 2003; 2004). Primary care facilities therefore appear to be well equipped to find elderly persons with subthreshold depression.

**Strengths and limitations**

The Amstel study is a large prospective cohort study with long follow-up times, and it incorporated a wide range of risk factors for late life depression. The findings are representative for urban populations of community dwelling older persons, and are highly relevant for primary care.

Another major strength of this study is that its outcomes can be cross-validated against those of a related study (Smit et al, 2006a; see Chapter 2.2). In the latter, a different cohort of older people was used, different measures were taken on the pertinent risk factors and depression status, and even a different methodology was used. Nevertheless, in both studies we arrived at virtually the same model for
indicated prevention. That is, the same combination of risk factors was identified as
important for selecting high-risk groups for cost-effective prevention of depression.
There was one exception: in the current study recent widowhood was identified as
an important risk indicator, in the other study it was not, but it is worth noting that
in the current study we had measured recent widowhood (which had occurred in the
last six months), whereas in the other study the variable is better described as ever
widowed, without reference to a particular time-frame. This difference in the
variables may account for the difference in study outcomes.

Nonetheless, a number of limitations of the current study deserve
mentioning. First of all, cohort studies with a limited number of measurements over
relatively long time intervals tend to over-represent chronic forms of disorder.
Between assessments, subjects may have had both onset and remission of shorter
episodes of depression. At the same time it may be concluded that, even with a
three-year interval, the set of risk factors employed in this study can identify a large
number of future cases of depression. Second, although the risk factors cover many
domains relevant to depression, biological (genetic) risk indicators are less well
represented in the dataset. Other factors that would ideally have been available are
medication use, substance abuse or dependence, and specific medical conditions
such as thyroid disorders that may be associated with depression onset. Third, it
should be concluded that this study does not prove that preventive interventions are
successful in community-living elderly. Available studies on indicated prevention
were mostly performed in younger adults. However, as other forms of treatment of
depression are also effective in the elderly, there would be no reason to doubt the
efficacy of preventive interventions in elderly subjects. Considering the potential
benefits of prevention there is an urgent need for randomised controlled trials to
address this matter.

Conclusions
The results of this study underscore the importance of preventive action for
depression as the mental disorder associated with the largest non-fatal disability.
And to answer the question put forward in the title of this chapter: Yes, we do know
where to begin. Indicated prevention of depression is the preferred strategy. Still, as
this involves screening patients for subclinical depression, selective prevention of
depression may prove to be a good alternative. It will only require the use of a
simple checklist of the relevant risk factors (recent widowhood, disability, presence
of a medical illness, female sex), and the expected health gains can also be
substantial. Either way, primary care facilities are equipped to perform adequate
case finding. Nonetheless, even though minimal contact, low-cost, evidence-based
interventions for persons with subthreshold depression are available, their cost-
effectiveness has to be established in proper randomised trials. Recent data however
indicate that this may be the case (Willemse et al, 2004; Smit et al, 2006a; 2006b).
It is therefore time to reconsider whether the exclusive focus on treatment of
common mental disorders should be adapted, and whether a more generous
allocation of personnel and resources for prevention may be more effective in reducing both individual suffering and the overall costs of health care.
This chapter is based on:


Of related interest:


Chapter 3.1

A guided self-help intervention to prevent depression in primary care patients, randomised trial

3.1.1 Introduction

Subthreshold depression is highly prevalent (Horwarth et al, 1992; Cuijpers et al, in 2004) with a considerable impact on the quality of life (Wells et al, 1992), resulting in a strongly increased service utilisation (Wagner et al, 2000), and has economic consequences (Berto et al, 2004; Greenberg and Birnbaum, 2005). Furthermore, people with subthreshold depression have an increased risk of developing major depressive disorder (Broadhead et al, 1990; Eaton et al, 1995; Cuijpers et al, 2004). A person can be considered to have subthreshold depression when he has clinically relevant depressive symptoms, without meeting the diagnostic criteria for a full-blown major depressive disorder. Despite the clinical relevance of this condition, few studies have examined the effectiveness of interventions in subthreshold depression (Muñoz et al, 1995; Barrett et al, 2001; Clarke et al, 1995, 2001).

In this chapter we present the results of a trial in which primary care patients with subthreshold depression were randomised to routine primary care alone or to primary health care augmented with an adjunctive self-help intervention. The latter was based on cognitive behavioural psychotherapy with some telephone support. The primary outcome is the incidence of major depression. We predicted that the combination of primary care with minimal-contact psychotherapy is more effective in reducing the incidence rate of full-blown depressive disorder than primary care alone.

3.1.2 Method

Participants and procedures

Study participants were recruited from 19 general practices in the Netherlands. People were eligible if they were 18 – 65 years old, and had subthreshold depression here defined as having at least one core symptom plus one, two or three current depressive symptoms according to the Instel screening instrument (Tiemens, 1999), but not meeting the DSM-IV (American Psychiatric Association, 1994) criteria for full-blown depressive disorder. Exclusion criteria were: (a) the presence of hearing or language difficulties; (b) receiving psychotherapeutic treatment by a mental health professional in the last year, or being on the waiting list for treatment of mental health problems; (c) suffering from a life-threatening illness, mental retardation, suicidal risk, psychotic symptoms, schizophrenia or dementia according to the patient’s general practitioner (GP); (d) meeting DSM-IV criteria for
depressive disorder, dysthymia, bipolar disorder, social phobia, agoraphobia or panic disorder in the last 12 months.

Participants were recruited in two steps (see Figure 3.1.1). In the first step, a research assistant approached all patients who were waiting to see their general practitioner (N= 5,276). This was done in four weeks per GP. The Instel screening questionnaire (see Measures) was filled in during a brief face-to-face interview in a separate room. From the 4,525 patients who gave informed consent for screening, 3,825 patients were screened. The other 700 patients were excluded on the basis of age (younger than 18, older than 65 years) or criteria (a) and (b) as cited above.

**Figure 3.1.1** Flow of participants in the trial

- **Approached** N= 5,276
- **Screened** N= 3,825
- **Diagnostic interview** N= 363
- **Randomised** N= 216
  - **Experimental condition** N= 107
    - Intervention
      - Did not start: N= 23
      - Dropped out: N= 40
      - Completed: N= 43
    - 12-months follow-up N= 83
  - **Control condition** N= 109
    - 12-months follow-up N= 94
- **No symptoms** N= 2,351
- **Other disorder** N= 559
- **Non-response** N= 699
In total, 1,018 patients were assessed as possibly having a subthreshold depression. After office hours, the research assistant consulted the general practitioner to determine if screen-positive patients met exclusion criteria (c). In step two, screen-positive patients who were willing to participate in the trial received a telephone interview with the Composite International Diagnostic Interview (CIDI). The interviewer was aided by a computer during this interview. This resulted in 363 baseline interviews. The main reasons for non-response (64%) were lack of interest or lack of time (47%), and no telephone contact with the patient (14%, after at least ten calls at different times and days). In addition, patients who met the CIDI/DSM-IV diagnostic criteria for a mood disorder, social phobia, agoraphobia or panic disorder in the last 12 months were excluded (N= 95). Patients meeting all inclusion criteria and who gave informed consent were randomised to the minimal-contact psychotherapy condition (N= 107) or to the care-as-usual (N= 109).

**Interventions**

The experimental intervention was minimal-contact cognitive-behavioural therapy for depression, based on the “Coping with Depression” course (Lewinsohn et al, 1984), in the Dutch version (Cuijpers, 2000). The main component is a self-help manual with instructions on cognitive-behavioural self-help in mood management skills. This intervention has proved to be effective in reducing depressive symptoms in several randomised clinical trials, both in group and individual format, and in minimal-contact format (Cuijpers, 1998). The manual also includes registration exercises and homework assignments. The intervention was augmented by one face-to-face interview (of less than one hour) with a prevention worker or clinician from a community mental health centre before the participant started reading the manual, and six short telephone calls of less than 15 minutes each. The first five calls took place once every two weeks, and the sixth call took place two months after the fifth. The telephone contacts were not of a psychotherapeutic nature, but were to support the participants in working through the manual. Participants receiving the minimal-contact psychotherapy were able to make use of all other types of health services during the intervention period, including those offered by their general practitioner.

The control intervention was care-as-usual by the general practitioner and other health care services; the former were requested to base their treatments on the Dutch primary care guideline for depression (Van Marwijk et al, 1994).

**Design**

We conducted this study as a pragmatic randomised controlled trial with the aim to test the hypothesis that the intervention combined with care-as-usual offers added value when compared with primary care alone under realistic conditions. In other words, this study was not designed as an efficacy study and beyond randomisation no other attempts at control – such as blinding and using attention placebo – were implemented. This was done to improve its external validity. We will return to this issue in the Discussion.
A power calculation indicated that 200 participants were needed per condition to test the unidirectional hypothesis of a superior effect in the treatment arm of the trial in a one-sided test at $\alpha=0.05$ and a power of $(1-\beta)=0.080$. This sample size would power the trial to detect a difference in the incidence rate of 10% or more (Hully and Cummings, 1988). Fewer eligible people ($N=216$) were randomised, with equal probability, to one of both trial arms. The randomisation was carried out centrally, using a blocked randomisation scheme stratified by general practice with the patient as unit of randomisation, with blocks of four patients. Data were collected at baseline, and at 4, 12 and 24 months after baseline. The 4-month follow-up was meant to collect data on patient satisfaction with the intervention. This paper focuses on the intervention effects at 12-months follow-up. The trial protocol was endorsed by an independent medical ethics committee.

**Measures**

Because no (Dutch) screening instrument existed for subthreshold depression, we used the Instel-screen (Tiemens, 1999), which was adapted for this purpose. The Instel starts with w core questions about loss of interest and feeling tense, with a negative predictive value of 99% for either CIDI depressive disorder or generalised anxiety disorder in primary care patients. A depression scale follows after a positive answer on one of the check questions. For detecting subthreshold depression, the positive predictive values for at least one core symptom plus, one, two, or three current depressive symptoms were 16%, 61%, and 56%, respectively.

The primary outcome of interest is the relative reduction in the incidence of DSM-IV depressive disorder at the 12-months assessment, according to the CIDI-Auto version 2.1 (WHO, 1997) in its Dutch version (Ter Smitten et al, 1998). The CIDI is a standardised diagnostic interview for the assessment of mental disorders, developed by the World Health Organization. It was designed for use by trained lay interviewers. The reliability has been demonstrated to be good to excellent and the validity has been demonstrated to be adequate (Andrews & Peters, 1998; Wittchen, 1994). The interviews were conducted by telephone, as several findings provide qualified justification for this mode of assessing psychiatric disorder (Rohde et al, 1997; Evan et al, 2004). The interviewers had undergone a prior 3-day training at the Dutch CIDI training centre, followed by a 1-day training in adhering to the interview protocol. During the data collection, the results of the interviewers were continuously monitored. The interviewers were blinded with respect to the randomisation status of the participants.

Secondary outcome measures were depressive symptoms and subjective functioning. Depressive symptom level was ascertained with the Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977) in its Dutch version (Bouma et al, 1995); this is a widely used self-report scale measuring the frequency of 20 depressive symptoms during the past week. The CES-D generates a total score that can range from 0 to 60, with a higher score indicating more
depressive symptoms. The Dutch translation has good reliability and validity (Bouma et al., 1995).

Subjective functioning was measured by the RAND 36-item Health Survey (RAND, 1992) in the Dutch version, which has good psychometric properties (Van der Zee and Sanderman, 1993). The RAND-36 is a generic self-report measure covering eight scales with respect to subjective functioning (see Table 3.1.1). Each scale can range from 0 to 100, with a higher score indicating better functioning.

CES-D data were collected by telephone at baseline, 12-months follow-up and 24-months follow-up (together with the CIDI-interview). The RAND-36 data were collected at the same measurement points, but with the help of paper and pencil questionnaires. The same procedure was followed for the collection of satisfaction data at the 4-month follow-up.

Analysis
All analyses on treatment effects were carried out in Stata SE/7.0 (StataCorp, 2001) while accounting for the cluster effect induced by the fact that groups of participants shared the same GP. As there were no significant differences between both treatment arms on socio-demographics and the baseline scores on the CES-D and the Rand-36, it was not necessary to control for confounders.

Clinical outcomes were analysed on an intention-to-treat basis. Missing diagnoses of DSM-IV depressive disorder (18%) were imputed with help of a linear probability model, with age, gender and CES-D baseline scores as predictors.

With regard to the central clinical end-point, we wanted to test the hypothesis that the incidence rate in the experimental group \( (I_E) \) was lower than the rate in the control group \( (I_C) \) after one year. In other words, we tested the hypothesis that the incidence rate ratio, \( \text{IRR} = \frac{I_E}{I_C} \), was lower than 1. This was done by regression the imputed CIDS/DSM-IV depression status at 12-months follow-up on the treatment dummy in a Poisson regression while adjusting for the cluster effect. In addition, the number needed to be treated (NNT) was calculated. Robust standard errors, P-values and 95% confidence intervals were obtained by the first-order Taylor-series linearisation method as implemented in Stata.

For continuous outcomes (CES-D and Rand-36), last-observation-carried-forward was used for imputing missing values at follow-ups. For the imputed continuous outcomes T-tests, adjusted for the cluster effect, were used to test the hypothesis of superior treatment effects in the experimental arm.

We evaluated the effectiveness of minimal-contact psychotherapy in one-tailed tests, because we were interested in the added value of minimal-contact psychotherapy in the context of routine primary care. Unless otherwise specified, we used the 5% significance level (P<0.05).

Furthermore, for the CES-D the standardised effect size (Cohen’s d) was calculated by dividing the mean difference of the 12-months follow-up score of control and intervention group by the 12-months follow-up score standard deviation of the control group (Hedges & Olkin, 1985).
In order to compare completers and non-completers of minimal-contact psychotherapy on the level of satisfaction with the intervention, the exact non-parametric Mann-Whitney U test was used.

### 3.1.3 Results

**Sample characteristics**

In both conditions, 66% were women, 78% lived with a partner, and 83% were employed. The mean age in the experimental condition was 39.4 (Sd= 11.4) and in the control condition this was 41.8 (Sd= 11.2). The mean number of years of education was 14.1 (Sd= 3.6) and 13.5 (Sd= 4.2), in both conditions respectively.

<table>
<thead>
<tr>
<th>CES-D depressive symptom level, mean (Sd)</th>
<th>Experimental (n= 107)</th>
<th>Control (n= 109)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical functioning</td>
<td>85.8 (19.1)</td>
<td>83.4 (19.9)</td>
</tr>
<tr>
<td>Social functioning</td>
<td>72.4 (22.7)</td>
<td>71.8 (20.6)</td>
</tr>
<tr>
<td>Role limitations (physical)</td>
<td>57.8 (42.1)</td>
<td>61.4 (39.5)</td>
</tr>
<tr>
<td>Role limitations (emotional)</td>
<td>67.0 (38.4)</td>
<td>65.0 (38.8)</td>
</tr>
<tr>
<td>Mental health</td>
<td>66.1 (14.2)</td>
<td>63.4 (17.4)</td>
</tr>
<tr>
<td>Vitality</td>
<td>55.1 (17.7)</td>
<td>53.5 (17.8)</td>
</tr>
<tr>
<td>Pain</td>
<td>73.5 (23.1)</td>
<td>74.0 (21.3)</td>
</tr>
<tr>
<td>General health</td>
<td>61.8 (19.9)</td>
<td>62.2 (20.5)</td>
</tr>
</tbody>
</table>

The CES-D and Rand-36 baseline scores for each group are presented in Table 3.1.1. No significant differences (at P<0.05 or even P<0.10) were found between both groups on these variables.

Overall, 177 (82%) people participated in the 12-month follow-up interview. There were no significant differences in follow-up rates between the two trial groups. Completers and non-completers of the measurements were compared by means of logistic regression analysis on socio-demographic variables, CES-D, and Rand-36 baseline scores. It was found that older participants and participants with more years of education were more likely to complete the interviews at follow-up (OR= 1.04, 95% CI= 1.01–1.08, P< 0.05; OR= 1.13, 95% CI= 1.00–1.27, P<0.05, respectively). Moreover, men and participants with less perceived mental health were less likely to complete the interviews at follow-up (OR= 0.44, 95% CI= 0.20–0.99, P<0.05; OR= 0.97, 95% CI= 0.95–1.00, P<0.05, respectively).
Effects on the incidence of depressive disorder

At 12 months, the incidence rate of major depressive disorder was 0.12 (13/107) for the minimal-contact psychotherapy and 0.18 (20/109) for the usual-care condition. The incidence rate ratio was 0.66 (95% CI= 0.40~1.09), and significant in the one-sided test (P= 0.049). The number of people needed to be treated (NNT) in order to avoid one case of major depression was 16.

Effects on depressive symptoms and subjective functioning

The mean and standard deviations of the CES-D and RAND-36 at pre-test and at 12 months are presented in Table 3.1.2. minimal-contact psychotherapy found to have a significant effect on the CES-D and on two scales of the RAND-36, i.e. physical functioning and mental health. The mean standardised effect size for the CES-D was 0.18.

Table 3.1.2 Secondary outcome measures after 12 months

<table>
<thead>
<tr>
<th></th>
<th>Experimental condition (N= 107)</th>
<th>Control condition (N= 109)</th>
<th>One-sided T-test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CES-D depressive symptom level</strong></td>
<td>9.4 (7.4)</td>
<td>11.1 (9.4)</td>
<td>1.98 0.032</td>
</tr>
<tr>
<td>RAND-36 subjective health</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical functioning</td>
<td>86.7 (19.0)</td>
<td>82.5 (20.5)</td>
<td>1.77 0.047</td>
</tr>
<tr>
<td>Social functioning</td>
<td>77.6 (22.2)</td>
<td>74.4 (21.3)</td>
<td>0.93 0.182</td>
</tr>
<tr>
<td>Role limitations (physical)</td>
<td>69.9 (38.3)</td>
<td>65.4 (41.3)</td>
<td>0.73 0.237</td>
</tr>
<tr>
<td>Role limitations (emotional)</td>
<td>75.1 (36.4)</td>
<td>69.5 (39.4)</td>
<td>1.13 0.137</td>
</tr>
<tr>
<td>Mental health</td>
<td>71.5 (14.2)</td>
<td>67.1 (18.8)</td>
<td>2.07 0.027</td>
</tr>
<tr>
<td>Vitality</td>
<td>57.7 (18.5)</td>
<td>55.6 (18.0)</td>
<td>1.05 0.153</td>
</tr>
<tr>
<td>Pain</td>
<td>77.1 (23.5)</td>
<td>73.6 (23.9)</td>
<td>0.87 0.198</td>
</tr>
<tr>
<td>General health</td>
<td>64.5 (20.6)</td>
<td>64.2 (22.0)</td>
<td>0.10 0.462</td>
</tr>
</tbody>
</table>

CES-D, Center of Epidemiological Studies Depression scale; RAND-36, 36-item Health Survey. M= mean, Sd= standard deviation. T= test statistic, P= p-value. All analyses were performed on an intention to treat basis.

Acceptability of the intervention

Twenty-four (22%) of the 107 participants randomised to the minimal-contact psychotherapy did not start with the intervention and another 40 (37%) discontinued the intervention. Of these, 23 (58%) dropped out before the third telephone call. Reasons for not starting or drop-out varied: lack of time, lack of motivation, no telephone contact possible with the subject, not feeling depressed, other problems that deserved priority according to the subject.
Completers and non-completers of the minimal-contact psychotherapy were compared by means of logistic regression analysis on socio-demographics, and baseline scores CES-D, and Rand-36. It was found that men were less likely to complete the intervention than women (OR = 0.26, 95% CI= 0.10–0.69, P<0.01). Furthermore, non-completers of the minimal-contact psychotherapy perceived a significant better mental health at baseline than completers (OR= 0.96, 95% CI= 0.93–0.99, P<0.05). There was no significant difference in the drop-out rates across the conditions ($\chi^2=2.74; \text{df}=1; P=0.10$).

The 4-months follow-up questionnaire gave information about the satisfaction with the experimental intervention (response 70%; N = 75). In general, participants were satisfied or very satisfied with the coaching from the community mental health centre, the telephone calls (number, length, spreading), the homework (difficulty, relevance) and the content of the self-help book. However, participants who discontinued the intervention were significantly less satisfied than those who completed the intervention or who were still working through the intervention at the time of completion of the questionnaire. This was true for the coaching from the community mental health centre (U = 201.50, P<0.05), and for the experienced relevance of the homework (U = 157.50, P<0.05).

3.1.4 Discussion

Limitations

This study had several limitations. First, the trial was underpowered. Notwithstanding its lack of power, the study indicates that the intervention reduces the incidence of depression significantly. Second, the initial response was relatively high (86% agreed to be screened), but 36% of the people who were identified as having a high risk for subthreshold depression gave informed consent to participate in the study. Although problems with recruitment and attrition are common in randomised trials of psychological interventions in general practice (Muñoz et al, 1995; Fairhurst & Dowrick, 1996), this raises questions about possible selection-bias in the sample. Of the participants randomised to the intervention, almost 60% did not start or complete the intervention. On the other hand, most participants (82%) did complete the interviews at follow-up, and observations that were missing were imputed in accordance with the intention-to-treat principle. Third, this study was conducted in the Netherlands, and the results cannot be generalised safely to other countries. Fourth, it was not possible to mask participants to the condition they were assigned to; this is true for most randomised trials of psychological interventions, but it may nevertheless have distorted the outcomes of this study. Fifth, we should be careful to emphasise that we compared two composite packages of interventions: the intervention consisted of a self-help manual + supporting phone calls + care-as-usual; the control condition is best described as care-as-usual delivered by GPs following the pertinent guideline for the treatment of depressive
disorder. This is to say that we can not attribute the effects to a single component of the intervention, but always to the whole package relative to the smaller package represented by care-as-usual. In future research we would like to see whether similar effects can be found when the intervention is offered as self-help without telephone support. Such as study would shed light on a possible Hawthorn effect induced by paying extra attention to the recipients of the intervention.

Because of these limitations, the results of this study should be considered with caution.

Reduction of incidence of major depression
The incidence of major depression among participants in the experimental condition (12%) was reduced by one third compared to the control condition (18%). This is an important finding. Two other trials found positive and significant effects of cognitive behavioural interventions on incidence of major depressive disorder among adolescents with depressive symptoms (Clarke et al, 1995; Clarke et al, 2001). Until now, only one study has examined the possibility of preventing major depression in primary care patients with subthreshold depression (Muñoz et al, 1995). This study had insufficient statistical power to find a significant effect. So the current trial is the first that shows that major depressive disorder can be prevented with an intervention plus minimal support in primary care patients with subthreshold depression.

Despite this positive finding, we need to address the question of whether we focused this intervention on the right population. An incidence rate of 18% is considerable, but it is questionable whether this is sufficient to offer people a preventive intervention. After all, more than 80% of the participants will not develop major depressive disorder within 1 year, and only 6% will benefit in terms of a prevented episode of major depression. Furthermore, almost 60% of the participants who were randomised to the experimental condition did not start the intervention, or failed to finish it. These participants showed a significant better mental health at baseline compared to the completers, and one of the reasons given for withdrawal was that the person did not feel depressed. The natural course of the symptoms of these participants might have been so positive that an intervention was not necessary, but this possible explanation must be seen as an hypothesis.

Effects on depressive symptoms
The effect of the experimental intervention on depressive symptoms was significant, but relatively small. A reduction in depressive symptoms of 0.18 standard units is usually considered to be small (Lipsey, 1990); psychotherapy and pharmacological treatments for depression usually find reductions of 0.45 standard units or more (Lipsey & Wilson, 1993) and this is also true for earlier studies of minimal-contact psychotherapy (Cuijpers, 1997). There are several characteristics of our trial that may be responsible for this small effect. First, trial participants were actively
recruited by screening, whereas most other trials have recruited people who were seeking care for their depression, or through announcements in papers and other media, which may have resulted in a selection of more motivated individuals. Second, most earlier trials of minimal-contact psychotherapy used waiting list comparison groups, while we used a usual care comparison group. Third, the small effects might be specific to the group of primary care patients with generally good prognosis. Finally, it is also possible that the small effects are in part caused by the fairly low participation rates in the psychotherapy condition.

Acceptability and suitability of the intervention

None of the previous trials examining the effects of preventive interventions on the incidence of new cases has made use of minimal-contact psychotherapy. This can be defined as a psychological therapy in which the patient takes home a standardised psychological treatment manual and works it through more or less independently, with only minimal support from professionals. In a meta-analysis of six randomised trials of minimal-contact psychotherapy for full-blown depression it was found that the effects in subjects with depressive symptoms are large and comparable with effects of traditional psychotherapy and antidepressive medications (Cuijpers, 1997).

This form of psychotherapy is also an interesting intervention for subthreshold depression in primary care, as it is brief and non-intrusive, and does not require much time or specific skills from the doctor. Because it can be assumed that subthreshold depression does not justify full psychiatric treatment for a major psychiatric disorder, a minimal intervention aimed at preventing the onset of a depressive disorder seems adequate. The participants in our study, who started the intervention and finished a considerable part of it, assessed it as positive and helpful.

In a systematic review of eight studies examining self-help treatments for anxiety and depressive disorders in primary care (Bower et al., 2001), it was concluded that most studies in this area did find some positive results, but also that more research in this area is necessary. Our study adds to the empirical evidence that self-help treatments are beneficial for some patients, but also that more research is needed before this approach can be confidently recommended as an evidence-based practice.

Offering this minimal intervention to a much more selective patient population might increase the effectiveness as well as the efficiency of the intervention. For instance, in a recent trial a preventive intervention was offered to subjects who not only had subthreshold depression, but also belonged to a high-risk group – adolescent children of depressed parents (Clarke et al., 2001). In that study the incidence rate dropped from 29% to 9%. Another possibility for selection is to offer the intervention on a stepped care base by starting with a short waiting time to exclude people who recover quickly, because ‘watchful waiting’ in itself can be an
effective strategy (Barrett et al, 2001). Finally, ‘open’ recruitment strategies should be used to reach motivated individuals.

Final remarks
Depression is one of the most important causes of disability, and is expected to be the second leading cause of disability world-wide in 2020 (Murray & Lopez, 1996). Reducing the burden of depression is possible, as was shown by the study described here. Perhaps of equal importance is the opportunity to offer a simple self-help treatment, which can be effective while consuming only small amounts of health care resources. The further development and research of preventive interventions in this area constitute a major challenge for prevention science.
This chapter is based on:


Of related interest:

Chapter 3.2

Psychological treatment of subthreshold depression, meta-analysis

3.2.1 Introduction

It is well established that subthreshold forms of depression are not only highly prevalent (Horwarth et al., 1992; Cuijpers et al., 2004), but also clinically relevant. Community studies in which the criteria for minor depression according to the DSM-IV have been used, show that the prevalence rates range from 5 to 10% (Cuijpers et al., 2004; Kessler et al., 1997). When subthreshold depression is defined as scoring above a cut-off score in self-rating depression scales, prevalence rates are much higher (Beekman et al., 1999).

Subthreshold depression has been found to have a considerable impact on the quality of life of patients (Cuijpers et al., 2004; Preisig et al., 2001; Rapaport and Judd, 1998) and results in increased utilization of medical services (Wagner et al., 2000). It is also associated with an increased mortality rate (Cuijpers and Smit, 2002; Cuijpers and Schoevers, 2004), is associated with substantial economic costs both in younger adults (Cuijpers et al., 2005) and older adults (Katon et al, 2004). Furthermore, subthreshold depression carries a high risk of developing a full-blown depressive disorder in both the short term (Cuijpers and Smit, 2004) and the long term (Fergusson et al., 2005).

Subthreshold depression can be defined from at least three different perspectives. In the first perspective, it is assumed that depressive symptomatology exists on a continuum with no symptoms at one end, major depression at the other, and subthreshold depression in between (Goldberg, 2000; Solomon et al., 2001; Kessler et al., 1997; Gotlib et al., 1995). There is indeed considerable empirical evidence indicating that depression may best be conceptualized as a continuum (Geiselman and Bauer, 2000; Goldberg, 2000; Angst et al., 2000), although the possibility of a latent qualitative difference between clinical depression and subclinical depressive symptoms cannot be ruled out (Solomon et al., 2001). In the second perspective, subthreshold depression is considered to be a condition with unique characteristics that distinguish it categorically from other depressive conditions (Fechner-Bates et al., 1994). The definition of minor depression in the DSM-IV and other diagnostic classification systems such as the ICD-10 or the Research Diagnostic Criteria, are examples of this approach. In the third perspective, subthreshold depression is regarded as a part of the prodromal phase of major depression, or are residual symptoms in people who have recently recovered from a major depression. All or nearly all subjects who develop major depression can be assumed to have initially passed through a period (however, brief) of...
subthreshold depression. Although the first two perspectives are mutually exclusive, the third perspective does not rule out the other perspectives.

From a clinical point of view, subthreshold depression is important for two reasons. First, as subthreshold depression is often an invalidating condition with considerable psychological suffering, treatment is frequently necessary. The goal of this treatment is to reduce depressive symptomatology and to improve quality of life.

The second reason why subthreshold depression is important from a clinical viewpoint, is the increased risk of developing major depression. In an earlier systematic review, we found that the incidence rate of major depression in subjects with subthreshold depression in community studies ranges from 0.01 to 0.15 new cases per 100 person years, compared with 0.00 to 0.05 in subjects without subthreshold depression (Cuijpers and Smit, 2004). In studies among medical patients with subthreshold depression the incidence rates range from 0.06 to 0.58, compared with 0.00 to 0.23 in subjects without subthreshold depression (ibid.). Virtually all of the many studies that have examined the incidence rates of major depression in subjects with subthreshold depression compared to those without, confirm that the incidence rate is greatly increased in subthreshold depression. Because of the increased risk of developing major depression, interventions in subthreshold depression are often conducted with the aim to prevent the onset of major depression.

In this systematic review, we examine the effects of psychological interventions aimed at subjects with subthreshold depression. We focus exclusively on studies in which subjects do have clinically relevant depressive symptoms, but do not meet criteria for major depressive disorder or dysthymia. We will examine whether the interventions used in these studies are capable of reducing depressive symptoms in subjects with subthreshold depression, but we will also investigate whether these interventions will reduce the incidence of new cases of major depressive disorder. As far as we know, no systematic review in this area has been conducted before now.

3.2.2 Method

Identification and selection of studies
Relevant studies were identified and retrieved in several ways. First, we conducted searches in the computerized databases of the scientific literature: Pubmed (from 1966 – April 2005), Psychinfo (from 1960 – April 2005) and the Cochrane database. We combined search terms indicative of subthreshold (subthreshold OR subclinical OR minor OR mild) depression (major depression, depression, depressive), and randomised trial (randomised OR randomized OR clinical OR trial OR experimental). Both key-words and text words were used. Unpublished and grey literature were searched in Dissertation Abstracts. Second, references of
reviews of minor depression and other related subjects were examined (Judd et al., 2002; Oxman and Sengupta, 2002; Ackerman et al., 2002; Banazak, 2000; Pincus et al., 1999; Beck and Koenig, 1996). Third, the references of retrieved papers were studied for referrals that were missed so far.

We included randomised trials of psychological interventions in subjects with clinically relevant depressive symptoms. The latter were defined as scoring above a cut-off of a self-rating depression questionnaire; scoring above a cut-off score on a clinician-rated instrument; or meeting criteria for minor depression according to the criteria described in the DSM, ICD, or Research Diagnostic Criteria, but excluding major depressive disorder or dysthymia. Studies had to use a standardised diagnostic interview (such as the DISC, CIDI, or SCAN) to exclude the presence of full-blown mood disorder at pre-test. We also included studies in which subjects with a depressive disorder were included, but stratified during randomisation over symptom-level, and when results were separately reported for subjects with subthreshold depression. No language restrictions or age limits were applied.

**Analyses**

We conducted separate analyses for the effects of the interventions on depressive symptom-level (a continuous outcome) and for the effects on the incidence of major depressive disorder (a dichotomous outcome).

**Effects on depressive symptom-level.** We calculated standardised effect sizes, \( d \), by subtracting the average post-test score of the control group, \( M_C \), from the average score of the experimental group, \( M_E \), and dividing this raw gain-score by the mean of the standard deviations of the experimental and control group, \( Sd_{CE} \). A standardised effect size of 0.5 thus indicates that the health-gain in the experimental group is larger than the control group by half a standardised unit. In the field of psychological interventions, effect sizes in the range of 0.56 to 1.2 are regarded as large, while effect sizes of 0.33 to 0.55 are moderate, and effect sizes of 0.00 to 0.32 are small (Lipsey and Wilson, 1993). In the calculations of effect sizes, only those instruments were used that explicitly measure depression (see Table 3.1.1, later). When in a primary study more than one depression measure was used, then we calculated the mean of the effect sizes, such that each primary study, or study-contrast, provided only one effect size to the meta-analysis. When means and standard deviations were not reported, then we used test statistics, \( \chi^2(2) \), \( F(2) \), \( T \), and \( P \)-values, to calculate effect sizes (cf. Wolf, 1985).

To compute the pooled mean effect size across the primary studies, we used the computer program Comprehensive Meta-Analysis (version 2.2.021). We conducted all analyses using both the fixed effects model and the random effects model (Clarke and Oxman, 1999). Because both methods yielded comparable outcomes, and heterogeneity was low, we report all results only of the fixed effects model. As indicators of heterogeneity, we calculated the Q-statistic and the \( I^2 \)-statistic, in the latter of which a value of 0% indicates no heterogeneity, and larger
values show increasing heterogeneity, with 25% as low, 50% as moderate, and 75% as high.

**Effects on incidence of major depression.** Because the follow-up period of the studies differed considerably, we based the calculation of the incidence rates on person-years. That is, we divided the number of new cases of mental disorder (the numerator) by the total amount of person-years (the denominator). Technically, this is known as the person-time incidence rate, or the incidence density rate, which is an appropriate measure of incidence when follow-up times differ between subjects (Rothman, 1988). For each study we calculated the incident rate ratio, IRR, of getting a major depressive disorder in experimental subjects relative to the rate in control subjects.

Again, we conducted all meta-analyses, using the same computer program, with both the fixed effects model and with the random effects model. In the meta-analyses, we first calculated overall relative risks with the DerSimonian and Laird (1986) method. And again, the outcomes of the two models were comparable. Since the Q-statistic and the $I^2$-statistic indicated a virtual absence of heterogeneity between study outcomes, we report the results only of the fixed effects model.

### 3.2.3 Results

**Description of studies**

We examined a total of 1,309 abstracts from the Cochrane database (1,047), Pubmed (205) and Psychinfo (57). We retrieved a total of 21 papers, of which 14 were excluded. Most papers (9) were excluded because these studies included subjects with major depression and/or dysthymia; three papers were excluded because no diagnostic interview was used to assess the presence of mood disorder; and the remaining two papers were excluded for varying reasons (no psychological intervention, and not properly randomised).

The seven included studies are summarized in Table 3.2.1. The studies examined a total of 700 subjects, with 343 subjects in the experimental conditions, and 357 in the control conditions. The number of subjects per study varied from 24 to 216.

In five of the seven studies, cognitive behaviour therapy was used as the intervention. In all of these studies, an adaptation of the “Coping with Depression” course was used, which is a psycho-educational intervention in which several mood management techniques are taught to the participant (Lewinsohn et al., 1984), and which has been found to be effective as a treatment for major depression in several studies (Cuijpers, 1998). In the two studies which did not use the “Coping with Depression” course, problem solving therapy (Lynch et al., 1997), and interpersonal counselling (Mossey et al., 1996) were used.
### Table 3.2.1A Characteristics of the selected studies

<table>
<thead>
<tr>
<th>Study, year</th>
<th>Age group, in years</th>
<th>Recruitment through</th>
<th>Minor depression at baseline defined as</th>
<th>Major depression at baseline excluded with</th>
<th>Major depression at follow-up assessed with</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarke, 1995</td>
<td>16 – 16</td>
<td>Screening at school</td>
<td>CES-D ≥ 24</td>
<td>K-SADS</td>
<td>CES-D</td>
</tr>
<tr>
<td>Mossey, 1996</td>
<td>&gt; 60</td>
<td>Health services</td>
<td>GDS ≥ 11</td>
<td>SCID</td>
<td>GDS</td>
</tr>
<tr>
<td>Lynch, 1997</td>
<td>&gt; 18</td>
<td>Primary care waiting rooms</td>
<td>MOS DSI &gt; cut off</td>
<td>DIS</td>
<td>HRSD</td>
</tr>
<tr>
<td>Clarke, 2001</td>
<td>13 – 18</td>
<td>Health services</td>
<td>CES-D ≥ 24, and parent treated for MDD in past year</td>
<td>K-SADS-E</td>
<td>CES-D</td>
</tr>
<tr>
<td>Allart, 2003</td>
<td>18 – 65</td>
<td>Media announcements</td>
<td>BDI ≥ 10</td>
<td>CIDI</td>
<td>BDI</td>
</tr>
<tr>
<td>Willemsen, 2004</td>
<td>18 – 65</td>
<td>Primary care waiting rooms</td>
<td>1 core plus 1, 2, or 3 depressive symptoms</td>
<td>CIDI</td>
<td>CES-D</td>
</tr>
<tr>
<td>Haringsma, 2005</td>
<td>&gt; 55</td>
<td>Media announcements</td>
<td>No inclusion criteria</td>
<td>MINI</td>
<td>CES-D</td>
</tr>
</tbody>
</table>

BDI: Becks Depression Inventory; CES-D: Centre of Epidemiologic Studies Depression scale; DSM: Diagnostic Statistical Manual; MDD: major depressive disorder; MOS DSI: MOS Depression Screening Inventory; GDS: Geriatric Depression Scale; CIDI: Composite International Diagnostic Interview; SADS: Schedule for Affective Disorder and Schizophrenia (E: Epidemiological version); SCID: Structured Clinical Interview for DSM-IV-TR; MINI: Mini mental status examination; DIS: Diagnostic Interview Schedule; HRSD: Hamilton Depression Rating Scale; CBCL-D: Child Behavioural Check List – Depression scale.

Most studies (6) used care-as-usual as the control condition. Two studies were directed at adolescents, three at adults, and two at elderly. The diagnostic instruments used to exclude subjects meeting diagnostic criteria for mood disorders included the CIDI (2 studies), the K-SADS (2 studies), the MINI, the SCID, and the DIS. In four studies, longer term follow-up measures (12 months post-test) were administered. The studies were conducted in the United States (4 studies) or the Netherlands (3 studies).

The methodological quality of all studies was high. All used designs with randomisation, well-validated measurement instruments, well-described and theoretically sound interventions, and adequate statistical analyses.
### Table 3.2.1B Characteristics of the selected studies (continued)

<table>
<thead>
<tr>
<th>Study and publication year</th>
<th>Sample size per condition</th>
<th>Conditions, Exp vs Cntrl</th>
<th>Delivery format of intervention</th>
<th>Follow-up times, in months</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarke, 1995</td>
<td>55</td>
<td>CBT vs CAU</td>
<td>15 groups sessions of 45 min</td>
<td>12</td>
<td>US</td>
</tr>
<tr>
<td>Mossey, 1996</td>
<td>35</td>
<td>IPC vs CAU</td>
<td>10 individual sessions of 1 hr</td>
<td>4</td>
<td>US</td>
</tr>
<tr>
<td>Lynch, 1997</td>
<td>11</td>
<td>PST vs CAU</td>
<td>6 groups sessions of 20 min</td>
<td>4</td>
<td>US</td>
</tr>
<tr>
<td>Clarke, 2001</td>
<td>43</td>
<td>CBT vs CAU</td>
<td>15 groups sessions of 1 hr</td>
<td>12, 24</td>
<td>US</td>
</tr>
<tr>
<td>Allart, 2003</td>
<td>61</td>
<td>CBT vs CAU</td>
<td>12 groups sessions of 2 hrs</td>
<td>6, 12</td>
<td>NL</td>
</tr>
<tr>
<td>Willemse, 2004</td>
<td>107</td>
<td>CBT vs CAU</td>
<td>Guided self-help with 1 individual contact and 6 brief telephone calls</td>
<td>4, 12</td>
<td>NL</td>
</tr>
<tr>
<td>Haringsma, 2005</td>
<td>31</td>
<td>CBT vs WL</td>
<td>10 groups sessions of 2 hrs</td>
<td>4</td>
<td>NL</td>
</tr>
</tbody>
</table>


**Effects on depressive symptom-level at post-test**

We could compare the effects of the psychological treatments to control conditions at the 4-months post-test in six studies (see, Table 3.2.1). The mean standardised effect size, \(d\), was 0.42 (95% CI= 0.23~0.60), indicating that psychological treatment resulted in a reduction in depressive symptom-level. This outcome was exactly the same in the random effects model and the fixed effects model. We have plotted the effect sizes and the corresponding 95% confidence intervals of the individual contrast groups in Figure 3.2.1. The Q-statistic was 1.47 (df=5; P=0.92) and the corresponding 0-hypothesis of no homogeneity could not be rejected. Accordingly, the \(I^2\) statistic indicated a heterogeneity of 0%.
Figure 3.2.1 Meta-analytic effects on symptom-level after 4-months

Diminishing effects over time
Over time the effects became smaller. The effects of the psychological treatments at the 6-months follow-up could be compared to care-as-usual in two studies. The mean effect size was 0.17 (95% CI= -0.11~0.45), which was not significant, while heterogeneity was moderate ($I^2=48.2\%$). At 12-month follow-up, 4 comparisons were available, resulting in a mean effect size of 0.16 (95% CI= -0.02~0.35). This was not significant at the $P < 0.05$ level, but it bordered on statistical significance with $P = 0.08$. Again, heterogeneity was very low with $I^2 = 0\%$. The results are summarised in Table 3.2.2.

Table 3.2.2 Meta-analytic effects of interventions over time

<table>
<thead>
<tr>
<th>Follow-up, in months</th>
<th>Number of comparisons</th>
<th>Number of subjects</th>
<th>Effect size, $d$</th>
<th>95% CI</th>
<th>P-value of $d$</th>
<th>Q-test</th>
<th>$I^2$-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>6</td>
<td>469</td>
<td>0.42</td>
<td>0.23~0.60</td>
<td>$P&lt;0.01$</td>
<td>1.47 ns</td>
<td>0.0%</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>227</td>
<td>0.17</td>
<td>-0.11~0.45</td>
<td>$P&gt;0.05$</td>
<td>1.93 ns</td>
<td>48.2%</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>533</td>
<td>0.16</td>
<td>-0.02~0.35</td>
<td>$P&lt;0.10$</td>
<td>1.04 ns</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

Ns: not significant

Meta Analysis

-2.00 -1.00 0.00 1.00 2.00
Trends towards Treatment

-2.00 -1.00 0.00 1.00 2.00
Trends towards Control
**Effects on the incidence of major depressive disorder**

We were able to examine the effects of psychological treatments on the incidence of new cases of depressive disorders in four studies. The incidence rate ratio (IRR) of developing a major depressive disorder in subjects who received the intervention was 0.70 (95% CI= 0.47~1.03), compared to subjects in the control condition. This was not significant at the P<0.05 level, but there was a trend (P = 0.07), suggesting that the risk of developing a major depression was lower in the intervention condition. Again, the resulting IRR was exactly the same when we used the random effects model and the fixed effects model. We have plotted the IRRs and 95% confidence intervals of the individual contrast groups in Figure 3.2.2. Again, heterogeneity was very low. The Q-statistic was 2.69 (df= 3; p= 0.44), and the $I^2$ statistic indicated a percentage of 0.

**Figure 3.2.2** Meta-analytic effects of the treatments on the incidence of depression

<table>
<thead>
<tr>
<th>Study name</th>
<th>Risk ratio Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarke, 1995</td>
<td>0.401 0.175 0.918</td>
<td>2.163 0.031</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clarke, 2001</td>
<td>0.796 0.337 1.881</td>
<td>-0.519 0.604</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allart, 2003</td>
<td>1.033 0.456 2.341</td>
<td>0.078 0.938</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Willems, 2004</td>
<td>0.705 0.352 1.408</td>
<td>-0.991 0.322</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.695 0.468 1.033</td>
<td>-1.801 0.072</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**3.2.4 Discussion**

**Main findings**

Our meta-analysis gives indications that psychological treatment for subthreshold depression has a significant effect on depressive symptom-level in the short term. The effect size of $d=0.42$ is in the medium range, but was obtained by offering only brief or light forms of psychological treatment.

With regard to the longer term outcomes, we found that the effects of psychological intervention drops to a small effect of $d= 0.16$ after one-year, which
was only marginally significant at P=0.08, and therefore only just suggestive of a trend toward superiority of psychological treatment over care-as-usual.

Perhaps more importantly, we found some indications that psychological interventions in subthreshold depression can be effective in reducing the incidence of full-blown depressive disorder by 30%, relative to those who were no recipients of the intervention but got only care-as-usual. In a two-sided test this effect bordered on significance with P= 0.07. However, it could well be argued that a one-sided test is more appropriate for testing the uni-directional hypothesis that an adjunctive psychological intervention combined with care-as-usual will generate superior effects as compared with care-as-usual alone (cf. Willemse et al., 2004). From this perspective, the results of the meta-analysis are not only clinically interesting, relevant for public health, but also statistically significant. In that light, and on the understanding that the number of included studies is only small, we consider this to be a promising outcome.

Another finding is that virtually no heterogeneity between the study outcomes could be found, and thus that the psychological treatments have comparable effects. Although this finding has to be considered with some caution, because the number of studies is small, it can be seen as an indication that the results, as obtained, are robust, and thus to provide both clinicians and their patients with some prognostic certainty.

**Limitations**

This study has several limitations. First, the number of studies which satisfied all the inclusion criteria, was small. Second, several basic elements of included studies differ from each other, including target groups, measures, and interventions. On the other hand, the low heterogeneity can be seen as an indication that the studies were comparable.

The studies also used different definitions of subthreshold depression. Most of them used a self-rating scale to assess the presence of clinically relevant depressive symptoms. It is not clear what the clinical status is of subjects that score highly on a self-rating scale but do not meet the criteria for a depressive disorder. However, because the subjects were willing to participate in the intervention, it may be assumed that the symptoms were severe enough to motivate participation.

For future research, it is important that clear criteria be developed for subthreshold depression which is clinically relevant. Minor depression, as defined in the Appendix of the DSM-IV, is an important step forward in this respect, although not all subjects with clinically relevant depressive symptoms will meet the criteria for minor depression.

The studies also differed in the type of treatment that was offered. However, most of them used a brief version of the “Coping with Depression” course (Lewinsohn et al., 1984). This intervention is attractive for this population, because it is an evidence-based psycho-educational intervention (Cuijpers, 1998), which can be easily adapted to the needs of specific populations.
It has been suggested that antidepressant medication should be used to treat subthreshold depression (Judd et al., 2004). However, the evidence supporting antidepressants as a treatment in this population is limited, and it is questionable whether patients will consider medication acceptable (Van Schaik et al., 2004). Brief psycho-educational interventions based on cognitive behavioural therapy, however, seem to be a more appropriate option for these problems.

**Conclusion**

Despite the limitations of this meta-analysis, we did find indications that psychological therapies are effective in the treatment of subthreshold depression, and may, in addition, offer some preventive protection against the risk of developing full-blown depressive disorder. However, more research in this area is direly needed, as definitions of subthreshold depression vary considerably, the target populations have not been defined consistently, and the number of randomised trials examining the (preventive) effectiveness of the interventions is still very limited.
Part IV
Is it affordable?

Chapter 4.1
Cost-effectiveness of preventing depression in primary care patients, randomised trial

This chapter is based on:

Chapter 4.1

Cost-effectiveness of preventing depression in primary care patients, randomised trial

4.1.1 Introduction

Depression is highly prevalent, compromises the quality of life, and has a substantial economic impact (Bijl et al, 1998; Wells et al, 1992; Löthgren, 2004). Several cost-effective interventions directed at depressed primary care patients are available (McCrone et al. 2004; Scott et al., 2003; Schulberg et al., 2002). However, according to one estimate the burden of depression can be averted for only 34% even under a hypothetical regimen of optimal care (Andrews et al., 2004), and less when, more realistically, actual practice is assumed (Ibid; Chisholm et al, 2004). This leaves a formidable gap between what the best treatments can offer and the needs of many. This gap calls for interventions other than curative interventions. Preventive interventions may have to play a role here (Smit et al, 2004; Smit et al, 2006a). A randomised clinical trial by Willemse et al. (2004) showed superior effectiveness of minimal contact psychotherapy over care-as-usual in preventing the onset of full-blown depressive disorder in primary care patients with subthreshold depression. Using the same trial data, we now investigate the cost-effectiveness of adjunctive minimal contact psychotherapy relative to routine primary care alone.

4.1.2 Method

The methods of this trial have been described in some detail in Chapter 3.1 (see also, Willemse et al., 2004). Here, we describe its main features and focus attention on the economic aspects.

Sample

Trial participants (aged 18 – 65 years) were recruited from nineteen general practices in the Netherlands. They were eligible when presenting subthreshold depression defined as having at least one core symptom plus one, two, or three current depressive symptoms according to the INSTEL screen (Tiemens et al, 1995). Exclusion criteria were the presence of full-blown DSM-IV (American Psychiatric Association, 1994) depressive disorder, dysthymia, bipolar disorder, social phobia, agoraphobia or panic disorder in the last 12 months as measured with the Composite International Diagnostic Interview (see Measures).

Participants were recruited in several steps (see Figure 3.1.1, in Chapter 3.1). Research assistants screened 3,825 patients who were waiting to see their general practitioner. In a next step 1,018 screen-positive patients were asked to give their
informed consent to participate in the trial. Of these, 363 were willing and received a computer assisted diagnostic interview with the CIDI. This was done to exclude patients with full-blown depression and other DSM-IV axis-I disorders as stated above.

The randomisation was carried out centrally, using blocked randomisation stratified by general practice with the patient as unit of randomisation, with blocks of 4 patients. Eligible participants who had given their informed consent were randomised, with equal probability, to the minimal contact psychotherapy condition (n=107) or to the care-as-usual condition (n=109). Of these, 83 and 94 were retained in the trial after twelve months. Fewer participants completed the economic questionnaire. At baseline, these were 99 and 102 for the minimal contact psychotherapy and care-as-usual conditions, respectively, of whom 75 and 87 completed the economic questionnaires at follow-up.

The study was conducted as a pragmatic trial. Only the interviewers were blinded with regard to the participants’ randomisation status. The trial protocol was approved by an independent medical ethics committee.

**Interventions**

The experimental intervention was cognitive-behavioural minimal contact psychotherapy for depression, based on the Dutch version (Cuijpers, 2000) of the “Coping with Depression” course (Lewinsohn et al, 1984). The main component was a self-help manual with instructions on mood management. The self-help therapy was guided by six short telephone calls with a prevention worker. The control intervention was care as routinely provided by GPs. All participants, in both conditions, could make use of all other types of health services during the intervention period.

**Clinical measures**

The participants’ DSM-IV depression status was assessed with the Composite International Diagnostic Interview, the Auto-version 2.1 (WHO, 1997), in its Dutch version (Ter Smitten et al, 1998). The CIDI is a standardised diagnostic interview for the assessment of mental disorders, developed by the World Health Organization. It was designed for use by trained lay interviewers, has high interrater and test-retest reliability and good validity for affective and anxiety disorders (Andrews & Peters, 1998; Wittchen, 1994). The interviews were carried out over the telephone. This should not have affected the results in any meaningful way (Evans et al, 2004; Rohde et al, 1997).

Depressive symptom level was measured using the Centre for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977), the Dutch version (Bouma et al, 1995), which is a widely used self-report scale measuring the frequency of 20 depressive symptoms during the past week. The CES-D generates a total score that can range from 0 to 60, with a higher score indicating more
depressive symptoms. The Dutch translation has good reliability and validity (Bouma et al, 1995).

**Measuring resource use**

For this study we adopted a societal perspective to include the costs of all types of health services and the costs that stem from production losses. The time frame of this study is restricted to one year. Therefore, we did not correct for inflation and did not discount costs. All costs are expressed in euro (€) for the reference year 2003 on a per capita basis for the period of one year.

Information on the participants’ use of health services was obtained with the Trimbos and Institute of Medical Technology Assessment Cost questionnaire for Psychiatry (TIC-P, Hakkaart - van Roijen et al., 2002). With this questionnaire patients register the number of general practitioner visits, sessions with psychiatrists, hospital days, etc. In addition, the number of work loss days (absenteeism from work) and the number of work cutback days (lesser efficiency at work while feeling ill) were also measured with help of the TIC-P.

**Cost of service use**

The intervention costs of minimal contact psychotherapy are € 124 for screening, and again € 124 for the intake, plus € 31 per additional contact over the telephone with a maximum of six calls. Participants had to pay € 25.50 for the self-help manual. The participants’ time for working through the self-help manual was valued at € 8.30 per hour, assuming that they would carry out their assignments after office hours. It should be noted that the intervention costs occurred only in the experimental group during the actual uptake of the intervention over four months.

Direct medical costs are the costs of treatments offered by a broad range of both formal and informal health service providers (see Table 2.1.1 in Chapter 2.1). Medical services were costed by multiplying the number of health service units (consultations, hospital days, etc.) by their standard cost price (Oostenbrink et al, 2004). To these we added the costs of anti-depressants, that is, the cost price per standard daily dose as reported in the Pharmaceutical Compass (at www.fk.cvz.nl), plus 6% tax, multiplied by the number of prescription days, plus the pharmacist’s dispensing costs of € 6.45 per prescription.

Direct non-medical costs arise when patients travel to health service providers, and pay for parking. These ‘out-of-pocket’ costs were valued at € 0.16 per km and € 2.50 per 1 hour parking time. To this we added the costs of the patients’ time spent in travel, waiting and in treatment at € 8.30 per hour.

Cost of production losses

Indirect non-medical costs arise when production losses occur due to illness. Three situations can be encountered here.

First, people can be absent from paid work. To evaluate a lost day in a paid job we used age and gender specific friction-costs obtained from Oostenbrink et al.
Friction costs represent the monetary counter-value of production losses that occur during absence from work with a limit to five months (Koopmanschap et al., 1995), because after that time, and given the current labour market in the Netherlands, they will be replaced by other labour force, and therefore cease to generate production losses. The gender and age specific friction costs are presented in Table 1.2.2, Chapter 1.2. Here, it should be noted that we could not employ the friction cost method in all its detail. The method requires that it is known if people stay ill beyond the five month limit, but we had no such data, and, by implication, we had to use a method more similar to the human capital method.

Second, production losses also occur when people are ill, try to work, and are then less efficient. We estimated the number of work cutback days as the number of days actually worked when ill, multiplied by an self-reported inefficiency score, which ranges between 0 and 1 (0 = as efficient as when in good health, 1 = totally inefficient). Again, we used friction costs to value these production losses.

Third, people may also be too ill to perform domestic tasks. These costs were valued at the free market price of domestic help at € 8,30 per hour.

**Analysis of clinical outcomes**

The analysis of clinical outcomes was carried out in accordance with the intention-to-treat principle. Use was made of the regression imputation procedure as implemented in the Stata 7.0 (StataCorp, 2001) for handling loss-to-follow-up. In the regression imputation model, baseline CES-D depression scores, age and gender were used as predictors, because they were significant predictors of depression status at follow-up. Since patients were recruited from 19 GPs, some degree of clustering in the data had occurred. Clustering violates the assumption of independence of observations, and may thus affect standard errors and P-values. This was handled with the help of so called ‘robust standard errors’, which were obtained using the first-order Taylor-series linearisation method as implemented in Stata. The incidence rate ratio (of the incidence rate in the minimal contact psychotherapy group over the incidence rate in the care-as-usual group) was obtained by regressing the imputed depression status at follow-up on the treatment dummy in a Poisson model, while taking into account the clustering effect. The statistical test was conducted at $\alpha < 0.05$, 1-sided, because inferior effectiveness of adjunctive minimal contact psychotherapy over care-as-usual alone was not expected.

**Analysis of costs**

The analysis of costs was also conducted in agreement with the intention-to-treat principle. Missing cost data at follow-up were imputed as before, but now with costs at baseline, and age and gender as predictors. Reported are the mean annual per capita costs of the intervention, plus the direct medical, direct non-medical and the indirect costs and some of their components (see Table 4.1.1). The 95%
confidence intervals of the economic outcomes were based on 2,500 bootstrap replications because cost data are non-normally distributed.

**Analysis of cost-effectiveness**

In the cost-effectiveness analysis health effects (depression free person-years) and costs in both treatment arms were computed by means of non-parametric bootstrapping (2,500 times) with respect to both incremental costs and incremental health effects. The comparison of the simulated differences in costs and health effects is presented in a cost-effectiveness plane (see Figure 4.1.1), with differences in costs on the vertical axis, and difference in health effects on the horizontal axis. In case an intervention appears at the top left-hand quadrant in the plane, higher costs are paid for lower effectiveness; the intervention is then unacceptable from a cost-effectiveness perspective, and conventional care remains the treatment of choice. In case the intervention appears in the lower right-hand quadrant, lower costs are then associated with health gains; the intervention dominates and is acceptable. In the other two quadrants, higher (or lower) cost levels have to be weighted against greater (or lesser) effectiveness.

A second way of illustrating the cost-effectiveness results, taking into account the uncertainty, is the cost-effectiveness acceptability curve (Van Hout *et al.*, 1994; cf. Barett & Byford, 2003). Such an acceptability curve represents the probability that the experimental minimal contact psychotherapy condition is more cost-effective than care-as-usual, given varying thresholds for the willingness to pay for a depression-free survival time of one year (see Figure 4.1.2).

**Sensitivity analyses**

It appeared that the total costs were dominated by the costs of production losses. Therefore, the analyses were repeated for the total costs, minus those of production losses. This gives an idea of the cost-effectiveness when the more narrow perspective on direct costs is used instead of the broader societal perspective.

### 4.1.3 Results

**Sample**

The participants were predominantly female (66%), living with a partner (78%), and employed (83%). The mean age was 41 years and the participants had received 14 years of education on average. At baseline their mean CES-D score on depressive symptomatology was 12.8, well below the cutoff of 16 at which people are considered to have clinically relevant depression. No significant difference was found between the study groups for these variables, indicating that randomisation had resulted in comparable groups (cf. Willemse *et al.*, 2004).
**Health effects**

At 12 months, the incidence rate of depressive disorder was 11.9% in the minimal contact psychotherapy condition versus 18.3% in the care-as-usual condition. The incidence rate ratio (IRR) was therefore 11.9 / 18.3 = 0.65, and the 0-hypothesis of inferior clinical effects in the minimal contact psychotherapy condition had to be rejected (IRR = 0.65; SE = 0.15; T = -1.82; P = 0.04, 1-sided), favouring the conclusion that adjunctive minimal contact psychotherapy is more successful than care-as-usual alone in reducing the incidence of depressive disorder (Willemse et al., 2004).

**Costs**

Over one year, the adjunctive psychotherapy group incurred the costs of the intervention (on average € 423 per recipient), but the intervention costs were partly compensated for by savings elsewhere in the medical sector: the mean difference of the direct medical costs were € 60 (SE = 555) in favour of care-as-usual, but this was not statistically significant (P = 0.914).

**Table 4.1.1** Annual per-capita costs (in euro) by condition.

<table>
<thead>
<tr>
<th></th>
<th>Experimental group</th>
<th>Control group</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SE)</td>
<td>Mean (SE)</td>
<td>Diff. (SE) P-value</td>
</tr>
<tr>
<td>Direct medical</td>
<td>1687 (305)</td>
<td>1627 (419)</td>
<td>60 (555) 0.914</td>
</tr>
<tr>
<td>. Intervention</td>
<td>423 (13)</td>
<td>0 (0)</td>
<td>423 (12) 0.000</td>
</tr>
<tr>
<td>. general practitioner</td>
<td>165 (30)</td>
<td>152 (14)</td>
<td>13 (32) 0.681</td>
</tr>
<tr>
<td>. Antidepressants</td>
<td>27 (9)</td>
<td>29 (10)</td>
<td>-2 (7) 0.789</td>
</tr>
<tr>
<td>. Other medical</td>
<td>1067 (291)</td>
<td>1442 (431)</td>
<td>-376 (521) 0.471</td>
</tr>
<tr>
<td>Direct non-medical</td>
<td>441 (59)</td>
<td>507 (77)</td>
<td>-66 (88) 0.453</td>
</tr>
<tr>
<td>Indirect non-medical</td>
<td>4638 (1634)</td>
<td>6481 (1393)</td>
<td>-1843 (1639) 0.261</td>
</tr>
<tr>
<td>. Work loss</td>
<td>2374 (807)</td>
<td>3279 (697)</td>
<td>-905 (833) 0.277</td>
</tr>
<tr>
<td>. Work cutback</td>
<td>2232 (823)</td>
<td>3175 (696)</td>
<td>-942 (796) 0.237</td>
</tr>
<tr>
<td>. Domestic</td>
<td>31 (13)</td>
<td>28 (9)</td>
<td>4 (17) 0.828</td>
</tr>
<tr>
<td>Total</td>
<td>6766 (1712)</td>
<td>8614 (1490)</td>
<td>-1849 (1715) 0.281</td>
</tr>
</tbody>
</table>

SE= standard error. SEs and P-values based on 2,500 bootstrap replications. Negative costs are savings.

Moreover, the out-of-pocket costs of the patients in the minimal contact psychotherapy condition were somewhat lower with € 441 against € 507 in the care-as-usual condition, representing a difference of € 66 in favour of the intervention.
Avoiding production losses resulted in further cost-savings. The mean cost of the production losses was € 4,638 in the intervention condition against a higher € 6,481 in the care-as-usual condition, resulting in an average cost saving of € 1,843.

The costs due to work loss days of € 2374 and € 3279 in the experimental and control conditions respectively, represent the average counter-value of 8.4 and 11.5 lost workdays.

Overall, the mean annual per capita total costs of minimal contact psychotherapy were € 6,766 which compares favourably with the € 8,614 in the care-as-usual group. The overall per-patient savings average € 1,849 when minimal contact psychotherapy is added to care-as-usual, but this is statistically not significant (95% BI -5,169 to 1,472; P = 0.281). Nevertheless, it is worth noting that there is a large probability that the costs of minimal contact psychotherapy are balanced by savings elsewhere. We return to this point shortly.

**Cost-effectiveness**

The incremental cost-effectiveness ratio, ICER, was calculated as

\[
 ICER = \frac{(C_1 - C_0)}{(E_1 - E_0)},
\]

where \(C\) = the average annual per capita costs, and \(E\) = the percentage of people who did not develop a depression in the experimental and control conditions (subscripted 1 and 0, respectively). In other words, the incremental cost-effectiveness ratio is the difference of mean costs between the conditions divided by the difference in effect. Substitution yields

\[
 (6,766 - 8,614) / (88.1 - 81.7) = -288.91.
\]

Hence, for each case of depression that can be avoided by offering the experimental treatment instead of care-as-usual, a saving is made of € 288.75.

The incremental cost-effectiveness ratio is surrounded by a certain amount of uncertainty. Figure 4.1.1 presents the cost-effectiveness plane for minimal contact psychotherapy versus care-as-usual. On the y-axis are the incremental costs; on the x-axis the incremental effects. Each dot (N=2,500) represents a bootstrap replication of the incremental cost-effectiveness ratio. Fifty nine percent of the dots are in the lower right-hand quadrant, indicating a 59% probability that minimal contact psychotherapy is the dominant intervention, because minimal contact psychotherapy generates better health effects against lower costs when compared to care-as-usual. On the other hand there is a 5% probability that minimal contact psychotherapy is inferior. There is a 10% probability that minimal contact psychotherapy is both less costly and less effective. Twenty-one percent of the dots fall in the upper right-hand quadrant, indicating that a health gain is produced, but at additional costs.
Figure 4.1.1 Cost-effectiveness plane, based on 2,500 bootstrap replications

Figure 4.1.2 presents the cost-effectiveness acceptability curve for minimal contact psychotherapy versus care-as-usual. The solid line intersects the y-axis at 0.70; when the willingness to pay for an averted depressive episode is absent (equal to € 0.00), then there is a 70% probability that minimal contact psychotherapy is more cost-effective than care-as-usual.

Generally, people are willing to pay for avoiding a depressive episode, and minimal contact psychotherapy will be regarded as good value for money given an, usually unknown, ceiling for this willingness to pay. Different ceilings are presented on the x-axis. As can be seen, when the willingness to pay is raised to € 10,000 per avoided depression, then minimal contact psychotherapy has a probability of 74% of being more cost-effective than its alternative; at € 20,000 the probability of an acceptable cost-effectiveness has risen to 80% and at € 30,000 it has reached a 83% probability of being more acceptable than care-as-usual alone.
Probability that the incremental cost-effectiveness ratio is acceptable (y-axis) given varying thresholds for willingness to pay (x-axis) for total costs (solid line) and direct medical costs (dotted line)

Sensitivity analysis
When the indirect costs related to the production losses are excluded, then the distribution of the bootstrapped cost-effectiveness ratios over the cost-effectiveness plane is as follows: (i) 41% of the ratios fall in the upper right-hand quadrant indicating that better effects are obtained against higher costs, (ii) 11% fall in the upper left-hand quadrant indicating that minimal contact psychotherapy is inferior, (iii) 7% fall in the lower left-hand quadrant indicating that minimal contact psychotherapy has worse clinical outcomes against lower costs, and (iv) 39% of the bootstrapped cost-effectiveness ratios fall in the lower right-hand quadrant, implying that minimal contact psychotherapy is the dominant intervention, as it generates better outcomes against fewer costs than care-as-usual.

Thus seen from the more narrow perspective which is solely focussed on medical costs, the minimal contact psychotherapy intervention has a probability of 46% of being more acceptable than its alternative when the willingness to pay equals zero (dotted line in Figure 4.2.1). When the willingness to pay is increased to € 10,000, € 20,000 and € 30,000, then the probability of minimal contact psychotherapy being more acceptable than care-as-usual increases to 61%, 70% and 75%, respectively.
4.1.4 Discussion

Main findings
The incidence of DSM-IV axis-I depression among participants in the care-as-usual condition was 18%. The incidence was significantly lower in the experimental condition, minimal contact psychotherapy, at 12% (Willemse et al., 2004). This represents a reduction in the incidence by one-third, and indicates superior effectiveness of minimal contact psychotherapy as compared to care-as-usual.

Not only is minimal contact psychotherapy more effective, this economic evaluation indicates that choosing adjunctive minimal contact psychotherapy over care-as-usual alone is likely to be the best treatment option, because there is a 70% probability that adjunctive minimal contact psychotherapy is preferable to care-as-usual alone when the costs of production losses are included in the analysis – even in the unlikely scenario that there is no willingness to pay for a depression free survival time of one year. Excluding the costs of production losses results in a comparable situation. From the cost-effectiveness perspective minimal contact psychotherapy then has a 46% probability of being more acceptable than care-as-usual. When the willingness to pay for avoiding a depressive episode is € 30,000, then minimal contact psychotherapy has a 75% probability of being the preferred option.

Limitations
This study has several limitations, some of which have already been pointed out in Chapter 3.1. First, although problems with attrition are quite common in randomised trials of psychological interventions in general practice (Fairhurst & Dowrick, 1996; Muñoz et al., 1995), the representativeness of the sample can be questioned. In recognition of this limitation, all analyses were conducted in accordance with the intention-to-treat principle, and imputation was used as a means to overcome the missing data problem due to loss-to-follow-up.

Second, it was not possible to blind participants for the condition to which they were assigned. This is true for most randomised trials of psychological interventions, but it may nevertheless have distorted the outcomes of our trial. Third, the study was conducted in the Netherlands, and the results cannot be safely generalised to countries that have (very) different primary care systems.

Fourth, the costs and effects were considered in the time-span of one year. We do not know how the cost-effectiveness of minimal contact psychotherapy is affected when a longer time horizon is used.

Fifth, the intervention is best seen as a package of component interventions and its relative cost-effectiveness is evaluated against routine primary care alone – which is yet another (ill-defined) package in itself. This is to say that the observed outcomes cannot be attributed to one of the components alone. For example, the superior effects of the treatment can either be ascribed to the self-help manual, or to the supporting telephone contacts, or to both synergistically. Therefore it is
unknown if the same clinical effects could have been produced without the telephone contacts, or could have been produced by the phone calls alone. This is an important issue in the context of a cost-effectiveness analysis. In our analysis we had to settle for an evaluation of (economic) input and (clinical) output relationships, and treat the throughput (the treatments) as a black box. This approach helped to answer the question whether adding minimally supported self-help to routine primary health care is clinically worthwhile and economically affordable under conditions that realistically mimic the primary health care setting as seen in the Netherlands. In the end of the day, this is the question that needs to be answered. Nevertheless, in future research one would do well to include an additional arm in the trial – one in which bibliotherapy is offered without the supporting telephone contacts. Such a design may help to shed light on a possible Hawthorn effect induced by the supporting telephone contacts.

Because of these limitations, the results of this study should be considered with some caution.

The wider context
Depression is one of the leading causes of disability (Ustun et al., 2004). However, according to one estimate, even under a hypothetical regimen of optimal (evidenced-based) care, the burden of depression can only be averted for 34% (Andrews et al., 2004). This leaves a formidable gap between what the best treatments can offer and the needs of many. This gap calls for interventions other than curative interventions alone.

Preventive interventions may have to play an adjunctive role here (Smit et al., 2004; Smit et al., 2006a). Reducing the burden of depression by means of a preventive intervention is possible, as was shown in a randomised clinical trial of Willemse et al. (2004) and in a meta-analysis of randomised prevention trials by Cuijpers et al. (2005).

Perhaps of equal importance is the opportunity to offer a low-cost, self-help treatment, which is effective while consuming small amounts of healthcare resources. In fact, this study showed that there is a 70% probability that minimal contact psychotherapy is more cost-effective than care-as-usual.

Our findings are in agreement with the reviews of both Churchill et al. (2001) and Schulberg et al. (2002), and the more recent cost-effectiveness analysis of McCrone et al. (2004). They found that psychological interventions based on cognitive behavioural therapy are cost-effective in depressed primary care patients. Cognitive behavioural therapy appears also to be cost-effective for relapse prevention in chronic depression (Scott et al., 2003). In contrast, other types of psychological interventions – specifically (psychodynamic) counselling and sometimes interpersonal therapy – have not shown similar effects and cost-effectiveness (Lave et al., 1998; Bower et al, 2000; Simpson et al, 2003). Now, our study adds the information that a self-help intervention based on cognitive behavioural therapy with minimal guidance is cost-effective in avoiding the onset of
full-blown depressive disorder in primary care patients with subthreshold depression.

**Directions for the future**
The “Coping with Depression” course (Lewinsohn *et al*, 1984) and its Dutch pendant (Cuijpers, 2000) can be used as a cost-effective adjunct to conventional primary care in order to reduce the incidence of depressive disorder. This choice is likely to result in health gains and economic benefits. Therefore, its dissemination seems appropriate. Nonetheless, three issues need more in-depth exploration. First, we need to know more about the cost-effectiveness of minimal contact psychotherapy in the long run. Second, minimal contact psychotherapy should perhaps be adapted for use over the internet. The latter approach may help to reduce the costs of minimal contact psychotherapy, while at the same time it may promote its use by a larger segment of the population. Third, we would like to better understand what component part of the intervention is especially responsible for the outcomes. Perhaps, it would be possible to generate the same effects with a stripped, and therefore less costly, version of the intervention.
Part IV

Is it affordable?

Chapter 4.2

Costs of childhood adversity: towards a risk factor epidemiology of costs in mental health

This chapter is based on:

Chapter 4.2

Costs of childhood adversity: towards a risk factor epidemiology of costs in mental health

4.2.1 Introduction

Burden-of-disease studies and cost-of-illness studies have done much to underscore the importance of psychiatry. The former helped to promote a better understanding of the formidable disease burden of mental disorders (Murray & Lopez, 1996; WHO, 2000; Ustun et al, 2004). The latter highlighted the notion that mental disorders are generating costs that are comparable to, or exceed, those of somatic illnesses (Berto et al, 2000; Greenberg & Birnbaum, 2005; Jöhnson & Olesen, 2004; Knerer et al, 2005; Marciniak et al, 2004). Together both types of studies helped to place psychiatry at the very frontier of public health.

In this study we take the health economic perspective, but want to think outside the box. If it is possible to calculate the costs of mental disorders, then it should also be possible to quantify the costs of their underlying risk factors (see Methods, later). Such an approach would give rise to what could be called a “risk-factor epidemiology of costs in mental health”.

We believe this “risk-factor epidemiology of costs in mental health” is novel and offers several advantages. First, attaching cost prices to risk factors makes for easy communication about “costly risk factors”, and will be of value to policy makers in public health. Second, identification of costly risk factors would direct attention to the very roots of disability. Third, it is likely that costly risk factors, such as long standing vulnerability factors, will not solely impact on a single disorder, but increase the risk of several disorders, both axis-I and axis-II disorders, comorbid conditions and subthreshold conditions. Thus a study of the costs of risk factors would bypass the very complex issue of diagnostic heterogeneity. Fourth, when costly risk factors can be identified, so can protective factors – i.e., factors that cushion the adverse economic impact of the risk factors. Thus, the proposed approach can be expanded to include both risk factors and protective factors, and become a (multivariate) risk factor epidemiology of costs in mental health. This would lend flexibility to this approach and make it suitable as a tool for answering research questions in the field of health economics. As a case in point, and perhaps most importantly, with its focus on costly risk factors, the proposed approach would help to shed light on areas where preventive psychiatry can be put to great advantage (see also Smit et al, 2006a, in this journal).

In this study we explore these ideas, using a large population-based cohort study of risk factors, mental disorders and costs. These costs are related to health service uptake and production losses due to absenteeism from work. The selected risk factors are exposure to childhood adversities and parental history of depression.
and anxiety. All selected risk factors can be seen as vulnerability factors with long-lasting effects on a wide array of disorders and comorbid conditions (Kessler et al. 1997; Merikangas et al. 1998; Bijl et al., 2001; De Graaf et al., 2002a; De Graaf et al., 2002b; Spataro et al., 2004). These risk factors are, therefore, likely to have a substantial impact on disability, quality of life, health service uptake, production losses, and the ensuing societal costs.

The studied protective factor is an above-average sense of internal locus-of-control, measured by the Pearlin & Schooler (1978) mastery scale, and henceforth called “mastery”. Mastery is known to be a protective factor in the pathogenesis of depressive disorder (Katz et al., 1994; Ormel et al., 1997; Zarit et al., 1999), and may thus cushion the adverse effect of the studied risk factors. Moreover, mastery is amenable to change under therapeutic interventions (McCullough, 1980; 2005), and this offers an important opportunity to strengthen mastery in people exposed to the risk factors.

We will first calculate the costs of the risk factors, and then evaluate to what degree better mastery levels are associated with lower costs in people who are exposed to the risk factors.

4.2.2 Methods

This study was conducted alongside the Netherlands Mental Health Survey and Incidence Study (NEMESIS), which is a large, population-based cohort study (Bijl et al., 1998). In NEMESIS, data were collected on risk factors, DSM-III-R axis-I mental disorders, and on the costs that arise when people are confronted with mental disorders. This design allows the prospective evaluation of the impact of risk factors (measured at baseline, \( t_0 \)) on health-related costs (measured one year later at follow-up, \( t_1 \)). Below, we describe, in more detail, the sample, the measures and the analysis.

Participants and procedure

The sampling procedures of the NEMESIS study have been described elsewhere in this journal (Vollebergh et al., 2001). In brief, a random, stratified, multistage sample was obtained in three steps. First, municipalities were stratified by urbanization, and 90 municipalities were drawn randomly and proportionately from the strata. Second, within each municipality, households were randomly drawn from the postal register. Finally, within each household the person with the most recent birthday was selected on condition that he or she was between 18 and 65 years and sufficiently fluent in Dutch to be interviewed. Eligible persons, who were not immediately available, were contacted later in the year. The response rate was 69.7% resulting in a sample of 7,076 people at \( t_0 \). The baseline sample followed the same multivariate distribution over age, sex, civil status and urbanity as the general Dutch population. Only men in the age bracket of 18-24 years were slightly underrepresented (De Graaf et al, 2000).
At first follow-up \( (t_1) \), which occurred one year after base line, 5,618 persons (79.4%) continued their participation. We evaluated the effect of attrition from \( t_0 \) to \( t_1 \) and found that it was not related to having a disorder at \( t_0 \) but again, the younger males were more likely to be lost to follow-up (De Graaf et al, 2000).

**Health-related costs**

Over the time span of a year most people are likely to generate some health-related costs. These costs may occur in different domains and are incurred by different parties. To be comprehensive, all the relevant domains must be included in an economic study. Below we present the different types of costs and describe how they have been calculated.

People may not feel well and stay at home for a period of time. When they have a job, then staying at home will result in production losses.

**Table 4.2.1** Monetary counter-value of production losses in paid labour by age and gender (in US$ per day) \(^a\)

<table>
<thead>
<tr>
<th>Age</th>
<th>Men US$/day</th>
<th>Women US$/day</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-24</td>
<td>178</td>
<td>175</td>
</tr>
<tr>
<td>25-34</td>
<td>285</td>
<td>260</td>
</tr>
<tr>
<td>35-44</td>
<td>355</td>
<td>292</td>
</tr>
<tr>
<td>45-54</td>
<td>395</td>
<td>297</td>
</tr>
<tr>
<td>55-64</td>
<td>416</td>
<td>317</td>
</tr>
<tr>
<td>65+</td>
<td>416</td>
<td>317</td>
</tr>
</tbody>
</table>

\(^a\) Oostenbrink et al. (2004).

Table 4.2.1 presents the average gender and age specific costs of losing one day of work in the Netherlands in the year 2003. When a person remains ill for several days, then the appropriate gender and age specific cost figure has to be multiplied by the number of work loss days. The number of work loss days is recorded in the NEMESIS database, and the corresponding costs calculations can thus be performed by simple multiplication.

When people are ill, they may also generate production losses in the domestic sphere. Family or friends may have to help out, or professional help must be hired. Either way, costs are incurred. The domestic production losses were monetized at US$ 9.00 per hour, which is the average wage of domestic help on the free market in the Netherlands in the year 2003.

People who do not feel well may visit their family doctor), and the doctor may start some treatment or refer to a specialist when the complaints warrant such action. Health service uptake will result in “direct medical costs”. During health service uptake, patients will travel, pay for parking, spend time travelling and waiting and in treatment, and these costs are called “direct non-medical costs”. Table 4.2.2 presents the integral direct medical and non-medical costs associated with health service uptake for a broad range of health services. Health service
uptake is recorded in the database of NEMESIS by service type and the number of units (visits, contacts, consults, sessions, hospital days). Again, the corresponding calculations can be performed by simple multiplication of the number of units by the appropriate integral cost price.

Finally, family doctors and other medical specialists may prescribe medication. The NEMESIS database has detailed data on the use of prescription drugs. Now the cost calculations are slightly more complex. The costs of pharmacy use are the costs of that particular brand of medication per standard daily dose (obtained from www.fk.cvz.nl), plus 6% tax, multiplied by the number of prescription days, plus the pharmacist’s dispensing costs of US$ 7.00 per prescription.

### Table 4.2.2 Direct medical and direct non-medical costs by health service type

<table>
<thead>
<tr>
<th>Health service type</th>
<th>Direct Medical</th>
<th>Direct Non-Medical</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unit Costs, US$</td>
<td>km, P, hrs</td>
</tr>
<tr>
<td>Medical doctor</td>
<td>Consult 22</td>
<td>1.8km, 1h</td>
</tr>
<tr>
<td>Medical specialist</td>
<td>Consult 107</td>
<td>7km, 2h</td>
</tr>
<tr>
<td>Regional mental health service</td>
<td>Contact 135</td>
<td>10km, 3h</td>
</tr>
<tr>
<td>Regional addiction service</td>
<td>Contact 135</td>
<td>10km, 3h</td>
</tr>
<tr>
<td>Mental Hospital – Outpatient</td>
<td>Consult 96</td>
<td>12km, 4h</td>
</tr>
<tr>
<td>Mental Hospital – Day care</td>
<td>Contact 136</td>
<td>12km, 4h</td>
</tr>
<tr>
<td>Mental Hospital – Inpatient</td>
<td>Day 272</td>
<td>8h</td>
</tr>
<tr>
<td>General Hospital – Outpatient</td>
<td>Consult 61</td>
<td>7km, 3h</td>
</tr>
<tr>
<td>General Hospital – Day care</td>
<td>Contact 249</td>
<td>7km, 4h</td>
</tr>
<tr>
<td>General Hospital – Inpatient</td>
<td>Day 366</td>
<td>8h</td>
</tr>
<tr>
<td>Teaching Hospital – Outpatient</td>
<td>Consult 109</td>
<td>12km, 3h</td>
</tr>
<tr>
<td>Teaching Hospital – Day care</td>
<td>Contact 249</td>
<td>12km, 4h</td>
</tr>
<tr>
<td>Teaching Hospital – Inpatient</td>
<td>Day 517</td>
<td>8h</td>
</tr>
<tr>
<td>Private practice psychotherapist</td>
<td>Session 83</td>
<td>5km, 2h</td>
</tr>
<tr>
<td>Social worker</td>
<td>Contact 49</td>
<td>7km, 3h</td>
</tr>
<tr>
<td>Physiotherapist</td>
<td>Contact 25</td>
<td>1.8km, 2h</td>
</tr>
<tr>
<td>Home care</td>
<td>Hour 33</td>
<td>0km, 0h</td>
</tr>
<tr>
<td>Informal care (family, friends)</td>
<td>Hour 9</td>
<td>0km, 0h</td>
</tr>
</tbody>
</table>


*b* Based on average distances (in km) and travel + waiting + treatment times (in hrs) for receiving treatment (cf. Oostenbrink et al, 2004).

*c* Costs of 1km = US$ 0.17, parking = US$ 2.72, 1h time = US$ 9.00 (cf. Oostenbrink et al, 2004).

*d* Valued as outpatient mental health services.

*e* From DFL 77.00 in 1993, indexed for 2003, converted into US$, and rounded.

To summarize, illness is associated with all sorts of down-stream costs such as direct medical and direct non-medical costs, and the costs that arise from production losses in both paid labour and in the domestic realm. The different cost components can be totalled into annual per-capita costs, which is the outcome variable in this study.

The costs were originally expressed in euro (€) for the year 2003, but have been converted in US$ for this paper. For conversion we used the purchasing power parities as obtained from the Organization for Economic Co-operation and Development (OECD; at [www.oecd.org/std/ppp](http://www.oecd.org/std/ppp)) where US$ 1.00 is equated with € 0.92 in the Netherlands, taking into account both the exchange rate and the relative purchasing power of the currencies. The cost calculations have been conducted in accordance with the Dutch guideline for health economic evaluations (Oostenbrink et al, 2004) which closely resembles other international guidelines (Langley et al, 1996; Siegel et al, 1997; Torrance et al, 1996). Elsewhere we have described similar costing procedures in greater technical detail (Smit et al, 2006b).

**Risk factors**

In this paper we have selected risk factors that occur early in the life-course and are known to increase the risk of mental disorder (Bijl et al., 2001; De Graaf et al, 2002a; De Graaf et al, 2002b; Kessler et al 1997; Merikangas et al, 1998).

A history of parental psychopathology, specifically “parental depression” and “parental anxiety” coded 1=presence, 0=absence of a depressed mood or anxiety in one or both parents. Parental problems were assessed in computer-assisted face-to-face interviews using questions like, “Did your biological father ever suffer from depression?” The variable “Multiple parental problems” was coded 1 when each of both parents had at least one mental problem, or one of the parents had two mental problems; otherwise it was coded 0.

“Childhood abuse” is captured by as a series of risk-indicators (1=occurred, 0=did not occur) of emotional neglect, psychological, physical and sexual abuse before the respondent reached the age of sixteen. Respondents were prompted to report these adversities with help of questions like: “Before you reached the age of sixteen, were you ever (-) neglected emotionally? (-) psychologically abused? (-) physically abused? (-) sexually abused?” During the interview definitions were given for each of these adversities. Emotional neglect, for example, was explained as follows: “(…) at home no one listened to you, your experiences and problems were ignored, you had the feeling that your parents were indifferent and did not support you.” Answer were placed on a scale ranging from never, once, more than once, often. In this study we used a low cutoff between “never” and “once”, thus in effect dichotomizing the scale into “never” and “ever”. The joint exposure to more than two forms of abuse or neglect defines “Multiple abuse”. 

131
**Protective factor: Mastery**

“Mastery”, also known as “internal locus of control”, was measured using the abbreviated (5 item) version of the Pearlin Mastery Scale (Pearlin & Schooler, 1978). It has a range from 1 to 5, higher scores indicating better mastery. For the analyses in this study, the scale is dichotomized at the median (1 = score below the 50th percentile on the scale, 0 = above 50th percentile; that is at the value of 4) in order to obtain a variable for low versus high mastery. We hypothesized that mastery will act as a protective factor: when a person’s mastery is intact, then it will significantly reduce the adverse effect of risk factors on the risk of becoming depressed. In other words, high mastery will reduce risks, whereas low mastery will increase risks. This effect modification may be understood in terms of the vulnerability-stress theory (Brown and Harris, 1978; Harris, 2000), which predicts that vulnerable people (here: people with low mastery) are more susceptible to the adverse effects of a stressor. Several studies have reported similar effects of mastery on the risk of disorder (Katz et al, 1994; Ormel et al, 1997; Zarit et al, 1999). Furthermore, mastery is amenable to change under psychological interventions (McCullough, 1980, 2005). Thus interventions may be able to strengthen mastery, and thus improve a patient’s capability to cope with adversities.

**Socio-demographic characteristics**

Socio-demographic characteristics are: “gender” (0=male, 1=female), “age” (in years, range 18 - 65), “education” (1= primary - lower vocational, 2= secondary - middle vocational, 3= higher vocational - academic), “living with a partner” (1=yes, 0=no), born in the Netherlands (1=Netherlands, 0=elsewhere), “occupational status” (1=with a paid job, 0=without), and “urbanity” (1=living in urban environments, 0=elsewhere).

**Analysis**

To account for initial non-response and drop-out, corrective post-stratification weights were used. After weighting, the sample followed exactly the same multivariate distribution over age, sex, civil status and urbanization as the population according to Statistics Netherlands (cf. www.cbs.nl).

To account for the possible non-normality of the cost data, we calculated sample errors, 95% confidence intervals, and P-values with help of non-parametric bootstrapping, with 1,000 bootstrap replications. In each bootstrap step, robust standard errors were obtained with help of the first-order Tailor-linearization method, as implemented in Stata version 8.2 (StataCorps, 2004). The latter was done to obtain correct 95% confidence intervals and P-values under weighting. For the remainder, the analyses were conducted in the following steps.

First, we estimated the average costs related to ill-health, irrespective of exposure to risk factors. The average costs were obtained with an intercept-only regression model. The regression model was weighted and its standard errors were based on robust techniques as outlined above.
In the second step the effect of each of the risk-factors was evaluated by entering one risk factor at a time in the model. In this way estimates of the excess costs were obtained that are attributable to each of the individual risk-factors. These costs are called “excess costs”, because we are looking at the “extra” or “added” costs that are generated when a person is exposed to a risk-factor. To be more precise, we refer to these costs as the unadjusted annual per capita excess costs (see Table 4.2.3, left hand panel). The same procedure was repeated, but now each risk-factor was accompanied by all the socio-demographic variables (see Measures). This was done to partial out the effects that might have been introduced by the socio-demographic characteristics. We refer to these costs as the adjusted annual per capita excess costs. The results of these analyses are presented in the right-hand panel of Table 4.2.3.

In the third step, we present the adjusted annual excess costs at population level (see Table 4.2.4). To this end, the adjusted annual per capita excess costs were multiplied with the number of people that were exposed to that particular risk-factor in every one million source population in the 18 – 65 age bracket.

In the fourth step, we calculated the societal costs, as before, first for the population segment with a below-average sense of mastery (left hand panel of Table 4.2.5), and then for the population segment with a higher than average sense of mastery (right hand panel of that table). In this way, the effect modification introduced by mastery can be visually inspected. To formally test the hypothesis that mastery is indeed a significant effect modifier a series of regression analyses were conducted with the appropriate interaction term of the risk factor, mastery and their constituent main effects while adjusting for socio-demographics. The P-values of the interaction terms are reported in Table 4.2.5.

It should be noted that all risk factors, mastery, and the socio-demographic characteristics were measured at \(t_0\), while the annual per-capita costs were assessed one year later at first follow-up \(t_1\). The temporal separation of putative causes and effects strengthens etiological inference.

4.2.3 Results

Sample
The study cohort consisted of 5,618 people. Of these, 49.2% were women. The mean age was 39.2 years (range 18 to 65). The distribution over educational levels was as follows: 5.5% elementary, 34.7% lower vocational, 30.0% secondary and middle vocational, and 29.8% higher vocational and academic. The vast majority, 93.5%, were born in the Netherlands, 68.5% were living together, 70.5% were employed (that is 81.2% of the men, and 59.5% of the women), and 82.7% lived in urban areas.
Average costs
The average annual per capita health-related costs are US$ 3804 (95% CI = 3510 ~ 4098), irrespective of exposure status. There is no appreciable difference in the costs between men (US$ 3814, 3411 ~ 4176) and women (US$ 3793, 3347 ~ 4281).

Excess costs of risk factors at individual level
Table 4.2.3 shows the annual per capita excess costs of being exposed to one of the risk factors. These excess costs are incurred over and above the base-rate cost of US$ 3,804, which was presented above. To give an example, the annual excess costs attributable to emotional neglect are US$ 2,612 per capita. However, a person who has been exposed to emotional neglect will generate costs equal to 3,804 + 2,612 = 6,416 dollars in that year, assuming that this person is not exposed to other risk factors. The table shows that the average excess costs of childhood adversities are somewhat higher than those of the parental risk-factors. Further, the unadjusted excess costs do not significantly differ from those that are adjusted for socio-demographics – as can be deduced from the substantial overlap of the 95% confidence intervals of the unadjusted and the adjusted estimates.

Table 4.2.3  Average per-capita excess costs (in US$ / year) of exposure to risk factors, not adjusted and adjusted for demographics

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>Not adjusted for demographics</th>
<th>Adjusted for demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Costs, US$/yr</td>
<td>95% CI*</td>
</tr>
<tr>
<td>Abuse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional</td>
<td>2612</td>
<td>1785 3437</td>
</tr>
<tr>
<td>Psychological</td>
<td>3285</td>
<td>2110 4460</td>
</tr>
<tr>
<td>Physical</td>
<td>3497</td>
<td>1965 5029</td>
</tr>
<tr>
<td>Sexual</td>
<td>2879</td>
<td>1411 4347</td>
</tr>
<tr>
<td>Multiple</td>
<td>3912</td>
<td>2735 5088</td>
</tr>
<tr>
<td>Parents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressed</td>
<td>1223</td>
<td>465 1982</td>
</tr>
<tr>
<td>Anxious</td>
<td>2638</td>
<td>1360 3916</td>
</tr>
<tr>
<td>Multiple</td>
<td>2323</td>
<td>784 3861</td>
</tr>
</tbody>
</table>

* 95% CIs based on 1,000 bootstrap replications and robust standard errors

Excess costs of risk factors at societal level
Table 4.2.4 presents the exposure rates of each of the risk factors. They range from 7% (sexual abuse) to 25.1% (emotional neglect). At population level the costs of the risk factors are equal to the per capita costs multiplied by the number of people exposed to the risk factor. This multiplication gives rise to the annual excess costs of each of the risk factors at societal level. In every one million people the excess costs of the risk factors vary between a relatively low 176.6 million US dollars for
multiple parental problems, which has a relatively low prevalence, to a staggering 635.8 million US dollar for emotional neglect, which has a much higher prevalence.

**Table 4.2.4** Exposure rates (in %) and annual costs of the risk factors (in millions of US$ per one million source population), adjusted for demographics

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>Exposure rates, %</th>
<th>Costs, mln US$/yr</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abuse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional</td>
<td>25.1</td>
<td>635.8</td>
<td>432.5 839.1</td>
</tr>
<tr>
<td>Psychological</td>
<td>12.9</td>
<td>390.1</td>
<td>243.6 536.8</td>
</tr>
<tr>
<td>Physical</td>
<td>8.7</td>
<td>280.0</td>
<td>150.2 409.6</td>
</tr>
<tr>
<td>Sexual</td>
<td>7.0</td>
<td>199.5</td>
<td>93.7 305.3</td>
</tr>
<tr>
<td>Multiple</td>
<td>13.7</td>
<td>507.2</td>
<td>346.9 667.5</td>
</tr>
<tr>
<td>Parents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressed</td>
<td>22.5</td>
<td>318.6</td>
<td>144.7 492.5</td>
</tr>
<tr>
<td>Anxious</td>
<td>10.9</td>
<td>298.2</td>
<td>153.8 442.6</td>
</tr>
<tr>
<td>Multiple</td>
<td>7.4</td>
<td>176.6</td>
<td>65.3 287.8</td>
</tr>
</tbody>
</table>

**Excess costs in the presence of a protective factor**

Table 4.2.5 shows how the costs of the risk factors depend on the level of mastery.

**Table 4.2.5** Costs at population level of the risk factors for people with low and high mastery (in millions of US$ per one million source population), adjusted for demographics

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>Low mastery Costs, mln US$/yr</th>
<th>High mastery Costs, mln US$/yr</th>
<th>P-value *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abuse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional</td>
<td>761.8</td>
<td>220.1</td>
<td>0.017</td>
</tr>
<tr>
<td>Psychological</td>
<td>449.0</td>
<td>134.0</td>
<td>0.046</td>
</tr>
<tr>
<td>Physical</td>
<td>385.9</td>
<td>14.0</td>
<td>0.166</td>
</tr>
<tr>
<td>Sexual</td>
<td>277.4</td>
<td>51.0</td>
<td>0.190</td>
</tr>
<tr>
<td>Multiple</td>
<td>652.5</td>
<td>127.4</td>
<td>0.019</td>
</tr>
<tr>
<td>Parents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressed</td>
<td>351.5</td>
<td>184.9</td>
<td>0.106</td>
</tr>
<tr>
<td>Anxious</td>
<td>503.0</td>
<td>35.0</td>
<td>0.008</td>
</tr>
<tr>
<td>Multiple</td>
<td>266.5</td>
<td>30.1</td>
<td>0.037</td>
</tr>
</tbody>
</table>

* P-value of the test that the interaction effect of low v high mastery on costs is significant

In the presence of poor mastery the costs associated with the risk factors rise. In the presence of high (intact) mastery the costs are much lower: these differences often
run into hundreds of millions of US dollars in every one million people, thus indicating the potential of substantial cost savings.

4.2.4 Discussion

Main findings
This study was set out to address the following questions: (1) What are the societal excess costs of risk factors that occur early in the life-course and are known to increase the risk of mental disorder? (2) How does presence of a putative protective factor interact with the adverse economic consequences of the selected risk factors?

The selected risk factors were childhood abuse (emotional, psychological, physical and sexual) and presence of mental problems in the parents (parental depression and anxiety). The protective factor was mastery (locus of control). The costs were expressed in US$ for the reference year 2003, and are given both at individual and population level, per one million people aged 18 - 65 years.

Now we can answer the questions as follows. The selected risk factors give rise to substantial per-capita costs. The annual excess costs fall in the range of US$ 1,223 to US$ 3,912 per exposed person, depending on the risk-factor. At population level these costs occur in many people, 7.0% to 25.1% of the studied population, which gives rise to societal costs in the range of US$ 176.6 million to 635.8 million annually in every one million people in the 18 - 65 age bracket. Mastery had a strong modifying effect on the costs associated with the long standing vulnerability factors. In the presence of below average mastery the risk-factors tend to generate more costs, but in the presence of above average mastery the costs associated with the risk-factors decrease, often to less than one-third of the average costs.

Cost estimates are lower bounds
For several reasons the costs reported here should be interpreted as conservative estimates. First, the sample is drawn from the non-institutionalized population. This has resulted in an under-representation of those who were hospitalized, which, in turn, has resulted in an under-estimate of the direct medical costs, because in several cost-of-illness studies it was shown that hospitalization was one of the major cost drivers (cf. Berto et al, 2000). Second, the medical costs that were considered in this study were limited to those related to mental health services and prescription drugs for mental disorders. Yet, it is well known that mental disorders also generate costs in non-mental health services (Greenberg et al, 1999). Third, we only present the short-term costs that are incurred over a one year period. Excluded are the costs of morbidity and mortality that may occur later in the life-course. Fourth, the cost calculations were based on self-reported medical consumption, which is known to result in an under-estimation of the actual health service uptake and corresponding costs (Van den Brink et al, 2004). Fifth, apart from work loss days due to absenteeism, work cutback (while at work and not feeling well) will also result in
production losses. The latter were not included in this study, and can be substantial (Brower et al, 1999; Kessler & Frank, 1997; Lim, Sanderson & Andrews, 2000). Finally, it is known that excessive alcohol consumption and some axis-II disorders are associated with crime-related costs (Andlin-Sobocki, 2004). In our study, these costs were not included. Therefore, the costs presented here are conservative lower bounds.

**Strengths and limitations**

This study has several strengths. Both the risk factors and the costs were assessed in a large epidemiological cohort study representative of the non-institutionalized population of 18 to 65 years in the Netherlands. In this way several risk factors and costs could be analyzed simultaneously. The cost calculations encompassed direct medical, direct non-medical and indirect costs. The costs of mental ill-health were adjusted for socio-demographic characteristics. By directing attention to the risk factors, we circumvented problems related to diagnostic heterogeneity. In this way it was possible to estimate the excess costs attributable to the distinct risk factors irrespective of the exact way these risk factors impacted on the incidence of disorders. It should further be noted that the risk factors were measured at baseline ($t_0$) and the costs one year later at first follow-up ($t_1$). This strengthens etiological inference. Finally, this is the first population-based cost-of-risk-factor study which also evaluated the effect of a protective factor.

Nevertheless, we must mention a number of important limitations. Despite the large sample, the 95% confidence intervals around the cost-estimates are broad. This is a common finding in health economic studies, because cost-data typically have large standard errors. The measures on parental psychopathology are not the best, because they were based on heteroanamnesis. Measurement error in the exposures may have attenuated the correlation between risk factor and the costs, which increases the likelihood that the cost have been underestimated. Finally, we must be careful in our interpretation of mastery as an effect modifier. In this population-based study we could observe that people with a relatively high level of mastery generate fewer costs when exposed to risk factors than people with relatively low levels of mastery. However, we could not observe if a change in mastery is followed by a change in costs. Hence, one must be careful when drawing conclusions from between-subject variability about within-subject change, and it is only the latter that is interesting from the clinical perspective. Therefore, we regard the idea that mastery acts as an effect modifier – and helps to cushion the adverse economic effects of risk factors – only as hypothetical. We find this hypothesis interesting for clinical and economic reasons, but still, it is only a hypothesis, and as such it will require further testing.

**Implications**

Burden-of-illness studies and cost-of-illness studies have placed psychiatry at the very frontier of public health: existing studies showed that mental disorders are
disabling, highly prevalent and have economic consequences equal to, and often surpassing, those of other medical conditions. However, both types of studies direct attention to one disorder at a time. In so doing they fail to highlight a singularly important aspect of many mental disorders: one and the same risk factor may play a role in the pathogenesis of more than a single disorder. Therefore, the underlying risk factors present society with a disease burden and economic costs that can be larger than those associated with each of the distinct disorders alone.

In this paper we demonstrated that this is true for a small number of risk factors that occur early in the life course: childhood adversities and parental history. Even in an adult population aged 18 – 65 years, with a mean age of 39 years, these early risk factors do not fail to lead to substantial production losses in both paid work and in the domestic sphere. Furthermore, they are associated with excessive health service uptake, and the corresponding patients’ out-of-pocket costs.

For these reasons it would be well to systematically address the question how best to curb the adverse consequences that these risk factors impose on society. Here preventive psychiatry may have to play a role. However, not all of the studied risk factors can be avoided, and for some of them it is difficult to see how they would be amenable to change – but it would, perhaps, be possible to cushion their adverse effects with help of a protective factor. Strengthening a sense of mastery in (adult) children of depressed and anxious parents, and in people who have been exposed to childhood adversities, may offer one strategic inroad: the average costs of the studied risk factors is, roughly speaking, one-third in people with an intact sense of mastery. This offers some hope and may also tell us something about the dollar value of prevention. Nevertheless, further evidence is needed for the possible role of strengthening mastery as a means to improve prognosis and reduce costs in the face of adversity.
Part V

Prevention of depression: summary and general discussion

Chapter 5.1

Prevention of depression: summary of preceding chapters
Chapter 5.1

Summary of preceding chapters

5.1.1 Introduction

The aim of the previous chapters was to answer four questions about prevention of depression: (1) do we need it? (2) do we know where to start? (3) is it effective? (4) is it affordable? Here, in the Summary we will return to these questions, and address each of them.

5.1.2 Do we need it?

Chapter 1.1

Individual suffering, however severe, does not make a major public health problem. Here a second factor has to come into play. It is only when individual suffering is multiplied by the vast number of prevalent cases that a disorder attains the status of a major problem in the realm of public health. Indeed, both severity and prevalence have combined to make depression one of the disorders with the largest burden of illness, world wide.

Although, depression is a condition that can be treated successfully, it has been shown in two studies (Andrews et al, 2004; Chisholm et al, 2004) that current curative interventions can only partially reduce the depression-related disease burden at population level. As a consequence there remains a formidable gap to be bridged between the needs of many and what treatment can offer. Therefore, we need to look for interventions other than curative treatment alone. In the formulation of Chapter 1.1, this major health problem has to be tackled along its width and breadth, and that is why prevention has to play an adjunctive role to treatment. This need is further underscored by the fact that the incidence of depressive disorder is very substantial relative to its prevalence: 49% of the prevalent cases are in fact new cases. Seen from a public health perspective, this calls for a reduction in the massive influx of new cases. Again, prevention may be the answer – provided that we know where to begin, that prevention is effective, and that it is affordable.

Chapter 1.2

The latter issue brings us to Chapter 1.2, which is a cost-of-illness study of the common mental disorders. Here we showed, by a conservative estimate, that the annual excess costs of depressive disorder amount to € 2,300 per capita, which gives us some idea of the potential cost-offset of successfully preventing new onsets. At population level, where the per capita costs have to be multiplied by the vast number of prevalent cases, the societal costs become staggering: depressive
disorder poses a formidable economic burden to society of more than €1,300,000,000 in the Dutch population aged 18 – 65 years, each year, every year. It stands to reason that a reduction in the incidence of depressive disorder will also bring about a reduction in the economic burden of the disorder. Thus Chapter 1.2 has two-fold take-home message. First, it might be worthwhile, from a strictly economic perspective, to further develop, evaluate, and implement preventive interventions in depression. Second, it holds the promise that prevention of mental disorder, in particular depressive disorder, is likely to become a cost-effective endeavour. This cost-of-illness study is in press at the Journal of Mental Health Policy and Economics.

5.1.3 Do we know where to start?

Background
Against the background sketched in Chapters 1.1 and 1.2, it becomes important to get clear-cut answers to the question, do we know where to begin with depression prevention?

The question is relevant, because far too many people are exposed to one or another possibly relevant risk factor; and, leaving ethic considerations aside, it would be a sheer impossibility, both logistically and economically, to provide all these people with selective prevention. Similarly, the number of people with some depressive symptoms who don’t meet the diagnostic criteria of depressive disorder can be very large too, depending, of course, on the exact definition of subthreshold depression. Moreover, vaguely defined groups, wherein the risk of actually becoming depressed may vary considerably, may give rise to justifiable ethical concerns with regard to the acceptability and appropriateness of offering preventive interventions. Hence, the relevance of the question, where to start?

Our approach to answering this question was to try to identify groups with a calculable ultra-high risk for the disorder. The idea was that ultra-high risk groups would at once be numerically small, and yet account for a large share of all the new cases of the disorder in the population. Furthermore, in small, ultra-high risk groups, with a large share in the incidence, prevention is likely to stand the best chance of becoming cost-effective. Expressed in epidemiological and statistical terms, we were looking for groups with a high incidence rate (IR) of depressive disorder, higher than the incidence rate in groups not exposed to certain risk factors. Hence, we were looking for groups with a high incidence rate ratio (IRR). At the same time, we would like these groups to be as small as possible, i.e. with a very low exposure rate (ER) of the risk factors. Both the exposure rate and the incidence rate ratio can be combined in another statistic, the population attributable fraction (AF). The AF is a useful statistic from a prevention perspective: it indicates by how many percent the incidence rate of the disorder in the population will be lowered, when the adverse effect of the risk factor can be completely blocked by some
preventive intervention. This makes the AF an index of potential health gain in a population under optimal preventive scenarios. It would not be realistic to assume that preventive interventions are completely successful in averting the impact of a risk factor. Therefore, it is readily understood that the AF puts an upper limit to the potential health gain. Nevertheless, the AF is a useful statistic when it comes to creating a hierarchy of risk factors that have value for prevention.

Chapter 2.1
This approach was taken in Chapter 2.1 in a somewhat explorative fashion. We made use of a population-based psychiatric cohort study, Nemesis, and calculated exposure rates (ERs), incidence rate ratios (IRRs), and population attributable fractions (AFs) for a large set of putative risk factors assumed to be predictive of first-ever onset of depressive disorder in people aged 18 – 65 years. For each risk factor we also calculated the number-needed-to-be-treated (NNT) as an index of the potential efficiency of a preventive intervention when it is successful in blocking the adverse effect of that particular risk factor. Using this methodology, we began to see the numerically small, high-risk groups, where prevention was likely to generate the largest health gains, in the most efficient way, for the fewest costs. This study was published in the Journal of Affective Disorders (Smit et al, 2004), and ended on a note to the effect that future research should compute the indices for health gain (IRR, AF) and effort (ER, NNT) not for a single risk factor at a time, but for multivariate sets of risk factors, such that these indices are maximised and minimised in their respective desired directions.

Chapter 2.2
It was precisely this multivariate approach that we took in Chapter 2.2. Here, we evaluated the effects of joint exposures on the indices of health gain and effort in another large population-based psychiatric cohort study of people aged 55 – 85 years, the Longitudinal Aging Study Amsterdam (LASA). This approach enabled us, in a much better way, to identify the ultra-high risk groups, that were numerically small, that accounted for the bulk of new cases of depressive disorder, and where prevention stood the best chance of being efficient and cost-effective. We also performed an ante hoc economic evaluation, demonstrating that prevention of late-life depression was indeed likely to be a cost-effective venture. The corresponding study was published in the Archives of General Psychiatry (Smit et al, 2006a), and ended with the observation that the new methodology employed in this study needed cross-validation in a further study.

Chapter 2.3
This cross-validation is contained in the last chapter of Part II of this study. Use was made of yet another psychiatric cohort study, viz., the Amsterdam Study of the Elderly (the Amstel study). It is worth noting that in this study depression was ascertained in a different way, and slightly different definitions of the putative risk
factors were used. We also made use of a different method of identifying ultra-high risk groups. This method is known as Classification and Regression Tree (CART) analysis. The “tree” in the name refers to the tree-like diagrams that are produced to systematically trace the effect of adding new risk factors in terms of ER, IRR, AF and NNT. The procedure of adding risk factors helps to ascertain what combinations help optimise the health gain and efficiency, and minimise the effort of generating these health gains.

It was interesting to see that this approach yielded near identical results to those found in Chapter 2.2: indicated prevention in later life is best directed at people who (-) have some depressive symptoms, (-) have a chronic somatic illness, (-) experience disability (-) are living alone, and (-) are female. There was one exception: in the Amstel study we found widowhood to be a significant predictor, while in the LASA study we did not. However, it should be noted that in the LASA study widowhood was defined as “ever widowed” (without reference to a specific time-frame), whereas in the Amstel study it was more accurately described as “recent widowhood” (in the last 6 months). Thus, the different definitions may account for the single difference in the studies’ outcomes.

On account of the overall congruence of both study outcomes, we are tempted to conclude that we may say with some confidence that we now do see where to begin: indicated prevention in groups that have some depressive symptoms offers the best start, and can be improved on when the target group is further selected with help of the afore-mentioned risk factors. The study also indicates that selective prevention may be an alternative; it offers the additional benefit that the corresponding target group can be easily identified in the primary care setting with help of a brief checklist based on a small set of easy-to-recognise risk factors, such as recent widowhood, medical illness, disability and female gender. It is these risk indicators, and in this combination, that carry the promise that prevention may yield substantial health gains in an efficient and presumably cost-effective way. This study, by Schoevers, myself and colleagues, has been published in the American Journal of Psychiatry (Schoevers et al, 2006).

5.1.4 Is it effective?

Background
The third part of the thesis addresses the question concerning the effectiveness of preventive interventions in avoiding new onsets of depressive disorder. To put this in context, the effectiveness of preventive interventions has been evaluated in literally hundreds of randomised trials, but the outcome was nearly always a reduction in symptom level, and only very rarely a reduction of incidence. Thus in effect, next to nothing was known about primary prevention of depression. The reason for this apparent omission is clearly explained in Cuijper’s 2003 article in the American Journal of Psychiatry where it was demonstrated that due to the often
very low incidence rates of mental disorders, differences in the incidence rates across experimental and control conditions can only be evaluated in trials that are based on astronomically large sample sizes. Logistic and financial constraints preclude this as a viable option. However, the sample size problem can be circumvented by conducting trials in high-risk groups with correspondingly high incidence rates.

Chapter 3.1
We followed this strategy in the study described in Chapter 3.1. This study was designed as a pragmatic randomised prevention trial in two parallel groups. The trial participants were general practice patients, who were recruited while waiting to see their doctor. They were eligible for inclusion in the trial if a number of criteria were met; the most important of these were having at least one depressive core symptom plus one, two, or three additional symptoms, not meeting the diagnostic criteria for DSM-IV depressive disorder as ascertained with the CIDI. Exclusion criteria were presence of full-blown DSM axis-I disorders, such as dysthymia and bipolar disorder. On the basis of the scientific literature we expected the risk of depressive disorder to be close to 20% per year in this population. We hypothesised that this risk could be lowered by one third if in addition to care-as-usual an adjunctive preventive cognitive-behavioural self-help intervention was offered with minimal guidance. To power the trial to detect such an effect, we needed to enrol 200 participants in each condition; a number sufficiently large to also compensate for some expected loss-to-follow-up. The central end-point of this trial was CIDI/DSM-IV depression status after 12 months.

Analyses were conducted while adhering to the intention-to-treat principle. For this, missing observations at follow-up were imputed. The analyses also accounted for a clustering effect in the data, which was induced by the fact that some patients were recruited from the same general practice. In total, 107 consenting participants were randomised to the intervention and 109 to the care-as-usual group.

The results were as follows. In the care-as-usual group the incidence rate of depressive disorder was 18%; and this was 12% in the treatment group, implying a reduction by 33%. The unidirectional 0-hypothesis that adjunctive self-help combined with care-as-usual is inferior to care-as-usual alone, had to be rejected at P<0.05 in a one-sided test, thus lending credibility to the idea that self-help as an adjunct is superior to primary care as usual alone.

This was good news, but there were two problems: recruitment had proved to be a difficult process, and the participation and completion rates in the intervention were low. This reflected unfavourably on the acceptability of the intervention, and may call for more stringent selection criteria of the participants, for example on the basis of the risk factors as identified in Chapters 2.1, 2.2 and 2.3, and especially with regard to the participants’ own perceived need to enrol in such an intervention.
The outcomes of the prevention trial were published in the British Journal of Psychiatry (Willemse et al, 2004).

**Chapter 3.2**

Our prevention trial (see Chapter 3.1) was one of the seven that ascertained the effectiveness of a psychological intervention in actually preventing the onset of full-blown depressive disorder in people with subthreshold depression. In Chapter 3.2 we pooled the outcomes of all seven studies meta-analytically.

Two types of outcomes were considered. The first is the reduction in depressive symptom level, and the corresponding effect size is Cohen’s d, which indicates by how many standard units the experimental group is removed from the control group in terms of health gain. The second outcome is the reduction in the risk of becoming a case of depressive disorder in the experimental group relative to the control group, which was ascertained with an outcome technically known as the incidence density ratio.

The outcomes of the meta-analysis were as follows. All seven studies examined a total of 700 subjects, with 343 subjects in the experimental conditions, and 357 in the control conditions. In six studies, the control condition was care-as-usual, and in one study it was a waiting list with unrestricted access to care-as-usual. Because heterogeneity was virtually absent, and outcomes of random effect models and fixed effect models were quite similar, we reported only the outcomes of the more simple fixed effect models. Symptom reduction was achieved by 0.42 standard units, in favour of the psychological interventions, and as measured immediately after the intervention. This effect size represents an effect of medium size as compared to what is found as the median of all effect sizes in the field of behavioural and psychotherapeutic interventions. However, we also noticed a drop in the effect size over time: after one year it had shrunk to a small effect of 0.16 standard units, which only bordered on statistical significance at P= 0.08, two-sided. The second outcome, which captured the relative risk reduction of becoming depressed, indicated that psychological interventions help to reduce the incidence of new cases by 30%, which was marginally significant at P= 0.07 in a two-sided test.

Our appreciation of these results is as follows. In clinical terms the effect sizes are in an order of magnitude that is in line of what may be expected from relatively low-key psychological interventions, but the observation that effects diminish over time must be seen as an important issue. The time-decay may indicate the need for booster sessions, and may further alert us that we have to understand prevention of mental disorders not in terms of a single-shot, magic-bullet, sort of approach, but rather more like safety belts in cars that need to be used every time we drive. In the same vein, the risk reduction of 30% looks good from the clinical perspective, but leaves one wondering whether this effect will survive over time. From the statistical point of view we are not overly concerned with P-values that were only bordering on significance, because the P-values would have been significant in an one-sided test, while, in this context, the use of a one-sided test can
be justified by the unidirectional hypothesis based on the idea that adjunctive interventions added to care-as-usual are almost certainly superior to care-as-usual alone. But not wishing to jump to conclusions prematurely, we suggest that all we have now are “indications” that preventive interventions of a psychological nature can be successful in preventing full-blown depressive disorder in people suffering from subthreshold depression. It is also clear that more prevention trials are required, because we need to know what groups in what settings benefit from what interventions and for how long.

5.1.5 Is it affordable?

Chapter 4.1
We conducted a cost-effectiveness analysis alongside the same prevention trial that was described in Chapter 3.1. In the economic evaluation we took the societal perspective. To this end, the following costs were considered. Direct medical costs arise in connection with the uptake of all types of health services, including this intervention. Direct non-medical costs include the patients’ out-of-pocket costs that are made when travelling to health services, paying for parking, plus the cost of the patients’ time while travelling, waiting and receiving treatment. Furthermore, we included the indirect non-medical costs, that is, the costs related to production losses in paid work due to work loss days and work cutback days, plus the costs of production losses in the domestic sphere. Costs were calculated for one year, and were therefore neither discounted nor corrected for inflation, as per the Dutch guideline for health economic evaluation.

The results were as follows. The adjunctive self-help intervention with minimal guidance combined with care-as-usual in primary care was 33% more successful in avoiding onsets of depressive disorder than care-as-usual alone (cf. Chapter 3.1). The mean integral costs are € 6,766 per capita per annum in the experimental group, which compares favourably with the sum of € 8,614 in the control condition. The incremental cost-effectiveness ratio indicated that offering the adjunctive self-help results in a modest cost saving of € 288.75 per depression-free survival year. However, the latter outcome was surrounded by a considerable degree of stochastic uncertainty.

A probabilistic approach, based on 2,500 bootstraps, indicated that adjunctive self-help had a 70% probability of being more acceptable from a cost-effectiveness perspective than routine primary care alone, even in the conservative scenario that there is no willingness to pay for the health gain of a depression-free survival year. A sensitivity analysis showed that production losses were the main cost-drivers. Excluding these costs, and thus limiting the economic evaluation to direct costs only, showed that adjunctive self-help has a 47% probability of dominating routine primary care in terms of the incremental cost-effectiveness ratio at a willingness to pay equal to zero. Raising the willingness to pay to a ceiling of €
30,000 per depression-free survival year resulted in a 72% probability that the adjunctive self-help intervention is more acceptable than routine primary care alone as seen from a cost-effectiveness point of view.

The outcomes of the study show that from a health economic point of view, adding the cognitive behavioural self-help intervention with minimal guidance to routine primary care has the benefit of producing health gains while at the same time it generates cost-offsets that compensate for the outlay involved in providing the self-help intervention in the first place. Thus, offering this intervention in the primary care setting appears to be an attractive option. Nevertheless, and as was pointed out before, we need to know more about the acceptability of the intervention, and its long term outcomes, before we can confidently recommend its broader implementation.

At the time when we wrote this paper, we were, to the best of our knowledge, the first to conduct an economic evaluation alongside a primary prevention trial in depressive disorder. However, before our paper appeared in print in the British Journal of Psychiatry (Smit et al, 2006), Lynch and colleagues published another cost-effectiveness study in the same research field, which appeared in the Archives of General Psychiatry, in 2005. In their paper they reached a similar conclusion to ours: a brief intervention to prevent depressive disorder in at-risk teens aged 13 to 18 years, was economically affordable given a variety of thresholds for the willingness to pay for a depression-free day. To summarise, we begin to see the first evidence that primary prevention of depression is an economically viable option. Nevertheless, there is a clear need to conduct more studies of this type, in order to better understand the outcomes of different interventions in different groups, and over longer time periods. We also need more in-depth efficacy studies to better understand what components of the intervention do really matter in producing the desired health effects. This knowledge can than be used to do a better job in designing cost-effective interventions.

Chapter 4.2

Many of the key concepts from the previous chapters were re-introduced in the last chapter, and brought together in a somewhat different framework. Again, the focus was on costs and risk factors, but this time we studied the costs of risk factors, thus putting an epidemiological slant on health economics. In so doing, we wanted to draw attention to the very roots of mental illness, disability, and the ensuing societal costs. In addition, we wanted to evaluate how the relationship between risk factor and costs was modified by a protective factor, mastery, which can be defined as the sense of being in control over one’s life. The idea was that an intact sense of mastery would help to cushion the adverse economic and health effects of being exposed to a risk factor. If this is indeed the case, then it may be of interest to pursue the concept of strengthening one’s sense of mastery as a strategic research topic.
For the analysis, we revisited the large population-based psychiatric cohort study, Nemesis, conducted among 5,618 people aged 18 – 65 years (cf. Chapters 1.2 and 2.1). The risk factors were measured at baseline and describe parental depression and parental anxiety as well as exposure to childhood abuse and emotional neglect before the respondents had reached the age of sixteen. These “early” risk factors are known to be predictive for the later onset of mental disorders. To facilitate etiological inference, the costs were measured one year later and encompass the integral medical and non-medical costs related to health care uptake, and the costs of production losses in both paid work and in the domestic sphere (cf. Chapter 1.2).

The results were as follows. The risk factors were associated with excess costs in the order of 2,000 – 3,000 US$ per exposed person per annum. Given the exposure rates of the selected risk factors in the population, the costs fall between 170 – 770 million US$ in every one million people. The effect of low versus high levels of mastery was in the hypothesised direction, with higher costs in exposed people with below average levels of mastery, and lower costs in exposed people who had above average mastery. In fact, the costs of the risk factors were down to roughly one-third of the average costs when people had an intact sense of mastery.

The results show that the selected risk factors are associated with formidable costs, even many years after exposure. The excess costs are incurred, year after year, and remain virtually the same when the effect of demographics variables is accounted for in multivariate analysis. This suggests that prevention directed at these risk factors may be economically rewarding. Still, it is difficult to see how prevention could help to contain the adverse effects of some of the selected risk factors, although strengthening mastery might prove to be the option we are looking for: it helps to reduce the costs associated with the risk factors to one-third of the original mean costs. However, at this point a word of caution is needed. One would much like to see analyses that show how a change in mastery is later followed by a change in costs, but these analyses could not be performed using the available data. Currently we are collecting longitudinal trial data on both mastery and costs, and we foresee that the proper analyses can be conducted in the near future. That research can now be guided by a new working hypothesis: with help of preventive interventions we may be able to induce a change in mastery, which, in turn, might be followed by changes in risk status and changes in the down-stream costs of mental disorder.
Part V
Prevention of depression: summary and general discussion

Chapter 5.2
General discussion
Chapter 5.2

General discussion

5.2.1 Introduction

In the General Discussion we will first review the methodology (and its underlying conceptual framework) that was used in this thesis, then we will return to the research questions and present their answers in the form of a number of propositions. Each of the answers came with some qualifications, and we will summarise these qualifications. The General Discussion will be concluded with a brief reflection on a number of issues that need, we believe, to be pursued in the future.

5.2.2 Conceptual and methodological insights

Conceptual issues
One way to summarise this thesis is to say that it offers both a conceptual outline and a methodology to the scientific underpinning of primary prevention of mental disorder. At the conceptual level, the idea is that population-based epidemiology and trial-based studies may help to shed light on the disease stages and the factors that promote or inhibit the transitions from one stage to another. Figure 5.2.1 illustrates this key-point.

Figure 5.2.1 Health conditions, transitions, prevention and treatment

- Universal prevention
  - General population
- Selective prevention
  - Vulnerable people
- Indicated prevention
  - People with early symptoms
- Maintenance therapy
  - Recovered people at risk of recurrence
- Relapse prevention
  - People in remission
- Treatment
  - People with the diagnosis
In brief, the idea is that the general population constitutes the target group of universal prevention. In this population, some may have inherited or acquired a risk status and thus form the target group for selective prevention. Of these, some may develop depressive symptoms. In these cases, the risk of becoming a case of full-blown depressive disorder is markedly increased, which makes it appropriate to offer indicated prevention based on early recognition and early intervention. Next, when people meet the diagnostic criteria of depressive disorder, they may look for and receive treatment. When treatment has been successful, or when people have remitted spontaneously, then they may still be at risk of relapse, especially in the presence of residual symptoms and prior history. At this stage some form of relapse prevention may be appropriate. Finally, people may have recovered partly or wholly from depressive disorder, but depression is known to often reoccur and people with previous recurrences may want to participate in some form of (collaborative) maintenance therapy to lower the risk for new onsets. Thus at each stage some form of intervention can be offered with the aim of curbing undesirable transitions and of promoting desirable transitions. For that, we need to know what factors play a role in each of these transitions.

**Methodological issues**

The broad idea is to employ population-based and trial-based studies to better understand what factors promote or inhibit the transitions. For this, population-based psychiatric cohort studies offer a good starting point: they help to generate clinical and economic hypotheses, which, in a second step, can be evaluated in trial-based studies. It is worth taking some time to better clarify the first – hypothesis generating – step, because here we need to be more precise. The idea is that existing hypotheses, however vague, can be “pre-tested” by conducting “virtual quasi-experiments” in cohort data by isolating interesting contrast (e.g., exposed versus unexposed) in selected at risk groups, and by studying the longitudinal impacts of these exposures on relevant clinical and economic endpoints. Thus “virtual quasi-experiments” can be conducted by statistical manipulation of available cohort data. In the Netherlands we are very fortunate to have high-quality longitudinal psychiatric epidemiological studies, which allow us to conduct these virtual quasi-experiments. This helps to “pre-test” hypotheses, to give them more substance and a greater specificity. In a next step, these “grounded hypotheses” help to guide research to areas where prevention is likely to become most cost-effective. This helps to rationalise our research agenda, which is necessary because prevention trials are time-consuming and expensive.

When taking this hypothesis generating approach, it is important to compute indices of health gain and costs, and we rekindled interest in some of these indices that have been around for a long time but were rarely used in the research field of mental health promotion and prevention, such as the population attributable fraction (AF). Measures like these provide a language to communicate about the outcomes
of the aforementioned “virtual quasi-experiments”. In addition, we thought about ways to use the universal language of hard currency to communicate about prevention and its consequences.

Once the “grounded hypotheses” are available, it is time to test them in randomised prevention trials. This is a necessary step, because the grounded hypotheses were derived from virtual quasi-experiments in which participants could not be randomised, and, as a consequence, selection-bias poses a real threat to the validity of the corresponding outcomes. To overcome that hurdle, properly designed randomised trials are needed.

Finally, when trials have been conducted, we may need to go one step further, and pool their results meta-analytically in order to raise power, increase precision, and to contribute to science’s overall aim: the accumulation of knowledge. Figure 5.2.2 illustrates these steps.

**Figure 5.2.2** Linking research types

<table>
<thead>
<tr>
<th>Cohort data</th>
<th>Trial data</th>
<th>Effect sizes from trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtual quasi-experiments</td>
<td>Hypothesis testing</td>
<td>Meta-analytic research synthesis</td>
</tr>
</tbody>
</table>

5.2.3 Answering the research questions

**Answers**

This thesis is not just a methodological exercise, but the methodology was applied to the case of depressive disorder on the understanding that depression is one of the major causes of disability in the general population, and we wanted to make some contributions to what ultimately should become evidence-based prevention of depression. Our contributions were the answers to the research questions that were posed at the very start of this thesis: is prevention of depressive disorder necessary, do we know where to start, is it effective, and is it economically affordable? Our answers can now be summarised in five statements:

1. The human and economic costs of depression are staggering
2. Prevention is the missing link in public mental health
3. Indicated prevention of depression is best directed at ultra-high risk groups
4. Indicated prevention of depression can be effective
5. Indicated prevention of depression can be economically affordable
As was shown in the previous chapters, all five propositions come with accompanying qualifications, and all have to be placed within their respective margins of uncertainty, but these five points pretty much represent the corollary of this thesis.

**Caveats**

Here we like to summarise the most salient qualifications and reservations that we have with regard to the five points.

We have few doubts about point 1, because the evidence about the burden of disease and the economic costs of depression is abundant, consistent across studies, and we believe, almost universally accepted. However, as was pointed out in Chapter 1.1, we would not be too surprised if the current estimates of the burden of depression should turn out to be underestimates, because the disability adjusted life year (DALY) of depression does not adequately reflect depression-related excess mortality other than suicide. If this assessment is correct, then this would strengthen our case.

When we say that prevention is the missing link in public mental health (point 2) then, obviously, we are not referring to the frenzy of activities and initiatives that make up the daily practice of prevention in mental health in the Netherlands. Also we are not referring to a lack of official recognition. On the contrary, it is very gratifying to see that the Netherlands Ministry of Health (VWS) has played a strong role in encouraging research and development in this field, and is now in the process of preparing the Second Policy White Paper on Prevention. If the signs are right (this was written in April, 2006), then prevention of depression will be explicitly included in this Policy White Paper, which would represent a historical break-through. Perhaps more important than official recognition – and now we get at the heart of “missing link” issue – will be the reception of the very concept of preventive psychiatry in clinical and medical thinking. For example, it would be a very significant development if primary prevention were recognised as a necessary first step in a stepped-care approach; however, this would imply that prevention has to be tied in with primary health care, which is not an easy task. In the same vein, it would be a significant improvement if relapse prevention and maintenance therapies for people at risk of recurrences of depressive disorder were to become an integral part of clinical and medical thinking and its practice. After all, prevention, cure and care are placed on a disease severity continuum, cannot be separated, and ideally will have to be organised as an unbroken chain along this continuum. Thus, when we say that prevention is the missing link in public mental health, then we are referring mainly to its organisational position among its sister disciplines. In the current state of affairs, primary prevention of depression is not firmly anchored in primary care, nor is it often recognised as a first step in a stepped-care approach. It is only very rarely mentioned in (multidisciplinary) clinical guidelines for the (collaborative) treatment of depression, and receives only a tiny fraction (about 3%) of the national budget for mental health care. At the
Trimbos Institute we are currently conducting field trials in which we experiment with an integrated model of prevention of late-life depression. This can be seen as an attempt to tie in with other health services at local level, improve early recognition of depression, and improve and rationalise referral streams of patients. It is also hoped that this will help to improve the organisational position of prevention, but outcomes of these studies have to be awaited. All in all, it is the lack of recognition of primary prevention as one of the players in the field of public mental health that renders its position obscure and marginal – although we would hasten to say that recognition is something that must be earned, and cannot be expected by right.

The most important qualification with regard to statement 3, that prevention is best directed towards ultra-high risk groups, is that we talk about indicated prevention. When people are aware of depressive symptoms, have complaints, and seek help for these complaints, then offering a preventive intervention to reduce the risk of further deterioration of their condition is likely to be ethically acceptable. However, for selective prevention another picture may emerge, and we may have ethical reservations about targeting high risk groups, who have no depressive complaints, are not looking for any kind of help, and are possibly unaware of their elevated risk status. We will return to this important issue shortly.

The cost-effectiveness of prevention of depression (points 4 and 5) can only be established for each and every intervention in well-defined risk groups in clearly defined settings. This will be a slow, piece-meal approach to gathering evidence. But in the light of our experiences so far, we believe we are on the right track, and have reason to be optimistic about the outcomes of future studies in this field. However, one of the most crucial areas in this research field will be the study of long-term clinical and economic outcomes, which may drastically alter the picture that we begin to see for the short-term outcomes.

In light of these qualifications and reservations, we do not regard the five propositions as definite answers, but we see them as directions in which we have to look for answers. These answers can be produced with more research, and it is our expectation that the ensuing studies will be successful in providing the much needed scientific underpinning of prevention of depression in particular, and of the common mental disorders in general.

### 5.2.4 Directions for the future

As yet, prevention of depression is an unfinished project. We have a long way to go before it can be described, with some confidence, as an integrated and respected discipline among its sister disciplines in public mental health. We need to know more about what interventions work best for whom. We need to know more about the long-term effects. We need to know more about the most acceptable and cost-effective delivery formats to reach the largest possible number of people in need of
it. The latter is an important issue, because the current coverage of depression prevention is somewhere in the region of several thousand people per year in the Netherlands, which amounts to less than one percent of the group that will become a case of depressive disorder in that year. To have any appreciable impact on the incidence rate, preventive interventions will have to be accessible on a much broader scale. Several options have been suggested and are described in the following sections.

_Towards evidence-based prevention_

First, we should persist in our efforts to upgrade the scientific status of prevention. Ultimately, we need to arrive at some established and recognised form of evidence-based prevention. As was said before, recognition cannot be expected by right, but must be earned. This requires effort. More preventive interventions will have to be evaluated in randomised trials, and once interventions have obtained the status of evidence-based interventions, measures need to be put in place to guarantee protocol adherence, and to monitor both compliance rates and outcomes in the recipients of the interventions. In this context, monitoring protocol adherence, compliance and outcomes would help not only to guarantee that professional standards are met, but also to generate the necessary feed-back to alert us to when and where improvements should be made. After all, it is our primary responsibility to offer interventions of high quality, and this responsibility will, of necessity, involve accountability.

_Selecting target groups: ethical considerations_

Second, we will have to pause, and think very hard about the population segments to which we want to offer preventive interventions. _Selective_ prevention aimed at people at elevated risk of developing depressive disorder, introduces the ethical problem that these people may be unaware of their risk status, may not look for help, and then it is likely to be inappropriate and unacceptable to confront them with their risk status and to start a (medical) treatment. _Indicated_ prevention, in people in the prodromal disease stage, may be more acceptable from an ethical perspective. These people have some depressive symptoms and may already be seeking help. The disadvantage of this approach is that the number of people in this category can be very large, depending, of course, on the exact definition of prodromal depression. At any rate, it is hard to see how, for example, 15% of the adult population could be offered preventive interventions. We would suggest offering preventive interventions only when the following criteria are met: (1) the participants have some depressive complaints, (2) the participants are willing to participate in the intervention, (3) effective and safe interventions are available, which are as un-intrusive as possible, (4) in addition to having some depressive symptoms, the participants have been exposed to risk factors suggestive of an elevated risk of developing depressive disorder in the near future. Criteria 1, 2 and 3 are in place to put ethical considerations upfront. Criterion 4 is invoked to target
prevention at the smallest possible group of people in need of it, which serves both ethical and economic interests.

Selecting target groups: a prognostic index
Third, it should be more fully understood that 50% of the people with a full-blown depressive disorder recover from it within three months. Ideally, prevention should be targeted only at those who have a poor prognosis. For that some sort of prognostic index should be created. As yet, this appears to be a poorly researched field, but we are in dire need of such a prognostic index.

Prevention or health promotion?
Fourth, we need to strengthen universal prevention with the aim to promote awareness in the general public about depression. It helps when the general population has a better understanding of depressive disorder, knows what it is, how to recognise it, and what people can do to improve outcomes of its milder forms, and to whom to turn to for help should self-help fail. Raising public awareness can be seen as an important aim of universal prevention, but consideration should be given as how to promote this awareness, and in whose hands this task should be placed. Some re-thinking of the concepts we use when promoting preventive interventions may be required. Currently, we persuade people to make use of preventive interventions through media announcements drawing attention to interventions for people “who feel depressed”. It might be better to adopt a positive health promotion approach, and to advertise interventions that “promote mental fitness”, or help to “improve self-esteem” or prepare people to “reach old age successfully”. This should not be a mere re-labelling of concepts. More fundamentally, it could, and perhaps should, become a way of thinking about prevention, and re-designing preventive interventions, with less emphasis on risks and preventable illnesses, and greater emphasis on strength, self-esteem, mastery, successful coping styles, a more adequate understanding of one’s self. In this context, we refer to the work of some of our colleagues, most notably that of Ernst Bohlmeijer, who is pursuing this line of research in his PhD thesis. As should already be understood, a better “marketing” of prevention is not just a matter of reaching more people; first and foremost it is a matter of reaching the right people. In this thesis we described a method that enables us to identify these groups. Generally speaking it is the more vulnerable people, with some (prodromal) symptoms of depression, with low incomes, who have received only a few years of formal education, have small personal networks, have experienced adversities early in their life-course, suffer illnesses, disability and functional impairments, and have sub-average levels of mastery, who, in principle, could benefit most from depression prevention. However, it is very likely that it is precisely these groups that will be hardest to reach and may be the least receptive to health promotion. To reach out to these groups will require that some form of targeted marketing is used.
**Linking up with other health services**

Fifth, improved marketing should be accompanied by efforts to also improve the organisational position of prevention within the realm of public health. Prevention of depression cannot be carried out in isolation, and consideration should be given to integrating it as a first step in a multidisciplinary collaborative stepped-care approach, where prevention, cure and care of depressive disorder are connected like links in a chain. Beyond that, the position of depression prevention should be expanded to become one element in an integrated network (not just a chain-like sequence) of health service providers, such that early recognition and referral of people to the appropriate services can be optimised. This would, for example, entail that early recognition of people with mild forms of depression is improved not only in primary care, but also in home help, paediatric consultation bureaus, nursing homes etc., enabling health care providers to do a better job of referring patients. All this should be reflected in interdisciplinary guidelines for collaborative evidence-based medicine.

**Improving the financial position of prevention**

Sixth, high-quality interventions, a stronger emphasis on health promotion and advocacy can only be realised when the financial position of prevention of mental disorders is drastically improved. Currently, only 3% of the national budget for mental health care is allocated to prevention of mental disorders. This might prove to be an irrational distribution of funds in the long run. In addition, the position of prevention in the context of the new health insurance scheme in the Netherlands should be considered. Under the new scheme only treatments directly offered to individuals (as opposed to collectives) qualify for re-imbursement via health insurances, which renders universal prevention and some forms of selective prevention fairly anomalous in this financial jigsaw puzzle.

**Prevention and the internet**

Seventh, we have reason to believe that the accessibility of prevention can be improved by offering prevention more often as web-based self-help interventions. This approach will allow people to make use of highly structured, state-of-the-art, preventive interventions in the privacy of their own homes, at times they find convenient, and at a pace that suits them best. Given the large number of people (11 million out of a population of 16 million in the Netherlands) that have access to the internet, we can expect web-based interventions to reach very large population segments. However, not all population segments will benefit equally from web-based self-help interventions, and some may benefit more from treatment under the guidance of a professional. The idea is that low-cost web-based self-help interventions will also help to free up resources that can then be re-directed to more labour intensive face-to-face interventions for those in need of those interventions. Currently, Heleen Riper and other colleagues are involved in a number of
randomised trials to ascertain the cost-effectiveness of web-based self-help interventions, and we already see promising results.

5.2.5 Postscript

In Albrecht Dürer’s Melancholia (title plate) we can see an hourglass. Its lower section is half filled with sand. A certain amount of time has elapsed. More time will slip through the narrow passage of the present and shower away to the past. Albrecht Dürer’s moonlit world of melancholia is an unfinished world. Likewise, prevention of depression is an unfinished project. In both instances, we are caught up in a world still under construction. None of these worlds is well understood. We know hardly anything about the evolutionary benefit of depression (cf. Watson and Andrews, 2002). Why did it come about? Why is it, in part, genetically transmitted? Why did it survive over the vast sweep of time of the human evolution to become such a common phenomenon? Does it offer adaptational advantages? Must we believe that it prompts us to reflect on important issues in our lives, allows us to conserve energy while we decide how to move on? Is it a way to signal to others that we need their help? Is it a way to carve out a niche amidst the turbulent social competition for support, services and exchange of goods? We can speculate about it, but do not know – not for sure.

Even in the much smaller timeframes of our individual lives, we do not know how we can benefit from depression, nor how we can re-arrange our lives such that depression will not destroy but make whole and better. Did Sir Winston Churchill reach the heights of his fame and glory despite, or because of, his depression? We don’t know. What is the relationship between depression and creativity? Why did Allessandro Botticelli create his most angelic paintings during the darkest moment in the history of Florence? Why did many prominent writers suffer from depression? Among many, Ludwig (1995) identifies the following: Virginia Woolf, George Eliot, Sylvia Plath, Lord George Byron, Count Leo Tolstoy, William Faulkner, Edgar Allen Poe, Joseph Conrad, Graham Greene, the venerable poet Percy Bysshe Shelley, and poor John Keats who once wrote, “I am in that temper that if I were underwater, I would scarcely kick to come to the top” (cf. Gilmore, 2002, Gittings, 1968; Lowell, 1925; Ludwig, 1995). Do genius and creativity feed off mental turmoil? Is there a link between mood and muse in composers such as Sergei Rachmaninoff and Robert Schumann? Again, we must admit that we have not a clue.

We know very little about both worlds, of depression and its prevention, so what can we say about the impact of one on the other? But this fusion of both worlds lies at the heart of the matter. When Ernest Hemingway was treated for depression in 1960, he feared that the treatment would rob him of his artistic capacities (cf. Meyers, 1985). He threatened his doctor with suicide, and told him, “If I can not exist on my own terms, then existence is impossible… This is how I’ve
lived, and this is how I *must* live – or not live.” Would the medical and scientific world of prevention rob depression of its force to engage us with the depths of our lives? We do not know – and that should make us careful.
Part V
Prevention of depression: summary and general discussion

Chapter 5.3
Summary in Dutch
Depressiepreventie: samenvatting

Inleiding

Dit proefschrift bestaat uit vier delen. In elk deel wordt een onderzoeksvraag beantwoord. De vragen zijn:
1. Is depressiepreventie nodig?
2. Op welke groepen dient depressiepreventie zich te richten?
3. Is depressiepreventie effectief?
4. Is depressiepreventie kosteneffectief?

Deze vragen zijn niet opzienbarend, liggen eigenlijk nogal voor de hand, maar worden zelden of nooit in onderlinge samenhang beantwoord. In het navolgende bieden wij eerst begripsverduidelijking, beschouwen in het kort waarom depressiepreventie relevant is, en geven aan hoe het proefschrift afgebakend is. Daarna komen de respectievelijke hoofdstukken aan de orde, waaraan nog enkele verbindende tekstdelen zijn toegevoegd.

Depressie

Depressie wordt gekenmerkt door een sombere stemming (dysphoria) en interesseverlies (anhedonia). Het lusteloze gevoel is voor het grootste deel van de dag aanwezig, gedurende meerdere dagen, en minimaal gedurende twee weken, en verlamt elke motivatie. Daarnaast zijn er meerdere klachten met een verstorende invloed op iemands functioneren, zoals een onregelde slaap (insomnia of hypersomnia), energiegebrek (anemia), problemen met de concentratie, toename of verlies van eetlust, piekeren over de dood en suïcide (APA, 1987). Depressie kent een ongunstige prognose. Gemiddeld duurt een depressieve episode zes maanden (Kruijshaar et al, 2005), terwijl bij 20% de depressie langer duurt dan twee jaar (Spijker et al, 2002). De kans op een nieuwe episode na herstel is zeer aanzienlijk met 85% binnen vijf jaar (Mueller et al, 1999) al zijn er ook beduidend lagere schattingen (Fombonne et al, 2001). Mensen met een depressie hebben een sterferisico dat een factor 1,65 hoger ligt ten opzichte van mensen zonder depressie (Cuijpers en Smit, 2002).

De stoornis gaat niet alleen gepaard met aanzienlijke functionele beperkingen, depressie is bovendien hoogprevalent. In Nederland lijden jaarlijks ongeveer 737.000 volwassenen aan depressie (Meijer et al, in voorbereiding). De combinatie van ernstige functionele beperkingen op het individuele niveau en het grote aantal depressieve personen leidt op macroniveau tot een grote maatschappelijke ziektelast. In Nederland gaan door depressie jaarlijks 157.700 gezonde levensjaren verloren (ibid). Volgens de World Health Organization is depressie wereldwijd één van de voornaamste veroorzakers van niet-fatale ziektelast (WHO, 2000). Mensen met een depressie doen een groter beroep op de
Relevante
de behandeling van depressie kan zich goed meten met behandelsuccessen elders in
de geneeskunde (Beekman e.a., 2006). Toch zijn er twee goede redenen om naast
behandeling ook aan preventie te denken.

Ten eerste is de jaarlijkse instroom van mensen die voor het eerst een
depressieve episode meemaken omvangrijk. In Nederland gaat het om een groep
van 358.500 mensen. Op jaarbasis vormen de nieuwe gevallen 49% van de
prevalente groep (zie Hoofdstuk 1.1). Dat betekent dat curatie alleen gelijk staat aan
dweilen met de kraan open. Vanuit het oogpunt van de openbare gezondheidszorg
lijkt het daarom verstandig ook iets te doen aan de omvangrijke jaarlijkse instroom,
et dat vraagt om primaire preventie.

De tweede reden is dat behandeling ontoereikend is op het macroniveau van
de samenleving. Dit heeft enkele oorzaken. Niet iedereen met een depressie wordt
als zodanig herkend, niet alle herkende gevallen ontvangen een optimale
behandeling, en niet elke patiënt is even therapietrouw. Zo wordt de
maatschappelijke ziektebelast van depressie, uitgedrukt in het aantal jaren dat met een
beperking wordt doorgebracht, met slechts 10% à 20% teruggebracht door

Het lijkt daarom verstandig om dit aanzienlijke maatschappelijke probleem
niet alleen via behandeling aan te pakken, maar ook door preventie. De preventie
dient dan gericht te zijn op het reduceren van de jaarlijkse instroom van nieuwe
gevallen. Die vorm van preventie heet primaire preventie.

Depressiepreventie
In de preventieve psychiatrie worden drie vormen van primaire preventie
onderscheiden: universele, selectieve en geïndiceerde preventie (Mrazek en
Haggerty, 1994).

Universele preventie richt zich op collectieven, zoals scholieren, alle
patiënten in de huisartsgeneeskunde, alle bewoners van verzorgingshuizen,
ongeacht hun individuele risicostatus. Universele preventie kan verschillende
vormen aannemen zoals publieksvoorziening, nationale screeningsdagen, en kan
ook bestaan uit interventies gericht op gezondheidsbevordering zoals het verbeteren
van sociale vaardigheden bij scholieren en het versterken van opvoedings-
vaardigheden van ouders.

Selectieve preventie richt zich op risicogroepen zoals weduwen, personen
met chronische ziekten, kinderen van depressieve ouders. Risicogroepen zijn de
beoogde ontvangers van gerichte preventieve interventies.

Geïndiceerde preventie richt zich op mensen bij wie al enkele depressieve
symptomen aanwezig zijn, maar bij wie nog geen sprake is van een volledige
depressieve stoornis. Hier neemt preventie de vorm aan van vroegherkenning

gevolgd door het interventiëren in een vroeg ontwikkelingsstadium van de stoornis. Dat ter voorkoming van de gang van kwaad naar erger.

Naast deze vormen van preventie kan nog gedacht worden aan terugvalpreventie, maar dat zou ook beschouwd kunnen worden als een onderdeel van een goede behandeling, en kan in elk niet beschouwd worden als primaire preventie. Terugvalpreventie valt daarmee buiten het bestek van dit proefschrift. Tegen het einde van deze samenvatting gaan wij nader in op de relatie tussen preventie, behandeling en nazorg.

Afbakening

Dit proefschrift is beperkt tot Nederlandse volwassenen en ouderen, en gaat niet in op preventie van depressie bij kinderen en jongeren. Het is vooral toegesplitst op geïndiceerde depressiepreventie. Wij zijn daarbij geïnteresseerd hoe geïndiceerde preventie (bij mensen met al enkele depressieve klachten) verder toegesplitst kan worden op ultrahoogrisicogroepen. Dat zijn groepen die niet alleen al enkele depressieve klachten hebben, maar ook zijn booggesteld aan risicofactoren waarvan bekend is dat zij voorspellers zijn van depressie. Deze insteek werd ingegeven vanuit de overweging dat het ethisch niet gepast en economisch onhaalbaar is om iedereen met een (mogelijk) beginnende depressie een preventieve interventie aan te bieden. Bovendien denken we dat in zulke ultrahoogrisicogroepen depressiepreventie de beste kans heeft om effectief en doelmatig te zijn.

In dit proefschrift worden een viertal vragen beantwoord. Twee vragen worden echter niet in dit proefschrift beantwoord. Er wordt niet ingegaan op de therapietheoretische vraag hoe depressiepreventie werkt, en er wordt niet, althans nauwelijks, ingegaan op de medisch-ethische vraag onder welke condities depressiepreventie gepast en acceptabel is. Die ethische vraag heeft wel een rol gespeeld bij de gekozen insteek, en laat zich vangen met de formule “minder waar mogelijk, meer waar nodig”. In de Antithese (zie aldaar) preludeerden we hier op. In de Discussie keren wij naar dit punt terug.

De door ons gekozen epidemiologische en gezondheidseconomische insteek, die gericht is op risicofactoren, corresponderende hoogrisicogroepen en kostenbeheersing, kent een alternatief. Er had ook gekozen kunnen worden voor gezondheidsbevordering met een accent op beschermende factoren, competentie-versterking, zelfredzaamheid, groei en positieve psychologie. Die lijn wordt gevolgd in het proefschrift van Ernst Bohlmeijer, en het was voor mij een genoegen om met hem aan die lijn mee te werken. Hiermee is meteen gezegd dat het voorliggende proefschrift beperkt is tot slechts één variant van preventie. Andere varianten blijven daarmee buiten beschouwing.

Een laatste opmerking over de beperkingen van het proefschrift. De vier onderzoeksvragen worden slechts gedeeltelijk beantwoord. Dat komt omdat dit werk in vordering is. Meer werk moet gedaan worden en meer werk wordt verricht. Het proefschrift is daarmee een momentopname van een langer traject, en wil laten zien hoe wij met deze materie omgaan.
Deel 1: Is depressiepreventie nodig?

Rationale voor depressiepreventie (Hoofdstuk 1.1)

Depressie is een ernstige stoornis. Om een vergelijking te maken: het gaat om een ziekteelast die net zo invaliderend is als bijna blindheid (Ruwaard en Kramers, 1997).

Individueel lijden, hoe erg ook, is echter niet meteen aanleiding om een ziekte een volksgezondheidsprobleem te noemen. Daarvoor is een tweede factor nodig. Pas wanneer individueel lijden vermenigvuldigd kan worden met een groot aantal mensen met de stoornis, kan een stoornis de status krijgen van een volksgezondheidsprobleem. Dit brengt ons op de tweede karakteristiek van depressie: de hoge prevalentie. Beide factoren dragen er gezamenlijk toe bij dat depressie één van de ziekten is met de meest omvangrijke maatschappelijke ziekteelast, wereldwijd (WHO, 2000; Ustün, 2004).

Maar hiermee wordt alleen gezegd dat depressie een ernstig probleem is, nog niet waarom de preventie ervan nodig is. Vanuit het oogpunt van de openbare gezondheidszorg zijn er twee redenen om niet alleen aan behandeling, maar ook aan preventie van depressie te denken.

Ten eerste wordt de behoefte aan preventie onderstreept door het feit dat de instroom van nieuwe gevallen van depressie op jaarbasis groot is. In Nederland bestaat die jaarlijkse instroom uit 287.000 mensen die in dat jaar voor het eerst van hun leven met een depressie geconfronteerd worden. Zij vormen 47% van de prevalente groep van 738.000 personen (Meijer et al, in voorbereiding). De massale instroom van nieuwe gevallen is kenmerkend voor depressie, en ziet er bijvoorbeeld voor schizofrenie er heel anders uit: daar zou men bij wijze van spreken met een zaklamp naar nieuwe gevallen in de bevolking moeten zoeken. Gelet op de massale instroom is depressie een volksgezondheidsprobleem dat zowel via de lengte als de breedte aangepakt moet worden: preventie om de massale instroom in te dammen en curatie om de ziekte duur te bekorten en herstel te bevorderen. Wij zwenken dan nog over de eveneens omvangrijke groep die na een eerdere depressie op een later tijdstip opnieuw een depressieve episode meemaakt: circa 30% van de mannen en 40% van de vrouwen lijden aan één of meer depressieve episoden gedurende hun leven, en gemiddeld zijn er 7,3 depressieve episoden (Kruijshaar et al, 2005). Niet alleen mensen die voor het eerst van hun leven een depressie krijgen, maar ook de mensen bij wie een herhaling dreigt, zouden baat kunnen hebben bij preventieve interventies.

De tweede reden is het relatief beperkte effect van curatie op het niveau van de populatie. Hoewel depressies succesvol behandeld kunnen worden, hebben twee studies (Andrews et al, 2004; Chisholm et al, 2004) aangetoond dat de maatschappelijke ziekteelast slechts zeer gedeeltelijk weggenomen kan worden door curatieve behandeldvormen: tussen de 10% en de 20% van de 157.700 gezonde levensjaren die jaarlijks door depressie verloren gaan in de Nederlandse volwassen bevolking (cf. Meijer et al, in voorbereiding). Dit heeft tot gevolg dat er een
gapende kloof blijft bestaan tussen wat curatieve interventies kunnen bieden en de behoefte van velen.

Om beide redenen zou naast curatie ook preventie een plek verdienen in het zorgbestel – vooropgesteld dat we weten waar we moeten beginnen, dat preventie effectief is, ethisch aanvaardbaar, en dat investeren in preventie economisch gezien doelmatig is.

**Kosten van ziekten (Hoofdstuk 1.2)**
Het laatste punt brengt ons bij Hoofdstuk 1.2, een kosten-van-ziektenstudie over psychische stoornissen die relatief veel voorkomen. In deze studie toonden wij aan dat depressie per persoon per jaar € 2300 aan extra kosten genereert. Dit geeft meteen een idee van de potentiële besparingen die depressiepreventie kan opleveren wanneer het ontstaan van een nieuwe depressieve episode bij iemand vermeden kan worden.

Op het niveau van de Nederlandse populatie in de leeftijdsgroep van 18 tot 65 jaar lopen deze kosten, mede door toedoen van de hoge prevalentie van depressie (5,8% van de bevolking op jaarbasis), op tot het forse bedrag van € 1.300.000.000 – een bedrag dat elk jaar terugkomt. Het ligt in de rede dat een reductie in de incidentie gepaard gaat met een navenante reductie in de economische kosten van deze stoornis.

Hoofdstuk 2.1 bevat dan ook een tweevoudige boodschap. Ten eerste zou het vanuit een puur economisch perspectief wel eens de moeite waard kunnen zijn om interventies gericht op de preventie van depressie te ontwikkelen, te evalueren, en te implementeren. Ten tweede is er een goede kans dat depressiepreventie kosteneffectief zou kunnen zijn. Het betreffende manuscript werd ter publicatie aangeboden aan het *Journal of Mental Health Policy and Economics*.

**Deel 2: Waar te beginnen?**

**Achtergrond**
Tegen de achtergrond die geschetst werd in de hoofdstukken 1.1 en 1.2 wordt het zinnig om een helder antwoord te krijgen op de vraag of we weten waar we met depressiepreventie moeten beginnen.

Deze vraag is relevant omdat er veel mensen blootgesteld zijn aan de ene of andere risicofactor en daarom in aanmerking komen voor selectieve preventie. Ook het aantal mensen dat wel enkele depressieve klachten heeft, en daarom in aanmerking komt voor geïndiceerde preventie kan, afhankelijk van de precieze definitie van “een aantal depressieve klachten”, al snel omvangrijk worden. Het wordt dan logistiek en economische onuitvoerbaar om al die mensen een preventieve interventie aan te bieden. Een medisch-ethische argument weegt ook mee. Het is moeilijk te rechtvaardigen om mensen met een verhoogd risico op depressie, maar die zich daarvan niet bewust zijn, een interventie aan te bieden. Ten
eerste is niet gezegd dat zij zonder interventie een depressie zullen ontwikkelen. Ten tweede is niet gezegd dat een preventieve interventie garantië biedt op het wegblijven van een depressie. Ondertussen worden deze mensen wel belast met onzeker makende kennis. Zo beschouwden medisch-ethische bezwaren opgeworven tegen preventie als deel bij geïndiceerde preventie omdat het dan gaat om groepen die al depressieve klachten ervaren. Deze mensen zullen niet verrast zijn om te horen dat ze depressieve klachten hebben; bij hen wordt dus geen nieuwe en belastende kennis aangedragen. Zij zullen bovendien last hebben van die klachten, en zoeken daarvoor wellicht ook hulp. In elk geval: de vraag aan wie preventieve interventies aangeboden kan worden, vraagt om zorgvuldige praktische, economische en ethische afwegingen.

De oplossing hiervoor zochten wij in het identificeren van groepen met een berekenbaar ultrahoog risico op het ontwikkelen van depressie. Het idee daarbij was dat zulke ultrahoogrisicogroepen numeriek klein zijn en toch verantwoordelijk voor het leeuwendeel van het aantal nieuwe depressiegevallen in de bevolking. In ultrahoogrisicogroepen zou depressiepreventie bovendien de beste kans hebben om kosteneffectief te worden.

Uitgedrukt in epidemiologisch en statistische termen waren wij op zoek naar groepen met een hoge incidentie (incidence rate, IR) van depressie, hoger dan de IR in groepen die niet aan bepaalde risicofactoren waren blootgesteld. Wij zochten, anders gezegd, groepen met een hoge incidence rate ratio (IRR). Gelijkstijdig wilden wij groepen samenstellen die zo klein mogelijk waren, ofwel groepen die gekenmerkt worden door een kleine blootstellingfractie (exposure rate, ER). Zowel de IRR als de ER kunnen in één statistiek gecombineerd worden, de populatie attributieve fractie (attributable fraction, AF). De AF heeft een handige gebruikswoorde vanuit het oogpunt van preventie: het leert ons met hoeveel procentpunten de incidentie van depressie in de bevolking gereduceerd wordt wanneer de ongunstige invloed van de betreffende risicofactor geheel geblokkeerd wordt door toedoen van een preventieve interventie. De AF is daarmee te beschouwen als een graadmeter voor de potentieel behaalbare gezondheidswinst in de populatie onder een hypothetisch preventiescenario.

Natuurlijk is het niet realistisch om aan te nemen dat preventie voor de volle 100% succesvol zou zijn in het blokkeren van de ongunstige invloed van een risicofactor. Het is daarom begrepen dat de AF een bovengrens definiëert van de theoretisch behaalbare gezondheidswinst. Dat neemt niet weg dat de AF een handige maat blijft voor het samenstellen van een hiërarchie van risicofactoren op basis waarvan we tot een optimale doelgroepselectie voor preventie kunnen komen. De volgende drie hoofdstukken zijn op deze redenering gebaseerd.

**Identificatie van groepen voor depressiepreventie (Hoofdstuk 2.1)**

 Deze benadering werd voor het eerst uitgeprobeerd in Hoofdstuk 2.1. We maakten gebruik van een grote epidemiologische cohortstudie, Nemesis, en berekenen de exposure rates (ER’s), de incidence rate ratios (IRR’s), en de population
attributable fractions (AF’s) voor een groep risicofactoren die wel eens een rol zouden kunnen spelen in het ontstaan van een depressie bij 18 tot 65 jarigen. Voor elke risicofactor berekende we ook de number-needed-to-be-treated (NNT), een index voor de potentiële efficiëntie van een interventie wanneer die succesvol zou zijn in het volledig indammen van de nadelige gevolgen van de betreffende risicofactor.

Met gebruikmaking van deze methode begonnen we kleine hoogrisico-groepen in beeld te krijgen waar depressiepreventie tegen de geringste inspanning, en dus tegen de laagste kosten, de grootst mogelijke gezondheidswinst zou genereren. Het waren vooral kwetsbare mensen met al enkele symptomen van depressie, een gering zelfvertrouwen, een hoge mate van neuroticisme en met een voorgeschiedenis van emotionele verwaarlozing, mishandeling, angststoornissen, overmatig alcoholgebruik of chronische lichamelijke ziekten bij wie het risico op depressie hoog was, en bij wie preventie het meest beloofd op te leveren.

Deze studie werd gepubliceerd in het Journal of Affective Disorders (Smit et al, 2004), en eindigde op een noot dat in toekomstig onderzoek de indexen voor gezondheidswinst (IRR, AF) en de indexen voor de inspanning om die gezondheidswinst te genereren (ER, NNT) niet voor afzonderlijk risicofactoren berekend moesten worden, maar voor groepen van risicofactoren – en wel zo dat de indexen verder in de gewenste richting geoptimaliseerd worden.

Kansen voor kosteneffectieve depressiepreventie (Hoofdstuk 2.2)
Het was precies deze multivariate aanpak die we in Hoofdstuk 2.2 volgden. Ditmaal evalueerden wij het effect van samengestelde blootstellingen van meerdere risicofactoren tegelijkertijd. De betreffende analyses voerden wij uit op de data van een andere grote epidemiologische studie, de Longitudinal Aging Study Amsterdam (LASA), een cohort van 2200 mensen van 55 tot 85 jaar. Ditmaal waren wij veel beter instaat om de numeriek kleine, ultra hoogrisicogroepen te identificeren waar preventie de beste kansen heeft om efficiënt en kosteneffectief te worden. Figuur 1 illustreert de gevolgde werkwijze.

Wij begonnen met de groep die enkele depressieve symptomen had, maar dat is een grote groep van 40% (ER) van de bevolking. Wanneer die mensen ook gekenmerkt worden door aanwezigheid van een tweede risicofactor (functionele beperkingen) daalt de ER tot 21%. De geselecteerde doelgroep wordt nu dus aanzienlijk kleiner, en wordt daarmee in logistiek opzicht beter behapbaar. Door meer risicofactoren aan het profiel toe te voegen, daalt de ER verder tot 8%. Gelijktijdig wij hoe ook de NNT, de maat voor efficiëntie, beter wordt: van een omslachtige 17 daalt die uiteindelijk tot onder de 5. Dat betekent dat er slechts 5 personen met een preventieve interventie benaderd moeten worden om bij één van hen een depressie te vermijden – althans wanneer de interventie volledig succesvol zou zijn in het ingedam houden van de nadelige effecten van de betreffende risicofactoren. Met steeds kleinere groepen (ER) wordt het steeds moeilijker om op populatie niveau gezondheidswinst te behalen. Wij zien daarom de attributieve
De attributieve fractie (AF) telkens wat zakken wanneer een nieuwe risicofactor aan het profiel wordt toegevoegd. Het komt er dus op aan juist die combinatie van risicofactoren te vinden die gezamenlijk leiden tot zo laag mogelijke waarden voor inspanning (ER en NNT), terwijl gelijktijdig de behaalbare gezondheidswinst (AF) zo hoog mogelijk blijft. Anders gezegd, de gezondheidswinst willen we maximaliseren en de inspanning minimaliseren. Dat is kosteneffectief.

**Figuur 1** Effect van het toevoegen van risicofactoren op de attributieve fractie AF (in %), exposure rate, ER (in %), en de number-needed-to-be-treated, NNT.

Langs deze lijnen voerden wij ook een *ante hoc* economische evaluatie uit, waarbij we aannemelijk maakten dat preventie van depressie in de latere levensloop waarschijnlijk kosteneffectief kan zijn wanneer die gericht wordt op de geselecteerde groepen. De betreffende studie werd gepubliceerd in de *Archives of General Psychiatry* (Smit et al, 2006a), en eindigde op een noot dat deze nieuwe methodologie validering behoefde in een andere studie.

**Depressiepreventie: waar te beginnen? (Hoofdstuk 2.3)**

De betreffende valideringstudie wordt in Hoofdstuk 2.3 beschreven. Er werd opnieuw van een psychiatrische cohortstudie gebruik gemaakt, ditmaal de *Amsterdam Study of the Elderly* (de Amstelstudie). Opgemerkt moet worden dat in deze studie depressie op een andere manier werd vastgesteld en dat ook de risicofactoren op een andere manier gedefinieerd waren. Wij maakten bovendien gebruik van een andere methode om hoogrisicogroepen te identificeren. Deze methode staat bekend onder de naam *Classification and Regression Tree* (CART) analyse. De *tree* in de naam verwijst naar boomachtige diagrammen die ontstaan wanneer systematisch wordt nagegaan wat het toegevoegd effect is van een volgende risicofactor in termen van veranderingen in de ER, IRR, AF en NNT. De
procedure van stapsgewijs nieuwe risicofactoren toevoegen helpt om na te gaan welke combinatie van risicofactoren leidt tot optimale waarden voor gezondheids-
winst, efficiëntie en kosteneffectiviteit, terwijl gelijktijdig de inspanningen en kosten zo laag mogelijk worden gehouden.

Het was interessant om te zien dat de CART methode vrijwel dezelfde resultaten liet zien als de eerdere studie (Hoofdstuk 2.2). Geïndiceerde preventie van depressie richt zich het best op mensen die al enige depressieve klachten hebben, die bovendien chronische lichamelijke ziekten hebben, beperkingen in het functioneren ervaren, die in een klein sociaal netwerk leven, en vrouw zijn. Er was één discrepantie. In de Amstelstudie zagen wij dat weduwschap een belangrijke voorspeller was voor het later ontstaan van depressie. Dit hadden wij in onze analyses van de LASA-data niet gezien. Hier moet echter opgemerkt worden dat in LASA weduwschap gedefinieerd was als “ooit verweduwd” (zonder verwijzing naar een tijdsperiode), terwijl het in de Amstelstudie gedefinieerd was als “recent verweduwd” (in de laatste 6 maanden). Het kan dus zijn dat het verschil in de uitkomsten samenhangt met het verschil in de definitie van weduwschap.

Op basis van de congruentie tussen de beide studies zijn we geneigd te concluderen dat we inderdaad beginnen te zien bij welke groepen we moeten beginnen met depressiepreventie. Geïndiceerde preventie neemt het best een aanvang bij mensen die al enkele symptomen van depressie hebben en kan dan meer toegepast worden door verder te selecteren op aanwezigheid van de voornoemde risicofactoren. De studie laat verder zien dat selectieve preventie een alternatief vormt wanneer het gericht wordt op mensen die recentelijk verweduwd zijn, een lichamelijke ziekte hebben, functionele beperkingen ervaren en vrouw zijn. De laatste aanpak biedt een additioneel voordeel: de betreffende doelgroep kan eenvoudig geïdentificeerd worden, namelijk aan de hand van een korte checklist bestaande uit de goed herkenbare risicofactoren. Bij geïndiceerde preventie ligt dat wel moeilijker omdat, bijvoorbeeld met hulp van een screener, nagegaan moet worden of iemand depressieve klachten heeft. Deze studie van Robert Schoevers, mijzelf en collega’s is in druk bij de American Journal of Psychiatry.

Deel 3: Is depressiepreventie effectief?

Achtergrond
In het derde deel van dit proefschrift gaan wij in op de vraag hoe effectief preventie is wanneer het er op aan komt om het ontstaan van de stoornis te voorkomen. Om één en ander in de juiste context te plaatsen, moeten we opmerken dat preventieve interventies in letterlijk honderdtallen gerandomiseerde studies werden geëvalueerd, maar de klinische eindterm in die studies was bijna altijd symptoomreductie en slechts zeer zelden reductie van de incidentie. Daarom was er nagenoeg niets bekend over de effectiviteit van primaire preventie. De reden voor deze ogenschijnlijke blunder wordt uitgelegd in Pim Cuijpers’ artikel dat in 2003 in de
Effectiviteit van depressiepreventie: trial (Hoofdstuk 3.1)

Deze strategie werd gedeeltelijk in hoofdstuk 3.1 gevolgd. De evaluatiestudie werd ontworpen als een gerandomiseerd experiment in twee parallelle groepen. De deelnemers aan het onderzoek waren huisartspatiënten die voor het onderzoek gerekruiteerd werden terwijl ze in de wachtkamer van de huisartspraktijk op hun afspraak met de arts wachten. Zij konden aan de studie deelnemen wanneer zij aan een aantal criteria beantwoordden. De belangrijkste criteria waren dat zij ten minste één kernsymptoom van depressie hadden (sombere stemming of interesseverlies) en verder nog één, twee of drie bijkomende symptomen van depressie, zonder echter te beantwoorden aan de diagnostische criteria voor een depressieve stoornis in de zin van de DSM-IV en zoals vastgesteld met de CIDI. Exclusiecriteria waren aanwezigheid van de overige stemmingsstoornissen zoals bipolaire stoornis en dysthyme stoornis. Op basis van de wetenschappelijke literatuur verwachten wij dat de 1-jaarsincidentie in deze populatie ongeveer 20% zou zijn, in plaats van het veel lagere percentage van 2,7% zoals dat in de algemene populatie wordt aangetroffen.

In deze hoogrisicogroep wilden wij de hypothese toetsen dat een zelfhulpinterventie om zelf beter om te gaan met de depressieve klachten zou helpen om de incidentie met 30% te reduceren ten opzichte van de controleconditie waarin alleen gangbare huisartszorg werd geboden. Om de studie voldoende onderscheidend vermogen (power) te geven, moesten 200 deelnemers per conditie aan de studie meedoen. Met dat aantal zou ook enige uitval na 12 maanden gecompenseerd kunnen worden. De centrale klinische eindterm van de studie was depressiestatus volgens de DSM-IV en gemeten met de CIDI.

De analyses werden uitgevoerd met in achtneming van het intention-to-treat principe. Ontbrekende waarden op de nameting werden daartoe geïmputeerd. In de analyse werd ook rekening gehouden met het feit dat meerdere huisartspatiënten geworven waren in eenzelfde huisartspraktijk, wat een clustering in de data veroorzaakt. In totaal werden 107 deelnemers naar de experimentele conditie (zelfhulp toegevoegd aan gangbare zorg) gerandomiseerd en 109 naar de controleconditie (alleen gangbare zorg).

De uitkomsten waren als volgt. In de controlegroep was de 1-jaarsincidentie 18% tegen 12% in de experimentele groep. Dat staat gelijk aan een relatieve risicoreductie van 33%. De unidirectionele hypothese dat de experimentele conditie
inferieur is aan de controleconditie moest worpwen worden bij $P<0.05$ in een 1-zijdige toets. Dit ondersteunt het idee dat de combinatie van zelfhulp en gangbare zorg in preventief opzicht superieur is aan gangbare zorg alleen.

Dit was goed nieuws, maar er waren ook twee problemen in beeld gekomen: zowel de werving als de participatiegraad van de deelnemers aan de interventie vielen tegen. Dit reflecteert ongunstig op de mate waarin de interventie door de groep huisarts patiënten geaccepteerd wordt. Dit zou kunnen betekenen dat de beoogde deelnemers beter geselecteerd moeten worden, bijvoorbeeld aan de hand van de risicofactoren die in de hoofdstukken 2.1, 2.2 en 2.3 werden geïdentificeerd. Ook moet er beter gelet worden op de subjectief ervaren zorgbehoeften van de doelgroep. Dat laatste weten we omdat veel uitvallers uit de interventie zeiden geen behoefte te hebben aan de interventie. Het onderzoek werd in 2004 gepubliceerd in de *British Journal of Psychiatry* door Godelief Willemse en collega’s.

**Effectiviteit van depressiepreventie: meta-analyse (Hoofdstuk 3.2)**

Onze preventiatrial (Hoofdstuk 3.1) vertegenwoordigde één van de in totaal zeven studies waarin de effectiviteit van psychologische preventieve interventies werd geëvalueerd op de reductie van de incidentie van depressie. In Hoofdstuk 3.2 werden de uitkomsten van deze zeven studies meta-analytisch samengenomen.

In de meta-analyse werden twee uitkomstmaten beschouwd. De eerste was de mate waarin de interventies er in slaagden om het depressieve klachten niveau naar beneden te brengen. De corresponderende effectgrootte, Cohen’s $d$, is een maat die aangeeft hoeveel standaardeenheden de experimentele groep verwijderde is van de controlegroep in termen van de behaalde klachtenreductie. De tweede uitkomstmaat is de relatieve risicoreductie van de experimentele groep ten opzichte van de controlegroep in termen van de incidentie van de depressieve stoornis. Dit werd vastgesteld met een maat die, technisch gesproken, de *incidence (density) rate ratio (IRR)* heet, een maat die rekening houdt met ongelijke follow-up-tijden.

De uitkomsten van de meta-analyse zijn als volgt. Aan de zeven studies deden in totaal 700 deelnemers mee, waarvan 343 in de experimentele (interventie) condities en 357 in de controlecondities. In zes studies bestond de controleconditie uit gangbare zorg, en in een zevende studie uit een wachtlijst waar de deelnemers verder gewoon toegang hadden tot gangbare zorg. Omdat er geen apprecieerbare heterogeniteit werd waargenomen tussen de studies, en omdat de uitkomsten van het random effectmodel en het fixed effectmodel zeer vergelijkbaar waren, werden alleen de uitkomsten van het meer eenvoudige fixed effectmodel gepresenteerd. Op de korte termijn, dat wil zeggen direct aan het einde van de interventie, werd een symptoomreductie gerealiseerd van gemiddeld 0,42 standaardeenheden in het voordeel van de interventies. Deze effectgrootte kan beschouwd worden als een middelgroot effect in vergelijking tot wat elders bij psychologische en gedragstherapeutische interventies aan effecten wordt gegenereerd (Lispey en Wilson, 1993). Verder werd geobserveerd dat de effectgrootte met het verstrijken van de tijd kleiner wordt: na één jaar was het gekrompen tot een klein effect van
0,16 standaardeenheden. Dit grensde slechts aan statistische significantie met 
P=0,08 in een 2-zijdige toets. De tweede uitkomstmaat, de relatieve risicoreductie, 
liet zien dat psychologische interventies helpen om de incidentie van depressie met 
30% te reduceren, wat eveneens slechts marginaal significant was met \( p=0,07 \) in 
een 2-zijdige toets.

In klinische termen zijn de gevonden effectgroottes in overeenstemming met 
wat verwacht mag worden van lichte psychologische interventies, maar het feit dat 
de effecten met de tijd kleiner worden is een belangrijke waarneming. De afname 
van het effect zou er op kunnen wijzen dat na en jaar een herhalingsessie van de 
interventie van belang zou kunnen zijn met het oog op effectbehoud. Het leert ons 
ook dat preventie van depressie misschien niet vergeleken mag worden met een 
eden malige inenting, maar dat een vergelijking met veiligheids gordels een betere is:
de interventie moet elke keer opnieuw uitgevoerd worden om er een preventief 
effect van te mogen verwachten. Bij een sterk remitterende stoornis zoals depressie 
valt dat ook te verwachten.

Vanuit statistisch oogpunt zijn er geen klemmende redenen om vergaande 
conclusies te verbinden aan de net niet significante P-waarden. In een 1-zijdige toets 
zouden deze namelijk significant geweest zijn, terwijl een 1-zijdige toets 
gerechtvaardigd kan worden door uit te gaan van de maatschappelijk relevante 
unidirectionele hypothese dat van gangbare zorg \( \text{plus} \) een preventieve interventie 
meer effect verwacht mag worden dan van gangbare zorg alleen. Maar als we geen 
voorbarige conclusies willen trekken, moeten we ons op het standpunt stellen dat er 
“aanwijzingen” zijn dat psychologische interventies succesvol kunnen zijn in 
termen van depressiepreventie. Daarbij is het duidelijk dat meer effectstudies 
gewenst zijn: wij zouden namelijk graag willen weten welke groepen, in welke 
settings baat hebben bij welke interventies, en voor hoe lang.

Deel 4: Is het kosteneffectief?

Kosteneffectiviteit van depressiepreventie: trial (Hoofdstuk 4.1)

Wij voerden een kosteneffectiviteitanalyse uit op de data van de trial die eerder 
besproken werd in Hoofdstuk 3.1. De economische evaluatie werd uitgevoerd 
vanuit maatschappelijk perspectief. Daartoe werden de volgende kostenposten 
beschouwd: (1) de direct medische kosten die namelijk ontstaan wanneer mensen 
gebruik maken van zorgvoorzieningen, met inbegrip van de interventie en 
geneesmiddelengebruik; (2) de directe niet-medische kosten zoals de kosten die 
patiënten zelf betalen voor hun vervoer en parkeergeld om zorg te verkrijgen; en 
(3) de indirecte niet-medische kosten die ontstaan door productieverliezen. Deze 
productieverliezen kunnen ontstaan in betaald werk door verzuim, maar ook 
wanneer iemand zich niet goed voelt, toch naar zijn werk gaat, en vervolgens 
minder efficiënt werkt. Ten slotte kunnen er productieverliezen ontstaan in de
huishoudelijke sfeer. De kosten werden berekend voor de tijdspanne van één jaar, en werden daarom niet gedisconteerd of voor inflatie gecorrigeerd.

De resultaten van de economische evaluatie zijn als volgt. De zelfhulp-interventie plus gangbare zorg was succesvoller in het voorkomen van nieuwe gevallen van depressie dan alleen gangbare zorg (zie Hoofdstuk 3.1). De integrale kosten waren gemiddeld € 6.766 per patiënt per jaar in de experimentele groep, wat gunstig afsteekt tegen de gemiddeld hogere kosten in de controlegroep van € 8.614. De kosteneffectiviteitratio toonde een bescheiden kostenbesparing aan van € 288,75 per gewonnen depressievrij levensjaar. Deze uitkomst was echter omringd door statistische onzekerheid.

Een probabilistische aanpak, gebaseerd op 2.500 bootstrap replicaties, liet zien dat toegevoegde zelfhulp een kans heeft van 70% om vanuit het oogpunt van kosteneffectiviteit acceptabel te zijn, zelfs in het economisch conservatieve scenario dat er geen bereidheid is om voor een gewonnen depressievrij levensjaar te betalen. Wanneer die bereidheid bijvoorbeeld een plafond heeft van € 30.000 per gewonnen depressievrij levensjaar, dan is kans dat de interventie vanuit het oogpunt van kosteneffectiviteit acceptabel is, toegenomen tot 82%.

Een gevoeligheidsanalyse toonde aan dat productieverliezen de belangrijkste kostencomponent vormen. Wanneer deze kostenpost buiten beschouwing wordt gelaten, en daarmee de economische evaluatie beperkt wordt tot een medisch perspectief, dan kan aangetoond worden dat de toegevoegde zelfhulpinterventie nog altijd een kans heeft van 47% om vanuit kosten-effectiviteitoogpunt acceptabel te zijn bij een willingness to pay gelijk aan € 0,00, oplopend tot circa 70% bij een plafond van € 30.000 per gewonnen depressievrij levensjaar.

**Postscript bij Hoofdstuk 4.1**

De economische evaluatie laat zien dat het aanbieden van additionele zelfhulp-interventie acceptabel is vanuit het oogpunt van kosten-effectiviteit: gemiddeld wordt meer gezondheidswinst gegenereerd en daarbij worden bovendien kleine economische besparingen gerealiseerd – althans de kans daarop is groot. Toch blijft het lage deelnamepercentage van huisartspatiënten aan de interventie een belangrijk punt van aandacht (zie Hoofdstuk 3.1). Sinds het verschijnen van het artikel zijn twee andere punten ook van belang gebleken.

Ten eerste liet een andere kosteneffectiviteitsstudie (Lynch et al, 2005) eveneens zien dat depressiepreventie kosteneffectief is. De beide economische studies ondersteunen elkaar dus.

Ten tweede, moeten we in het licht van eigen vervolgonderzoek (Cuijpers et al, in voorbereiding) daaraan toevoegen dat er een verlies aan effectbehoud optreedt na twee jaar. Dit effectverlies betekent waarschijnlijk dat we er alleen in slagen het ontstaan van depressie uit te stellen, maar niet te vermijden. Nu betekent uitstel van het ontstaan van depressie ook gezondheidswinst, maar het leert ons andermaal dat depressiepreventie niet vergeleken kan worden met een einmalige inenting. In een remitterende stoornis zoals depressie is zoiets waarschijnlijk sowieso niet te
verwachten, en lijkt het waarschijnlijker dat depressiepreventie bij voortdurend zal vragen om een afgewogen en telkens terugkerende vorm van ziektemanagement. In de algemene nabeschouwing komen we hierop terug.

*Kosten van risicofactoren: een andere benadering (Hoofdstuk 4.2)*

In het laatste hoofdstuk wordt een aantal kernbegrippen uit de vorige hoofdstukken opnieuw gebruikt, maar deze worden ditmaal in een ander kader bijeengebracht. Opnieuw valt het accent op kosten en risicofactoren, maar nu berekenen wij de kosten van risicofactoren. De risicofactoren zijn ouderlijke psychopathologie en blootstelling aan jeugdtrauma. Het zijn risicofactoren die vaak aan de grondslag liggen van psychische stoornissen zoals depressie. Met deze aanpak beogen wij dat er gezondheids-economische uitspraken gedaan kunnen worden over de directe oorzaken van psychische stoornissen. Wij wilden daarbij bovendien evalueren hoe een beschermende factor de nadelige effecten modificeert van de risicofactoren. De beschermende factor is *mastery*, het gevoel dat iemand greep heeft op zijn eigen leven. Het idee daarbij was dat een meer dan gemiddeld hoge mastery helpt om de nadelige gevolgen van de risicofactoren in te perken. Als dit inderdaad het geval is, dan zou dat er voor pleiten om mastery te versterken in mensen die blootgesteld zijn aan de betreffende risicofactoren.

Voor deze studie maakten wij opnieuw gebruik van Nemesis, de epidemiologische cohortstudie onder 5.618 volwassenen van 18 – 65 jaar (vergelijk hoofdstukken 1.2 en 2.1). Blootstelling aan de risicofactoren werden op de voor meting vastgesteld en betrof depressie en angst in de ouders van de respondenten, en blootstelling aan emotionele verwaarlozing, of geestelijke, lichamelijke en seksuele mishandeling toen de respondent 16 jaar of jonger was. Om causale verbanden met enige zekerheid te kunnen aantonen, werd blootstelling aan de risicofactoren op de eerste meting vastgesteld en de kosten van die blootstellingen een jaar later. De kosten omvatten integrale behandelkosten, kosten die door patiënten gemaakt worden in het kader van het verkrijgen van zorg, en de kosten van productieverliezen door ziekteverzuim in betaalde arbeid en in de huishoudelijke sfeer (zie ook Hoofdstuk 2.1).

De resultaten zijn als volgt. De geselecteerde risicofactoren zijn geassocieerd met extra kosten in de orde van de 1.200 ~ 3.500 US$ per persoon per jaar. In de bevolking is het percentage dat werd blootgesteld aan de risicofactoren groot (tussen de 7% en de 25.1%) en daarom zijn de kosten op bevolkingsniveau zeer omvangrijk: tussen de 170 ~ 636 miljoen US$ per 1 miljoen volwassen. Het effect van lage versus hoge mastery was in de voorspelde richting met lagere kosten van de risicofactoren bij hoge mastery en hogere kosten bij lagere mastery.

De resultaten laten zien dat de meerkosten van de risicofactoren in de populatie zeer omvangrijk zijn. Deze kosten dienen zich jaarlijks aan, elk jaar opnieuw. Dit suggereert dat preventieve interventies gericht op deze risicofactoren wel eens kosteneffectief kunnen worden. Het is echter niet direct duidelijk hoe preventie er in moet slagen de nadelige effecten van deze risicofactoren beperkt te
houden, maar het versterken van mastery in blootgestelde mensen zou één optie kunnen zijn. In aanwezigheid van een hoge mastery waren de kosten meestal eenderde van de gemiddelde kosten. Dat wijst, in aanzet, op een schokdempende werking van mastery. Juist op dit punt moeten we echter een belangrijke beperking van dit onderzoek noemen. We hadden graag willen evalueren hoe een verandering in mastery later in de tijd gevolgd wordt door een verandering in kosten, maar zo’n analyse kon niet op de beschikbare data worden uitgevoerd. We verwachten zulke analyses in de nabije toekomst te kunnen uitvoeren omdat we dit soort data nu aan het verzamelen zijn in de context van enkele trials. Dat toekomstige onderzoek kan nu geleid worden door de werkhypothese dat met gerichte preventieve interventies mastery versterkt kan worden en dat daarmee gunstige veranderingen geïndiceerd kunnen worden in risicostatus en ziektegerelateerde kosten.

Nabeschouwing

Conceptueel kader

Preventieve interventies komen in soorten en maten, en richten zich op diverse doelgroepen. Wij schetsen daarom een conceptueel kader om de onderscheiden vormen van preventie te positioneren en de relatie tussen preventie, behandeling en nazorg te verduidelijken (zie Figuur 2).

Figuur 2 Gezondheidscondities, transities, preventie en behandeling

Het algemene idee is dat mensen zich in een bepaalde gezondheidsconditie bevinden en van daaruit een transitie kunnen doormaken naar een andere gezondheidsconditie. Vanuit de openbare volksgezondheid bekeken is het uiteraard
wenselijk om gunstige overgangen te bevorderen (gezondheidsbevordering) en de kans op ongunstige transities te reduceren (preventie).

Universele preventie richt zich op het algemeen publiek, ongeacht risicostatus en eventuele aanwezigheid van eerste symptomen. Voorbeelden van universele preventie zijn publieksvoorlichting, programma’s bij scholieren gericht op gezondheidsbevordering, en screeningsprogramma’s in de eerstelijnszorg.

Selectieve preventie richt zich op groepen die gekenmerkt worden door kwetsbaarheidsfactoren (zoals een lage opleiding, wonen in een achterstandswijk, een hoog neuroticisme, een lage zelfwaardering) en risicofactoren (zoals kindermishandeling, ouders met psychische problemen, ernstige lichamelijke ziekten, verlies van een partner). Dit zijn factoren waarvan bekend is dat zij de kans op het ontstaan van een depressie vergroten.

Geïndiceerde preventie richt zich op mensen die al vroege tekenen hebben (klachten, symptomen) van een naderende stoornis. Deze vorm van preventie bestaat uit vroegeherkening en het vroegtijdig interveniëren in het ziektebeloop. Het heeft als doel de gang van kwaad naar erger in gunstige zin te beïnvloeden, beoogt daarom in de eerste plaats symptoomreductie en wil uiteindelijk het ontstaan van de stoornis voorkomen.

Ideaal gesproken sluit preventie aan op de zorg voor mensen met de diagnose. Met behandeling wordt beoogd de ziekteduur te verkorten en herstel te bevorderen. Bij mensen die aan het herstellen zijn kan zich terugval voordoen. Terugvalpreventie wil dit voorkomen en wordt dikwijls beschouwd als een vast onderdeel van elke goede behandeling. Ook na herstel kunnen mensen opnieuw de stoornis terugkrijgen. Er wordt dan van een herhaling gesproken. Nu kunnen twee gezichtspunten worden ingenomen. Of dit wordt als een taak voor nazorg beschouwd, of de onderscheiden vormen van preventie komen hiervoor in aanmerking. Al met al is het wenselijk om het hele transitieproces te flankeren door een ketenbenadering van preventie, zorg en nazorg. Daarvoor is kennis nodig van de factoren die de overgangskansen tussen de gezondheidscondities beïnvloeden.

Methodologische kantekeningen

Op hoofdlijnen was het idee om data van grootschalige epidemiologische bevolkingsstudies en de data van gecontroleerd effectonderzoek te gebruiken om beter zicht te krijgen op de factoren die de overgangen tussen de gezondheidscondities beïnvloeden.

Epidemiologische cohortstudies in de algemene bevolking vormen hiervoor een gunstig vertrekpunt: zij helpen met het empirisch onderbouwen en aanscherpen van hypothesen. Het idee daarbij is dat de data van cohortstudie zich lenen tot het uitvoeren van “virtuele quasi-experimenten”. Zo kunnen we risicogroepen selecteren, in die groepen blootstelling aan een theoretisch interessante risicofactor isoleren, en de longitudinale impact ervan op klinische en economische eindpunten evalueren. Anders gezegd, de “virtuele quasi-experimenten” worden uitgevoerd door statistische manipulatie van cohortdata. In Nederland mogen we ons gelukkig
prijzen met het bezit van meerdere uitmuntende psychiatrische cohortstudies die de algemene populatie weerspiegelen. Dit helpt met het van te voren empirisch onderbouwen en aanscherpen van hypothesen. Dat is belangrijk omdat op die manier de onderzoeksaagenda gerationaliseerd kan worden, waardoor de ontwikkeling van interventies en het onderzoek naar hun kosteneffectiviteit efficiënt ingezet kan worden, namelijk op die gebieden waar we beste kansen verwachten voor kosteneffectieve preventie.

Tijdens de hypothese onderbouwende onderzoeksphase is het van belang graadmeters (indexen) te berekenen voor de behaalbare gezondheidswinst van preventie, zoals de attributieve fractie (AF). Dit soort indexen helpt om te kunnen communiceren over de uitkomsten van de genoemde virtuele quasi-experimenten. Daarnaast zochten we ook naar manieren om de opbrengsten van preventie uit te kunnen drukken in de universele taal van harde valuta.

Wanneer de “empirisch onderbouwde hypothesen” beschikbaar zijn, dan is het tijd om deze te toetsen in gerandomiseerd effectonderzoek. Dit is een noodzakelijke stap omdat die hypothesen alleen in virtuele quasi-experimenten bestudeerd waren, waar het niet mogelijk is om deelnemers aan de studie te randomiseren waardoor de kans op selectiebias levensgroot is. Om die hindernis te overwinnen zijn er als nog gerandomiseerde effectstudies nodig.

Ten slotte, wanneer ook die gerandomiseerde effectstudies zijn uitgevoerd, dan kan het nodig zijn om nog één stap verder te gaan, de uitkomsten van effectstudies meta-analytisch te combineren, zo de nauwkeurigheid en het onderscheidend vermogen (power) te verbeteren, en bij te dragen aan dat ultieme doel van de wetenschap: de accumulatie van replicaerbare kennis. Figuur 3 illustreert dit.

**Figuur 3** Aaneenschakelen van onderzoeksvermen

**Beantwoording van de onderzoeksvragen**
1. De menselijke en economische kosten van depressie zijn torenhoog
2. Preventie is de ontbrekende schakel in de geestelijke gezondheidszorg
3. Geïndiceerde preventie wordt het best gericht op ultrahoogrisicogroepen.
4. Geïndiceerde preventie kan effectief zijn
5. Geïndiceerde preventie kan kosteneffectief zijn.
Deze stellingen gaan gepaard met enkele kanttekeningen.

Kanttekeningen
Wij hebben weinig twijfel over het eerste punt dat depressie aanzienlijke kosten veroorzaakt. Vele studies hebben de omvang van de ziektelast en de economische last aangetoond.

De stelling dat preventie de ontbrekende schakel is, moeten we nuanceren. In het veld worden vele activiteiten ontplooid en daarvoor bestaat ook officiële erkenning van de zijde van het Ministerie van VWS. Toch zou de positie van preventie verder versterkt kunnen worden, vooral ten opzichte van samenwerkende partijen zoals de thuiszorg, de zorgaanbieders in de eerste lijn, het maatschappelijk werk, de verschillende consultatiebureaus en nog tal van andere partijen. De verbinding met die partijen kan en moet sterker. Daarvoor is het nodig dat preventie meer prestige verwerft, onder andere door bewezen effectieve interventies volgens professionele standaarden aan te bieden waarbij bovendien inzichtelijk gemaakt wordt wat preventie oplevert. Preventie heeft verder al jaren last van een krap financieel Jasje: slechts 3% van het budget voor de geestelijke gezondheidszorg gaat naar preventie, en dat vraagt om aandacht, zeker nu er een stelselwijziging plaatsvindt in de financiering.

De belangrijkste kwalificatie bij de derde stelling, dat preventie vooral op ultrahoogrisicogroepen gericht moet zijn, is de inperking dat die stelling vooral betrekking heeft op geïndiceerde preventie, dus bij mensen met al enkele symptomen. Dit zullen in het algemeen personen zijn die zich bewust zijn van hun depressieve klachten, en het ligt in de rede dat zij daar hulp voor zoeken of voor zouden willen zoeken. Dat kan anders liggen bij selectieve preventie, bij mensen die wel blootgesteld zijn aan risicofactoren, maar nog geen klachten hebben. Preventie bij deze groep, kan stuiten op medisch-ethische bezwaren, want het is niet gezegd dat zij zonder preventie de stoornis krijgen, evenmin dat preventie hen kan vrijwaarden van depressie, terwijl ze wel met de kennis dat zij tot een risicogroep behoren moeten leven. Hier past terughoudendheid.

De kosteneffectiviteit van depressiepreventie (punten 4 en 5) dient elke keer per interventie apart te worden vastgesteld voor onderscheiden doelgroepen. Dit wordt een langzame, want stapsgewijze, aanpak waarin het bewijsmateriaal bij stukjes en beetjes verzameld wordt. Gelet onze eigen ervaringen tot dus ver, zijn we optimistisch over de uitkomsten van dergelijk onderzoek, maar we beseffen dat de klinische en economische uitkomsten die over de langere termijn gemeten worden wel eens een ander plaatje kunnen laten zien dan wat we tot dus ver zagen. Onderzoek met een lange tijds horizon is daarom nodig.

In het licht van deze kanttekeningen beschouwen we de vijf stellingen niet als definitieve antwoorden. Het zijn eerder zoekrichtingen en agendapunten voor verder onderzoek.
Richtingen voor de toekomst

Er is nog een lange weg te gaan voordat preventie een solide positie heeft temidden van haar zusterdisciplines in de geestelijke gezondheidszorg, en er staan nog vele vragen open. Graag zouden we willen weten voor welke groepen welke interventies het beste werken. En we willen meer weten over de effecten over de langere termijn. Wij willen vooral weten hoe de interventies door meer mensen benut kunnen worden, want het huidige bereik is gering (circa 1% van de groep die uiteindelijk in het betreffend jaar een depressie krijgt) waardoor de behaalde gezondheidswinst ruim achterblijft bij die behaalbare. Om het bereik te verbeteren, zijn meerdere voorstellen gedaan, waarvan we een achtal willen noemen.

1. De positie en het aanzien van preventie verder versterken door een wetenschappelijke onderbouwde en goed geprotocolleerde werkwijze.

2. Bij het vergroten van het bereik van depressiepreventie gaat het niet alleen om het bereiken van meer mensen; het gaat vooral om het bereik van de juiste mensen. Dit proefschrift reikte methoden aan om die doelgroepen te identificeren.

3. Toch zouden we nog preciezer willen weten wie de mensen zijn die een preventieve interventie nodig hebben, want bij 50% van de mensen die een depressie ontwikkelen is die depressie binnen drie maanden weer verdwenen. Preventie zou daarom vooral op die mensen gericht moeten worden waarbij de prognose ongunstiger is. Daartoe zou een prognostische index ontwikkeld moeten worden.

4. De risicofactor benadering die in dit proefschrift werd gekozen zou geflankeerd en ondersteund moeten worden door een aanpak die gericht is op gezondheidsbevordering, waarbij het accent meer valt op het versterken van competenties en weerbaarheid. Depressiepreventie wordt zo aantrekkelijker voor doelgroepen.

5. Preventie zou er goed aan doen zich in de eerste instantie te concentreren op de medisch-ethisch minder omstreden geïndiceerde preventie, en dat aanbieden op een zo minst ingrijpende manier waar mogelijk, maar steviger waar nodig; dus bijvoorbeeld volgens de werkwijze van getrapte zorg.

6. Bovendien zou depressiepreventie meer geïntegreerd moeten worden met andere zorgvoorzieningen en zorgaanbieders met als doel te komen tot een betere vroegherkenning en doorverwijzing.

7. De financiële positie van de GGZ-preventie was niet sterk, en is nu met vraagtekens omringd. Dit behoeft aandacht bij GGZ-preventieafdelingen, gemeentelijke overheden, zorgverzekeraars, en in de beleidsvoornemens van VWS.

8. Preventie zou goed kunnen meeliften op de interessante ontwikkelen van e-mental health, waarbij (deels) geautomatiseerde interventies via het internet worden aangeboden. Dit zou het bereik op een kosteneffectieve manier kunnen vergroten.
Het is te hopen dat depressiepreventie een groter bereik krijgt waarbij vooral de mensen die er behoefte aan hebben bereikt worden op een manier die verantwoord en betaalbaar is.
References


Bijl D, Van Marwijk HWJ, Beekman ATF, De Haan M, Van Tilburg W. (2003) A Randomized Controlled Trial to Improve the Recognition, Diagnosis and Treatment of Major Depression in Elderly People in General Practice: Design, First Results and Feasibility of the West Friesland Study. Primary Care Psychiatry 8(4):135-40.


Treatment of Mood Disorders in Late Life. Archives of General Psychiatry 60(7):664-72.


assessment of mortality and disability from disease, injuries and risk factors

Nelson EC, Health AC, Madden PA, Cooper ML, Dinwiddie SH. (2002)
Association between self-reported childhood sexual abuse and adverse
psychosocial outcomes: results from a twin study. Archives of General
Psychiatry 59:139-145.

Costing: Methods and standard costs for economic evaluations in health care
(in Dutch). Diemen: Health Insurance Board.

Ormel J. (1980) Moeite met leven of moeilijk leven.[Difficulty with life, or a

Chronic medical conditions and mental health in older people: disability and
psychosocial resources mediate specific mental health effects. Psychological


HM. (Eds.). Depression: Neurobiological, psychopathological and
therapeutic advances. Chichester: John Wiley & Sons.

Osborn DP, Fletcher AE, Smeeth L, Stirling S, Nunes M, Breeze E, Siu-Woon NgE,
Scale Scores in a Representative Sample of 14 545 People Aged 75 and Over
in the United Kingdom: Results From the MRC Trial of Assessment and
Management of Older People in the Community. International Journal of
Geriatric Psychiatry 17(4):375-82.

of Geriatric Psychiatry 10:256-264.

Behaviour 19:2-21.

Penninx BWJH, Geerlings SW, Deeg DJH, Van Eijk JTM, Van Tilburg W,
Beekman ATF. (1999) Minor and major depression and the risk of death in


Pincus HA, Davis WW, McQueen LE. (1999) 'Subthreshold' Mental Disorders. A
Review and Synthesis of Studies on Minor Depression and Other 'Brand


Wittchen H-U, Robins LN, Cottler LB, Sartorius N, Burke JD, Regier DA and participants in the Multicentre WHO/ADAMHA Field Trials. (1991) Cross-


Acknowledgement

To finally make a dash for a PhD at my age does suggest a number of personality traits, none of which are too impressive. A series of quick guesses would produce the following list: (-) a great talent for bad planning, (-) absence of a focussed ambition, (-) coupled with a sharply focussed fear to fail, (-) more specifically, a deeply felt apprehension in the face of an immensely learned Doctoral Committee, and finally, (-) an unwavering respect for Peter’s principle: the statistical certainty that hubris will eventually spell catastrophe. To tell the truth, in my case it was a bit of everything. And yes, it took an effort to get where I am today. Only, the effort was mainly of others, hardly my own. My debts are therefore numerous.

To begin, I would like to thank the many people who helped me to better understand science and did much to promote an interest in it. In this respect I owe debts to Professor Lieteke Vught-Tijssen, who showed me that science is pure scholarship – reading, and then more reading – coupled with intellectual curiosity. Professor David Ingleby, with whom I conducted my first research, and who introduced me to meta-analysis. Professor Harm ‘t Hart, who doesn’t know but was a role model for me, because he is a most gentle gentleman, a firm believer in sharing knowledge, in the open universality of universities. I am especially indebted to Professor Peter van der Heijden, who saw something in me and made me research fellow at the Department of Methods and Statistics of the University of Utrecht. It is from him that I learned much about the intricacies of conducting research. I wish to mention Ger van Gils who made a living from conducting research. With him and Peter, science was turned into a commercial and often hilarious enterprise, which added quite another dimension to it.

Later, at the Trimbos Institute, I felt supported by Professor Henk Rigter, who was the first to direct my attention to health economics, and Professor Marianne Donker, who did the same, but in a more practical way. There were more people who promoted interests that became valuable later on. Rob Bijl, Ph.D, Ron de Graaf, Ph.D, and Professor Wilma Vollebergh, who initiated me in psychiatric epidemiology; and Erik van Ameijden, Ph.D, and Inge Spruit, Ph.D, who did the same for addiction epidemiology. In addition, I would like to mention some of my former and present colleagues at the Trimbos Institute, each of whom sparked off an interest in one or another direction, and made our work worthwhile, in particular: Neeltje Batelaan, Dorien Beusekom, Linda Bolier, Wim Brunenberg, Barbara Conijn, Guus Cruts, Marja Depla, Saskia van Dorselaer, Christina van der Feltz, Gerdien Franx, André Gageldonk, Karel Geelen (RIP), Ireen de Graaf, Margreet ten Have, Lourens Henkelman, Hazel Hull, Ad Kaasenbrood, Margriet van Laar, Paula van Lammeren, Lex Lemmers, Hans Kroon, Jolanda Meeuwis, Magda van der Meijden, Peter Meulenbeek, Karin Monshouwer, Raymond Niessink, Simone Onrust, Barbara van Os, Lyna Poliak, Jan Spijker, Bea Tiemens, Jaap Toet, Jacqueline Verdurmen, Ineke Voordouw, Jaap van Weeghel, John Wennink, Godelief Willemse, Rianne van der Zande, and Annemiek Zeegers. If it has not
already been very obvious, I would like to re-confirm my fondness and affection for Ernst Bohlmeijer and Heleen Riper, who make working at the Trimbos Institute such a pleasure, because they inspire, and do much to give our work a sense of direction and purpose; and Jeannet Kramer who with unfaltering honesty always lifts my mood. I wish to thank Detta Wijdh and Carola Vos for secretarial support, and Roisin de Jong for English language editing. Paul Anzion needs be mentioned for kindly checking the Dutch summary for sins against grammar, and giving indispensable help with the production of this book.

I have had the benefit and pleasure of learning much from professor Frans Rutten and his co-workers at the institute of Medical Technology Assessment, Erasmus Medical Centre: Leona Hakkaart - van Rijoen, Marc Koopmanschap, and Jan Oostenbrink. Special thanks to both Marc and Jan for their instructive advice and co-authoring some of the papers in this thesis.

At the EMGO institute, the Program of Common Mental Disorders, I would like to mention Professor Dorly Deeg who has been successful in bringing together a team of very fine researchers over a long period of time, and who has been instrumental to the ongoing LASA study. She read the draft versions of many of my papers and gave advice that helped to push matters much further. In the same team of researchers, I like to mention Hannie Comijs, Agnieska Ederveen, and Robert Schoevers who were inspiring and keen co-authors of articles – some of which are included in this thesis, and it is thanks to them that these papers are among the best.

At the Department of Clinical Psychology of the Free University (VU) in Amsterdam I wish to mention first and foremost Annemieke Van Straten, present and future co-author of many papers. Here I was happy to make the acquaintance of Professor Jack Dekkers, Professor Mark van der Gaag, Professor Ad Kerkhof, Josien Schuurmans and Niels Smits with whom I hope to work more closely in the future.

It has been hideously hard work but rewarding to write with Casper Schoemaker and Susan Meijer (both of the National Institute for Public Health and the Environment, RIVM) a report for the Ministry of Health on much the same topic as the one covered in this thesis. Both books were written in the same time. It was therefore inevitable that some amount of cross-fertilisation occurred – I hope, and believe, for the better of both books.

In a world where hard cash is at a premium, I thank the following institutes for generously funding some of the studies underpinning this thesis: The Netherlands Foundation of Mental Health (NFG, Utrecht); the Netherlands Health Research Council (ZonMw, The Hague), and the Ministry of Health (VWS, The Hague), with special thanks to Roland Gorter, Ben Roelands, Sue van Soest (NFG), Margreet Bloemers, Nanny van den Brink, Willem de Regt, Herma Schuttenbelt, Lucienne Willems (ZonMw), Jaap Penninga, Wim Brunenberg, Thijske Fraanje, Sabine Geerts, Laetitia Kuijpers, Brenda Langezaal, Nicole van Loy and Aaf Tiems (VWS) for their keen and stimulating interest in preventing mental disorders.

204
I give many thanks to my employers, the Trimbos Institute and the Department of Clinical Psychology of the Vrije Universiteit for enabling me to take sabbatical leave for working on this thesis. In this context, many thanks to Professor Pim Cuijpers (VU), Heleen Riper (Ti) and Professor Jan Walburg (Ti) for their generous support.

This dissertation would not possess the degree of accuracy and subtlety without the kind and helpful comments of, and discussions with, the learned members of the Doctoral Committee: Professor Anton van Balkom (VU Medical Centre, Amsterdam), Professor Lex Bouter (Rector Magnificus of the Vrije Universiteit, Amsterdam), Professor Clemens Hosman (Director of the Prevention Research Centre at the Universities of Nijmegen and Maastricht), Professor Isaac Marks (Professor Emeritus of the Institute of Psychiatry, Kings College, London), Professor Frans Rutten (Director of the institute of Medical Technology Assessment at the Erasmus University, Rotterdam) and Professor Philip Spinhoven (Department of Clinical Psychology and the Medical Centre, Leiden University, Leiden). In this context I also like to recall, with much pleasure, the several thought-provoking conversations with Professor Jan Swinkels (University of Amsterdam, AMC).

The best are kept for the last. I like to gratefully acknowledge Professor Aartjan Beekman, my first thesis supervisor. He was quick to see that I could and should write this thesis, and encouraged me to do so from the very moment we first met. He has been extremely helpful with the conceptual aspects of this work, took good care of epidemiological and statistical detail, and inspired me to complete the road from population-based epidemiology towards testing the hypotheses in clinical trials. To top it off, he improved the punch, passion and polish of many articles. I value most his drive, enthusiasm and aspirations, all of which are coupled with a fine sense of direction in the vast landscape that we call science.

Professor Pim Cuijpers, my second thesis supervisor, was in equal measure a driving force behind the genesis of this thesis. More than helping me with this thesis, he helped me to shape my career, and did so with astonishing alacrity and energy. First, he inspired me to take the step from psychiatric epidemiology to prevention – not to replace one with another, but to combine the two of them. Then he had an endless series of research questions addressing issues crucial to prevention and for which epidemiology could be helpful. Later he offered me the opportunity to work with him at the Department of Clinical Psychology at the VU. I found it immensely enjoyable to work with both Pim and Aartjan and to address issues that are relevant from a public health point of view, and also interesting from an academic perspective.

I owe many thanks to Heleen Riper and Jeannet Kramer for standing by me, most literary, when the moment will arrive to defend this thesis. I will rely on my “nimphae” in no small measure!

Foremost I want to thank my parents for always supporting me, in many ways, all the way, through their love, their passion and compassion for the many things my brothers and I have learned to hold dear. And there is Yvy, who shares
my life, and shows me, day after day, that life can be good, and that happiness is close at hand.

Finally, I want to acknowledge, with many thanks, all the people who volunteered to participate in the several studies on which this thesis is based: general practitioners, their patients and the literally thousands of people who took part in the large-scale epidemiological cohort studies. It is they who have been fundamental to this work, and it is to them that this work is dedicated. If, with the help of all, we have marginally contributed to the wellbeing of some, then, I feel, we have done our bit.
Curriculum vitae

Hans-Filip Engelbert (Filip) Smit studied anthropology and graduated in 1990. After his graduation professor Peter van der Heijden invited him to become a research associate at the Department of Methodology and Statistics at the Faculty of Social Sciences of the University of Utrecht. Here, he participated in a series of studies for the Ministry of Transport, the Home Office, and the Ministry of Justice. During his collaboration with Van der Heijden he developed an interest in meta-analysis, econometrics, and population size estimators based on capture-recapture techniques.

As of September 1993 he worked as a research specialist at the Trimbos Institute, which is Netherlands Institute of Mental Health and Addiction. Currently, he is working as a senior research fellow at the Institute’s centre of Prevention and Brief Intervention. In September 2004 he took up a similar position at the Department of Clinical Psychology at the Vrije Universiteit in Amsterdam. Both jobs are combined. He conducts research in psychiatric and addiction epidemiology, and health economics.

His professional experience record can be summarised as follows: he (co)authored over 50 publications in peer reviewed journals, is currently involved in some 20 research projects; collaborated in a total of 27 successful grant acquisitions of research projects; was project leader of 8 grant acquisitions, was awarded 7 grants, is participating researcher in 7 Ph.D. thesis projects, and is ad hoc reviewer for several journals among them the Archives of General Psychiatry (AGP), the British Journal of Psychiatry (BJP) and the Journal of Affective Disorders (JAD).

He lives in Utrecht with Yvonne (Yvy) Sittrop and a cat. With them he shares his love for classical music, renaissance painting, English literature, travel, wine, and the history of Central Asia.
Publications

Submitted in 2006


Smits N, Smit F, De Graaf R, Cuijpers P. Assessing optimal cut-off scores of mental health screeners on the basis of cost-benefit analysis.


Comijs HC, Beekman ATF, Smit F, Van Tilburg T, Deeg DJH. The association between recent life-events and depression in older persons with and without childhood adversities.

Cuijpers P, Smit F, Van Straten A. Psychological treatment of subthreshold depression: A systematic review


In press 2006


Printed in 2006


Printed in 2005


Printed in 2004


Printed in 2003


**Printed in 2002**


**Printed in 2001**


**Printed in 2000 and before**


Book (chapter) and reports (selection)


**Presentations**


In a Serbian-orthodox prayer Saint George is addressed as deliverer of captives, defender of the poor, and healer of the infirm.
1. The human and economic costs of depression are staggering (this thesis)

3. Prevention is the missing link in public mental health (this thesis)

3. Prevention should become the first step in a stepped-care approach (logic)

4. Prevention of depression is best directed at ultra-high risk groups (this thesis)

5. Prevention of depression can be cost-effective (the smoking gun)

6. We do not know how people benefit from depression exactly (antithesis)

7. It is true: we need far more research (recommendation)

8. Prevention needs to be evidence-based (priority)

9. Prevention should be directed at people with a poor prognosis (new directions)

10. Mastery might be the alchemists’ gold of preventive psychiatry (future research)

11. Prevention of depression could be made more amusing (why not?)
This book is about preventing depressive disorder. Four questions are addressed:

1. Do we need depression prevention?
2. Do we know where to start?
3. Is it effective?
4. Is economically affordable?

Building carefully on population-based epidemiological studies and randomised trials, it is argued that this highly prevalent and disabling disorder is best tackled not only by curative interventions directed at promoting recovery, but also by prevention to avoid new onsets of depression in the first place. Avoiding new onsets would help to avoid human suffering, maintain the quality of life of many, and is likely to avoid the substantial costs associated with the disorder.

Prevention is best directed at ultra-high risk groups: people who have some early depressive symptoms and, in addition, have been exposed to risk factors predictive of the onset of depression. These ultra-high risk groups are numerically small, but account for the majority of new cases, making prevention at once effective, efficient, and economically affordable – that is, in theory.

Of course, the proof is in the eating of the pudding. Several randomised prevention trials indicate that the risk of depression can be reduced by 30%. Moreover, an economic evaluation indicated that prevention of depression is likely to be cost-effective. It is concluded that prevention of depression is still in its infancy, but the signs are promising. Prevention of depression is likely to be an endeavour well worth the effort.

Filip Smit (fsmit@trimbos.nl) is research fellow at the Trimbos Institute, which is the Netherlands Institute of Mental Health and Addiction, in Utrecht. He holds a similar position at the Department of Clinical Psychology of the Vrije Universiteit, in Amsterdam. The thesis was written at the Faculty of Medicine of the same university.