Supplementary online material

Methods

Participants and sampling procedure

The longitudinal Zurich Cohort Study comprised a cohort of 4547 persons (2201 males; 2346 females) representative of the canton of Zurich in Switzerland, who were screened in 1978 with the Symptom Checklist 90-Revised (SCL-90-R) [1] when men were 19 and women 20 years old. In Switzerland, every male person must undertake a military screening test at the age of 19. Therefore, conscripts within a defined catchment area comprise its respective, complete male age group. With the consent of military authorities, but independent of their screening procedure, we randomly screened 50% of all conscripts of the canton of Zurich (refusal rate: 0.3%). Female participants were identified from the complete electoral register of the canton of Zurich. Again, 50% of them were randomly selected and received questionnaires by mail (refusal rate: 25%). In order to increase the probability of the development of psychiatric syndromes, a stratified subsample of 591 subjects (292 males; 299 females) was selected for interview, with two-thirds consisting of high-scorers defined by the 85th percentile or more of the Global Severity Index (GSI) of the SCL-90-R and a random sample of those with scores below the 85th percentile of the GSI. Such a two-phase approach consisting of initial screening and subsequent interview with a stratified subsample is an established procedure in epidemiological research [2]. For more information, see also reference [3]. The seven interview waves were conducted in 1979 at age 20/21 (n=591), 1981 at age 22/23 (n=456), 1986 at age 27/28 (n=457), 1988 at age 29/30 (n=424), 1993 at age 34/35 (n=407), 1999 at age 40/41 (n=367), and 2008 at age 49/50 (n=335). Across the whole 30-year observation period the final retention rate was 57%. The participant flow is shown in Figure 1. The initial distribution above and below the 85th percentile of the GSI did not change over the seven interview waves, although dropouts were more common among extremely high or low scorers on the GSI [4]. We repeated those dropout analyses for the

last interview in 2008 and found that dropouts did not differ significantly in their socio-economic status and education at onset of the study from subjects who remained in the study. Neither was there a difference in initial psychopathologic impairment according to the nine SCL-90-R subscales, but there were slightly more dropouts among men (OR=1.82; 95%-CI=1.31-2.53; p<0.001). Reasons for dropout were mostly unknown and, therefore, were not assessed systematically.

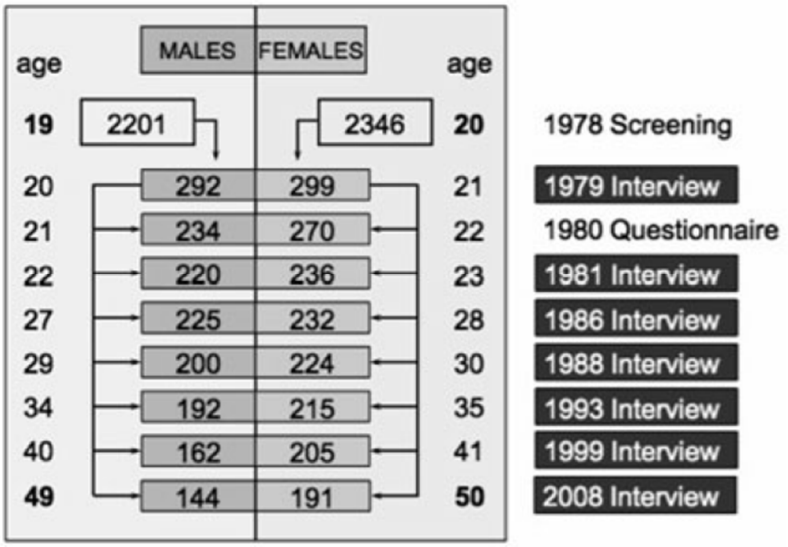

Figure 1: Participant flow

The Zurich Cohort Study was approved by the ethics committee of the canton of Zurich (KEK) as fulfilling all ethical, legal and data privacy protection requirements and is in accordance with the declaration of Helsinki of the World Medical Association. All participants gave their written informed consent. Due to strict data protection requirements based on Swiss legislation effective at the time of study inception, we cannot make the data freely available. However, full analysis scripts and original data outputs can be obtained from the first author upon request for validation and reproduction of the results.

Instruments and measures

Interviews were conducted with the "Structured Psychopathological Interview and Rating of the Social Consequences of Psychological Disturbances for Epidemiology" (SPIKE) [3]. This semi-structured interview was specifically developed for epidemiological surveys in psychiatry and captures socio-demography, environmental conditions, somatic syndromes, psychopathology and social functioning. Its excellent reliability and validity have been reported elsewhere [5]. All interviews were administered by trained and supervised clinical psychologists or psychiatrists. The SPIKE assesses a broad spectrum of psychopathological symptoms according to severity and duration within the last 12 months. These data further allow for defining 12-month prevalence rates of psychiatric diagnoses according to DSM criteria: 1979-1986 by DSM-III, 1988-1993 by DSM-III-R, and 1999-2008 by DSM-IV [6]. This was the case for major depression, generalised anxiety disorder, panic disorder, all phobias, and substance-use disorder. Obsessive-compulsive disorder was always diagnosed according to DSM-III-R. For the present study we collapsed episodes of unipolar and bipolar depression into a diagnosis of mood disorder. Panic disorder, specific phobia, social phobia and agoraphobia as well as obsessive-compulsive disorder were combined into an umbrella diagnosis of anxiety disorder. Mood and/or anxiety disorder were further defined as affective disorders.

Depression within the past 12 months was graded according to severity. Major depression was defined as the most severe form (coded as 3). Sub-threshold depression was defined by 3-4 of 9 criterial symptoms with a minimum duration of 2 weeks or by 5 or more symptoms lasting less than 2 weeks (coded as 2). All other forms were defined as depression symptoms (coded as 1). Depression was considered absent (coded as 0) if no symptoms were reported. Consequences of depression, such as distress/suffering, impairment and treatment seeking, were not included in the sub-threshold definition of depression, but were previously used as validators of the clinical relevance of this syndrome, alongside a positive family history of depression, age at onset and course (e.g. [7, 8]). Severe depression was defined as meeting criteria for major depression and reporting a distress of 70 or more on a

visual analogue scale ranging from 0 to 100. Assessing depression continuously is necessary because in this cohort it has previously been shown that up to 36% of persons with sub-threshold depression and up to 25% of persons with depressive symptoms sought psychiatric treatment [8]. Twelve-month prevalence of suicidality was carefully evaluated in a separate section, that is, independent of depression. The following items were assessed: 1) I wouldn't mind being dead; 2) I had transient thoughts of harming myself; 3) I had serious persisting thoughts of harming myself; 4) I had clear and precise ideas of how to commit suicide; and 5) I made a suicide attempt. Items 1) and 2) were categorized as transient suicidality, and items 3), 4) and 5) as severe suicidality. Twelve-month prevalence of AD use was rated as present if participants reported the use of monoamine oxidase inhibitors (MAOI), tricyclics, SSRI, or other AD.

Statistical analysis

All associations were estimated with Generalized Estimating Equations (GEE). These statistical models were introduced to fit regression analyses that account for within-subject correlation, which is an inherent part of longitudinal studies that rely on repeated outcome measures [9]. The GEE approach uses weighted combinations between a predictor variable and repeated outcomes that account for varying observations, e.g. a disorder being present vs. absent, within a person across time. This modelling approach makes GEE particularly useful for epidemiologic research [10, 11]. GEE use all available data and estimate missing values under the assumption of Missing Completely at Random (MCAR). Prerequisite to the application of GEE was therefore a thorough missing value analysis. According to Little's MCAR test, the outcome was missing completely at random over time (χ^2=132.4, df=146, p=0.784). We used antidepressant medication at time point t to predict depression severity at time point t+1 (i.e., antidepressant use at age 20/21 related to depression at age 22/23, antidepressant use at age 22/23 related to depression at age 27/27, and so on). As detailed in the results, these analyses were adjusted for various time-invariant (e.g. sex, education

level) and time-variant confounders (e.g. concomitant mental disorders at time point t). In these models the within-subject covariance was specified with the “unstructured” correlation type to avoid having any constraints on the covariance structure. Since depression severity was a categorical outcome, we fitted these models with ordered multinomial distribution and cumulative logit link-function.

References

1. Derogatis LR. Symptom Checklist 90, R-Version manual I: scoring, administration, and procedures for the SCL-90. Baltimore, MD: Johns Hopkins Press; 1977.
2. Dunn G, Pickles A, Tansella M, Vazquez-Barquero JL. Two-phase epidemiological surveys in psychiatric research. Br J Psychiatry 1999;174:95-100.
3. Angst J, Dobler-Mikola A, Binder J. The Zurich study--a prospective epidemiological study of depressive, neurotic and psychosomatic syndromes. I. Problem, methodology. Eur Arch Psychiatry Neurol Sci 1984;234(1):13-20.
4. Eich D, Ajdacic-Gross V, Condrau M, Huber H, Gamma A, Angst J, et al. The Zurich Study: participation patterns and Symptom Checklist 90-R scores in six interviews, 1979-99. Acta Psychiatr Scand Suppl 2003;108(s418):11-4.
5. Angst J, Gamma A, Neuenschwander M, Ajdacic-Gross V, Eich D, Rossler W, et al. Prevalence of mental disorders in the Zurich Cohort Study: a twenty year prospective study. Epidemiol Psichiatr Soc 2005;14(2):68-76.
6. Angst J, Paksarian D, Cui L, Merikangas KR, Hengartner MP, Ajdacic-Gross V, et al. The epidemiology of common mental disorders from age 20 to 50: results from the prospective Zurich cohort Study. Epidemiol Psychiatr Sci 2016;25(1):24-32.
7. Angst J, Sellaro R, Merikangas KR. Depressive spectrum diagnoses. Compr Psychiatry 2000;41(2 Suppl 1):39-47.

8. Hengartner MP, Angst F, Ajdacic-Gross V, Rossler W, Angst J. Treated versus non-treated subjects with depression from a 30-year cohort study: prevalence and clinical covariates. Eur Arch Psychiatry Clin Neurosci 2016;266(2):173-80.

9. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. Biometrics 1988;44(4):1049-60.

10. Hanley JA, Negassa A, Edwardes MD, Forrester JE. Statistical analysis of correlated data using generalized estimating equations: an orientation. Am J Epidemiol 2003;157(4):364-75.

11. Twisk JWR. Applied longitudinal data analysis for epidemiology: a practical guide. Cambridge: Cambridge University Press; 2003.

Supplementary Table: Frequency of depression severity and AD use within depression category from 1979 (at age 20/21) to 2008 (at age 49/50)

Assessment wave	Depression severity N (%)	AD use N (%)
1979	None: 155 (26.2)	0
	Symptoms: 305 (51.6)	4 (23.5)
	Sub-threshold: 87 (14.7)	10 (58.8)
	Major depression: 44 (7.4)	3 (17.6)
1981	None: 201 (44.1)	0
	Symptoms: 158 (34.6)	5 (38.5)
	Sub-threshold: 53 (11.6)	4 (30.8)
	Major depression: 44 (9.6)	4 (30.8)
1986	None: 230 (50.3)	0
	Symptoms: 113 (24.7)	0
	Sub-threshold: 67 (14.7)	2 (50.0)
	Major depression: 47 (10.3)	2 (50.0)
1988	None: 232 (54.7)	0
	Symptoms: 105 (24.8)	1 (9.1)
	Sub-threshold: 39 (9.2)	1 (9.1)
	Major depression: 48 (11.3)	9 (81.8)
1993	None: 243 (59.7)	0
	Symptoms: 81 (19.9)	8 (38.1)
	Sub-threshold: 28 (6.9)	1 (4.8)
	Major depression: 55 (13.5)	12 (57.1)
1999	None: 183 (49.9)	0
	Symptoms: 99 (27.0)	17 (44.7)
	Sub-threshold: 35 (9.5)	2 (5.3)
	Major depression: 50 (13.6)	19 (50.0)
2008	None: 189 (56.4)	0
	Symptoms: 88 (26.3)	16 (39.0)
	Sub-threshold: 17 (5.1)	3 (7.3)
	Major depression: 41 (12.2)	22 (53.7)
Total (counts)	None: 1433 (47.2)	0
	Symptoms: 949 (31.2)	51 (35.2)
	Sub-threshold: 326 (10.7)	23 (15.9)
	Major depression: 329 (10.8)	71 (49.0)