

# Opportunities for Cost-effective Prevention of Late-Life Depression

## *An Epidemiological Approach*

Filip Smit, MSc; Agnieska Ederveen, MD; Pim Cuijpers, PhD; Dorly Deeg, PhD; Aartjan Beekman, MD, PhD

**Context:** Clinically relevant late-life depression has a prevalence of 16% and is associated with substantial societal costs through its disease burden and unfavorable prognosis. From the public health perspective, depression prevention may be an attractive, if not imperative, means to generate health gains and reduce future costs.

**Objective:** To target high-risk groups for depression prevention such that maximum health gains are generated against the lowest cost.

**Design:** Population-based cohort study over 3 years.

**Setting:** General population in the Netherlands.

**Participants:** Twenty-two hundred community residents aged 55 to 85 years. Of these, 1925 were not depressed at baseline.

**Main Outcome Measure:** The onset of clinically relevant depression was measured with the Center for Epidemiological Studies Depression Scale. For each of the risk factors (and their combinations), we calculated in-

stances of potential health gain and the effort (costs) required to generate those health gains.

**Results:** One in every 5 cases of clinically relevant late-life depression is a new case. Consequently, depression prevention has to play a key role in reducing the influx of new cases. This is best done by directing prevention efforts toward elderly people who have depressive symptoms, experience functional impairment, and have a small social network, in particular women, as well as people who have attained only a low educational level or who suffer from chronic diseases.

**Conclusions:** Directing prevention efforts toward selected high-risk groups could help reduce the incidence of depression and is likely to be more cost-effective than alternative approaches. This article further shows that we have the methodology at our disposal to conduct ante hoc cost-benefit analysis in preventive psychiatry. This helps set a rational research and development agenda before testing the cost-effectiveness of interventions in time-consuming and expensive trials.

*Arch Gen Psychiatry.* 2006;63:290-296

**Author Affiliations:** Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht (Mr Smit and Drs Cuijpers and Beekman), and Departments of Psychiatry (Drs Ederveen, Deeg, and Beekman) and Clinical Psychology (Dr Cuijpers), Vrije Universiteit Amsterdam, the Netherlands.

**L**ATE-LIFE DEPRESSION IS CHARACTERIZED by high prevalence, unfavorable prognosis, reduced quality of life, and excess mortality.<sup>1-8</sup> It is also associated with substantial societal costs.<sup>9-12</sup>

Late-life depression is further characterized by a large annual influx of new cases because we have found that 1 in every 5 cases is a new case. 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.<sup>13</sup>

In this context, it should be noted that depression is a treatable condition.<sup>14-16</sup> However, according to a recent estimate, the total disease burden associated with depression can only be reduced about 50%, even under a hypothetical regime of optimal evidence-based treatment.<sup>17</sup> This is another reason why prevention has to play an important role in public health. Re-

cently, 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.<sup>18</sup>

However, developing preventive interventions and testing their cost-effectiveness in randomized trials is time-consuming and expensive.<sup>19</sup> 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 article is to describe a methodology that could help identify cost-effective preventive interventions at the earliest possible stage and apply this methodology to the case of late-life depression.

The methodology of identifying high-risk groups for prevention is not new,<sup>20,21</sup> 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 often not 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.

## METHODS

### SUBJECTS AND PROCEDURES

The analyses were based on the data of the first 2 waves of the Longitudinal Aging Study Amsterdam. The sampling and procedures of this study have been described in the ARCHIVES in detail.<sup>22</sup> At baseline, we interviewed 3056 community residents aged 55 to 85 years. Participating subjects gave their informed consent and underwent face-to-face interviews in their homes. The random sample was stratified by age and sex. The older age strata and men were oversampled in anticipation of higher attrition rates among these groups during the course of the study. After 3 years (mean  $\pm$  SD time, 1115  $\pm$  59 days), 2200 subjects (72%) were successfully reinterviewed. 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 sex, lower education, functional limitations, chronic diseases, and cognitive decline, but not depression status at baseline.<sup>22</sup> Corrective weights were used to account for the joint effect of intentional oversampling and accidental attrition (see the subsection "Analysis").

### DEPRESSION

Depression was ascertained with the Center of Epidemiological Studies Depression Scale (CES-D).<sup>23</sup> It consists of 20 items and its total score has a range between 0 and 60. Scores greater than or equal to 16 indicate clinically significant levels of depressive symptoms.<sup>24</sup> In the remainder of this article, CES-D scores of 16 or greater will be referred to as "depression." At this cutoff, sensitivity is 100% and specificity is 88% for major depressive disorder in the elderly Dutch population.<sup>25</sup> Measurements were taken at baseline ( $t_0$ ) and at first follow-up ( $t_1$ ). A person was deemed to be an incident case when 3 criteria were met: (1) absence of depression at  $t_0$  (CES-D < 16), (2) presence of depression at  $t_1$  (CES-D  $\geq$  16), and (3) significant change between  $t_0$  and  $t_1$  (change of CES-D score  $\geq$  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. For the third criterion, we chose a minimum change of 5 CES-D points because it represents, in clinical terms, a medium to large change<sup>26</sup> and has the advantage that it has also been used in other studies.<sup>22</sup> 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.<sup>27</sup> 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<sup>28</sup> and a recently published review on risk indicators of late-life depression,<sup>29</sup> the following putative risk indicators were included. The demographics were female sex (1 = female, 0 = male); age older than 65 years, the age at which 30% of the sample makes a significant transition in their life because of retirement (1 = older than 65 years,

0 = younger); low education (dichotomized into 1 = elementary school, 0 = high school and higher); and living in an urban environment (1 = living in Amsterdam, 0 = living elsewhere).

We also included chronic illnesses<sup>30</sup> (dichotomized, 1 = 2 or more, 0 = 1 or none), such as diabetes mellitus, chronic obstructive lung disease, cardiac disease, arthritis of knee or hip, and cancer, and cognitive impairment<sup>31</sup> (1 = Mini-Mental State Examination score < 24, 0 = Mini-Mental State Examination score of 24-30). Earlier studies have indicated that it is not so much the presence of chronic medical conditions that predict the onset of depression, but rather the functional limitations that may stem from them, the subjective appraisal of one's health, and the degree by which one's sense of mastery (locus of control) is affected.<sup>4,32,33</sup> Therefore, the following measures were also included: functional limitations<sup>34</sup> (1 = 1 or more, 0 = none), self-rated poor health<sup>35</sup> (1 = poor health; 0 = sometimes good/sometimes bad, fair, good, or excellent health), and low mastery<sup>36</sup> (1 = score below the 50th percentile on the scale, 0 = score above 50th percentile). Depressive symptoms (1 = CES-D scores between 5 and 15; 0 = CES-D < 5, ie, below 50th percentile) at baseline were also relevant because they can act as precursors of a CES-D score greater than 16. The distribution of the CES-D is as follows: 25% of the population falls in the range of 0 to 2, 50% in 0 to 5, 75% in 0 to 10, and 90% in the range of 0 to 16.

Finally, social vulnerability was assessed by 2 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  $t_0$  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 oversampling of men and the older age strata and loss to follow-up. This was done by weighting the data such that the multivariate distribution over sex and age in the sample was exactly the same as in the general Dutch population in the age range of 55 to 85 years as reported by Statistics Netherlands for the year 2002. To obtain correct 95% confidence intervals and *P* values under weighting, all variance-related statistics were obtained with help of the first-order Taylor-series linearization method as implemented in Stata SE version 7.0.<sup>37</sup> Weighted numbers are reported, rounded to the nearest integer, throughout the remainder of this article. The (weighted) analyses were based on the 1925 people at risk of becoming depressed, ie, 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 (**Table 1**).

For each risk indicator, we got the incidence rate ratio (IRR) 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. Incidence rate ratio 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 regres-

sion while adjusting for competing risks.<sup>38</sup> When converted into a percentage, the AF denotes by how many percentage points the current incidence rate of depression in the population would be reduced if the adverse effect of the risk indicator is completely blocked.<sup>20,39</sup> This equals the maximum possible impact of a completely successful preventive intervention. Because 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 of the potential health gain in the population.

**Table 1. Exposure Status (in Percent and 95% CI) in the Group at Risk and the Incident Group (Weighted Analyses)**

Independent Variable	Group at Risk, % (95% CI), (n = 1925)	Incident Cases, % (95% CI), (n = 158)
Female sex	51.6 (49.3 to 53.9)	72.9 (66.1 to 79.7)*
Age >65 y	52.2 (49.8 to 54.5)	66.7 (58.9 to 74.6)*
Low education	36.4 (34.2 to 38.7)	53.9 (46.0 to 61.7)*
Urban environment	24.7 (22.7 to 26.7)	29.4 (22.3 to 36.5)
≥2 Chronic diseases	32.8 (30.6 to 34.9)	51.7 (43.8 to 59.5)*
Functional limitations	28.5 (26.4 to 30.5)	52.1 (44.2 to 60.0)*
Self-rated poor health	0.7 (0.3 to 1.0)	2.6 (0.0 to 5.3)
Cognitive impairment	4.5 (3.6 to 5.4)	8.2 (4.1 to 12.3)
Depressive symptoms	40.3 (38.0 to 42.6)*	66.1 (58.6 to 73.5)*
Low mastery	53.8 (51.5 to 56.2)	69.2 (61.8 to 76.5)*
Small social network	45.5 (43.1 to 47.8)	58.5 (50.6 to 66.3)*
Ever widowed	19.9 (18.2 to 21.7)	33.4 (26.2 to 40.6)*

Abbreviation: CI, confidence interval.  
\*Value is significant at  $\alpha < .05$ .

Although it is possible to adjust the AF statistic for interventions that are not completely effective,<sup>21</sup> it is also understood that we need not correct the AF statistic for the purpose of this article: a measure of relative performance is good enough for ranking risk indicators by their utility for prevention. We 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. 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 to avoid 1 new case of late-life depression. Again we do not expect that preventive interventions are completely successful, so it is 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 ER, the strength of association between risk indicator and outcome (IRR), the maximum achievable health gain (AF), and the 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 against 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**). Then, using conventional backward-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 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

**Table 2. Complete Multivariate Model of the Risk Indicators for 1925 Subjects (Weighted Analysis)**

Risk Indicator (at $t_0$ )	ER, %	IRR	95% CI	AF, %	95% CI	NNT
Female sex	51.6	1.79	1.24 to 2.60*	32.5	13.2 to 47.6*	26
Age >65 y	52.2	1.21	0.82 to 1.78	13.1	-14.0 to 33.7	83
Low education	36.4	1.38	0.99 to 1.92	15.8	0.0 to 29.4	38
Urban environment	24.7	1.25	0.89 to 1.75	7.3	-3.8 to 17.2	67
≥2 Chronic diseases	32.8	1.55	1.11 to 2.16*	20.6	5.2 to 33.4*	27
Functional limitations	28.5	1.52	1.06 to 2.18*	21.1	3.5 to 35.5*	25
Self-rated poor health	0.7	2.07	0.80 to 5.34	3.7	-1.6 to 8.7	6
Cognitive impairment	4.5	1.32	0.76 to 2.29	2.8	-1.1 to 7.0	24
Depressive symptoms	40.3	2.09	1.46 to 2.97*	39.4	22.2 to 52.8*	17
Low mastery	53.8	1.25	0.88 to 1.78	15.0	-8.6 to 33.4	70
Small social network	45.5	1.34	0.98 to 1.84	16.0	-1.0 to 30.0	42
Ever widowed	19.9	1.06	0.76 to 1.50	2.3	-10.1 to 13.3	80
Total AF				86.2	77.9 to 91.4*	

Abbreviations: AF, attributable fraction; CI, confidence interval; ER, exposure rate; IRR, incidence rate ratio; NNT, number needed to be treated;  $t_0$ , baseline.  
\*Value is significant ( $P < .05$ ).

**Table 3. Parsimonious Model\***

Risk Indicator	IRR	95% CI	AF, %	95% CI	NNT
Female sex	1.85	1.30 to 2.64	34.0	15.4 to 48.5	23
Low education	1.41	1.03 to 1.92	16.6	1.5 to 29.4	34
≥2 Chronic diseases	1.57	1.14 to 2.16	20.7	6.2 to 33.0	26
Functional limitations	1.75	1.26 to 2.43	26.5	11.5 to 38.9	18
Depressive symptoms	2.18	1.56 to 3.04	40.3	24.5 to 52.8	16
Small social network	1.52	1.12 to 2.06	21.6	5.8 to 34.7	30
Total AF			82.8	74.3 to 88.5	

Abbreviations: AF, attributable fraction; CI, confidence interval; IRR, incidence rate ratio; NNT, number needed to be treated.  
\*All parameters are statistically significant at  $P < .05$ .

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 maximizing and minimizing is depicted in the **Figure**.

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 savings of a future prevention. The method of this ante-hoc health-economical evaluation is straightforward but best illustrated with real data.

## RESULTS

### INCIDENCE

One hundred fifty-eight people (weighted N) became incident cases of depression in a sum total of 5643 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 5861 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 2200 people. Of these, 1925 (87.5%) were at risk of becoming depressed, of whom 158 (8.2%) became incident cases after 3 years at  $t_1$ . Table 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 with the group at risk (Table 1).

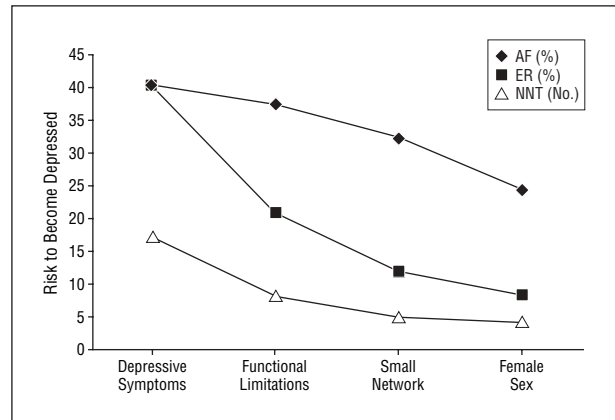
### A MODEL WITH ALL RISK INDICATORS

Table 2 shows the ER, the IRR, the population AF, and the NNT for each of the risk indicators after adjusting for the effects of all other risks in the model. The factors female sex, having 2 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).

### 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  $\alpha < .05$ ) 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 3.

In the parsimonious model, a smaller number of risk indicators was retained (Table 3). These were female sex, low education, having 2 or more chronic diseases, experiencing functional limitations, having an above-average number of depressive symptoms, and having a small social network.



**Figure.** Effect of adding extra risk indicators on the attributable fraction (AF), exposure rate (ER), and number needed to be treated (NNT).

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), 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 efforts target people who are exposed to combinations of risk indicators (**Table 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 4, it is clear that when these subsyndromatically depressed people also experience impairment, then the IRR and AF values rise and ER and NNT values drop even further (row 5 in Table 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 and has a risk of becoming depressed that is higher by a factor of 4.5; if the adverse effect of the exposure to all 3 risk indicators could be blocked completely, then the incidence rate of depression in the population would drop by 32.3%.

Preventive interventions are likely to become even more cost-effective when they target people who are in addition suffering from chronic diseases (row 15), have attained only low education levels (row 14), or are women (row 13). 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.

This minimization/maximization process is depicted graphically in the Figure. It shows how consecutively adding a risk indicator to the risk profile impacts the AF, ER, and NNT. In general, one would like to have the AF curve as high as possible (to maximize the health benefit in the



**Table 4. Additive Effect of Joint Exposures**

Joint Exposure	ER, %	IRR	AF, %	NNT
Depressive symptoms				
+ Female sex	29.3	3.00	37.2	9
+ Low education	20.6	3.06	31.8	8
+ Chronic diseases	20.4	2.86	28.7	8
+ Functional limitations	20.8	3.58	37.5	7
+ Small network	24.8	3.16	36.2	8
Depressive symptoms and functional limitations				
+ Female sex	14.3	3.60	28.7	6
+ Low education	11.1	3.79	25.9	5
+ Chronic diseases	12.4	3.16	23.0	7
+ Small social network	11.7	4.53	32.3	5
Depressive symptoms and functional limitations and a small network				
+ Female sex	8.0	4.60	24.2	4
+ Low education	6.7	4.59	21.9	4
+ Chronic diseases	6.5	4.25	19.7	4

Abbreviations: AF, attributable fraction; ER, exposure rate; IRR, incidence rate ratio; NNT, number needed to be treated.

population) while, at the same time, it is better to have the ER and NNT curves as low as possible (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 because 40.3% of the population should then receive the intervention. Furthermore, the efficiency of such an approach is not optimal, because—as the NNT indicates—16 people of this target group must receive the intervention to avoid 1 new case of depressive disorder. This can be improved. In a next step, there are several risk indicators to choose 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 is selected that is associated with the best IRR, AF, ER, and NNT values.

#### QUANTIFYING THE HEALTH GAIN

Consider the risk profile in row 13 of Table 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  $2.8 \times (1 - 0.24) = 2.1$  cases per 100 person-years. In every 1 million older adults, this would represent roughly 5950 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,<sup>18</sup> then 1785 onsets will be avoided. This still represents a substantial health gain in a source population of 1 million.

#### COST-BENEFIT CONSIDERATIONS

Avoiding new onsets has economical ramifications. Katon and colleagues<sup>40</sup> 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 1785 onsets would thus result in a cost saving of at least \$1.9 million for 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 do 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 with the costs and benefits of an intervention. This type of ante hoc cost-benefit calculation can be used to select interesting preventive interventions for further cost-effectiveness research.

#### COMMENT

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 also guide research efforts toward more promising research areas in preventive psychiatry.

#### 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 AF and a low 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 identify cost-effective preventive interventions at a very early stage of the costly and time-consuming cycle of development and evaluation of pre-

ventive 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,<sup>20,21</sup> 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 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, they 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 sex; low education; having 2 or more chronic illnesses; experiencing functional limitations; having above-average symptom levels of depression not exceeding the threshold for clinically relevant depression, ie, CES-D scores above 5 and below 16; and, finally, having a small social network.

Third, the combined effect of being exposed to 3 or 4 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 a regional level. Even based on conservative assumptions, a preventive intervention could prevent 1785 new onsets in every 1 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 1785 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 that 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 that is important for setting a rational research and development agenda for depression prevention.

The limitations of this study consist of the measurement of the exposures, which was not very detailed. 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 important issue 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, such as depressive symptoms, from those that are not. It is also worth noting that some risk indicators are not modifiable, such as chronic illnesses, but their adverse psychological effects might be contained. Finally, there are risk indicators, such as female sex, that are not modifiable but are valuable from the perspective of identifying groups at risk—which was the principal aim of this article.

### DIRECTIONS FOR FUTURE RESEARCH

In future research, thought should be given to the quantification of success rates of preventive interventions; the quantification of impacts on the incidence not only of major depressive disorder, but of the whole spectrum of anxiety and mood disorders; and the accommodation of cost-benefit considerations in this sort of analysis. It would also be valuable to cross-validate the methodology presented in this article by comparing it with alternative methodologies, such as Classification and Regression Tree analysis.<sup>41</sup> We are currently conducting such a study. It is recommended that in the future, the methodology presented in this article or related methodologies<sup>17,42</sup> should be further developed and employed to direct both research and prevention where they are likely to be most cost-effective.

**Submitted for Publication:** March 11, 2004; final revision received August 6, 2004; accepted April 27, 2005.

**Correspondence:** Filip Smit, MSc, Trimbos Institute, Netherlands Institute of Mental Health and Addiction, PO Box 725, 3500 AS, Utrecht, The Netherlands (fsmit@trimbos.nl).

**Funding/Support:** This study is financially supported by grant 5753 from the Netherlands Fund of Mental Health, Utrecht. It is based on the data that were collected in the context of the Longitudinal Aging Study Amsterdam,

which is financed by the Netherlands Ministry of Welfare, Health, and Sports, the Hague.

**Acknowledgment:** We acknowledge, with many thanks, Jan Poppelaars, MSc, and Mariette Westendorp, MSc, for their contributions to data collection.

## REFERENCES

1. Beekman AT, Penninx BW, Deeg DJ, de Beurs E, Geerlings SW, van Tilburg W. The impact of depression on the well-being, disability and use of services in older adults: a longitudinal perspective. *Acta Psychiatr Scand.* 2002;105:20-27.
2. Cole MG, Bellavance F, Masour A. Prognosis of depression in elderly community and primary care populations: a systematic review and meta-analysis. *Am J Psychiatry.* 1999;156:1182-1188.
3. Cuijpers P, Smit F. Excess mortality in depression: a meta-analysis of community studies. *J Affect Disord.* 2002;72:227-236.
4. Geerlings SW, Beekman AT, Deeg DJ, van Tilburg W. Physical health and the onset and persistence of depression in older adults: an eight-wave prospective community-based study. *Psychol Med.* 2000;30:369-380.
5. Geerlings SW, Beekman AT, Deeg DJ, Twisk JW, van Tilburg W. The longitudinal effect of depression on functional limitations and disability in older adults: an eight-wave prospective community-based study. *Psychol Med.* 2001;31:1361-1371.
6. Geerlings SW, Beekman AT, Deeg DJ, Twisk JW, van Tilburg W. Duration and severity of depression predict mortality in older adults in the community. *Psychol Med.* 2002;32:609-618.
7. Penninx BW, Geerlings SW, Deeg DJ, van Eijk JT, van Tilburg W, Beekman AT. Minor and major depression and the risk of death in older persons. *Arch Gen Psychiatry.* 1999;56:889-895.
8. Stek ML, van Exel E, van Tilburg W, Westendorp RG, Beekman AT. The prognosis of depression in old age: outcome six to eight years after clinical treatment. *Aging Ment Health.* 2002;6:282-285.
9. Beekman AT, Deeg DJ, Braam AW, Smit JH, van Tilburg W. Consequences of major and minor depression in later life: a study of disability, well-being and service utilization. *Psychol Med.* 1997;27:1397-1409.
10. Hunkeler EM, Spector WD, Fireman B, Rice DP, Weisner C. Psychiatric symptoms, impaired function, and medical care costs in an HMO setting. *Gen Hosp Psychiatry.* 2003;25:178-184.
11. Unutzer J, Patrick DL, Simon G, Grembowski D, Walker E, Rutter C, Katon W. Depressive symptoms and the cost of health services in HMO patients aged 65 years and older: a 4-year prospective study. *JAMA.* 1997;277:1618-1623.
12. Von Korff M, Ormel J, Katon W, Lin EH. Disability and depression among high utilizers of health care: a longitudinal analysis. *Arch Gen Psychiatry.* 1992;49:91-100.
13. Smit F, Beekman A, Cuijpers P, de Graaf R, Vollebergh WA. Selecting key variables for depression prevention: results from a population-based prospective epidemiological study. *J Affect Disord.* 2004;81:241-249.
14. Cuijpers P. Psychological outreach programs for depressed elderly: a meta-analysis of effects and drop-out. *Int J Geriatr Psychiatry.* 1998;13:41-48.
15. McCusker J, Cole M, Keller E, Bellavance F, Berard A. Effectiveness of treatments of depression in older ambulatory patients. *Arch Intern Med.* 1998;158:705-712.
16. Thase ME, Greenhouse JB, Frank E, Reynolds CF III, Pilkonis PA, Hurley K, Grochocinski V, Kupfer DJ. Treatment of major depression with psychotherapy or psychotherapy-pharmacotherapy combinations. *Arch Gen Psychiatry.* 1997;54:1009-1015.
17. Andrews G, Sanderson K, Corry J, Lapsley HM. Using epidemiological data to model efficiency in reducing the burden of depression. *J Ment Health Policy Econ.* 2000;3:175-186.
18. Cuijpers P, Van Straten A, Smit F. Preventing the incidence of new cases of mental disorders: a meta-analytic review. *J Nerv Ment Dis.* 2005;193:119-125.
19. Cuijpers P. Examining the effects of prevention programs on the incidence of new cases of mental disorders: the lack of statistical power. *Am J Psychiatry.* 2003;160:1385-1391.
20. Miettinen OS. Proportion of disease caused or prevented by a given exposure, trait, or intervention. *Am J Epidemiol.* 1974;99:325-332.
21. Morgenstern H, Bursic ES. A method for using epidemiologic data to estimate the potential impact of an intervention on the health status of a target population. *J Community Health.* 1982;7:292-309.
22. Beekman AT, Geerlings SW, Deeg DJ, Smit JH, Schoevers RS, de Beurs E, Braam AW, Penninx BW, van Tilburg W. The natural history of late-life depression: a 6-year prospective study in the community. *Arch Gen Psychiatry.* 2002;59:605-611.
23. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas.* 1977;1:385-401.
24. Berkman LF, Berkman CS, Kasl SV, Freeman DH, Leo L, Ostfeld AM, Coronini-Huntly J, Brody JA. Depressive symptoms in relation to physical health and functioning in the elderly. *Am J Epidemiol.* 1986;124:372-388.
25. Beekman AT, Deeg DJ, van Limbeek J, Braam AW, de Vries MZ, Van Tilburg W. Criterion validity of the Center for Epidemiologic Studies Depression scale (CES-D): results from a community-based sample of older adults in the Netherlands. *Psychol Med.* 1997;27:231-235.
26. Lipsey MW, Wilson DB. The efficacy of psychological, educational and behavioral treatment. *Am Psychol.* 1993;48:1181-1209.
27. Jacobson NS, Truax P. Clinical change: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol.* 1991;59:12-19.
28. Brown GW, Harris TO. *Social Origins of Depression.* London, England: Tavistock; 1978.
29. Cole MG, Dendukuri N. Risk factors for depression among elderly community subjects: a systematic review and meta-analysis. *Am J Psychiatry.* 2003;160:1147-1156.
30. Kriegsman DM, Penninx BW, van Eijk JT, Boeke AJ, Deeg DJ. Self-reports and general practitioner information on the presence of chronic diseases in community dwelling elderly: a study on the accuracy of patients' self-reports and on determinants of inaccuracy. *J Clin Epidemiol.* 1996;49:1407-1417.
31. Folstein MF, Folstein SE, McHugh PR. Mini-mental state: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12:189-198.
32. Ormel J, Kempen G, Penninx B, Brilman E, Beekman A, van Sonderen E. Chronic medical conditions and mental health in older people: disability and psychosocial resources mediate specific mental health effects. *Psychol Med.* 1997;27:1065-1077.
33. Zarit SH, Femia EE, Gatz M, Johansson B. Prevalence, incidence and correlates of depression in the oldest old: the OCTO study. *Aging Ment Health.* 1999;3:119-128.
34. Van Sonsbeek JLA. Methodological and substantial aspects of the OECD indicator of chronic functional limitations. *Maandbericht Gezondheid.* 1988;88:4-17.
35. Central Bureau of Statistics. *Health Interview Questionnaire.* Heerlen, the Netherlands: Central Bureau of Statistics; 1989.
36. Pearlin LI, Schooler C. The structure of coping. *J Health Soc Beh.* 1978;19:2-21.
37. Stata statistical software release 7.0. College Station, Tex: Stata Corporation; 2001.
38. Greenland S, Drescher K. Maximum likelihood estimation of the attributable fraction from logistic models. *Biometrics.* 1993;49:865-872.
39. Rothman KJ, Greenland S. *Modern Epidemiology,* 2nd ed. Philadelphia, Pa: Lippincott-Raven; 1998.
40. Katon WJ, Lin E, Russo J, Unutzer J. Increased medical costs of a population-based sample of depressed elderly patients. *Arch Gen Psychiatry.* 2003;60:897-903.
41. Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Ann Behav Med.* 2003;26:172-181.
42. Issakidis C, Sanderson K, Corry J, Andrews G, Lapsley H. Modeling the population cost-effectiveness of current and evidence-based optimal treatment for anxiety disorders. *Psychol Med.* 2004;34:19-35.