doi:10.1093/fampra/cml071 Family Practice Advance Access published on 20 January 2007

The prognosis of depression in older patients in general practice and the community. A systematic review

Els Licht-Strunk[a], Daniëlle AWM van der Windt[a,c], Harm WJ van Marwijk[a], Marten de Haan[a] and Aartjan TF Beekman[b]

Licht-Strunk E, van der Windt DAWM, van Marwijk HWJ, de Haan M, Beekman ATF. The prognosis of depression in older patients in general practice and the community. A systematic review. *Family Practice* 2007; **24**: 168–180.

Background. Little is known about the prognosis of depression in older patients in general practice or the community.

Objectives. To summarize available evidence on the course and prognostic factors of depression in older persons.

Methods. We conducted a systematic, computerized search of Medline and PsycINFO. Manual search of references of included studies were done. Studies potentially eligible for inclusion were discussed by two reviewers. Methodological quality was independently assessed by two reviewers. Data regarding selection criteria, duration of follow-up, outcome of depression and prognostic factors were extracted.

Results. We identified 40 studies reporting on four cohorts in general practice and 17 in the community. Of all, 67% were of high quality. Follow-up was up to 1 year in general practice and up to 10 years in the community. Information on treatment was hardly provided. About one in three patients developed a chronic course. Five cohorts used more than two measurements during follow-up, illustrating a fluctuating course of depression. Using a best evidence synthesis we summarized the value of prognostic indicators. General practice studies did not provide strong evidence for any factor. Community studies provided strong evidence for an association of baseline depression level, older age, external locus of control, somatic co-morbidity and functional limitations with persistent depression.

Conclusion. Within the older population, age seems to be a negative prognostic factor, while older people are more likely to be exposed to most of the other prognostic factors identified.

Introduction

Depression is a common disorder in older age. We found a prevalence of major depression of 14% and of minor depression of 10% in older patients visiting GPs in The Netherlands.[1] Depression in older patients is associated with disability, morbidity and mortality.[2] Most depressive patients are diagnosed and treated in general practice. Several treatments, both medical and psychological, have shown to be effective in treating depression in older patients.[3] However, it is unclear which patients will have a self-limiting course and who will benefit most from treatment. In order to improve mental health care it is important to identify patients at high risk of persistence of depression. This could help to focus the limited resources available in general practice to those patients in whom treatment is most urgently needed. Furthermore, it can prevent treatment with its adverse side effects for those who do not need it.

The aim of the present study was to carry out a systematic search of the literature summarizing the available evidence regarding course and prognostic factors of depression in older persons (55 years and older). Data of studies carried out in specialized psychiatry settings cannot easily be translated to general practice, due to its selection of more serious depression.

Received 10 July 2006; Revised 10 November 2006; Accepted 4 December 2006.
[a]Department of General Practice and [b]Department of Psychiatry, Institute for Extramural Medicine (EMGO), VU University Medical Center, Amsterdam, The Netherlands and [c]Primary Care Sciences Research Centre, Keele University, Staffordshire, UK; Correspondence to Els Licht-Strunk, Department of General Practice, Institute for Extramural Medicine (EMGO), VU University Medical Center, Amsterdam, The Netherlands; Email: e.licht@vumc.nl.

Therefore, we aimed our research at studies in general practice and the community. Previous research has shown that 76% of depressed adults recover within 1 year.[4] We hypothesized that depression outcome is worse in older age categories compared to this overall estimate. Furthermore, we hypothesized that this might be explained by prognostic factors that are more prevalent among older age groups.

Methods

Identification and selection of the literature

We conducted a systematic, computerized search of Medline (1966 through December 2005) and PsycINFO (1967 through December 2005) based on recommendations by Haynes *et al.*[5] Key words and MeSH headings relating to depression, longitudinal design, age (55 years and older) and setting (general practice or community) were used. For details see Box 1.

All citations (n = 1826) were screened by one reviewer (EL-S). Studies potentially eligible for inclusion were discussed by two reviewers (EL-S and DW) during a consensus meeting (n = 162, 8.9%). The reference lists of all selected publications were checked to retrieve relevant publications which had not been identified by the computerized search. Experts (ATFB and HM) were consulted to identify missing cohort studies. The publications had to meet the following selection criteria.

- The study enrolled patients diagnosed with depression. Depression could be defined as depressive disorder according to DSM-IV[52] criteria or as clinically relevant depressive symptoms not fulfilling DSM criteria ('cohort').
- The setting of the cohort was in general practice or the community ('setting').
- Subjects were 55 years or older at baseline ('age').
- The study is a prospective cohort study, presenting at least one follow-up measurement including results on depressive symptoms ('longitudinal data collection').
- The study included an outcome measurement of depression, either using DSM criteria or clinically relevant depressive symptoms ('design').
- Results were published as a full report before December 2005.

BOX 1 *Key words and medical subject headings used for literature search*

depression, depressive disorder, Depression-Emotion, Major-Depression, aged, middle aged, old*, agin*, elderly*, Geriatric-Psychiatry, morbidity, mortality, cause of death, prognos*, predict*, course*, longitudinal, follow-up, followup, cohort*, survival, cohort studies, prospect*, family medicine, general practi*, family practi*, family physician*, primary care, primary health care, family doctor*, communit*, population*, human, not case report, not case study, not clinical case report.

Quality assessment

The methodological quality of each of the studies was assessed independently by two reviewers (EL-S and DW). A standardized checklist of predefined criteria was used, which is a modified version of the checklists by Kuijpers *et al.*[6] and is based on theoretical considerations and methodological aspects described by Hudak *et al.*[7] and Altman[8] (Table 1). Disagreement among the reviewers was resolved during a consensus meeting. The list contains items regarding the study population, response to follow-up, treatment, outcome, prognostic factors and data presentation. A detailed explanation of each criterion is given in the Appendix. Each methodological quality criterion was rated as positive, negative (sufficient information, but potential bias) or inconclusive (insufficient information presented). A total score was calculated by summing the number of positively scored criteria (range 0–16). A priori we chose to consider a study of 'high quality' when it scored more than 10 points (>60% of the maximum attainable score) and of 'low quality' when it scored 10 or lesser points.

TABLE 1 *Criteria list for assessing the methodological quality of prognostic cohort studies on depression in older persons*

Criteria		Score
Study population		
A	Inception cohort	+/–/?
B	Description of study population	+/?
C	Definition of depression	+/–/?
D	Number of subjects in study population ⩾100	+/–
Response		
E	Response rate ⩾75%	+/–/?
F	Information about non-responders versus responders	+/–/?
Follow-up (extent and length)		
G	Prospective data collection	+/–/?
H	Follow-up of at least 6 months	+/–/?
I	Dropouts/loss to follow-up <20%	+/–/?
J	Information about completers versus loss to follow-up/dropouts	+/–/?
Treatment		
K	Description of possible treatment in cohort	+/–/?
Outcome		
L	Standardized assessment of depression outcome	+/–/?
Prognostic factors		
M	Standardized assessment of potential prognostic factors	+/–/?
Data presentation		
N	Frequencies of most important outcome measures presented	+/–/?
O	Frequencies of most important prognostic factors presented	+/–/?
P	Influence of prognostic factors presented	+/–/?

The symbol '+' indicates positive (sufficient information and a positive assessment);'-', negative (sufficient information, but potential bias) and '?', unclear (insufficient information).

Data extraction

For each study, we extracted data regarding study population, design, setting, outcome measures, prognostic factors and strength of association with a poor outcome of depression. The results were stratified by setting (general practice and community). The associations between prognostic factors and outcome were often expressed by relative risks (RRs) or odds ratios (ORs). In some studies, mean differences in baseline scores were presented for participants with depression at follow-up compared to those without depression. If not provided by the publication, but sufficient data were available, we calculated the univariate association between prognostic factors and outcome in terms of RRs or ORs with 95% confidence intervals (CIs). Univariable or, if available, multivariable associations were presented in tables.

Analyses

The prognosis of depression can be defined in different ways. Clinical remission is usually defined by a score on a depression rating scale below a preset cut-off score for depression.[9] However, remissions are often followed by recurrences. Another method for defining the course is to calculate the proportion of time the patients are depressed. In the present review, we will present the different course types as presented in the original articles. We will first report the results of successive measurements of depression outcome. Secondly, we will report the results of studies with more than two follow-up measurements of depression.

The studies in this review used a wide variety of methods to diagnose depression, which limited the possibilities of a quantitative analysis (statistical pooling of results). Furthermore, studies included a wide variety of prognostic indicators. Therefore, we decided to perform a qualitative analysis (best evidence synthesis) to summarize the available evidence for the predictive value of the prognostic indicators. In this analysis, the number of studies evaluating a specific factor, the methodological quality of these studies and the consistency of results were taken into account. Prognostic factors reported in different papers on the same cohort were counted once. Findings were consistent if $\geqslant$75% of the studies reporting on a factor showed the same direction of the association. We defined five levels of evidence which are based on Sackett *et al.*[10] and Ariëns *et al.*[11] (Table 2).

TABLE 2 *Levels of evidence for prognostic factors for unfavourable outcome of depression*

Level of evidence	
Statistical significant associations	
Strong	Consistent associations found in at least two high-quality cohorts
Moderate	Consistent associations found in one high-quality cohort and at least one low-quality cohort
Weak	Association found in one high-quality cohort or consistent associations found in at least three low-quality cohorts
Inconclusive	Association found in less than three low-quality cohorts
Inconsistent	Inconsistent findings irrespective of study quality
Associations without statistical significance	
Inconclusive	Non-significant associations found in at least two studies
Insufficient	Insufficient evidence: only one study presenting non-statistical significant association, irrespective of study quality

Results

Selection of studies

In Figure 1 we present a flow chart of our study selection. The electronic search resulted in 1826 citations, of which 67 articles were considered eligible for the review based on their abstract. Reviewing the full text resulted in the inclusion of 35 articles. Five additional articles were identified by reference checking, including one new cohort. Papers not reporting on potential prognostic factors were included if they did report data on depression outcome. Finally, 40 papers were included reporting data on 4 primary care and 17 community cohorts.

Methodological quality

The results of the quality assessment are presented in Table 3. The two reviewers agreed on 84% of all criteria. The overall quality score ranged from 7 to 14 points, with a mean of 10.5. Four of six primary care studies (67%) and 23 of 34 (67%) community studies were considered to be of relatively high quality using our cut-off of 10 points. In all, 23 of 40 articles did not present data on baseline characteristics of the depressed cohort and only 10 articles reported information on non-responders or dropouts, mainly due to the fact that in many studies the depressed subgroup was part of a larger cohort. Fifteen of 40 articles presented data on less than 100 depressed patients. Some information on treatment offered to depressed study participants was presented in 14 articles.

Course of depression

All included studies presented data on the course of depression. A variety of different diagnostic instruments were used, which can be divided into methods diagnosing depression according to the DSM criteria[52] and instruments that identify clinically relevant depressive symptoms (Table 4). The follow-up in primary care was 6–12 months. When using DSM criteria for depression, short-term persistence ($\leqslant$1 year) was 22.7–51.3%.[14,17] Using clinical measures for depression, short-term persistence was 14.8–47.6%.[13,17]. All primary care cohorts included patients who were

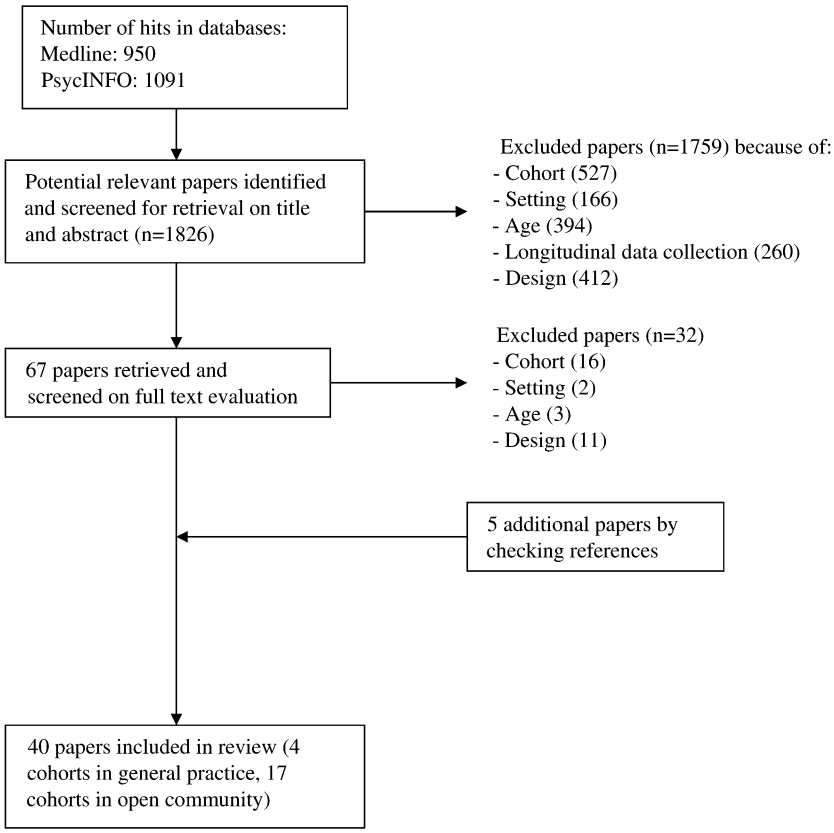

FIGURE 1 *Study selection*

screened during practice visits. No study used patients diagnosed by their GP.

The follow-up in the community studies varied between 1 and 10 years. Depression according to DSM criteria showed a short-term persistence (≤1 year) of 41.3%,[19] intermediate-term persistence (1–3 years) of 32.3–54.4%[24,30] and long-term persistence of depressive disorder of 32.3–54.3%.[24,30] Only three studies reported on the course of dysthymic disorder according to DSM criteria. After 1 year, 47.2% was still depressed,[19] after 5 years 52.4%[26] and after 6 years 52%.[23] The short-term persistence of depressive symptoms not fulfilling DSM criteria was 33.8–65%,[47,42] intermediate-term persistence was 32.3–50.4%[51,33] and long-term persistence was 13.6–61.5%.[42,46]

Course measured with more than two follow-up assessments of depression

We identified only six cohorts presenting more than two follow-up assessments of depression; two situated in general practice and four in the community (Table 5).[13,14,23,29,30,51] Regardless of the setting, the duration of follow-up or the availability of diagnoses according to DSM or not, the results showed that about one in three patients developed a chronic course. Studies with three follow-up assessments reported remission in about one in two patients. However, one study reporting results of 14 follow-up assessments showed a remission rate of 23%.[23]

Comparing the results of studies with more than two follow-up measurements with those reporting only one outcome assessment showed not only a smaller proportion of chronically depressed patients in studies with repeated measurements but also a smaller proportion of remitted patients. This may reflect the fluctuating course of depressive symptoms, which is missed in studies with few measurements.

Prognostic factors predicting poor outcome of depression

For three primary care and eight community cohorts, data have been presented on the association between potential prognostic factors and a poor outcome of depression. Not all studies provided enough data to compute ORs or RRs with CIs. However, for our best evidence synthesis, we were able to use data even if studies only presented *P*-values. In Table 6 we presented ORs or RRs where possible. Non-significant associations were summarized. An extended version of Table 6 can be found in the supplementary material online.

The prognostic value of the severity of depression at baseline was studied in two primary care cohorts, both of relatively good quality. A significant association with poor outcome was reported for one cohort.

TABLE 3 *Results of methodological assessment of prognostic cohort studies on depression in older patients*

First author	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Quality score (total '+')	Score (%)
Primary care																		
Callahan *et al.*[12]	+	+	+	+	–	+	+	+	+	?	?	+	+	+	+	+	13	81
Callahan *et al.*[13]	+	+	+	+	+	?	+	+	?	?	?	+	+	+	+	+	12	75
van Marwijk *et al.*[14]	+	?	–	+	+	+	+	+	–	?	+	–	+	+	+	+	11	69
Schulberg *et al.*[15]	+	+	+	–	–	+	+	+	–	?	?	+	+	+	+	?	10	63
van Marwijk *et al.*[16]	+	+	+	–	+	?	+	+	–	?	+	+	?	+	–	–	9	56
Lyness *et al.*[17]	+	?	+	–	?	?	+	+	?	+	?	+	+	+	?	?	8	50
Community																		
Forsell *et al.*[18]	+	?	+	+	+	?	+	+	+	+	+	+	+	+	+	+	14	88
Kivelä *et al.*[19]	+	+	+	+	?	?	+	+	+	+	+	+	+	+	+	+	14	88
Geerlings *et al.*[20]	+	+	+	+	–	+	+	+	–	+	?	+	+	+	+	+	13	81
Kivelä *et al.*[21]	+	+	+	+	?	?	+	+	+	?	+	+	+	+	+	+	13	81
Lenze *et al.*[22]	+	+	+	+	?	?	+	+	–	–	+	+	+	+	+	?	12	75
Beekman *et al.*[23]	+	+	+	+	?	?	+	+	–	+	?	+	+	+	+	+	12	75
Wilson *et al.*[24]	+	?	+	+	+	?	+	+	+	?	+	+	+	?	+	+	12	75
Kivelä *et al.*[25]	+	+	+	+	?	?	+	+	+	?	+	+	+	+	?	+	12	75
Kivelä *et al.*[26]	+	+	+	–	?	?	+	+	+	?	+	+	+	+	+	+	12	75
Kivelä *et al.*[27]	+	+	+	+	?	?	+	+	+	+	?	+	?	+	+	+	12	75
Schoevers *et al.*[28]	+	?	+	+	–	+	+	+	–	?	?	+	+	+	+	+	11	69
Kivelä *et al.*[29]	+	+	+	–	?	?	+	+	?	?	+	+	+	+	+	+	11	69
Sharma *et al.*[30]	+	+	+	+	?	?	+	+	–	?	?	+	+	+	+	+	11	69
Beekman *et al.*[31]	+	?	+	+	+	?	+	+	–	+	?	+	+	+	?	+	11	69
Braam *et al.*[32]	+	?	+	+	+	?	+	+	–	?	?	+	+	+	+	+	11	69
Geerlings *et al.*[33]	+	?	+	+	–	+	+	+	?	?	?	+	+	+	+	+	11	69
Schoevers *et al.*[34]	+	?	+	+	?	?	+	+	?	+	?	+	+	+	+	+	11	69
Kennedy *et al.*[35]	+	+	+	+	?	?	+	+	?	?	+	?	+	+	+	+	11	69
Zarit *et al.*[36]	+	?	+	+	–	+	+	+	?	+	?	+	+	+	?	?	10	63
Denihan *et al.*[37]	+	?	+	+	?	?	+	+	+	?	+	+	+	+	?	?	10	63
Green *et al.*[38]	+	?	+	+	–	?	+	+	+	?	+	+	?	+	?	+	10	63
Kivelä *et al.*[39]	+	?	+	+	?	?	+	+	+	?	+	+	+	+	?	?	10	63
Copeland *et al.*[40]	+	?	+	+	?	?	+	+	+	?	+	+	+	+	?	?	10	63
Beekman *et al.*[41]	+	?	+	–	?	+	+	+	+	?	?	+	+	+	?	?	9	56
Pitkälä *et al.*[42]	+	+	+	–	?	?	+	+	+	?	–	+	+	+	?	?	9	56
Oiji *et al.*[43]	+	?	+	–	+	?	+	+	+	?	+	+	?	+	?	?	9	56
Forsell *et al.*[44]	+	?	+	–	?	?	+	+	+	?	+	+	+	+	?	?	9	56
Kua *et al.*[45]	+	?	+	–	?	?	+	+	+	?	+	+	+	+	?	?	9	56
Frojdh *et al.*[46]	+	+	+	–	–	+	+	+	–	?	?	+	?	+	?	?	8	50
Braam *et al.*[47]	+	?	+	–	–	?	+	+	?	?	?	+	+	+	?	+	8	50
Ben Arie *et al.*[48]	+	?	–	–	?	?	+	+	+	?	+	+	+	+	?	?	8	50
Haynie *et al.*[49]	+	?	+	–	–	+	+	+	?	?	?	+	+	+	?	?	8	50
Penninx *et al.*[50]	+	?	+	+	?	?	+	+	?	?	?	+	+	+	?	?	8	50
Musil *et al.*[51]	+	?	+	–	–	?	+	+	?	?	?	+	+	+	?	?	7	44

Consequently, the best evidence synthesis indicates weak evidence for the predictive value of baseline depression severity. Likewise, weak evidence was found for better overall functioning at baseline. For primary care settings, we found no strong evidence for any prognostic factor. In community studies strong evidence (i.e. significant associations with poor outcome in at least two high-quality cohorts) was found for older age, the presence of chronic somatic diseases, the presence of functional limitations, higher baseline depression level and an external locus of control. The best evidence synthesis showed moderate evidence for an effect of religion on the prognosis of depression. Weak evidence for an association was found for lower education in men, drinking beer, presence of comorbid generalized anxiety disorder and pain, personal and family history for depression, diurnal variation of symptoms, low self-perceived health in women, loneliness and life dissatisfaction. The extended version of Table 6 presents the associations found in the individual studies.

Discussion

The available studies in the community and general practice suggest that the prognosis of late-life depression is poor in 20–50% of those afflicted, regardless of the way depression was defined at baseline and regardless of the duration of follow-up. Compared with adults (18–64 years of age), in whom 76% have been found to recover within 1 year,[4] this suggests that the prognosis deteriorates with ageing. Our second

TABLE 4 *Course of depression in primary care and in community (data of two measurements).*

First author	Study quality (%)	Number of patients	Age; mean (SD)	Female (%)	Setting	Dropout	Diagnostic criteria	Length of follow-up	Outcome measures
Primary care									
DSM criteria									
van Marwijk[14,a]	69	46 MD 88 BD	⩾65; mean?	?	Patients consulting GPs, The Netherlands	26%	DIS GDS-30 ⩾11 Non-structured interview by GP	6 months	MD T0: MD 51.3% BD T0: MD 11.7%
Schulberg[15]	63	36	⩾60; mean 66.9 (7.1)	78%	Patients visiting internal medicine centers, Pennsylvania, USA	24%	CES-D ⩾11, plus psychiatric interview	6 months	MD: 38.5%
Callahan[13]	75	292	⩾60; mean?	81%	Patients visiting an academic primary care practice, USA.	?	CES-D ⩾16	9 months	47.6%
van Marwijk[14,a]	69	46 MD 88 BD	⩾65; mean?	?	Patients consulting GPs, The Netherlands	26%	DIS GDS-30 ⩾11 Non-structured interview by GP	1 year	MD T0: MD 46.2% CD: 33.3% BD T0: MD 10%
Lyness[17]	50	22 MD 14 MI 27 SD	⩾60; mean 71. 1 (7.5)	59%	Patients visiting internal medicine clinics, USA	?	CES-D, SCID, HAM-D DS: CES-D ⩾21 MD: SCID >4 symptoms MI: SCID 2–4 symptoms SD: HAM-D >10	1 year	MD T0: MD 22.7%, MI 0%, SD 45.4% MI T0: MD 14.3%, MI 14.3%, SD 21.4% SD T0: MD 7.4%, MI 7.4%, SD 37.0%
Community									
DSM criteria									
Kivelä[21,b]	75	42 MD 199 dysth.	⩾60; mean: male 69.6 (7.2) and female 71.3 (7.7).	MD 69.0% Dysth. 67.3%	Survey in Ähtäri, Finland	MD 0.0% Dysth. 1.7%	ZSDS HAM-D	1 year	Dysth.: relapse 4.5%, continuously ill 47.2% MD: relapse 13.2%, continuously ill 16.7%
Wilson[24]	75	483	⩾65; mean?	?	Drawn from register of the Liverpool Family Practitioner Committee, UK	15%	GMS-Agecat Depression levels 3–5	2 years	Depression: 32.3%
Oiji[43]	56	43	⩾65; mean?	69.8%	Survey in Nagai City, Japan	10%	SCID-NP Non-structured interview by psychiatrist 'Clinically depressed'	2 years	Depression 34.4%
Forsell[18]	88	78 MD 39 dysth.	⩾75; mean?	?	Community Stockholm, Sweden	1.7%	CPRS	3 years	MD: MD 37.8%, dysth.: 2.7% Dysth.: MD 8.3%, dysth.: 25.0%
Schoevers[34]	69	236	65–84. mean?	?	Non-institutional individuals, Amsterdam, The Netherlands	?	GMS-Agecat Depression levels 3–5 SC levels 1–2	3 years	Depression: 51.7% SC: 19.1%
Copeland[40,c]	63	123 MD 114 SC	⩾65; mean?	?	Random sample from GPs' lists, Liverpool, UK	MD 6.3% SC 19.4%	GMS-Agecat MD; SC	3 years	MD: SC 17.1%, MD 40.2% SC: SC 16.5%, MD 21.5%

TABLE 4 *Continued*

First author	Study quality (%)	Number of patients	Age; mean (SD)	Female (%)	Setting	Dropout	Diagnostic criteria	Length of follow-up	Outcome measures
Denihan[37]	63	127	⩾65; mean?	?	Individuals on practice list of seven GPs, Dublin, Ireland	16.5%	GMS-Agecat Depression levels 3–5	3 years	Persistent or relapsed case-level depression: 34.9%
Sharma[30,c]	69	120	⩾65; mean?	?	Random sample from GPs' lists, Liverpool, UK	17%	GMS-Agecat	3 years	Depression: 54.4%
Wilson[24]	75	483	⩾65; mean?	?	Drawn from register of the Liverpool Family Practitioner Committee, UK	18%	GMS-Agecat Depression levels 3–5	4 years	Depression: 32.3%
Kivelä[26,b]	75	42	⩾60; mean: male 73.2 (6.3) and female 73.0 (7.3)	69.0%	Survey in Ähtäri, Finland	8.7%	ZSDS HAM-D	5 years	Depression: 52.4%
Sharma[30,c]	69	120	⩾65; mean?	?	Random sample from GPs' lists, Liverpool, UK	27%	GMS-Agecat	5 years	Depression: 54.3%
Kua[45]	53	35 MD	⩾65; mean?	?	Open community, Singapore	MD 13.3%	GMS-Agecat	5 years	MD: MD 38.5%, SC 19.2%
		28 SC				SC 12.0%	MD levels 3–5 SC levels 1–2		SC: MD 13.6%, SC 18.2%
Kivelä[21,b]	81	199	⩾60; mean: male 69.6 (7.2) and female 71.3 (7.7)	67.3%	Survey in Ähtäri, Finland	1.7%	ZSDS HAM-D: dysth.	5 years	Continuously ill: 36.9%
Beekman[23,d]	75	277	55–89; mean 71.8 (8.8)	65%	Population register, stratified for age and sex, The Netherlands	38%	CES-D ⩾16, DIS	6 yrs	MD: CI 35%, C 35% Dysth.: CI 20%, C 52% DD: CI 14%, C 77% SD: CI 36%, C 25%
Depressive symptoms									
Musil[51]	44	62	⩾65; mean?	?	Selected from Medicare lists, Cuyahoga County, Ohio, USA	?	CES-D ⩾16	9 months	Depression 48.4%
Pitkälä[42]	53	98	75, 80 and 85.	72.4%	Open population, Helsinki, Finland	7.1%	ZSDS ⩾45: 'lowered mood'	1 year	Lowered mood 33.8%
Braam[47,d]	50	48[g]	55–89; mean?	50.0%	Population register, stratified for age and sex, The Netherlands	?	CES-D ⩾16	1 year	Chronic depression (CES-D ⩾16 at all measures): 65%
Musil[51]	44	62	⩾65; mean?	?	Selected from Medicare lists, Cuyahoga County, Ohio, USA	?	CES-D ⩾16	1.5 years	Chronic depression 32.3% Recurrent depression 9.7%
Kennedy[35]	69	211	⩾65; mean?	?	Representative sample of Medicare recipients residing in a Bronx community, USA	?	CES-D ⩾16	2 years	Persistent depression: 46.0%
Zarit[36]	59	109	84, 86, 88, and 90	?	Population-based cohort, including people living in specialized housing for elderly, Jönköping, Sweden	?	CES-D (11 items) ⩾9: significant depression	2 years	Depressed 74.6%

Haynie[49]	47	20	⩾80; mean?	?	Twins, of which one twin was randomly selected, Sweden	?	CES-D ⩾16	2 years	Clinic depression 45.0%
Lenze[22]	75	377	⩾65; mean 72 (5.3)	76%	Random sample, four communities in USA	29.2%	CES-D (10 items) ⩾10	3 years	Temporarily depressed 68.4% Persistent depression: 31.6%
Geerlings[33,d]	69	327	55–89; mean?	?	Population register, stratified for age and sex, The Netherlands.	?	CES-D ⩾16	3 years	Persistence of depression 50.4%
Ben-Arie[48]	47	23	⩾65; mean?	?	Community psychosocial survey, Cape Town, South Africa	13.0%	CATEGO-tentative diagnosis of depressive disorder or additional clinical information	3.5 years	Clinically depressed 45.0%
Fröjd[46]	47	71	⩾65; mean 75.5 (11.6)	66.2%	Open population, Karlstad, Sweden.	38.1%	HSCL-25 ⩾1.75 of mean: 'depression in need of treatment'	6 years	'High score' 61.5%
Pitkälä[42]	53	98	75, 80 and 85.	72.4%	Population based, Helsinki, Finland	11.2%	ZSDS ⩾45: lowered mood	10 years	10 year: lowered mood 13.6%

Instruments for DSM diagnoses of depression: DIS, Diagnostic Interview Schedule; SCID, Structured Clinical Interview for DSM-III-R; GMS-Agecat, Geriatric Mental State, Automated Geriatric Examination for Computer Assisted Taxonomy; CPRS, Comprehensive Psychopathological Rating Scale.
Instruments for diagnosing depressive symptoms, not according to DSM: GDS-30, Geriatric Depression Scale (30 items); CES-D, Center for Epidemiologic Studies Depression Scale; HAM-D, Hamilton Rating Scale for Depression; HSCL, Hopkins Symptom Checklist; ZSDS, Zung Self-Rating Depression Scale.
MD, major depression; dysth., dysthymic disorder; SC, subcases; DD, double depression: MD and dysth.; SD, subthreshold depression; DS, depressive symptoms; MI, minor depression; BD, borderline depression: not fulfilling depression criteria and GDS-30 ⩾11 or considered depressed by GP; CD, chronic depression: depressed at 6 and 12 months; T0 (baseline); CI, chronic intermittent; C, chronic course.
Data derived from cohort of the following:
[a]van Marwijk *et al.*
[b]Kivelä *et al.*
[c]Sharma *et al.*
[d]Longitudinal Aging Study Amsterdam.
[e]Number with dropouts excluded.

TABLE 5 *Course of depression related to the number of measurements*

First author	Number of measurements	DSM diagnosis	Two measurements remission (%)	Repeated measurements	
				Course types	% (*n*)
General practice					
Callahan[13]	0, 6, 9 months	No	53.1	Chronic (HHH)	35.7 (70)
				Intermittent (HLH)	11.2 (22)
				Remission (HHL + HLL)	53.1 (104)
van Marwijk[14]	0, 6, 12 months	Yes	53.8	Chronic (HHH)	33.3 (13)
				Intermittent (HLH)	12.8 (5)
				Long-term remission (HHL)	17.9 (7)
				Short-term remission (HLL)	35.9 (14)
Community					
Musil[51]	0, 9, 18 months	No	58.1	Chronic (HHH)	32.3 (10)
				Intermittent (HLH)	9.7 (6)
				Long-term remission (HHL)	16.1 (10)
				Short-term remission (HLL)	41.9 (26)
Sharma[30]	0, 3, 5 years	Yes	45.7	Chronic (HHH)	34.8 (16)
				Intermittent (HLH)	19.6 (9)
				Long-term remission (HHL)	19.6 (9)
				Short-term remission (HLL)	26.1 (12)
Kivelä[29]	0, 1, 5 years	Yes	55.6	Chronic (HHH)	31.1 (47)
				Intermittent (HLH)	13.2 (20)
				Long-term remission (HHL)	22.5 (34)
				Short-term remission (HLL)	33.1 (50)
Beekman[23]	Every 5 months during 6 years	No	51 (3 years)	Chronic	32[a] (90)
				Chronic intermittent	32[a] (90)
				Remission with recurrence	12[a] (34)
				Remission	23[a] (63)

[a]Chronic: depressed >80% observations; chronic intermittent: more than one remission, followed by recurrence of symptoms; remission with recurrence: remission with later in the study a relevant increase of symptoms; remission: relevant decline of symptoms and remaining non-depressed.
H, score above cut-off for depression; L, score below cut-off for depression.

TABLE 6 *Prognostic factors and strength of association for unfavourable outcome of depression in older patients in primary care or community*

Prognostic factor	Strength of association (95% CI)	Quality score >60%	Quality score ⩽60%	Level of evidence
Primary care				
Demographic characteristics				
Age	Not significant[13,14]	0/2		Inconclusive
Female gender	Not significant[13,14]	0/2		Inconclusive
White race	Not significant[13]	0/1		Insufficient
Living alone	Not significant[14]	0/1		Insufficient
Living independent	Not significant[14]	0/1		Insufficient
⩽8 years of education	Not significant[13]	0/1		Insufficient
Co-morbidity (somatic or psychiatric)				
Mean number of diagnoses at baseline	Not significant[13]	0/1		Insufficient
Alcoholism	Not significant[13]	0/1		Insufficient
Dementia	Not significant[13]	0/1		Insufficient
Cognitive functioning	MMSE scores: not significant[15]	0/1		Insufficient
Depression characteristics				
Baseline depression level	CES-D baseline score (per point), MD 0.39 (0.33–0.45)*, adjusted for physical health[12] Not significant[15]	1/2		Weak
Quality of life				
Better overall functioning	MD 5.9 (0.6–11.2), $P = 0.03$[15]	1/1		Weak

TABLE 6 *Continued*

Prognostic factor	Strength of association (95% CI)	Quality score >60%	Quality score ⩽60%	Level of evidence
Community				
Demographic characteristics				
Age	MD 3.5 years (1.68–5.32)*,[35] Age ⩾75 ($P = 0.002$)*,[23] Age ⩾75 ($P < 0.05$)*,[30] Not significant[22,25,34,37]	2/6	0/1	Strong
Female gender	RR 0.76 (0.58–0.99)*,[25] OR 3.80 (1.07–13.5)*,[47] Not significant[22,30,34,35,37]	1/6	1/1	Inconsistent
White race	Not significant[22]	0/1		Insufficient
Marital status: widowed/ divorced/unmarried	Not significant[25,30,34,35,47]	0/4	0/1	Inconclusive
Living with spouse	Not significant[25,35]	0/2		Inconclusive
Education	Male: ⩽ primary school: 2.06 (1.27–3.35)*,[25] Female: not significant[25] Not significant[31,34]	1/1 0/1 0/2		Weak, only in men
Socio-economic status	Not significant[22,25,30,34,35,47]	0/5	0/1	Inconclusive
Religion	Jewish versus not Jewish: RR 1.41 (1.02–1.93)*,[35] Catholic versus not Catholic: not significant[35] None versus any: OR 5.85 (1.52–22.6)*,[47]	1/1	1/1	Moderate
Stressful events	Bereavement: $P < 0.01$[37] Social or health stress factors: not significant[25,30]	0/2	1/1	Inconclusive
Social support	Not significant[25,30,31,34,37]	0/5		Inconclusive
External locus of control	OR 1.15 (1.06–1.24)*,[31] Poor self-appreciation OR 4.6 (1.08–19.3)*,[29]	2/2		Strong
Drinking beer	OR 11.7 (1.49–91.1)*,[29]	1/1		Weak
Smoking	Not significant[25,38]	0/2		Inconclusive
Co-morbidity (somatic or psychiatric)				
Chronic somatic diseases	Any versus none: OR 1.45 (1.12–1.87)*,[31] Serious illness: RR 1.52 (1.15–2.01)*,[35] Number of chronic medical conditions: $P < 0.001$*,[37] Not significant[25,30,34]	3/6		Strong
Generalized anxiety disorder	Anxiety score: $P = 0.01$*,[30] Not significant[28]	1/2		Weak
Dementia	Not significant[18,23]	0/2		Inconclusive
Depression characteristics				
Baseline depression level	HRSD score: MD 2.6 (0.8–4.4)*,[19] Depression score: $P < 0.05$*,[30] Not significant[31]	2/3		Strong
Depression type	Not significant[30]	0/1		Insufficient
Diurnal variation of symptoms	OR 3.9 (1.10–13.5)*,[29]	1/1		Weak
Personal history of depression	RR 1.35 (1.06–1.71)*,[34] Not significant[25]	1/2		Weak
Family history of depression	$P < 0.001$*,[37] Not significant[34,38]	1/3		Weak
Sleep disturbance	Not significant[35]	0/1		Insufficient
Quality of life				
Functional limitations	More than one versus none: OR 2.78 (1.16–6.66)*,[20] ADL disability: RR 1.44 (1.14–1.83)*,[34] Not significant[35,37,38]	2/5		Strong
Pain	OR 2.56 (1.49–4.42)*,[33]	1/1		Weak
Low self-perceived health	Female: $P = 0.047$*,[25] Men: not significant[25] Not significant[35]	1/1 0/1 0/1		Weak, only in women
Loneliness	OR 12.8 (4.01–37.3)*,[38]	1/1		Weak
Life dissatisfaction	OR 14.5 (3.86–60.96)*,[38]	1/1		Weak

MD, mean difference; CES-D, Center for Epidemiologic Studies Depression Scale; MMSE, mini-mental state exmaination; ADL, activities of daily living; *$P < 0.05$.

[a]Adjusted for sex, age and index depression score.

finding was that unfavourable outcome of depression in the elderly is associated with increasing age. The third finding was that somatic co-morbidity and functional limitations, which are more prevalent in the elderly, are associated with poor depression outcome. These findings support our hypotheses.

Course of depression

There is an ongoing discussion on the definition of outcome of depression. To analyse the effect of treatment, clinical trails usually measure clinical response rates defined by a 50% or greater reduction from baseline scores on depression rating scales. This response may be different from recovery, for patients may suffer from residual symptoms. It is known that residual symptoms are an important risk for relapse. Furthermore, remissions are often followed by recurrences. Therefore, a fluctuating or chronic intermittent course type is best studied in designs with more than two measurements.[23] However, only 6 of the 21 cohorts presented data on more than two measurements of depression.

The included studies showed that about one in three depressed patients developed a chronic course and about one in three patients had a short-term remission. This proportion hardly decreased with increasing length of follow-up and did not differ between the two settings or the use of different diagnostic criteria. A previous review in 1999 found four studies in primary care and eight studies in the community.[53] They used a different search strategy and definition of primary care and general practice. Yet, our conclusions were in concordance with their findings. Compared to younger patients, older patients seem to have a higher risk of recurrent episodes.[54] These findings emphasize the importance of identifying factors predicting poor depression outcome.

Prognostic factors

Only three studies carried out in general practice presented evidence on prognostic factors, and only weak evidence could be found for associations between these factors and poor outcome. These findings are insufficient to support management of depression in daily practice. However, based on eight community studies strong evidence could be found for the following prognostic factors: higher age, chronic somatic co-morbidity, more functional limitations, higher baseline depression level and an external locus of control. In a previous review Cole *et al.*[55] reported on several studies presenting prognostic factors, but did not systematically summarize this evidence.

We found six studies on the association between functional limitations and poor depression outcome; three community studies found no significant association, two community studies found a significant association with poor outcome. However, one general practice study found a significant association between better overall functioning and poor outcome. This finding in general practice is inconsistent with community studies and seems less plausible from a theoretical point of view.

Kivelä *et al.*[25] performed subgroup analyses, showing an association between poor depression outcome and education in men and low self-perceived health in women. These results illustrate that more research should be aimed at identifying the predictive value of prognostic factors across clinically relevant subgroups.

Limitations

In our best evidence synthesis we also included the results of studies presenting *P*-values only, without risk estimates (RRs or ORs). The advantage of this method is that we used all available evidence for each prognostic factor. The absence of statistical significance may be due to a lack of power or to the absence of an association. Table 3 showed that one general practice[15] and two community cohorts[47,29] presenting prognostic factors included less than 100 participants. Furthermore, several studies investigated more than one prognostic factor simultaneously (range 1–12) resulting in reduced power. This makes it difficult to interpret non-significant associations in studies only presenting *P*-values. When can we state with confidence that a factor has no association with the outcome? For the present review, we hoped to identify prognostic factors predicting poor outcome in depression in older patients that are relevant to daily practice. Therefore, we are particularly interested in those factors with (strong) evidence for an association.

None of the studies presented sufficient data on the treatment for depression. This makes it impossible to assess which patients were diagnosed with depression nor whether treatment may have improved outcome. We may hypothesize that more severe depression will probably be better recognized and more often treated than minor or subthreshold depression. However, due to the shortcomings in reporting treatment and the absence of standardization of treatment, we could not test this hypothesis. Furthermore, all general practice studies included patients who had been screened during a scheduled visit for any reason to their GP. We found no studies in which a diagnosis made by the GP was used as inclusion criterion, which may limit generalizability of the results to daily practice.

Implications for daily practice

Our results emphasize the need for adequate treatment in a large proportion of depressed older patients in general practice, but not in all. We found strong evidence for an association between several prognostic factors and poor outcome of depression in community studies. Future research should validate these factors in general practice settings. Identifying patients at risk for a chronic course of depression may help to design stepped care programs with tailor-made interventions for depression. Until then,

clinicians should be aware of several factors that may be associated with poor outcome in depressed elderly.

Supplementary data

Supplementary Table 6 is available at *Family Practice* online (http://fampra.oxfordjournlas.org/).

Acknowledgements

We would like to thank Ingrid Riphagen, Medical Information Specialist, Vrije Universiteit, Library VUmc, Amsterdam, for her help with the search of the literature.

Declaration

Funding: None.
Ethical approval: None.
Conflicts of interest: None.

References

[1] Licht-Strunk E, van der Kooij KG, van Schaik DJ *et al*. Prevalence of depression in older patients consulting their general practitioner in The Netherlands. *Int J Geriatr Psychiatry* 2005; **20:** 1013–1019.

[2] Cuijpers P, Smit F. Excess mortality in depression: a meta-analysis of community studies. *J Affect Disord* 2002; **72:** 227–236.

[3] Frazer CJ, Christensen H, Griffiths KM. Effectiveness of treatments for depression in older people. *Med J Aust* 2005; **182:** 627–632.

[4] Spijker J, de Graaf R, Bijl RV, Beekman AT, Ormel J, Nolen WA. Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Br J Psychiatry* 2002; **181:** 208–213.

[5] Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *J Am Med Inform Assoc.* 1994; **1:** 447–458.

[6] Kuijpers T, van der Windt DA, van der Heijden GJ, Bouter LM. Systematic review of prognostic cohort studies on shoulder disorders. *Pain* 2004; **109:** 420–431.

[7] Hudak PL, Cole DC, Frank JW. Perspectives on prognosis of soft tissue musculoskeletal disorders. *Int J Rehabil Res.* 1998; **21:** 29–40.

[8] Altman DG. Systematic reviews of evaluations of prognostic variables. *Br Med J* 2001; **323:** 224–228.

[9] Dawson MY, Michalak EE, Waraich P, Anderson JE, Lam RW. Is remission of depressive symptoms in primary care a realistic goal? A meta-analysis. *BMC Fam Pract* 2004; **5:** 19.

[10] Sackett DL, Straus SE, Richardson WS. *Evidence-Based Medicine. How to Practice and Teach EBM. Guidelines.* Edinburg: Churchill Livingstone, 2000.

[11] Ariëns GA, van Mechelen W, Bongers PM, Bouter LM, van der Wal G. Physical risk factors for neck pain. *Scand J Work Environ Health* 2000; **26:** 7–19.

[12] Callahan CM, Wolinsky FD, Stump TE, Nienaber NA, Hui SL, Tierney WM. Mortality, symptoms, and functional impairment in late-life depression. *J Gen Intern Med* 1998; **13:** 746–752.

[13] Callahan CM, Hui SL, Nienaber NA, Musick BS, Tierney WM. Longitudinal study of depression and health services use among elderly primary care patients. *J Am Geriatr Soc.* 1994; **42:** 833–838.

[14] van Marwijk H, Hermans J, Springer MP. The prognosis of depressive disorder in elderly primary care patients. An exploratory observational study: course at 6 and 12 months. *Scand J Prim Health Care* 1998; **16:** 107–111.

[15] Schulberg HC, Mulsant B, Schulz R, Rollman BL, Houck PR, Reynolds CF III. Characteristics and course of major depression in older primary care patients. *Int J Psychiatry Med* 1998; **28:** 421–436.

[16] van Marwijk HW. Depressie bij ouderen in de huisartsenpraktijk. [Depression in the elderly in family practice]. *Tijdschr Gerontol Geriatr* 1997; **28:** 69–75.

[17] Lyness JM, Caine ED, King DA, Conwell Y, Duberstein PR, Cox C. Depressive disorders and symptoms in older primary care patients: one-year outcomes. *Am J Geriatr Psychiatry* 2002; **10:** 275–282.

[18] Forsell Y, Jorm AF, Winblad B. The outcome of depression and dysthymia in a very elderly population: results from a three-year follow-up study. *Aging Ment Health* 1998; **2:** 100–104.

[19] Kivelä SL, Pahkala K. The prognosis of depression in old age. *Int Psychogeriatr* 1989; **1:** 119–133.

[20] 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.

[21] Kivelä SL, Kongas-Saviaro P, Pahkala K, Kesti E, Laippala P. Five-year prognosis for dysthymic disorder in old age. *Int J Geriatr Psychiatry* 1993; **8:** 939–947.

[22] Lenze EJ, Schulz R, Martire LM *et al*. The course of functional decline in older people with persistently elevated depressive symptoms: longitudinal findings from the cardiovascular health study. *J Am Geriatr Soc.* 2005; **53:** 575.

[23] Beekman AT, Geerlings SW, Deeg DJ *et al*. The natural history of late-life depression: a 6-year prospective study in the community. *Arch Gen Psychiatry* 2002; **59:** 605–611.

[24] Wilson KC, Copeland JR, Taylor S, Donoghue J, McCracken CF. Natural history of pharmacotherapy of older depressed community residents. The MRC-ALPHA Study. *Br J Psychiatry* 1999; **175:** 439–443.

[25] Kivelä SL, Koengaes-Saviaro P, Kesti E, Pahkala K. Five-year prognosis for depression in old age. *Int Psychogeriatr* 1994; **6:** 69–78.

[26] Kivelä SL. Long-term prognosis of major depression in old age: A comparison with prognosis of dysthymic disorder. *Int Psychogeriatr* 1995; **7:** 69–82.

[27] Kivelä SL, Pahkala K, Laippala P. A one-year prognosis of dysthymic disorder and major depression in old age. *Int J Geriatr Psychiatry* 1991; **6:** 81–87.

[28] Schoevers RA, van Tilburg W, Beekman ATF, Deeg DJH. Depression and generalized anxiety disorder: co-occurrence and longitudinal patterns in elderly patients. *Am J Geriatr Psychiatry* 2005; **13:** 39.

[29] Kivelä SL, Viramo P, Pahkala K. Factors predicting chronicity of depression in elderly primary care patients. *Int Psychogeriatr* 2000; **12:** 183–194.

[30] Sharma VK, Copeland JR, Dewey ME, Lowe D, Davidson I. Outcome of the depressed elderly living in the community in Liverpool: a 5-year follow-up. *Psychol Med* 1998; **28:** 1329–1337.

[31] Beekman ATF, Deeg DJH, Geerlings SW, Schoevers RA, Smit JH, van Tilburg W. Emergence and persistence of late life depression: A 3-year follow-up of the longitudinal aging study Amsterdam. *J Affect Disord* 2001; **65:** 131–138.

[32] Braam A, Beekman A, Deeg D, van Tilburg W. Gereformeerde depressie of depressies bij gereformeerden? Gegevens uit ouderenonderzoek [Calvinist depression or depression in Calvinists? Data from a study of the elderly]. *Psyche en Geloof* 2000; **11:** 114–129.

[33] Geerlings SW, Twisk JWR, Beekman ATF, Deeg DJH, van Tilburg W. Longitudinal relationship between pain and

depression in older adults: Sex, age, and physical disability. *Soc Psychiatry Psychiatr Epidemiol* 2002; **37:** 23–30.

[34] Schoevers RA, Beekman ATF, Deeg DJH, Hooijer C, Jonker C, van Tilburg W. The natural history of late-life depression: results from the Amsterdam Study of the Elderly (AMSTEL). *J Affect Disord* 2003; **76:** 5–14.

[35] Kennedy GJ, Kelman HR, Thomas C. Persistence and remission of depressive symptoms in late life. *Am J Psychiatry* 1991; **148:** 174–178.

[36] 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.

[37] Denihan A, Kirby M, Bruce I, Cunningham C, Coakley D, Lawlor BA. Three-year prognosis of depression in the community-dwelling elderly. *Br J Psychiatry* 2000; **176:** 453–457.

[38] Green BH, Copeland J-RM, Dewey ME, Sharma V. Factors associated with recovery and recurrence of depression in older people: a prospective study. *Int J Geriatr Psychiatry* 1994; **9:** 789–795.

[39] Kivelä SL. Depression in the aged. *Psychiatria Fennica* 1999; **30:** 132–144.

[40] Copeland JR, Davidson IA, Dewey ME *et al*. Alzheimer's disease, other dementias, depression and pseudodementia: prevalence, incidence and three-year outcome in Liverpool. *Br J Psychiatry* 1992; **161:** 230–239.

[41] Beekman AT, Deeg DJ, Smit JH *et al*. Dysthymia in later life: a study in the community. *J Affect Disord* 2004; **81:** 191–199.

[42] Pitkälä K, Kahonen-Vare M, Valvanne J, Strandberg TE, Tilvis RS. Long-term changes in mood of an aged population: repeated Zung-tests during a 10-year follow-up. *Arch Gerontol Geriatr* 2003; **36:** 185–195.

[43] Oiji A, Muraoka Y, Nadaoka T *et al*. One-year prognosis of depressive illness in the elderly population in Japan. *Psychiatry Clin Neurosci* 1998; **52:** 391–395.

[44] Forsell Y, Jorm AF, Winblad B. Outcome of depression in demented and non-demented elderly: observations from a three-year follow-up in a community-based study. *Int J Geriatr Psychiatry* 1994; **9:** 5–10.

[45] Kua EH. The depressed elderly Chinese living in the community: a five-year follow-up study. *Int J Geriatr Psychiatry* 1993; **8:** 427–430.

[46] Frojdh K, Hakansson A, Karlsson I, Molarius A. Deceased, disabled or depressed—a population-based 6-year follow-up study of elderly people with depression. *Soc Psychiatry Psychiatr Epidemiol* 2003; **38:** 557–562.

[47] Braam AW, Beekman AT, Deeg DJ, Smit JH, van Tilburg W. Religiosity as a protective or prognostic factor of depression in later life; results from a community survey in The Netherlands. *Acta Psychiatr Scand* 1997; **96:** 199–205.

[48] Ben Arie O, Welman M, Teggin AF. The depressed elderly living in the community. A follow-up study. *Br J Psychiatry* 1990; **157:** 425–427.

[49] Haynie DA, Berg S, Johansson B, Gatz M, Zarit SH. Symptoms of depression in the oldest old: a longitudinal study. *J Gerontol B Psychol Sci Soc Sci.* 2001; **56B:** 111–118.

[50] Penninx BWJH, Deeg DJH, van Eijk JT, Beekman ATF, Guralnik JM. Changes in depression and physical decline in older adults: a longitudinal perspective. *J Affect Disord* 2000; **61:** 1–12.

[51] Musil CM, Haug MR, Warner CD. Stress, health, and depressive symptoms in older adults at three time points over 18 months. *Issues Ment Health Nurs.* 1998; **19:** 207–224.

[52] American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV).* Washington: American Psychiatric Association, 1994.

[53] Cole MG, Bellavance F, Mansour A. Prognosis of depression in elderly community and primary care populations: a systematic review and meta-analysis. *Am J Psychiatry* 1999; **156:** 1182–1189.

[54] Mitchell AJ, Subramaniam H. Prognosis of depression in old age compared to middle age: a systematic review of comparative studies. *Am J Psychiatry* 2005; **162:** 1588–1601.

[55] 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.

Appendix

Explanation of the criteria from Table 1

A	Patients were identified with current depression (inception cohort). Positive if the interval between diagnosis of depression and baseline assessment was 6 weeks or less.
B	Positive if criteria were formulated for at least age, gender and setting.
C	Positive if depression was diagnosed using structured, validated instruments.
D	Positive if the number of subjects with depression in the study population was at least 100 at baseline.
E	Positive if response rate for case finding was ⩾75%.
F	Positive if information was presented about patient/disease characteristics of responders and non-responders or if there was no selective response.
G	Positive if a prospective design was used, also positive in case of a historical cohort in which the determinants had been measured before outcome was determined.
H	Positive if the follow-up period was at least 6 months.
I	Positive if the total number of participants was ⩾80% on the last moment of follow-up compared to the number of participants at baseline. Participants who died during follow-up were excluded from this analysis.
J	Positive if demographic/clinical information (patient/disease characteristics such as age, sex and other potential prognostic predictors) was presented for completers and those lost to follow-up/dropouts at the main moment of outcome measurement, or no selective dropouts/lost to follow-up, or no dropouts.
K	Positive if treatment subsequent to inclusion in cohort is fully described or standardized. Also positive in case no reatment was given.
L	Positive if standardized questionnaires or diagnostic interviews were used regarding at least one of the following three outcome measures for each follow-up measurement: (i) depression diagnosis; (ii) depressive symptoms; and (iii) remission or recurrence.
M	Positive if standardized questionnaires or objective measurements were used at baseline of at least one of the following four clusters of potential prognostic factors: (i) sociodemographic variables [gender, age, marital status, race, social economic status, education level and urbanicity]; (ii) clinical characteristics of depression (baseline depression level, number of episodes of depression, age of onset of first depressive episode, duration of depressive symptoms, family history of depression, treatment and mental health care); (iii) psychosocial factors (social support, stressful life events, locus of control and personality); (iv) general health [co-morbidity (i.e. anxiety disorder or chronic somatic disease), functional impairment and cognition].
N	Positive if frequency, percentage or mean and median (inter-quartile range) was reported for the most important outcome measures.
O	Positive if frequency, percentage or mean and median (inter-quartile range) was reported for baseline values of the most important prognostic factors.
P	Positive if univariate estimates were provided or could be calculated for the association of a prognostic factor with outcome.