Eur Arch Psychiatry Clin Neurosci (2008) 258:476–488 DOI 10.1007/s00406-008-0821-z

ORIGINAL PAPER

Manfred M. Fichter · Gabriele Kohlboeck · Norbert Quadflieg

The Upper Bavarian longitudinal community study 1975–2004. 2. Long-term course and outcome of depression

A controlled study

Received: 5 December 2007 / Accepted: 11 April 2008 / Published online: 27 June 2008

Abstract *Objective* The study describes course and outcome over 25 years in depressed and non-depressed men and women from a large community study. Outcome measures covered psychopathology, disability, and impaired functioning. *Method* A depressive syndrome (depressed mood and three additional depressive symptoms) was defined and compared to a control condition without depressive symptoms in the seven days preceding baseline assessment. Assessments focused on three time points: baseline survey, 5-year follow-up, and 25-year follow-up. Self-rating scales as well as expert-rating interviews yielded data on a wide range of social and psychopathological risk factors and outcome measures. *Results* Among participants of all three waves ($N = 838$), the baseline prevalence for depressive syndrome was 18.1%. Depressive symptoms manifest at the first wave had substantial impact over the 25-year study. Subjects with a depressive syndrome were predisposed for later adverse mental health outcomes, more disability in social domains and reduced functionality. No long-term increase or decrease of the prevalence of the depressive syndrome was observed. *Conclusion* There is a persistent and long lasting impact of depressive syndrome, irrespective of diagnostic status, in the general population. Our results underscore the importance of sub-syndromal depressive syndrome when estimating the risk of future mental disorders and functional impairment in the long-term.

Key words depression · longitudinal community study · long-term course of illness · prediction of outcome · chronicity

M.M. Fichter, MD, Dipl-Psych (✉) · G. Kohlboeck, PhD
N. Quadflieg, Dipl-Psych
Department of Psychiatry and Psychotherapy
University of Munich (LMU)
Nussbaumstr. 7
80336 Munich, Germany
Tel.: +49-8051/68-3510
Fax: +49-8051/68-3532
E-Mail: MFichter@Schoen-Kliniken.de

M.M. Fichter, MD, Dipl-Psych
Klinik Roseneck
83209 Prien, Germany

EAPCN 821

Introduction

Most general population studies on depression have used a cross-sectional design, were limited to short time periods and focused on the current period or on point prevalence in different locations. Evidence about the outcome of depression is largely based on the results of studies performed a small selective sample of patients referred to psychiatric professionals [55]. Existing literature has emphasized the need for long-term follow-up studies of community-based samples in order to examine the stability, comorbidity, and long-term outcome of depression in the population [32, 36, 37]. Lee [32] has argued that there is still a shortage of representative prospective follow-up studies of depressive disorders covering longer periods and presenting data from community settings, primary and secondary care as well as inpatient settings. Longitudinal approaches are of importance because understanding the natural course of depression is essential for prevention and treatment programs. An unselected and representative cohort study mirrors the mental health in a population, so risk factors in association with outcome can be studied and several outcomes of one determinant can be assessed [44]. The only prospective epidemiological studies based on community samples that have assessed the same individuals at two or more time points over a longer time period are the epidemiologic catchment area survey (ECA [13]), the

Swedish Lundby study [19], the Stirling County study [42], and the Finnish UKKI study [33].

These long-term follow-up studies indicated a tendency towards stable prevalence rates over time: Murphy et al. [42] report a stable point prevalence of depression in several cross-sections over a 40-year period. Similar results were obtained by Hagnell and Gräsbeck [21] who demonstrated that depression as sole diagnosis remained quite stable over the time of 25 years. The 13-year follow-up of the epidemiologic catchment area program of the National Institute of Mental Health presents evidence for a chronic course of depressive syndromes or major depressive disorders [7]. Similarly, the Finnish results indicated no significant changes in psychiatric disorders, except for an increase in psychoses [33]. In the Zurich Cohort study of young adults, high rates of chronicity showed that 10–20% never recovered from the index depressive episode during a 10-year follow-up period [3].

Results of short-term follow-up studies yielded evidence of at least a moderate degree of stability of depression and, to a lesser extent, anxiety [12, 61]. Wittchen et al. [61] found a considerable degree of fluctuation not only in the diagnostic status and severity of specific disorders. They suggested that in adolescence, syndromes as well as full diagnoses of mental disorders have a strong tendency to wax and wane over time.

Clinical studies of the United States National Institute of Mental Health (NIMH) collaborative program on the Psychobiology of Depression examined the long-term course and outcome of depression and suggested that, after approximately 16 years, about 60% of patients had been readmitted at least once [28, 31].

The interest in studying sub-syndromal depressive symptoms is currently increasing. In the 15-year follow-up of the Zurich Cohort Study of Young Adults, Angst and Merikangas [4] found that the presence of sub-threshold depression was related to substantially higher risk of later major depression. In the NIMH collaborative Depression Study (CDS), Judd et al. [24] suggested that—during the long-term course of depression—major, minor, dysthymic, and subsyndromal symptoms alternate in the same patient and that symptomatic periods are interspersed with symptom free episodes. Other studies showed that depressive symptoms were associated with limitations in six domains of functioning (physical functioning, role functioning, social functioning, number of days in bed, current health, and being free from pain when compared with patients with no chronic conditions in a general medical outpatient sample [7, 58].

Numerous unresolved issues make it difficult to draw firm conclusions for depression regarding its course, associated impairment, and consequences. Although epidemiological studies have used of the same diagnostic instrument over time to assess chronological patterns, there is still substantial and sometimes confusing variation in findings attributable to differing case definitions and discrepancies in birth cohorts, lengths of follow-up intervals, and age ranges of the samples. Definitions of anxiety or depression used in the Lundby study are mainly based on the global clinical impression of a psychiatrist, but do not use the criteria of current standardized diagnostic systems for assessment. To date, little is known on longitudinal trends regarding the impairment or severity associated with mental disorders.

The Upper Bavarian study (UBS) is a longitudinal study of mental disorders in a representative sample of 1,668 subjects in rural South Eastern Bavaria, Germany [15, 16, Fichter et al. in this issue]. In the present study the longitudinal course of depressive syndromes was investigated over a time period of 25 years. The aims of the study were (1) to describe the longitudinal course and outcome of depression in a representative community sample using the same assessment procedure at each time, (2) to report on possible associations between depression at first assessment and mental disorders at later assessments, (3) to investigate whether depression at an earlier time point constitutes a risk factor for increased disability, reduced functionality, and higher mortality; and (4) to investigate the stability of the prevalence of identically defined depressive syndromes over time.

Methods

Sample

The study population and methods have been described in detail elsewhere [16, Fichter et al. in this issue]. In brief, the UBS study began in 1975 (t_1) [11], and random samples were drawn from the community registers in three selected areas. Of the original representative community sample of 1,536 adults aged 15 and above, 1,342 study participants were interviewed 5 years later (second assessment period or "wave", t_2: 1980–1984; [15]). At the 25-year follow-up (third wave, t_3: 2001–2004) of the 1,342 study participants interviewed at first and second wave, 390 individuals (29.1%) had died. Interviewers met with 838 individuals (88% of those alive): 61 individuals (6.4%) refused participation and 53 (5.5%) persons could not be traced or contacted. This paper reports on 838 study participants (373 men [44.5%], 465 women [55.5%]; "three waves longitudinal sample") who constitute the full sample for this paper.

Design

Data were gathered prospectively during three assessment periods or "waves" at the original survey in 1975–1979, and in two follow-ups 5 and 25 years later.

Instruments

The semi-structured "Standardized Psychiatric Interview" (SPI; [17] was employed at all three assessment periods. The SPI included subjective ratings of ten areas and 12 additional expert-ratings of abnormalities observed during the interview. Severity of each psychiatric symptom of the SPI was rated on a 5-point scale

(0 = no impairment to 4 = severe impairment) for the current (last 7 days) status. The 22 ratings for each patient were added and weighted (subjective rating × 1, expert rating × 2) to give an overall index of general psychopathology, and objective symptoms and subjective symptom scores were also summarized in two separate scales. Mean inter-rater reliability between wave 2 and wave 3 for SPI interviews was Kappa = 0.83 ± 0.01 (mean ± standard deviation). For more details on instructions to ensure identical assessment criteria for scoring SPI symptoms, see Fichter et al. (in this issue).

At each wave, the self-rating complaint list ("Beschwerdenliste"; von Zerssen [56] assessed 24 somatic, psychological and psychophysiological symptoms on a scale ranging from 0 to 3 with higher scores indicating more frequent complaints. Social class was determined according to Kleining and Moore [29]. At wave 1 poverty and poor housing standards were rated according to questions about real estate ownership, number of bedrooms, and one's satisfaction with his or her living conditions. For assessment of alcohol consumption, abuse and dependence the Munich Alcoholism Test (MALT [14]) was employed at the second and third wave. At wave 2 social maladjustment and dysfunction were assessed by the Social Interview Schedule (SIS [8]). At the third wave interview completion of the SPI was followed by the computer-assisted CIDI-Interview [46, 60]. If there were indications of psychoses or psychotic symptoms, the subjects were evaluated thoroughly with the psychosis sections of the Structured Clinical Interview for DSM-IV (SCID-I [51]). Additional psychiatric diagnoses not formally covered by the CIDI were established in expert consensus conferences. At the third wave special efforts were directed at the collection of data on the respondents' impairments and disabilities by employing the German version of the World Health Organization Disability Assessment Schedule II (WHODAS II [62]) and the Global Assessment of Functioning scale (GAF; DSM-IV [2]). Somatic illnesses and complaints were assessed by a checklist for somatic diseases and were documented according to ICD-10 based on self-report. The study was approved by the ethics committee of the medical department at the University of Munich (LMU).

Definition of depressive syndrome

Our definition of depressive syndrome was guided by the description of criterion A of the DSM-IV Major Depressive Episode:

(1) Depressed mood was identified by a severity rating of at least 1 (="slight impairment") in one of two items from the SPI ("depressive mood", "depressed").
(2) Significant weight loss required a severity rating of at least 1 in the complaint list item "weight loss".
(3) Insomnia or hypersomnia were identified by a severity rating of at least 1 in either the SPI item "sleep disturbance" or the complaint list items "sleeplessness" or "excessive sleepiness", or by the report on consumption of sleeping pills.
(4) Psychomotor agitation or retardation was confirmed by a severity rating of at least 1 for one of two relevant SPI objective symptoms ("slow, lacking spontaneity", "elated, euphoric").
(5) Fatigue or loss of energy was measured by a severity rating of at least 1 in the SPI question "fatigue" or the complaint list question "faintness".
(6) Feelings of worthlessness or inappropriate guilt were assessed by a severity rating of at least 1 in the SPI objective symptom "depressive thought content" or the complaint list item "ruminative depressive thoughts".
(7) Diminished ability to think or concentrate was measured by a severity rating of at least 1 in the corresponding SPI item.

Criteria for depressive syndrome were met if depressed mood and, in addition, at least three of the other six depressive symptoms described above were present in the last seven days preceding the baseline interview. As controls we chose an extreme, non-depressed group. Control group participants needed to be free from any depressive symptoms and thus at baseline assessment have ratings of "0" in all seven symptoms described above.

Statistical analysis

Means with standard deviations and frequencies of categorical variables are presented, with t tests and chi^2-tests for group comparisons. Fisher's exact test was used to compare proportions between groups, when any cell count in a 2 × 2 table was less than five. Effect sizes (Cohen's d) helped determine the relevance of statistically significant differences. Longitudinal course was analyzed by ANOVAs with assessment period as repeated measures factor and depressive syndrome versus non-depressed controls as a between-subjects factor. Age at baseline was included as covariate in these analyses. Logistic regression analyses established odds ratios for mental disorders at 25-year follow-up. Comparisons of prevalence of depressive syndromes between the same age cohorts at three different time points were made by 2-way contingency table analyses using chi^2-tests. Missing data varied slightly between measures, and numbers of cases are reported for each analysis. The statistical software used was SPSS for Windows, Version 13 [53].

Results

Prevalence of depressive mood, depressive thought and depressive syndrome at three-time points

As shown in Table 1 (right side) the 7-days point prevalence (total group, N = 838) of *depressive syndrome* (identically defined and assessed at all time points) was 18.1% at wave 1, 14.0% at wave 2 and 16.1% at wave 3. For "*depressive mood*" it was 30.7% at wave 1, 21.4% at wave 2 and 23.1% at wave 3. For "*depressive thought content*" it was 19.6% for wave 1, 15.0% at wave 2 and 13.6% at wave 3. Thus, there definitely was no *increase* over time in the prevalence of these variables measured with the same instrument (SPI-interview) at each time point. See Table 1 for prevalence rates for the total sample and all three variables.

Prevalence rates were significantly higher at the first wave (chi^2-test, t_1 vs. t_2, total sample of men and women, $P < 0.05$). When wave 2 and 3 were compared there were hardly any differences in statistical significance, indicating stability in prevalence over this 25-year time period.

Regarding "depressive mood", females showed a drop in prevalence after the 1970s in all age groups except at ages 55–74. No such effects were found for the more elaborately defined depressive syndrome. Males generally exhibited lower rates of "depressive mood" and depressive syndrome, with a similar pattern of prevalence rates over time, but no significant differences between time points.

A total of 251 persons (30%) endorsed none of our criteria of depressive syndrome at wave 1 and were therefore assigned to the control condition. Significantly more women than men suffered from a depressive syndrome.

Table 1 *Seven-days point* prevalence of depressive mood and depressive syndrome in cohorts of the same age group at three time points

	Females								
	t_1 (1975–1979)	t_2 (1980–1984)	t_3 (2001–2004)	Statistics					
				t_1 vs. t_2 $df = 1$		t_1 vs. t_3 $df = 1$		t_2 vs. t_3 $df = 1$	
	N (percentage of age group)	*N* (percentage of age group)	*N* (percentage of age group)	Chi^2	OR $CI_{(95\%)}$	Chi^2	OR $CI_{(95\%)}$	Chi^2	OR $CI_{(95\%)}$
Depressive syndrome									
Total sample	106 (22.8%)	85 (18.3%)	96 (20.6%)	NS	0.76 0.55–1.04	NS	0.88 0.65–1.20	NS	1.16 0.84–1.61
Age group 35–54 years	57 (29.2%)	42 (21.2%)	30 (19.9%)	NS	0.65 0.41–1.03	NS	0.60 0.36–0.99	NS	0.92 0.54–1.56
Age group 55–74 years	17 (27.0%)	29 (25.7%)	35 (17.6%)	NS	0.93 0.46–1.88	NS	0.58 0.30–1.12	NS	0.62 0.35–1.08
Age group 35–74 years	74 (28.7%)	71 (22.8%)	65 (18.6%)	NS	0.74 0.50–1.07	8.1**	0.57 0.39–0.83	NS	0.77 0.53–1.12
Depressive mood[c]									
Total sample	178 (38.4%)	121 (26.1%)	133 (28.7%)	15.5***	0.57 0.43–0.75	9.4**	0.65 0.49–0.85	NS	1.14 0.85–1.52
Age group 35–54 years	81 (41.8%)	60 (30.5%)	37 (24.5%)	4.9*	0.61 0.40–0.93	10.5**	0.45 0.28–0.72	NS	0.74 0.46–1.20
Age group 55–74 years	31 (49.2%)	36 (31.9%)	54 (27.3%)	NS	0.95 0.53–1.70	NS	0.76 0.45–1.30	NS	0.80 0.48–1.33
Age group 35–74 years	112 (43.6%)	96 (31.0%)	91 (26.1%)	9.9*	0.58 0.41–0.82	19.6***	0.45 0.32–0.64	NS	0.80 0.56–1.10
Depressive thought content[d]									
Total sample	115 (24.7%)	81 (17.4%)	84 (18.1%)	6.9**	0.65 0.47–0.89	5.8*	0.67 0.49–0.92	NS	1.04 0.74–1.46
Age group 35–54 years	61 (31.3%)	31 (15.7%)	23 (15.2%)	12.3**	0.41 0.25–0.67	11.1**	0.39 0.23–0.68	NS	0.96 0.54–1.73
Age group 55–74 years	15 (23.8%)	31 (27.4%)	37 (18.6%)	NS	1.21 0.59–2.47	NS	0.73 0.37–1.44	NS	0.60 0.35–1.04
Age group 35–74 years	76 (29.5%)	62 (20.0%)	60 (17.1%)	6.3*	0.60 0.41–0.88	12.3**	0.50 0.34–0.73	NS	0.83 0.56–1.23

	Males						Females + males[b]		
	t_1 (1975–1979)	t_2 (1980–1984)	t_3 (2001–2004)	Statistics[a]			t_1 (1975–1979)	t_2 (1980–1984)	t_3 (2001–2004)
	N (percentage of age group)	*N* (percentage of age group)	*N* (percentage of age group)	OR $CI_{(95\%)}$	OR $CI_{(95\%)}$	OR $CI_{(95\%)}$	*N* (percentage of age group)	*N* (percentage of age group)	*N* (percentage of age group)
Depressive syndrome									
Total sample	46 (12.3%)	32 (8.6%)	39 (10.5%)	0.69 0.43–1.11	0.83 0.53–1.31	1.20 0.74–1.96	152 (18.1%)	117 (14.0%)	135 (16.1%)
Age group 35–54 years	22 (13.3%)	14 (7.3%)	11 (7.9%)	0.51 0.25–1.04	0.55 0.26–1.19	1.08 0.47–2.45	79 (21.9%)	56 (14.4%)	41(14.1%)
Age group 55–74 years	2 (9.1%)	3 (7.0%)	20 (10.6%)	0.76 0.12–4.86	1.19 0.26–5.47	1.59 0.45–5.60	19 (22.4%)	32 (20.5%)	55 (14.2%)
Age group 35–74 years	24 (12.8%)	17 (7.3%)	31 (9.5%)	0.53 0.28–1.02	0.71 0.40–1.25	1.33 0.72–2.47	98 (22.0%)	88 (16.2%)	96 (14.2%)
Depressive mood[c]									
Total sample	78 (21.0%)	58 (15.6%)	60 (16.1%)	0.80 0.48–101	0.72 0.50–1.05	1.04 0.70–1.54	256 (30.7%)	179 (21.4%)	193 (23.1%)
Age group 35–54 years	34 (20.6%)	30 (15.7%)	18 (12.9%)	0.72 0.42–1.24	0.57 0.31–1.06	0.79 0.42–1.49	116 (32.2%)	90 (23.2%)	55 (18.9%)
Age group 55–74 years	7 (33.3%)	5 (11.9%)	34 (18.1%)	0.27 0.07–0.99	0.44 0.17–1.18	1.63 0.60–4.46	38 (44.7%)	41 (26.3%)	88 (22.7%)
Age group 35–74 years	41 (22.0%)	35 (15.0%)	52 (15.9%)	0.63 0.38–1.03	0.67 0.42–1.05	1.07 0.67–1.70	154 (34.6%)	131 (24.1%)	143 (21.1%)
Depressive thought content[d]									
Total sample	49 (13.1%)	45 (12.1%)	30 (8.0%)	0.91 0.59–1.40	0.58* 0.36–0.94	0.64 0.39–1.04	164 (19.6%)	126 (15.0%)	114 (13.6%)
Age group 35–54 years	21 (12.7%)	21 (11.0%)	7 (5.0%)	0.85 0.44–1.61	0.36* 0.15–0.88	0.43 0.18–1.04	82 (22.8%)	52 (13.4%)	30 (10.3%)
Age group 55–74 years	5 (22.7%)	5 (11.6%)	19 (10.1%)	0.45 0.11–1.75	0.38 0.13–1.15	0.85 0.30–2.43	20 (23.5%)	36 (23.1%)	56 (14.5%)
Age group 35–74 years	26 (13.9%)	26 (11.1%)	26 (8.0%)	0.77 0.43–1.38	0.53* 0.30–0.95	0.69 0.39–1.22	102 (22.9%)	88 (16.2%)	86 (12.7%)

OR odds ratio, CI95% confidence interval

[a]All tests were non-significant in males for depressive syndrome and depressive mood and depressive thought content

[b]Detailed statistics available from the senior author: For total sample (both sexes, all age groups) prevalence rates did not differ between waves except for >Depressive syndrome< t_1 vs. t_2 ($chi^2 = 5.1$, $df = 1$, $P < 0.05$), >Depressive mood < t_1 vs. t_2 ($chi^2 = 18.3$, $df = 1$, $P < 0.001$) and t_1 vs. t_3 ($chi^2 = 12.0$, $df = 1$, $P < 0.001$), >Depressive thought content< t_1 vs. t_2 ($chi^2 = 5.7$, $df = 1$, $P < 0.05$) and t_1 vs. t_3 ($chi^2 = 10.4$, $df = 1$, $P < 0.01$)

[c]Depressive Mood t_1, t_2, t_3 = severity of at least 1 (“slight impairment”) in SPI depressive mood or SPI depressed

[d]Depressive thought content t_1, t_2, t_3 = severity of at least 1 (“slight impairment”) in SPI depressive thought content

* $P < 0.050$; ** $P < 0.010$; *** $P < 0.001$; *NS* not significant; “–” not applicable

There were 152 cases of depressive syndrome at t_1. The number of new cases of depressive syndrome was 67 at t_2 and an additional 70 at t_3. Thus between all three time points a total of 289 (34.5% of $N = 838$) cases of depressive syndrome were observed.

At baseline, mean age was 39.4 ± 12.7 years in depressed persons and 34.1 ± 12.7 years in controls ($t = 4.06$, $df = 401$; $P = 0.002$). Women with depressive syndrome were significantly older (41.2 ± 12.6 years) than female controls (34.2 ± 13.3 years; $t = 3.99$, $df = 217$, $P = 0.000$). There was no significant age difference in men (34.0 ± 12.3 years vs. 35.3 ± 12.2 years, depressed and controls, respectively).

At the 5-year follow-up, mean age for women and men, combined, was 44.5 ± 12.8 years (depressed) and 39.1 ± 12.7 years (controls; $t = 4.07$, $df = 401$; $P = 0.000$). Ages at 25-year follow-up were 64.7 ± 12.8 years and 59.6 ± 12.6 years (depressed and controls, respectively; $t = 3.92$, $df = 401$; $P = 0.000$).

As shown in Table 2, more depressed subjects were divorced or widowed. Depressed and control subjects did not differ in school education, social class, and the frequency of poor living conditions. Moderate or severe poverty was encountered more often in depressed subjects. Regarding the SIS scales, only "social contacts" showed significant differences, with depressed subjects reporting more severe difficulties (cf. Table 1).

Table 2 also includes data on individuals who endorsed some items on depression but did not qualify for our definition of depressive syndrome. This very mildly depressed group appears to be more similar to controls than to depressive syndrome. Chi^2-tests including this group indicated the same variables to be significant. In addition SIS work condition was also significantly different between groups ($chi^2 = 6.1$; $df = 2$; $P = 0.047$).

Course of psychopathology

For both women and men, the course of several indicators of psychopathology differed significantly according to "depressed" or "control" status (Table 3).

A general pattern observed in both sexes was that the depressed subjects started with rather high scores, which declined or remained nearly stable over time. Controls exhibited low scores at baseline and showed an increase over time. Differences between depressed and control subjects at the same time point were significant for most scores, with men showing less significant differences at the 25-year follow-up. Generally, men's ratings were lower than women's.

Table 2 Sample characteristics at baseline assessment

	Depressive syndrome ($n = 152$)		Depressive symptoms ($n = 435$)		Controls ($n = 251$)		Chi^2-test/Fisher's exact test depressive syndrome vs. controls		
	n	(%)	*N*	(%)	*n*	(%)	Chi^2	*P*	*df*
Gender ($N = 838$)							23.3	0.000	1
Male	46	30.3	189	43.4	138	55.0			
Female	106	69.7	246	56.6	113	45.0			
Legal status ($N = 838$)							15.7	0.004	4
Single	32	21.1	122	28.0	69	27.5			
Married	103	67.8	277	63.7	174	69.3			
Separated	0	0	2	0.5	2	0.8			
Divorced	9	5.9	17	3.9	3	1.2			
Widowed	8	5.3	17	3.9	3	1.2			
Poverty ($N = 838$) (evaluation interviewer-rating)							16.8	0.000	1
None/mild	136	89.5	419	96.3	248	98.8			
Moderate/severe	16	10.5	16	3.7	3	1.2			
SIS income and housing conditions maximum severity ($N = 829$)							–	NS	1
Adequate or slightly inadequate	82	56.2	245	56.7	153	61.0			
Clearly inadequate/strongly inadequate	64	43.8	187	43.3	98	39.0			
SIS work conditions or role satisfaction for housewives, retired and jobless ($N = 548$)							–	NS	1
No or mild difficulty	50	62.5	212	75.7	130	69.1			
Moderate or severe difficulty	30	37.5	68	24.3	58	30.9			
SIS social contacts ($N = 828$)							8.3	0.004	1
No or mild difficulty	66	45.2	260	60.3	151	60.2			
Moderate or severe difficulty	80	54.8	171	39.7	100	39.8			
SIS leisure time ($N = 829$)							–	NS	1
No or mild difficulty	89	61.0	308	71.3	175	69.7			
Moderate or severe difficulty	57	39.0	124	28.7	76	30.3			

NS not significant; "–" not applicable

Table 3 Course of psychopathology in depressed and non-depressed individuals from the community

	Females								Males							
	Baseline		5-year follow-up		25-year follow-up		Gen. linear models[d] GLM ($df = 2$)		Baseline (D)		5-year follow-up		25-year follow-up		GLM ($df = 2$)	
	Depressive syndrome M (SD)	Controls M (SD)	Depressive syndrome M (SD)	Controls M (SD)	Depressive syndrome M (SD)	Controls M (SD)	F Age (A) at baseline[a]	F $T \times D$[b]Time (T) by Depress (D)	Depressive syndrome M (SD)	Controls M (SD)	Depressive syndrome M (SD)	Controls M (SD)	Depressive syndrome M (SD)	Controls M (SD)	F Age at baseline (A)	F $T \times D$ time by depress (A)
SPI total Psychopathology	1.72$^{\#\#\#,c}$ (0.78)	0.15 (0.19)	1.00$^{\#\#\#}$ (0.80)	0.38 (0.43)	0.89$^{\#\#\#}$ (0.81)	0.47 (0.57)	10.83**	67.86***	1.48$^{\#\#\#}$ (0.74)	0.11 (0.14)	0.67$^{\#\#\#}$ (0.61)	0.27 (0.32)	0.54 (0.56)	0.38 (0.51)	5.56*	89.41***
SPI subjective Psychopathology	0.91$^{\#\#\#}$ (0.43)	0.08 (0.10)	0.51$^{\#\#\#}$ (0.41)	0.19 (0.23)	0.45$^{\#\#\#}$ (0.37)	0.24 (0.26)	6.81*	76.64***	0.76$^{\#\#\#}$ (0.39)	0.06 (0.09)	0.34$^{\#\#\#}$ (0.32)	0.13 (0.16)	0.28$^{\#}$ (0.28)	0.18 (0.27)	6.71*	83.76***
SPI objective Psychopathology	0.40$^{\#\#\#}$ (0.23)	0.04 (0.06)	0.25$^{\#\#\#}$ (0.24)	0.09 (0.14)	0.23$^{\#\#\#}$ (0.26)	0.12 (0.19)	10.68*	36.93***	0.36$^{\#\#\#}$ (0.22)	0.03 (0.05)	0.17$^{\#\#\#}$ (0.17)	0.07 (0.11)	0.13 (0.17)	0.10 (0.15)	NS	57.81***
Complaint list Total score	21.41$^{\#\#\#}$ (11.27)	2.19 (2.65)	19.30$^{\#\#\#}$ (11.0)	8.28 (8.23)	19.03$^{\#\#\#}$ (11.30)	10.44 (7.59)	NS	23.86***	19.02$^{\#\#\#}$ (10.40)	2.64 (3.24)	15.28$^{\#\#\#}$ (9.32)	6.73 (6.43)	15.25$^{\#\#}$ (10.07)	10.37 (8.77)	6.48*	24.86***
SPI depressive Mood	1.26$^{\#\#\#}$ (0.64)	0[e]	0.50$^{\#\#\#}$ (0.74)	0.18 (0.47)	0.51$^{\#\#\#}$ (0.79)	0.18 (0.52)	NS	50.94***	1.26$^{\#\#\#}$ (0.71)	0[e]	0.30$^{\#\#\#}$ (0.63)	0.05 (0.25)	0.17 (0.49)	0.12 (0.44)	NS	98.23***
SPI depressive Thought content	1.02$^{\#\#\#}$ (0.75)	0[e]	0.43$^{\#\#\#}$ (0.72)	0.11 (0.34)	0.32$^{\#}$ (0.63)	0.17 (0.50)	NS	38.37***	0.96$^{\#\#\#}$ (0.76)	0[e]	0.24$^{\#\#}$ (0.57)	0.05 (0.22)	0.17 (0.49)	0.09 (0.38)	NS	59.07***
SPI anxiety and Worry	1.09$^{\#\#\#}$ (0.86)	0.12 (0.32)	(0.74$^{\#\#\#}$ (0.76)	0.27 (0.50)	0.46$^{\#\#}$ (0.68)	0.23 (0.48)	NS	23.60***	0.65$^{\#\#\#}$ (0.77)	0.03 (0.17)	0.43$^{\#\#}$ (0.62)	0.17 (0.43)	0.22 (0.42)	0.20 (0.55)	NS	15.52***
SPI phobias	0.64$^{\#\#\#}$ (0.45)	0.21 (0.76)	0.27$^{\#}$ (0.56)	0.13 (0.41)	0.33 (0.61)	0.33 (0.53)	NS	10.55***	0.33$^{\#\#}$ (0.56)	0.09 (0.33)	0.02 (0.15)	0.04 (0.19)	0.30$^{\#\#}$ (0.59)	0.10 (0.33)	NS	5.75**

SPI standardized psychiatric interview, *M* mean, *SD* standard deviation
[a]Covariate of GLM
[b]Interaction between time and group with age at baseline as covariate
[c]t-test Depressive Syndrome Group vs. Controls # = $p < 0.050$; ## = $p < 0.010$; ### = $p < 0.001$
[d]For repeated measurements
[e]By definition
* $P < 0.050$; ** $P < 0.010$; *** $P < 0.001$; *NS* not significant; ''–'' not applicable

Table 4 Prevalence (12 months) of psychiatric disorders at 25-year follow-up

	Females					Males				
	Depressive syndrome ($N = 106$) *n* (%)	Controls ($N = 113$) *n* (%)	Chi[2i]	OR	CI 95%	Depressive syndrome ($N = 46$) *n* (%)	Controls ($N = 138$) *n* (%)	Chi[2]	OR	CI 95%
Organic, including symptomatic psychiatric disorders	10 (9.4%)	7 (6.2%)	NS	NS	–	0	3 (2.2%)	NS	NS	–
Substance use disorders[a]	7 (6.6%)	1 (0.9%)	5.08*	7.92	0.96–65.49	7 (15.2%)	12 (8.7%)	NS	NS	–
Alcohol abuse/dependence	3 (2.8%)	1 (0.9%)	NS	NS	–	7 (15.2%)	11 (8.0%)	NS	NS	–
Illegal substance abuse/dependence	1 (0.9%)	0	NS	NS	–	0	1 (0.7%)	NS	NS	–
Possible psychotic disorders[b]	4 (3.8%)	0	4.34*	–	–	0	1 (0.7%)	NS	NS	–
Any mood disorder[c]	20 (18.9%)	7 (6.2 %)	8.13**	3.52	1.42–8.72	4 (8.7%)	6 (4.3%)	NS	NS	–
Bipolar disorder	0	0	–	–	–	0	0	–	–	–
Unipolar depression	18 (17.0%)	4 (3.5%)	10.94**	5.57	1.82–17.07	4 (8.7%)	5 (3.6%)	NS	NS	–
Dysthymia	9 (8.5%)	1 (0.9%)	7.26**	10.39	1.29–83.50	0	2 (1.4%)	NS	NS	–
Moderate to severe mood disorders	9 (8.5%)	0	10.1**	–	–	4 (8.7%)	3 (2.2%)	4.02*	4.29	0.92–19.92
Any anxiety disorder[d]	22 (20.8%)	9 (8.0%)	7.36**	3.03	1.32–6.92	7 (15.2%)	8 (5.8%)	4.09*	2.92	1.00–8.55
Panic disorder[e]	4 (3.8%)	0	4.34*	–	–	0	1 (0.7%)	NS	NS	–
Phobia[f]	18 (17.0%)	8 (7.1%)	5.13*	2.69	1.11–6.47	7 (15.2%)	6 (4.3%)	6.21*	3.95	1.25–12.44
Generalized anxiety disorder	2 (1.9%)	1 (0.9%)	NS	NS	–	0	1 (0.7%)	NS	NS	–
Obsessive-compulsive disorder	1 (0.9%)	0	NS	NS	–	0	0	–	–	–
Somatoform disorders[g]	11 (10.4%)	7 (6.2%)	NS	NS	–	5 (10.9%)	4 (2.9%)	4.71*	4.09	1.05–15.93
Pain disorder	7 (6.6%)	5 (4.4%)	NS	NS	–	4 (8.7%)	2 (1.4%)	5.74*	6.48	1.15–36.61
Eating disorders[h]	0	0	–	–	–	0	0	–	–	–
Any of the above	52 (49.1%)	24 (21.2%)	18.68***	3.57	1.98–6.44	16 (34.8%)	27 (19.6%)	4.46*	2.19	1.05–4.59

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$; *NS* not significant; "–" not applicable
OR odds ratio, CI 95% confidence interval
[a]Abuse or dependence (without nicotine)
[b]Schizophrenia and other psychotic disorders
[c]Major depressive disorder, dysthymic disorder, bipolar I or bipolar II disorders, single hypomanic episode
[d]Anxiety disorders including obsessive–compulsive disorders, excluding post-traumatic stress disorder
[e]With or without Agoraphobia;
[f]Agoraphobia without history of panic disorder, social phobia, specific phobias, anxiety disorder NOS;
[g]Somatization disorder, undifferentiated somatization disorder, hypochondriasis, pain disorder;
[h]Anorexia nervosa, atypical anorexia nervosa, bulimia nervosa, atypical bulimia nervosa
[i]Chi[2]-tests no depressive syndrome versus controls, $df = 1$

Mental illness outcome at the 25-year follow-up

The 12-month prevalence of psychiatric disorders at the 25-year follow-up is shown in Table 4. As expected, depressed subjects who were depressed at wave 1 also reported more psychiatric disorders at later time points than subjects not depressed at wave 1.

Unipolar depression was found more frequently than dysthymia. Only for women did a depressive syndrome at baseline increase the risk for depression 25 years later. Concerning anxiety disorders, depressed subjects showed a higher risk of developing phobias than controls in both sexes. Somatoform and pain disorders did not develop significantly more often in depressed women than in non-depressed women, but did appear more often in depressed men than in non-depressed men. An earlier depressive syndrome may be an indicator of later psychotic disorder for women but not for men.

Disability outcome at the 25-year follow-up

Table 5 presents data on disability and impairment at the 25-year follow-up in both sexes separately for subjects below and above age 45 at the baseline assessment.

Very few differences emerged for depressed and non-depressed men. In women, however, except for WHODAS-II "self-care" and "life activities", all indicators of disability or impaired functioning showed the same pattern. Women below age 45 who suffered from a depressive syndrome at baseline showed more impairment 25 years later than non-depressed controls; these analyses showed largely medium effect sizes. In women who were at or beyond age 45 at baseline assessment, no significant differences emerged for depressed and non-depressed subjects. It seems that the impairing long-term effect of depression is balanced by higher age.

Results on the very mildly depressed group

In Table 2 a very mildly depressed group was additionally described with severity of depression lying between the depressive syndrome and non-depressed controls. The latter groups are the focus of this paper and these results are described as an additional information. Including the mildly depressed individ-

Table 5 Disability and functioning at 25-year follow-up

Outcome at 25- year follow-up	Age at baseline	Females							Males						
		Depressive Syndrome (N = 105)		Controls (N = 111)		t test			Cohen's effect size	Depressive syndrome (N = 45)		Controls (N = 136)		t test			Cohen's effect size
		Mean	SD	Mean	SD	t	P	df	d	Mean	SD	Mean	SD	t	P	df	d
WHODAS II	<45 years	1.43	(0.74)	1.15	(0.26)	3.12	0.002	139	0.54	1.21	(0.34)	1.12	(0.16)	2.03	0.044	141	0.42
Total Disability Score	≥45 years	1.66	(0.87)	1.62	(0.94)	–	NS	73	0.04	1.48	(0.62)	1.55	(0.96)	–	NS	36	−0.08
(N = 397)	Total	1.53	(0.80)	1.27	(0.56)	2.78	0.006	214	0.38	1.27	(0.43)	1.21	(0.48)	–	NS	179	0.13
WHODAS II	<45 years	1.39	(0.78)	1.12	(0.29)	2.86	0.005	139	0.49	1.15	(0.39)	1.07	(0.20)	–	NS	141	0.31
Understanding and Communication Score	≥45 years	1.56	(0.90)	1.61	(1.14)	–	NS	73	−0.05	1.17	(0.26)	1.54	(1.17)	–	NS	36	−0.37
(N = 395)	Total	1.46	(0.83)	1.24	(0.65)	2.16	0.032	214	0.30	1.16	(0.36)	1.16	(0.58)	–	NS	179	0.00
WHODAS II	<45 years	1.55	(0.89)	1.21	(0.43)	2.96	0.004	139	0.52	1.22	(0.38)	1.24	(0.39)	–	NS	141	−0.05
Getting around Score	≥45 years	2.10	(1.13)	1.94	(1.15)	–	NS	73	0.14	1.78	(1.10)	1.88	(1.28)	–	NS	36	−0.08
(N = 397)	Total	1.80	(1.04)	1.40	(0.75)	3.25	0.001	214	0.44	1.36	(0.67)	1.36	(0.71)	–	NS	179	0.00
WHODAS II	<45 years	1.16	(0.69)	1.01	(0.10)	–	NS	139	0.33	1.07	(0.39)	1.02	(0.09)	–	NS	141	0.24
Self care-score	≥45 years	1.37	(0.94)	1.49	(1.11)	–	NS	73	−0.12	1.57	(1.27)	1.43	(1.18)	–	NS	36	0.12
(N = 396)	Total	1.25	(0.82)	1.14	(0.60)	–	NS	214	0.15	1.19	(0.72)	1.10	(0.55)	–	NS	179	0.15
WHODAS II	<45 years	1.32	(0.81)	1.07	(0.20)	2.75	0.007	139	0.46	1.14	(0.56)	1.09	(0.29)	–	NS	141	0.13
Getting along with People	≥45 years	1.38	(0.92)	1.46	(1.15)	–	NS	73	−0.08	1.33	(0.44)	1.31	(0.87)	–	NS	36	0.03
(N = 392)	Total	1.35	(0.86)	1.17	(0.62)	–	NS	214	0.24	1.19	(0.53)	1.13	(0.47)	–	NS	179	0.12
WHODAS II	<45 years	1.37	(0.85)	1.18	(0.51)	–	NS	138	0.28	1.16	(0.41)	1.08	(0.26)	–	NS	127	0.27
Life activities	≥45 years	1.59	(0.89)	1.49	(0.98)	–	NS	67	0.11	1.50	(1.41)	1.27	(1.00)	–	NS	22	0.20
(N = 362)	Total	1.46	(0.87)	1.25	(0.66)	1.97	0.051	207	0.27	1.23	(0.73)	1.10	(0.44)	–	NS	151	0.25
WHODAS II	<45 years	1.61	(0.83)	1.23	(0.43)	3.55	0.001	139	0.61	1.35	(0.44)	1.20	(0.30)	2.27	0.025	141	0.44
Participation in Society	≥45 years	1.69	(0.87)	1.52	(0.76)	–	NS	73	0.20	1.50	(0.60)	1.50	(0.61)	–	NS	36	0.00
(N = 397)	Total	1.65	(0.84)	1.31	(0.55)	3.54	0.000	214	0.48	1.38	(0.48)	1.26	(0.40)	–	NS	179	0.28
Global Assessment of	<45 years	75.31	(17.74)	86.61	(9.67)	4.88	0.000	140	-0.83	80.12	(15.27)	85.65	(10.80)	2.35	0.020	142	−0.46
Functioning Score (0–100)	≥45 years	74.28	(18.92)	78.32	(21.76)	–	NS	73	-0.20	81.45	(9.59)	80.59	(20.52)	–	NS	36	0.05
N = 399 (last 7 days)	Total	74.85	(18.20)	84.54	(14.08)	4.40	0.000	215	-0.60	80.44	(14.01)	84.66	(13.34)	–	NS	180	−0.31
Number of somatic diagnoses	<45 years	3.93	(2.12)	2.55	(2.13)	3.77	0.000	137	0.65	2.94	(2.65)	2.32	(1.92)	–	NS	134	0.29
(point prevalence 7 days)	≥45 years	4.13	(2.44)	3.56	(1.97)	–	NS	72	0.25	4.55	(2.42)	3.12	(2.27)	–	NS	35	0.62
N = 386	Total	4.02	(2.26)	2.80	(2.12)	4.06	0.000	211	0.56	3.34	(2.66)	2.48	(2.01)	–	NS	171	0.39

Cohens d effect sizes: small effect (≥0.15 and <0.40), medium effect (≥0.40 and <0.75), large effect (≥0.75)
WHODAS II, world health organization disability assessment schedule II; rating from 1 (none) to 5 (severe disability)
SD standard deviation

uals in the analyses showed that the level and course of psychopathology did not materially differ from the results described in Table 3 for the more extremely defined groups. Generally, very mildly depressed subjects scored between the other two groups with scores being rather similar to non-depressed controls. Prevalence rates for psychiatric disorders as listed in Table 4 in very mildly depressed subjects also laid between rates of non-depressed controls and individuals with depressive syndrome. Overall rate of any disorder was 26.5% in females (49.1% in depressive syndrome and 21.2% in controls) and 25.0% in males (34.8% in depressive syndrome and 19.6% in controls). Disability and functioning (cf. Table 5) in very mildly depressed subjects generally also was very similar to controls.

Discussion

The strength of our study consists in (1) the prospective longitudinal approach with three cross-sectional assessments, (2) a long follow-up period of 25 years, (3) a large sample size of 838 subjects with assessments at all three cross-sections, (4) a reasonable diagnostic uniformity, (5) the use of standardized/structured interviews and self-rating scales, and (6) a participation rate (88% at 25-years follow-up) that is higher than commonly found in such long-term follow-ups (70–80% [48]). This is of relevance because subjects who refuse to participate in a follow-up study tend to exhibit more psychiatric disorders [15] and non-responders seem to have a strong tendency to distort the results in psychiatric population studies [20].

There are some limitations to our study: (1) Due to the long interval of 20 years between the second and third assessments, we may have missed some relevant information. Similarly, no information is available on the 390 individuals who died between the original survey and the third investigation. (2) As our sample was followed over 25 years, our study participants grew older and were no longer representative of the general population of present Upper Bavaria. A cohort which was originally selected to be representative of a population may develop in another direction than the population at large with respect to socio-economic stratum or other demographic characteristics [44]. (3)

The sample investigated in the 2000s was an ageing cohort. An under-estimation of prevalence of psychiatric disorders may derive from the fact that data could be gathered for surviving subjects only. Some studies found evidence for an elevated mortality risk in depressive subjects [38, 47, 54]. Potentially, cases were lost to follow-up because they had died before the assessment in the 2000s. (4) Nettelbladt et al. [44] discussed that interrater reliability over long time periods is difficult to achieve, as field workers in later waves may focus on other aspects and have different concepts and thresholds for establishing a "case" even when the same diagnostic system is applied. Similarly, Murphy et al. [43] reported that the concepts people form of psychiatric symptoms also change over time; this cannot be ruled out for our procedures. We found an interrater reliability of kappa 0.83 for the SPI between wave 2 and wave 3. Our finding of no significant changes in depressive syndrome between wave 2 and wave 3 probably confirms that assessment criteria at both waves were the same. (5) Our definition of depressive syndrome did not meet DSM-III or DSM-IV criteria but rather describes a sub-threshold depressive disorder.

Including an additional group of individuals who endorsed some items on depression but did not qualify for our definition of depressive syndrome did not materially change the findings of our study. These very mildly depressed individuals were more similar to non-depressed controls than to individuals with depressive syndrome. However, care should be taken to pay attention to even small indications of depression as very mildly depressed subjects showed slightly higher but non-significantly different psychopathology compared to non-depressed controls.

The 7-day point prevalence of depression as defined here was higher (18.1%) than that reported by other epidemiological studies, with rates for affective disorders varying from approximately 0.8% [22] to about 10.3% in the National Comorbidy Survey [26]. These studies, however, mostly reported prevalence rates for a more severe depression based on DSM-criteria for major depressive episode. Our findings, which are based on more broadly defined criteria, still virtually confirm the lifetime prevalence rates for depressive syndrome of 16% reported by Chen et al. [7] for the 13-year follow-up in the Baltimore Epidemiologic Catchment Area Study. Judd et al. [23] defined sub-syndromal depressive symptoms in the ECA Study as at least two or more current depressive symptoms, present for most or all of the time, lasting for at least 2 weeks, in individuals who did not meet criteria for major depression or dysthymia, and reported a rate of 11.8% of depressive syndromes.

One important research question is whether there are changes in the prevalence of depression over defined time periods. In our prospective study, using the same assessment methods at all three time points, there was a slight decrease in prevalence of our three indicators for depression from wave 1 to wave 2. There was stability in prevalence over the 20-year time interval from wave 2 to wave 3. This finding of relatively stable rates over time were confirmed by other prospective longitudinal epidemiologic studies such as the Baltimore ECA Study [13] the Lundby Study [35], the Stirling County Study [41], and the Zurich Study [4, 36]. Murphy et al. [41] reported a stable overall rate of depression (±5%) over the period from 1952 to 1992. These data are consistent with the Zurich Study of young adults: data showed a considerable tendency for individuals to meet criteria for multiple depressive subtypes over a 15-year period [4]. The persistence of subthreshold-level depression from early to mid-adulthood thus has been demonstrated [36]. The Baltimore Epidemiologic Catchment Area 23-year follow-up [13] found only slight variations in the prevalence of depression (2.8% in 1981; 2.4% in 1993; 2.7% in 2004).

There is some, albeit weak, evidence during the early phase of the Lundby Study (a prospective community study in Sweden) of *increasing* depression rates over time. In the 25-year follow-up of a complete sample from a geographically defined area in Sweden ("Lundby Study" [19]), the annual incidence of depressive disorder increased from the period of 1947–1957 to the period of 1957–1972. Hagnell et al. [19] calculated standardized age and sex incidence rates for each 5-year period as well as the cumulative probabilities of ever developing a depressive illness up to age 80. A tenfold increase in the risk of depression was found for the young adult group (aged 20–39 years) from 1957 until 1972 as compared with the period from 1947 to 1957. Hagnell et al. [19] suggested that we might be entering an "age of melancholy". They speculated that the apparent increase could be partly attributable to the growing awareness of depression among physicians and in the general population, to better knowledge of therapeutic options, and to a diminishing social stigma of mental illness. Murphy et al. [43] argued that the difference in research results may be due to the methods used. In the early Lundby approach, psychiatrists interviewed the subjects without a structured or standardized interview. The validity of the psychiatric diagnoses can therefore be questioned, particularly regarding their sensitivity to short-lived episodes. A very recent re-analysis of data from the Lundby Study [35] demonstrated that the recurrence rates of depression were similar in the time periods before and after 1972. Thus, Mattison et al. [35] suggested that the course of depressive disorder had not changed markedly in the later period (1997) in the Lundby study.

Results of the more recent prospective community EDSP Study by Wittchen and Perkonigg [59] found a two-fold increase of depressive disorders for adolescents and young adults in Germany in the 4-year time period from 1994/1995 to 1998/1999. Both studies

were carried out in Germany and covered the 1990s. However, subjects in our sample in 1994 were adults (age 33+) while the EDSP Study participants at that time were adolescents aged 14–15 years. Thus, their data supplement ours and point to a critical age was the prevalence of psychiatric disorders may (temporarily) increase.

Another type of study uses the assessments of independent samples at two or more different time points. For example, the National Comorbidity (Replication) Survey conducted assessments roughly a decade apart and reported data on a comparison of the original sample of the 1990–1992 National Comorbidity Survey (NCS [26]) with a new sample of the 2001–2002 National Comorbidity Survey Replication (NCS-R [18]): the prevalence of current major depressive disorder decreased from 10.1 to 8.7%. However, because the two surveys used different diagnostic criteria and assessment methods, the two studies' prevalence rates were not directly comparable, so that the change over time could not be assessed reliably [9]. Data from two large cross-sectional surveys of representative samples of the U.S. population (NLAES; NESARC) conducted 10 years apart showed that rates of major depression rose markedly over the past decade [9]. From 1991–1992 to 2001–2002, the prevalence of major depression among US adults increased from 3.33 to 7.06%. These results are not consistent with the stable rates found in prospective studies. Differences in findings may be due to the fact that results of the Baltimore ECA Study, Stirling County Study, Lundby Study and UBS Study are based on "survivors" (disregarding attrition due to death and other causes) with repeated assessments of the same population, whereas Compton et al. [9] compared two different age cohorts.

A third type of study is of retrospective nature and assesses subjects once by asking them questions about their psychopathology during years or even decades previous to the assessment. Some studies with this type of design indicated that prevalence rates of depression have dramatically increased throughout the twentieth century in the United States [26, 30, 34]. However, methodological concerns about the reliability of reporting symptoms of depression and other psychopathology precisely for past decades have cast some doubt on the validity of these findings [42, 49].

One aspect of our research question about the stability of depression over time concerns the difference found in age cohorts and gender. In our study, no effect of time and age cohort was found for the prevalence of depressive syndromes, except for a slight decrease of the prevalence of depressive syndromes from wave 1 to wave 3 in women aged 35–74 years. No such effect was found for male respondents. Our finding confirmed results of other prospective community studies such as the Midtown Manhattan study [52] and the Finnish UKKI Study [33]. Using no specific diagnostic classification system, Srole [52] found a higher percentage of impairment in the 40–59 year old age group in 1954 (22.4%) than in 1974 (13.7%). Later, Lehtinen et al. [33] reported that in the same Finnish age groups, prevalence of psychopathology in later phases of the study was lower than in earlier phases. Especially, prevalence of "neuroses" in the youngest birth cohort (born 1945–1954) was only half as high as the rate in the cohort born 10 years earlier. The authors hypothesized that the cohort born after the Second World War may have been mentally healthier than the older cohorts.

These findings are in conflict with studies in the USA [9, 13], Canada (Stirling County; [43]), and Sweden [19]. For gender and age distribution, Murphy et al. [43] found that study year and age interacted significantly with regard to depression, resulting in a higher rate among younger women in 1992 than in 1952 and 1970. The rates of depression in men remained very similar over time both overall and in terms of age distribution. The point prevalence rate among women younger than 45 years was about twice (8.2%) what it had been for women of similar age in the two earlier cross-sections (4.8, 3.2%). Murphy et al. [43] concluded that the evidence about depression among younger women may be seen as a "birth cohort effect". All women in our youngest female sample (35–54 years) were born after World War II and were of similar age, in contrast to the women under 45 years of age in the 1992 Stirling County sample. However, we could not find a rise in depression in this age group. Differences may be explained by the longer follow-up interval of 40 years in the Stirling county study. In addition, definition of depression in the Stirling county study included different depression symptoms assessed for a duration of one month, such as poor spirits, many ailments, early morning fatigue, apprehension about suffering a nervous breakdown, not working very much. This definition is, however, quite different from our depressive syndrome criterion.

Similar to the Stirling County Study findings and the Baltimore ECA follow-up [13] our study reports a higher prevalence of depression among those aged 30–44 compared to younger and older age groups. Eaton et al. [13], using slightly differing definitions of depression, reported rising 12-month prevalence rates of depressive disorders for females aged 30–44 between 1981 (6.3%), 1993 (5.1%), and 2004 (11.0%). The prevalence of depression in males was found to be stable or even declining over this period in each age group. The Canadian study (1970–1992) and the ECA study (1993–2004) took place in chronologically adjacent time periods, which means that the younger females in the Baltimore cohort would be middle aged. Thus, the rise in rate of depression was found to be consistent between these North American samples [13]. Eaton et al. [13] suggested that the chronicity of depressive disorder is rising among females in late middle age.

In contrast to the findings the Stirling County Study in Canada and the Baltimore ECA Study, Compton et al. [9] found no increase in prevalence of depression specific to a particular age cohort at a particular point in time. They observed consistent increases in major depression (from 3.3% in 1991/1992 to 7.1% in 2001/2002; 12-month prevalence) among both genders and all age and racial-ethnic subgroups, except in Hispanic men and women. In the Lundby study [19] the probability for depression increased considerably among young adult men (20–39 year group). Evidence from retrospective studies of community samples [30] suggested an increase in lifetime rates of major depression for their younger age cohort, usually for those born after 1940, and a decrease in the lifetime prevalence for the older cohort. Again, subjects' imprecise recall may have been relevant.

In summary, findings from prospective community studies with repeated assessments of the same sample published in the last three decades generally suggest that there were no marked changes in the rates of depression over longer time periods as the subjects grew older. Based on these studies, there are no indications that depressive syndromes generally have dramatically increased over the past few decades.

In contrast to the studies from the US and Canada, no rise in prevalence of depression in middle-aged females could be found in Europe [33, 35]. One possible explanation for the rise in prevalence for females is the increased availability and use of treatment for depression [27], and widespread consumer advertisements for anti-depressants during the last decade in North America [5], which have increased the recognition of depression, and lowered the stigma attached to reporting the disorder [13]. We conclude that this phenomenon did not affect the German population or other European countries to this extent, as there is no direct advertising of antidepressants to consumers. An expert survey in Germany [10] showed a high public awareness of psychiatric illness in the general population, presumably because the media have dealt with the issue quite frequently. However, it also seems possible that there is still more stigma associated with depression in Germany than in the US and that our subjects were often not that willing to report depressive symptoms. Most probably for that reason, our study ascertained no increasing depression rates in middle-aged women.

In our study, the presence of a baseline depressive syndrome was associated with more severe psychopathology at all three later time points, even 25 years later, compared to "no depression". Controlling for age did not alter these findings. This finding is confirmed by other long-term studies [7, 35, 36, 39, 40, 43]. The Stirling Study, for example [39], reports a poor outcome for 82% of depressed respondents, with poor outcome defined in terms of increased probability to die or to suffer from a chronic or recurrent illness. Moreover, depression and poverty tended to be chronic [40]. In the Zurich Study [36], a remarkably long-term stability of anxiety and depression was reported, and yet another one [7] found that 23% of subjects with a depressive syndrome developed some form of depressive disorder over the following 13 years. The Lundby Study [35] also established high rates of recurrence for men (42%) and women (46%) over a follow-up period of 30–39 years. As with our subclinical level of depression in depressive syndrome, other studies supported the notion that the risk for developing later major depression is high, e.g. >30% [4], or that the outcomes of subsyndromal and clinical levels of depression did not differ significantly [58]. A previous depressive episode at least doubles the risk for recurrence [25] or, more exactly, carries a 59% risk of experiencing one more depressive episode within the two following years, with each subsequent episode adding 16% to the risk of recurrence (NIMH [50]).

The CIDI-interview at t_3 yielded no diagnosis of bipolar disorder in the 12 months preceding the interview. Actually, the CIDI identified no lifetime bipolar disorder in the depressive syndrome group and controls neither. Only in the group with very mild depression (not qualifying for depressive syndrome) were there three individuals with a lifetime diagnosis of bipolar disorder. These individuals, too, had no bipolar disorder in the 12 months preceding t_3 assessment. It seems that bipolar disorder is an exceedingly rare diagnosis in a rural population. Possibly we may have missed bipolar cases at any wave due to these individuals having been institutionalized at the time of assessment.

Additional analyses not reported in this paper and focusing on birth cohort by period (time point of assessment) interaction resulted in only minor significant findings. We concluded from these findings that no major interaction of birth cohort and period was present in our data.

Our finding that depressive syndrome at a younger age is related to later rates of disability and reductions in functioning is consistent with considerable evidence from cross-sectional community and primary care studies [1, 6, 23, 45, 57]. In the cross-national study of the Epidemiology of Mental Disorders (ESEMED), Alonso et al. [1] reported that dysthymia and major depressive episode were among the five mental disorders with the strongest impact on WHODAS-II disability and emphasized that the reduced cognitive, emotional and motivational functioning associated with depression affects the highest mental capacities in human beings. Similarly, the ECA study [23] found that subsyndromal depressive symptoms were associated with reduced functionality, high impairment and disability; the authors conclude that depression constitutes a very important and costly public health issue. Interestingly, the effect persists even after controlling for severity of physical

disease—psychopathology such as major depression, panic disorder, generalized anxiety disorder and neurasthenia, was still strongly associated with increased disability [45]. Strong support for the disability hypothesis of depression comes from a large study on 11,242 outpatients (Medical Outcomes Study [57]). Partial or full depressive syndrome led to worse physical, social, and role functioning, worse perceived current health, and greater bodily pain. Depressive symptoms were uniquely associated with poor functioning, and this effect was at least commensurate with or greater than the debilitating effects of eight major chronic medical conditions.

Our study has extended these findings by showing that depressive symptoms from mild to clinical severity exert a substantial impact over long time periods in the general population. Future studies should focus on further investigating the long-term outcome of different depressive subtypes in the community and on effective treatments aimed at a better relapse prevention in order to improve the prognosis for depression. Effective treatments need to have a considerable effect on the long-term course of depression and reduce the overall human and economic burden of this chronic and recurring disorder.

Acknowledgments This research was funded by the Deutsche Forschungsgemeinschaft, Bonn, grant number Fi 333/4-1. We especially thank Dipl.-Psych. Eva Bolz, Dipl.-Psych. Helen Castellanos, Dipl.-Psych. Andrea Dusold, Dipl.-Psych. Andreas Hampel, Dr. Gabriele Kohlböck, Dipl.-Psych. Birgit Schaider, Dipl.-Psych. Nicole Sevignani and Dipl.-Psych. Nadia Sosnowsky for their sensitive way of conducting interviews with the subjects, and Susanne Hedlund (Ph.D.) for her support in formulating this manuscript.

References

1. Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H, de Girolamo G, Graaf R, Demyttenaere K, Gasquet I, Haro JM, Katz SJ, Kessler RC, Kovess V, Lepine JP, Ormel J, Polidori G, Russo LJ, Vilagut G, Almansa J, Arbabzadeh-Bouchez S, Autonell J, Bernal M, Buist-Bouwman MA, Codony M, Domingo-Salvany A, Ferrer M, Joo SS, Martinez-Alonso M, Matschinger H, Mazzi F, Morgan Z, Morosini P, Palacin C, Romera B, Taub N, Vollebergh WA (2004) Disability and quality of life impact of mental disorders in Europe: results from the European study of the epidemiology of mental disorders (ESEMeD) project. Acta Psychiatr Scand Suppl 420:38–46
2. American Psychiatric Association (APA) (2000) Diagnostic and Statistical Manual of Mental Disorders, 4th edn (text revision). American Psychiatric Association, Washington, DC
3. Angst J (1996) Comorbidity of mood disorders: a longitudinal prospective study. Br J Psychiatry 168:58–67
4. Angst J, Merikangas K (1997) The depressive spectrum: diagnostic classification and course. J Aff Dis 45:31–40
5. Bell RA, Kravitz RL, Wilkes MS (2000) Direct-to-consumer prescription drug advertising, 1989–1998. A content analysis of conditions, targets, inducements, and appeals. J Fam Pract 49:329–335
6. Broadhead WE, Blazer DG, George LK, Tse LK (1990) Depression, disability days and days lost from work in a prospective epidemiologic survey. JAMA 264:2524–2528
7. Chen LS, Eaton WW, Gallo JJ, Nestadt G, Crum R (2000) Empirical examination of current depression categories in a population-based study: symptoms, course, and risk factors. Am J Psych 157:573–580
8. Clare AS, Cairns VE (1978) Design, development and use of standardized interview to assess social maladjustment and dysfunction in community studies. Psychol Med 8:589–604
9. Compton WM, Conway KP, Stinson FS, Grant BF (2006) Changes in the prevalence of major depression and comorbid substance use disorders in the United States between 1991–1992 and 2001–2002. Am J Psychiatry 163:2141–2147
10. DAK Gesundheitsreport (2005) DAK Versorgungsmanagement, Hamburg
11. Dilling H, Weyerer S (1984) Psychische Erkrankungen in der Bevölkerung bei Erwachsenen und Jugendlichen. In: Dilling H, Weyerer S, Castell R (eds) Psychische Erkrankungen in der Bevölkerung. Enke Verlag, Stuttgart, pp 1–122
12. Eaton WW, Ritter C (1988) Distinguishing anxiety and depression with field survey data. Psychol Med 18(1):155–166
13. Eaton WW, Kalaydjian A, Scharfstein DO, Mezuk B, Ding Y (2007) Prevalence and incidence of depressive disorder: the Baltimore ECA follow-up, 1981–2004. Acta Psychiatr Scand 116:182–188
14. Feuerlein W, Ringer C, Küfner H, Antons K (1977) Diagnose des Alkoholismus: per Münchner Alkoholismustest (MALT). MMW Munch Med Wochenschr 119:1275–1282
15. Fichter MM (1990) Verlauf psychischer Erkrankungen in der Bevölkerung. Springer, Heidelberg
16. Fichter MM, Narrow WE, Roper MT, Rehm J, Elton M, Rae DS, Locke BZ, Regier DA (1996) Prevalence of mental illness in Germany and the USA: comparison of the upper Bavarian study and the epidemiologic catchment area program. J Nerv Ment Dis 184:598–606
17. Goldberg DP, Cooper P, Eastwood RR, Kedward MB, Shepherd M (1970) A standardized psychiatric interview for use in community surveys. Br J Prev Soc Med 24:18–23
18. Greenberg PE, Kessler RC, Birnbaum HG, Leong SA, Lowe SW, Berglund PA, Corey-Lisle PK (2003) The economic burden of depression in the United States: how did it change between 1990 and 2000? J Clin Psychiatry 64:465–1475
19. Hagnell O, Lanke J, Rorsman B, Öjesjö L (1982) Are we entering an age of melancholy? Depressive illness in a prospective epidemiological study over 25 years. Psychol Med 12:279–289
20. Hagnell O (1989) Repeated incidence and prevalence studies of mental disorders in a total population followed during 25 years. The Lundby Study, Sweden. Acta Psychiatr Scand (Suppl) 348:61–77
21. Hagnell O, Gräsbeck A (1990) Comorbidity of anxiety and depression in the Lundby 25-year prospective study: the pattern of subsequent episodes. In: Maser JD, Cloninger CR (eds) Comorbidity of mood and anxiety disorders. American Psychiatric Press Inc., Arlington, pp 139–152
22. Hagnell O, Öjesjö L, Otterbeck L, Rorsman B (1994) Prevalence of mental disorders, personality traits and mental complaints in the Lundby Study. Scand J Soc Med (Suppl) 50:7–72
23. Judd LJ, Paulus MP, Wells KB, Rapaport MH (1996) Socioeconomic burden of subsyndromal depressive symptoms and major depression in a sample of the general population. Am J Psych 153:1411–1417
24. Judd LL, Schettler PJ, Akiskal HS (2002) The prevalence, clinical relevance, and public health significance of subthreshold depressions. Psychiatr Clin North Am 25(4):685–698
25. Kennedy N, Abbott R, Paykel ES (2003) Remission and recurrence of depression in the maintenance era: long-term outcome in a Cambridge cohort. Psychol Med 33(5):827–838
26. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, Wittchen HU, Kendler KS (1994) Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the national comorbidity survey. Arch Gen Psychiatry 51(1):8–19

27. Kessler RC, Demler O, Frank RG, Olfson M, Pincus HA, Walters EE, Wang PS, Wells KB, Zaslavsky AM (2005) Prevalence and treatment of mental disorders, 1990–2003. N Engl J Med 352:2515–2523
28. Kiloh LG, Andrews G, Neilson M (1988) The long-term outcome of depressive illness. Br J Psych 153:752–757
29. Kleining G, Moore H (1968) Soziale Selbsteinstufung (SSE). Ein Instrument zur Messung sozialer Schichten. Köln Z Soziol Sozialpsychol 20:502–552
30. Klerman GL, Weissman MM (1989) Increasing rates of depression. JAMA 15(261):2229–2235
31. Lee AS, Murray RM (1988) The long-term outcome of Maudsley depressives. Br J Psych 153:741–751
32. Lee AS (2003) Better outcomes for depressive disorders? Psychol Med 33:769–774
33. Lehtinen V, Lindholm T, Veijola J, Vaisanen E, Puukka P (1991) Stability of prevalences of mental disorders in a normal population cohort followed for 16 years. Soc Psychiatry Psychiatr Epidemiol 26(1):40–46
34. Lewinsohn PM, Rohde P, Seely JR, Fischer SA (1993) Age-cohort changes in the lifetime occurrence of depression and orther mental disorders. J Abnorm Psychol 102:110–120
35. Mattison C, Bogren M, Horstmann V, Munk-Jörgensen P, Nettelbladt P (2007) The long-term course of depressive disorders in the Lundby study. Psychol Med 37:883–891
36. Merikangas KR, Zhang H, Avenevoli S, Acharyya S, Neuenschwander M, Angst J (2003) Longitudinal trajectories of depression and anxiety in a prospective community study. Arch Gen Psychiatry 60:993–1000
37. Murphy JM (1980) Continuities in community-based psychiatric epidemiology. Arch Gen Psychiatry 37(11):1215–1223
38. Murphy JM, Monson RR, Olivier DC, Sobol AM, Leighton AH (1987) Affective disorders and mortality. A general population study. Arch Gen Psychiatry 44(5):473–480
39. Murphy JM (1990) Depression in the community: findings from the Stirling County study. Can J Psychiatry 35(5):390–396
40. Murphy JM, Olivier DC, Monson RR, Sobol AM, Federman EB, Leighton AH (1991) Depression and anxiety in relation to social status. A prospective epidemiologic study. Arch Gen Psychiatry 48(3):223–229
41. Murphy JM, Monson RR, Olivier DC, Zahner GE, Sobol AM, Leighton AH (1992) Relations over time between psychiatric and somatic disorders: the Stirling County study. Am J Epidemiol 136(1):95–105
42. Murphy JM, Laird NM, Monson RR, Sobol AM, Leighton AH (2000) A 40-year perspective on the prevalence of depression from the Stirling County study. Arch Gen Psychiatry 57:209–215
43. Murphy JM, Horton NJ, Laird NM, Monson RR, Sobol AM, Leighton AH (2004) Anxiety and depression: a 40-year perspective on relationships regarding prevalence, distribution, and comorbidity. Acta Psychiatr Scand 109:355–375
44. Nettelbladt P, Bogren M, Mattison C, Öjesjö L, Hagnell O, Hofvendahl E, Toraker P, Phugra D (2005) Does it make sense to do repeated surveys? The Lundby Study, 1947–1997. Acta Psychiatr Scand 111:444–452
45. Ormel J, Vonkorff M, Ustun TB, PINI S, Korten A, Oldehinkel T (1994) Common mental disorders and disability across cultures. Results from the WHO collaborative study on psychological problems in general health care. JAMA 272(22):1741–1748
46. Robins LN, Wing J, Wittchen HU, Helzer JE, Babor TF, Burke J, Farmer A, Jablensky A, Pickens R, Regier DA, Sartorius N, Towle LH (1989) The composite international diagnostic interview: an epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. Arch Gen Psychiatry 45:1069–1077
47. Rorsman B, Hagnell O, Lanke J (1982) Mortality in the Lundby study. Natural death in different forms of mental disorder in a total population investigated during a 25-year period. Neuropsychobiology 8(4):188–197
48. Rothman KJ, Greenland S (1998) Modern epidemiology. 2nd edn. Lippincott-Raven Publishers, Philadelphia
49. Simon GE, VonKorff M (1992) Reevaluation of secular trends in depression rates. Am J Epidemiol 135:1411–1422
50. Solomon DA, Keller MB, Leon AC, Mueller TI, Lavori PW, Shea MT, Coryell W, Warshaw M, Turvey C, Maser JD, Endicott J (2000) Multiple recurrences of major depressive disorder. Am J Psychiatry 157:229–233
51. Spitzer RL, Williams JBW, Gibbon M (1985) Instructional Manual for the Structured Clinical Interview for DSM-III-R (SCID, 7/1/1985 Revision). Biometrics Research Department, New York State Psychiatric Institute, 722 West, 168th Street, New York, 10032. Deutsche Bearbeitung: Wittchen HU, Zaudig M, Schramm E, Sprengler P, Mombaur W, Klug J, Horn R "Strukturiertes Klinisches Interview für DSM-III-R" (SKID). Interviewheft und Manual, Beltz Verlag, Weinheim
52. Srole L (1975) Measurement and classification in sociopsychiatric epidemiology: Midtown Manhattan study (1954) and Midtown Manhattan restudy (1974). J Health Soc Behav 16:347–364
53. SPSS for Windows (2004) Release 13. SPSS Inc., Chicago
54. Tsuang MT, Woolson RF (1977) Mortality in patients with schizophrenia, mania, depression and surgical conditions. A comparison with general population mortality. Br J Psychiatry 130:162–166
55. Van Weel-Baumgarten EM, Schers HJ, Van den Bosch WJ, Van den Hoogen HJ, Zitman FG (2000) Long-term follow-up of depression among patients in the community and in family practice settings: a systematic review. J Family Practice 49:1113–1120
56. Von Zerssen D, Koeller DM (1976) Die Beschwerdenliste (BL). Beltz Verlag, Weinheim
57. Wells KB, Stewart A, Hays RD, Burnam MA, Rogers W, Daniels M, Berry S, Greenfield S, Ware J (1989) The functioning and well-being of depressed patients: results from the Medical outcomes study. JAMA 262:914–919
58. Wells KB, Burnam MA, Rogers W, Hays R, Camp P (1992) The course of depression in adult outpatients. Results from the medical outcomes study. Arch Gen Psych 49:788–794
59. Wittchen HU, Perkonigg A (1996) Epidemiologie psychischer Störungen. In: Ehlers A, Hahlweg K (eds) Enzyklopädie der Psychologie. Grundlagen der Klinischen Psychologie, Göttingen
60. Wittchen HU, Pfister H (1997) Diagnostisches Expertensystem für Psychische Störungen DIA-X. Swets Test Services. Frankfurt
61. Wittchen HU, Lieb R, Pfister H, Schuster P (2000) The waxing and waning of mental disorders: evaluating the stability of syndromes of mental disorders in the population. Compr Psychiatry (Suppl 1) 41(2):122–132
62. World Health Organization (2000) World health organization: disability assessment schedule WHODAS II Phase 2 field trials. Health services research. World Health Organization, Geneva, pp 678–687