Psychological Medicine (2010), **40**, 899–909. © Cambridge University Press 2009
doi:10.1017/S0033291709991036 Printed in the United Kingdom

ORIGINAL ARTICLE

How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective *versus* retrospective ascertainment

T. E. Moffitt[1,2]*, A. Caspi[1,2], A. Taylor[3], J. Kokaua[4], B. J. Milne[2], G. Polanczyk[1,2] and R. Poulton[5]

[1] *Duke University, Durham, NC, USA*
[2] *Institute of Psychiatry, King's College London, UK*
[3] *UK Office of National Statistics, London, UK*
[4] *New Zealand Ministry of Health, New Zealand*
[5] *Dunedin Multidisciplinary Health and Development Research Unit, University of Otago, New Zealand*

Background. Most information about the lifetime prevalence of mental disorders comes from retrospective surveys, but how much these surveys have undercounted due to recall failure is unknown. We compared results from a prospective study with those from retrospective studies.

Method. The representative 1972–1973 Dunedin New Zealand birth cohort ($n=1037$) was followed to age 32 years with 96% retention, and compared to the national New Zealand Mental Health Survey (NZMHS) and two US National Comorbidity Surveys (NCS and NCS-R). Measures were research diagnoses of anxiety, depression, alcohol dependence and cannabis dependence from ages 18 to 32 years.

Results. The prevalence of lifetime disorder to age 32 was approximately doubled in prospective as compared to retrospective data for all four disorder types. Moreover, across disorders, prospective measurement yielded a mean past-year-to-lifetime ratio of 38% whereas retrospective measurement yielded higher mean past-year-to-lifetime ratios of 57% (NZMHS, NCS-R) and 65% (NCS).

Conclusions. Prospective longitudinal studies complement retrospective surveys by providing unique information about lifetime prevalence. The experience of at least one episode of DSM-defined disorder during a lifetime may be far more common in the population than previously thought. Research should ask what this means for etiological theory, construct validity of the DSM approach, public perception of stigma, estimates of the burden of disease and public health policy.

Received 14 April 2009; Revised 30 June 2009; Accepted 4 July 2009; First published online 1 September 2009

Key words: Epidemiology, longitudinal, prevalence, psychiatry, retrospective.

Introduction

How common is the experience of a mental disorder? How many individuals in the population will experience a diagnosable mental disorder during their lifetimes? The answer to this perennial question of the lifetime population prevalence of mental disorder has many implications: for etiological theories, for service-delivery policy, for public perceptions of the stigma of mental disorder, and for understanding the burden of mental disorder on economic productivity (Insel & Fenton, 2005). Important information about the lifetime population prevalence of mental disorder in the USA has been provided by epidemiological surveys such as the Epidemiological Catchment Area (ECA) study (Robins & Regier, 1991), the National Comorbidity Survey (NCS; Kessler *et al*. 1994), the National Comorbidity Survey Replication (NCS-R; Kessler *et al*. 2005*a*) and the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC; Compton *et al*. 2007; Hasin *et al*. 2007). Surveys offer a key advantage over official mental-health service records for estimating population prevalence because they are able to count individuals who experience disorder but never use mental-health services. However, surveys have the disadvantage of estimating lifetime prevalence using the retrospective method, in which respondents retrospect over past years of their lives to recall whether they have ever experienced mental disorder symptoms. Such retrospective ascertainment is known to undercount lifetime prevalence

* Address for correspondence: T. E. Moffitt, Suite 201 Grey House, 2020 West Main Street, Duke University Box 104410, Durham, NC 27708, USA.
(Email: t.moffitt@iop.kcl.ac.uk or terrie.moffitt@duke.edu)

(Kessler *et al*. 2005*a*) because respondents underreport past disorder symptoms (Simon & VonKorff, 1995), but the extent of this undercounting is unknown.

Longitudinal cohort studies provide an alternative, complementary method for ascertaining lifetime prevalence. Like surveys, longitudinal studies count cases of disorder irrespective of service use. However, they estimate lifetime prevalence of disorder using the prospective method, in which researchers follow a representative cohort of individuals over their life course while undertaking repeated diagnostic assessments. The proportion of cohort members who experience disorder is counted as it accumulates. Because the prospective method reduces the undercounting that results from the retrospective method, there is reason to expect that prospective ascertainment could yield lifetime prevalence estimates that are higher than the existing retrospective ascertainments, and more representative of true lifetime prevalence (Costello *et al*. 2003; Wells & Horwood, 2004; Jaffee *et al*. 2005; Moffitt *et al*. 2007).

The extent of undercounting by retrospective studies could be estimated by comparing results from retrospective and prospective methods. The ideal design would compare prospective *versus* retrospective lifetime prevalence measures of common disorders in the same cohort followed longitudinally from childhood to adulthood. To our knowledge no such study has been undertaken, and completing one could take years. As such, we report a comparison of lifetime prevalence rates derived from a prospective longitudinal study *versus* from three retrospective surveys. Furthermore, for depression, studies show that half of respondents with a documented prior episode do not recall it when interviewed years later (Kendler *et al*. 1993; Andrews *et al*. 1999; Wells & Horwood, 2004), but it is not known if this finding applies to other disorders. Thus, we studied anxiety and substance disorders in addition to depression.

The question of the lifetime prevalence of mental disorders in the population has continued unanswered for many years. After the ECA study revealed rates of mental disorder that were higher than expected by many mental-health professionals, this surprising information prompted concerns about the true rate of disorder in the population and raised questions about the validity of ascertaining disorders using standardized interviews in household surveys (Regier *et al*. 1998; Brugha *et al*. 1999). The NCS and NCS-R were carefully designed to address many of these concerns, but their prevalence rates still seemed too high to many experts (Pincus *et al*. 1998; Regier, 2000; Narrow *et al*. 2002). The NCS-R lifetime prevalence of 46% prompted questions about how many of the identified cases were trivially mild (Insel & Fenton, 2005), but analyses attested that nearly 60% were serious or moderate (Kessler *et al*. 2005*b*).

Vigorous debates have been stimulated by survey reports that a higher than expected number of individuals in the population have a diagnosable mental disorder during their lifetime. Some researchers perceive a large unmet need for mental-health care (NAMHC, 1993; Mechanic, 2003; Insel & Fenton, 2005), whereas others counter that DSM definitions over-medicalize normal behavior (Horwitz & Wakefield, 2007; Parker, 2007). Some researchers propose to correct too-high prevalences downward by requiring more evidence of severity (Narrow *et al*. 2002) whereas others counter that diagnosing mild disorders constitutes a prevention opportunity (Kessler *et al*. 2003*b*; Hickie, 2007).

However, in the heat of the debate little attention has been paid to a disconcerting possibility: that the higher than expected lifetime rate of disorder derived through retrospective surveys may represent an undercount. Based on this realization, the provocative hypothesis has recently been put forward that if individuals were followed for enough years through their lives, almost everyone in the population might have at least one episode of a common disorder such as depression (Andrews *et al*. 2005). To date, progress toward resolving ongoing debates about the lifetime prevalence of mental disorder is hampered because the epidemiological evidence base relies solely on retrospective surveys. This article aims to add prospective data to the evidence base.

We report the cumulative prevalence of DSM-defined (APA, 1994) disorders during the 15-year period from age 18 to 32 years in the prospective longitudinal Dunedin (New Zealand) Study, as compared to retrospective lifetime prevalence for the same age group in the NCS, the NCS-R and the New Zealand Mental Health Survey (NZMHS; Oakley Browne *et al*. 2006). We chose these comparisons for the following reasons. First, we compare the Dunedin Study to the NZMHS because both represent the same nation (yet differ on method of ascertaining lifetime prevalence). Second, we compare the Dunedin Study to the NCS-R, because the NCS-R is considered a contemporary gold-standard source on prevalence. We present NZMHS and NCS-R data to show that the USA and New Zealand do not differ on disorder prevalence (both used the same measure and methods as part of the World Mental Health Surveys; Degenhardt *et al*. 2008). Third, we include the NCS because the NCS-R and NZMHS prevalence rates for substance disorders have been criticized (Hasin & Grant, 2004; Grant *et al*. 2007; Kessler & Merikangas,

2007); thus the NCS rates remain a key standard of comparison for substance disorders.

Depression, anxiety disorders, alcohol dependence and cannabis dependence were chosen for study because they are common in the population and they had been diagnosed in all four studies. Age 18 years was our starting point because the youngest participants in the NCS and NCS-R were aged 18 when diagnosed for adult disorders. Age 32 was our ending point because it is the oldest age that the Dunedin cohort has been diagnosed for adult disorders so far. In addition, under-reporting of past disorder probably affects older survey respondents most. By examining the youngest participants of the surveys we used the surveys' best lifetime prevalence estimates for our comparison. Although age 18–32 does not represent the entire life course, it constitutes the peak age-of-onset window for common disorders.

The Dunedin Study's prospective estimate of lifetime prevalence represents a cumulative count of cases diagnosed during the course of the study, each of which was ascertained in a past-year assessment. Therefore, an essential first step in this research was to compare past-year prevalence rates from the Dunedin Study to those from the NZMHS, NCS-R and NCS. If past-year rates seemed similarly accurate across the three studies but the prospective method yielded higher lifetime rates than the retrospective method, the data would support initial inferences about how much lifetime disorder is underestimated when diagnosis relies on long-term recall.

Method

Samples

Longitudinal participants were members of the Dunedin Multidisciplinary Health and Development Study (Moffitt *et al*. 2001). Of infants born in Dunedin, New Zealand, between April 1972 and March 1973, 1037 children (91% of eligible births; 52% male) participated in the first follow-up at age 3, constituting the base sample for the longitudinal study. Participants represent the full range of socio-economic status in the general population of New Zealand's South Island and were primarily white. Participants attended the Research Unit for a full day of individual data collection. The Otago Ethics Committee approved each phase of the study. Study members gave written informed consent before participating. Assessments were undertaken at ages 3, 5, 7, 9, 11, 13, 15, 18, 21, 26 and most recently at age 32 when we assessed 96% of the 1015 Study members still alive in 2004–2005. This article examines participants who were assessed for mental disorders at ages 18 ($n=930$), 21 ($n=961$), 26 ($n=976$) and 32 years ($n=962$).

Detailed sample descriptions are available elsewhere for the NCS (Kessler *et al*. 1994), NSC-R (Kessler *et al*. 1994, 2004) and NZMHS (Oakley-Browne *et al*. 2006; Wells *et al*. 2006; Degenhardt *et al*. 2008; www.hcp.med.harvard.edu/wmh). All three were national stratified multistage clustered area probability samples of household residents. Numbers of participants are shown in Tables 1 and 2.

Measures

Disorder at ages 18, 21, 26 and 32 years in the Dunedin cohort was measured using the Diagnostic Interview Schedule (DIS-III, DIS-IV; Robins *et al*. 1989, 1995). The DIS was administered in private at the research unit, by trained interviewers with tertiary qualifications and clinical experience in a mental health-related field such as family medicine, clinical psychology or psychiatric social work (i.e. not lay interviewers). Interviewers were kept blind to cohort members' prior data. At ages 18 and 21, diagnoses were made according to the then-current DSM-III-R (APA, 1987) and at ages 26 and 32 diagnoses were made according to the DSM-IV (APA, 1994). In addition to symptom criteria, diagnosis required impairment ratings >2 on a scale from 1 (some impairment) to 5 (severe impairment). Each disorder was diagnosed regardless of the presence of other disorders. Variable construction details, reliability and validity, and evidence of life impairment for diagnoses have been reported previously (Feehan *et al*. 1994; Newman *et al*. 1996; Kim-Cohen *et al*. 2003; Moffitt *et al*. 2007). The reporting period at each assessment was the past 12 months.

Disorder among respondents aged 18–32 years in the NZMHS, NCS-R and NCS was assessed as described previously (Kessler *et al*. 1994, 2004, 2005*a*; Oakley-Browne *et al*. 2006; Wells *et al*. 2006). NCS-R and NZMHS used CIDI version 3.0 to make DSM-IV diagnoses. NCS used CIDI version 1.1 to make DSM-III-R diagnoses. NZMHS data were provided by the New Zealand Ministry of Health and accessed by author J.K. Public domain data were downloaded from www.icpsr.umich.edu/cocoon/SAMHDA/STUDY/06693.xml for NCS-1 and www.icpsr.umich.edu/cocoon/cpes/ncsr/sections/all/sections.xml for NCS-R. (Past-year and lifetime diagnosis variables accessed from websites are given in the online Appendix for readers who use NCS and NCS-R data.)

Statistical analyses

Past-year and lifetime prevalence rates and the confidence intervals (CIs) around them are reported for all four studies. NZMHS CIs were calculated by Taylor Series Linearization using SURVEYFREQ in SAS 9.1.3

Table 1. *Past-year prevalence of common adult mental disorders, by age at diagnostic interview, for informants aged 18 to 32 years. The Dunedin Study is compared to the New Zealand Mental Health Survey (NZMHS) and the two US National Comorbidity Surveys (NCS and NCS-R)*

	Dunedin cohort past-year				Dunedin past-year (mean)	NZMHS past-year	NCS-R past-year	NCS past-year
Disorder	18 years ($n=930$)	21 years ($n=961$)	26 years ($n=976$)	32 years ($n=962$)	18–32 years	18–32 years ($n=3173$)	18–32 years ($n=2676$)	18–32 years ($n=3340$)
Any anxiety								
%	24.3	20.3	24.3	22.2	22.8	19.4	21.9	18.1
(no. of cases)	(226)	(195)	(237)	(214)		(634)	(582)	(591)
[CI]					[21.0–24.5]	[17.4–21.5]	[20.1–23.8]	[16.2–20.2]
Panic								
%	0.9	0.6	3.9	1.7	1.8	2.4	3.2	2.1
(no. of cases)	(8)	(6)	(38)	(16)		(94)	(90)	(71)
[CI]					[1.2–2.1]	[1.8–3.0]	[2.6–4.0]	[1.5–3.0]
Specific phobia								
%	7.4	8.4	7.1	6.1	7.2	9.3	10.4	8.8
(no. of cases)	(69)	(80)	(69)	(59)		(327)	(302)	(277)
[CI]					[6.5–8.8]	[8.0–10.5]	[9.0–12.0]	[7.4–10.4]
Social phobia								
%	13.8	9.7	10.7	8.8	10.8	6.9	8.9	8.4
(no. of cases)	(128)	(92)	(104)	(85)		(327)	(241)	(292)
[CI]					[9.2–11.8]	[8.0–10.5]	[7.7–10.2]	[7.1–9.3]
GAD								
%	1.8	1.9	5.5	7.7	4.2	3.2	3.3	3.3
(no. of cases)	(17)	(18)	(54)	(74)		(107)	(91)	(86)
[CI]					[3.3–4.6]	[2.4–3.9]	[2.6–4.1]	[2.6–4.3]
Depression								
%	17.3	16.8	16.5	16.3	16.7	9.7	10.1	11.3
(no. of cases)	(161)	(161)	(161)	(157)		(317)	(280)	(369)
[CI]					[15.2–18.2]	[8.2–11.1]	[9.1–11.1]	[9.7–13.3]
Alcohol dependence								
%	10.8	18.4	13.6	8.1	12.7	2.9	2.5	10.3
(no. of cases)	(100)	(176)	(133)	(78)		(118)	(54)	(356)
[CI]					[10.7–13.6]	[2.2–3.6]	[1.6–3.7]	[9.0–11.9]
Cannabis dependence								
%	6.6	9.6	9.4	5.4	7.8	2.2	1.0	4.2
(no. of cases)	(61)	(91)	(92)	(52)		(78)	(26)	(139)
[CI]					[6.0–8.4]	[1.6–2.9]	[0.6–1.6]	[3.3–5.3]

GAD, Generalized anxiety disorder; CI, confidence interval.

n's for NCS-1, NCS-R and NZMHS are unweighted.

Certain anxiety disorders were only assessed in Part II of the NZMHS (the long form), thus the total *n* for any anxiety is reduced to 2057. Certain anxiety disorders, alcohol dependence and cannabis dependence were only assessed in Part II of the NCS-R (the long form), thus the total *n* for these variables is reduced to 1728.

Diagnoses followed DSM-III-R in Dunedin at ages 18 and 21 and in the NCS. Diagnoses followed DSM-IV in Dunedin at ages 26 and 32 and in the NCS-R and NZMHS. The DSM version does not seem to affect prevalence in this 18–32 years age group.

Dunedin any anxiety includes panic, specific or social phobia, GAD, agoraphobia, obsessive–compulsive disorder (OCD) and post-traumatic stress disorder (PTSD). NZMHS any anxiety includes panic, specific or social phobia, GAD, agoraphobia, OCD and PTSD (identical to the Dunedin composite). NCS-R any anxiety includes panic, specific or social phobia, GAD, agoraphobia, PTSD and adult/child separation anxiety disorder (according to Kessler *et al*. 2005). NCS any anxiety includes panic, specific or social phobia, GAD and agoraphobia (according to Kessler *et al*. 1994).

Dunedin and NZMHS cannabis dependence includes only cannabis. NCS and NCS-R variables available in the public domain files are for drug dependence, which includes primarily cannabis cases, but also a very small minority of cases dependent on other drugs.

Table 2. *Cumulative lifetime prevalence of common adult mental disorders from age 18 to 32 years. Adult disorders accumulated across four prospective assessments at ages 18, 21, 26 and 32 in the Dunedin Study are compared against lifetime prevalence up to age 32 based on retrospective recall in the New Zealand Mental Health Survey (NZMHS) and the two US National Comorbidity Surveys (NCS and NCS-R)*

Disorder	Dunedin Study ($n=1000$)	NZMHS ($n=3173$)	NCS-R ($n=2676$)	NCS ($n=3340$)
Any anxiety				
%	49.5	28.3	33.1	25.4
(no. of cases)	(495)	(920)	(853)	(840)
[CI]	[46.4–52.6]	[25.8–30.9]	[30.8–35.5]	[23.1–27.8]
Panic				
%	6.5	3.2	4.7	3.0
(no. of cases)	(65)	(121)	(130)	(103)
[CI]	[5.0–8.0]	[2.5–3.9]	[3.9–5.6]	[2.3–4.0]
Specific phobia				
%	18.8	12.7	13.3	10.8
(no. of cases)	(188)	(440)	(385)	(349)
[CI]	[16.4–21.2]	[11.2–14.2]	[11.8–14.9]	[9.2–12.6]
Social phobia				
%	27.9	10.8	13.6	14.3
(no. of cases)	(279)	(388)	(369)	(469)
[CI]	[25.1–30.7]	[9.5–12.1]	[12.2–15.0]	[12.6–16.2]
GAD				
%	14.2	6.1	5.6	4.7
(no. of cases)	(142)	(223)	(159)	(142)
[CI]	[12.0–16.4]	[5.0–7.1]	[4.5–6.7]	[3.8–5.9]
Depression				
%	41.4	18.5	19.0	16.9
(no. of cases)	(414)	(621)	(526)	(582)
[CI]	[38.3–44.5]	[16.7–20.2]	[17.5–20.6]	[14.9–19.0]
Alcohol dependence				
%	31.8	6.3	6.4	16.9
(no. of cases)	(318)	(262)	(151)	(590)
[CI]	[28.9–34.7]	[5.2–7.4]	[5.2–7.9]	[15.1–18.9]
Cannabis dependence				
%	18.0	9.8	3.9	9.7
(no. of cases)	(180)	(354)	(96)	(334)
[CI]	[15.6–20.4]	[8.4–11.1]	[3.2–4.9]	[8.5–11.0]

GAD, Generalized anxiety disorder; CI, confidence interval.

n's for NCS-1, NCS-R and NZMHS are unweighted. The Dunedin Study denominator is *n* assessed at one or more of four study phases = 1000.

Certain anxiety disorders were only assessed in Part II of the NZMHS (the long form), thus the total *n* for any anxiety is reduced to 2057. Certain anxiety disorders, alcohol dependence and cannabis dependence were only assessed in Part II of the NCS-R (the long form), thus the total *n* for these variables is reduced to 1728.

Lifetime prevalence in Dunedin was the sum of past-year cases from assessments at ages 18, 21, 26 and 32. This period thus covered 15 years. Lifetime prevalence in the NZMHS, NCS-R and NCS was based on retrospective recall by respondents aged 18–32 years. The 18–32 years group were able to recall disorder back to childhood. Presuming that the survey respondents could recall back to age 10; then 18-year-olds could recall an 8-year period, 19-year-olds could recall a 9-year period, and so forth, until 31-year-olds could recall a 21-year period and 32-year-olds could recall a 22-year period. Therefore, the average recall period for the three surveys was 15 years, which is the same as the Dunedin Study period.

(SAS Institute, 2004). NCS-R and NCS CIs were estimated using Stata 9.1 (StataCorp, 2005), taking into account their sampling designs as stipulated by the disseminator of the data sets to match methods previously published from these studies (ICPSR, 2006, 2007). The Dunedin Study's past-year prevalence was calculated as the mean past-year prevalence averaged across the four assessment years (ages 18, 21, 26 and 32). NZMHS, NCS-R and NCS past-year prevalence represented groups of respondents aged 18–32 years. Dunedin lifetime prevalence was calculated as the percentage of cohort members ever past-year diagnosed, among cohort members who were interviewed at any of the four assessments ($n=1000$). NZMHS, NCS-R and NCS lifetime prevalence represented retrospective diagnoses among the group of respondents aged 18–32 years. Past-year prevalence from Table 1 was divided by lifetime prevalence from Table 2 to generate the ratio of individuals in each study who ever had the disorder at any time during the 15-year assessment window to individuals who were also diagnosed in the past year.

Results

Past-year prevalence (*Table 1*)

For the composite category of any anxiety disorder and for individual anxiety disorders of panic, specific phobia, social phobia and generalized anxiety, point prevalence estimates were similar and CIs for all anxiety disorders generally overlapped. This suggested no notable discrepancies among the four studies (Cummings & Finch, 2005). For example, the last four columns of Table 1 show that the mean past-year prevalence of any anxiety disorder was 22.8% in the Dunedin cohort, 19.4% in NZMHS, 21.9% in NCS-R and 18.1% in NCS.

For depression and substance disorders, Dunedin cohort past-year rates generally exceeded NZMHS, NCS-R and NCS rates. One partial explanation is sample completeness. The Dunedin Study assesses 96% of an identified birth cohort whereas the three surveys had much less complete samples. The Dunedin Study identifies disordered individuals who are missing from surveys because they are in prison, in hospital, institutionalized, homeless, not found, or refuse surveys. By contrast, NZMHS, NCS-R and NCS excluded individuals in institutions, and approximately 30% of community respondents approached refused these surveys. Such difficult-to-recruit groups are known to have elevated rates of disorder. We tested this by quantifying the 'level of effort' required to locate and recruit each Dunedin Study member for assessment at age 32 (operationalized as the number of contacts). The rate of depression was 20% among the 20% most difficult-to-recruit cohort members as compared to 14% among the 80% who were relatively easier to recruit. Likewise, rates of alcohol dependence were 12% *versus* 7%, and rates of cannabis dependence were 10% *versus* 5%. (Of note, the prevalence of anxiety disorders was unaffected by sample completeness, which is consistent with no difference between Dunedin and the surveys in past-year anxiety prevalence.) Sample completeness implies that it is reasonable for Dunedin past-year rates of depression and substance disorders to exceed survey rates somewhat.

For substance dependence the NZMHS and NCS-R rates were much below the Dunedin rates. However, in part, low rates of substance dependence in NZMHS and NCS-R (1–3%) have been attributed to CIDI 3.0 gate questions that inadvertently stopped the interview too early for some respondents (Grant *et al*. 2007; Kessler & Merikangas, 2007). By contrast, for example, the past-year prevalence of alcohol dependence for this age group is similar in NCS (10%), Dunedin, (12%) and NESARC (9%) (Hasin *et al*. 2007). Finally, another reason that the Dunedin past-year rate of cannabis dependence exceeded the surveys' rates is that longitudinal study members may be more forthcoming about drug-use behaviors that are illegal, as compared to research-naive survey participants interviewed for the first time, because participants who have been interviewed repeatedly learn to trust the study's confidentiality guarantee. Dunedin's prevalence of cannabis dependence has been verified by the Christchurch New Zealand longitudinal study (Boden *et al*. 2006). Taken together, these explanations imply reasonable confidence in the Dunedin past-year rates of substance disorders.

Lifetime prevalence (*Table 2*)

Lifetime prevalence rates in the prospective Dunedin Study were approximately double the retrospective NZMHS and NCS prevalence rates, for every disorder. Dunedin lifetime prevalence rates were also double the NCS-R rates, for every disorder except panic and specific phobia (even for those, prospective prevalence exceeded retrospective prevalence). CIs for the prospective rates did not overlap with those for retrospective rates, indicating that the discrepancies were statistically significant at $p<0.01$ (Cummings & Finch, 2005).

Past-year-to-lifetime ratio (*Table 3*)

In the NZMHS, NCS-R and NCS, one-half to two-thirds of respondents who ever had an episode of

Table 3. *Ratios of past-year prevalence to lifetime prevalence of adult disorders up to age 32. Prospective 12-month-to-lifetime ratios in the Dunedin Study are compared against retrospective 12-month-to-lifetime ratios from the New Zealand Mental Health Survey (NZMHS) and the two US National Comorbidity Surveys (NCS and NCS-R)*

	Ratio: 12-month prevalence/lifetime prevalence			
Disorder	Dunedin Study 18–32 (cohort $n=1000$)	NZMHS 18–32 (cohort $n=3173$)	NCS-R 18–32 (cohort $n=2676$)	NCS 18–32 (cohort $n=3340$)
Any anxiety	0.46	0.69	0.66	0.71
Panic	0.28	0.75	0.69	0.70
Specific phobia	0.38	0.73	0.78	0.81
Social phobia	0.39	0.63	0.66	0.58
GAD	0.30	0.52	0.59	0.70
Depression	0.40	0.52	0.53	0.67
Alcohol dependence	0.40	0.46	0.39	0.61
Cannabis dependence	0.43	0.22	0.26	0.43
Mean ratio	0.38	0.57	0.57	0.65

GAD, Generalized anxiety disorder.

disorder also had an episode during the year they were interviewed for the survey (NZMHS and NCS-R mean ratio = 0.57; NCS mean ratio = 0.65). In the Dunedin cohort, the corresponding ratio was lower; approximately one-third of participants (mean ratio = 0.38).

Discussion

We found that prospective estimates of the lifetime prevalence of DSM-defined disorders markedly exceed retrospective estimates. Our comparison of prospective *versus* retrospective data held constant the 15-year age window from 18 to 32 years, and the historical period when data were collected, from 1990 to 2005. Lifetime prevalence was almost identical in the NZMHS and the NCS-R, ruling out the possibility that cultural or ethnic differences between New Zealand and the USA could account for our findings.

Three findings suggest that lifetime prevalence of disorder is higher than previously estimated by retrospective surveys. First, the percentage of people who experienced lifetime disorder to age 32 was approximately doubled in prospective data as compared to retrospective data. Second, prospective assessment resulted in a mean of only 38% of lifetime cases having disorder during the past year whereas retrospective measurement of lifetime disorder resulted in higher means of 57% (NZMHS, NCS-R) and 65% (NCS) of lifetime cases having disorder during the past year. Third, prospective measurement yielded lifetime estimates that suggest the experience of certain DSM-defined disorders by age 32 may be very common indeed: anxiety disorder (49.5%), depression (41.4%), alcohol dependence (31.8%) and cannabis dependence (18.0%).

We initially compared past-year prevalence rates in the Dunedin Study *versus* past-year prevalence rates in the NZMHS, NCS-R and NCS (Table 1). This comparison was required because only if past-year prevalences in the Dunedin cohort seemed reasonably valid could we later infer that the lifetime prevalences derived as a count of past-year cases were likewise reasonable. There were methodological differences among the studies but despite these differences, the past-year prevalence of disorder in the Dunedin Study was similar to the past-year prevalence in the NZMHS, NCS-R and NCS, or somewhat higher for expected reasons (such as the Dunedin Study's more complete sample). Dunedin lifetime prevalence rates from age 18 to 32 reflect a cumulative count of these past-year cases.

The Dunedin Study's cumulative prevalence of individuals experiencing disorder between ages 18 and 32 years was approximately double the counterpart prevalence in the NZMHS, NCS-R and NCS (Table 2). For example, a count of individuals ever diagnosed at any Dunedin Study assessment revealed that 41.4% of the cohort experienced at least one episode of depression between ages 18 and 32 *versus* 18.5% in the NZMHS, 19.0% in the NCS-R and 16.9% in the NCS. This discrepancy does not arise from cultural differences between the USA and New Zealand because the NCS-R and NZMHS lifetime depression rates match. However, even if the Dunedin Study's past-year depression data are doubted, the identical prospective/retrospective discrepancy was observed for anxiety and substance disorders. Past-year rates of anxiety

disorders were almost the same in all four studies, but the prospective lifetime rates were double the rates from the three retrospective surveys. Furthermore, past-year rates of substance disorders were similar in Dunedin and NCS, but Dunedin prospective lifetime rates were double NCS retrospective rates.

If the past-year diagnoses added up to yield the Dunedin cohort's lifetime prevalence are acceptably valid, then the surviving explanation for the discrepancy must implicate the fundamental difference between prospective *versus* retrospective measurement: recall failure. Research into depression supports the notion that recall failure is substantial. Studies show that half of hospitalized depression cases ceased to be lifetime cases when interviewed with the CIDI 25 years later (Andrews *et al*. 1999), 10% of depression cases diagnosed at baseline ceased to be lifetime cases when reinterviewed only 3 years later (Newman & Bland, 1998), and half of longitudinal cohort members who previously reported depression did not recall their episodes by age 21 years (Wells & Horwood, 2004). Another indicator of recall failure is that, among primary-care patients interviewed up to age 65, most recalled their first depression episode as occurring within 5 years of the interview, which was deemed implausible given depression's peak age of onset in young adulthood (Simon *et al*. 1995). One analysis indicated that plausible rates of recall failure (concealing 2–4% of depression cases per year) accumulated across the lifetime could account for retrospective surveys' low prevalence (Patten, 2003). An analysis that modeled Dutch and Australian national surveys to correct for recall failure estimated the lifetime prevalence of depression to be 30% in men and 40% in women (Kruijshaar *et al*. 2005). Our findings suggest that this amount of recall failure applies beyond depression, to other disorders.

Who are all these people who experience disorder and then forget it? It is possible that people who under-report have only mild disorder, but this is unlikely to be the full explanation because marked under-reporting occurs among individuals hospitalized for depression (Andrews *et al*. 1999). A check of the data for Dunedin cohort members with lifetime disorder revealed that many had not experienced disorder that was chronic or recurrent, at least not up to age 32. Of lifetime cases, 53% of those with anxiety, 60% of those with depression, 61% of those with alcohol dependence and 57% of those with cannabis dependence had been diagnosed with the disorder at only one of our past-year assessments. (Of lifetime cases, the percentages of cases diagnosed twice and diagnosed three or more times were respectively: 47% and 22% for anxiety disorder, 40% and 12% for depression, 39% and 12% for alcohol, and 42% and 18% for cannabis.) That half of lifetime cases were diagnosed only once in our longitudinal study suggests a hypothesis: retrospective surveys may undercount primarily individuals who have relatively short-term disorder or single episodes. Testing this hypothesis requires undertaking retrospective interviews in a prospectively studied cohort, to reveal which prospectively diagnosed cases go undetected retrospectively. Lacking retrospective lifetime interviews in the Dunedin Study, we could not carry out this test. However, our comparison of prospective *versus* retrospective past-year-to-lifetime ratios is relevant (Table 3). In prospective studies, many more respondents have lifetime disorder than have past-year disorder, yielding a low past-year-to-lifetime ratio. This finding is expected. By contrast, in many retrospective surveys almost as many respondents have past-year disorder as have lifetime disorder, yielding a higher past-year-to-lifetime ratio. It is implausible that most respondents who report that they ever in their lives had an episode also happen to have an episode during the year they are interviewed for a survey (Kessler *et al*. 2002). This implausible result from retrospective surveys could be explained if respondents who have long-standing, chronic or recurrent disorder are particularly likely to remember and report symptoms from the long-distant past, whereas respondents who experienced short-term, single episodes of disorder are likely to forget them, regardless of severity (Simon & VonKorff, 1995). Supporting evidence comes from a two-wave study of depression that revealed that short illness duration is associated with unreliable reporting of lifetime depression (Foley *et al*. 1998).

Limitations

The data we used to estimate prospective prevalence come from one cohort in New Zealand. However, similarly high cumulative prevalence rates have been reported by researchers who have followed adolescent cohorts to young adulthood while conducting repeated diagnostic assessments (using different standardized interview instruments) in North Carolina (Costello *et al*. 2003), New York (Jaffee *et al*. 2005) and Oregon (Lewinsohn *et al*. 1993), and in a 25-year longitudinal study of Australian teachers (Wilhelm *et al*. 2006). That lifetime prevalence accumulates with repeated longitudinal measurement was confirmed by follow-ups in the ECA study (Regier *et al*. 1998) and the NCS (Kessler *et al*. 2007). For example, when the NCS-1 sample was followed up, NCS-1 lifetime depression prevalence was 21%, but this rose to 29% by adding NCS-2 (Kessler *et al*. 2007).

Our study has three additional design limitations. However, all three indicate that the true lifetime prevalence of mental disorder may in fact be higher

than we have been able to estimate in the Dunedin Study. First, NZMHS, NCS-R and NCS 18–32-year-old respondents could retrospectively report disorder they recalled as having occurred before age 18, whereas our Dunedin Study cumulative lifetime count began only at age 18. (The three surveys did not have respondents under age 18 to make Table 1's comparisons of past-year prevalence. Therefore, although the Dunedin Study has juvenile diagnoses we could not include them in this article.) As an example, adding depression diagnoses made before age 18 increases the Dunedin lifetime prevalence of depression from 41% to 44%. This limitation has the net effects of lowering the Dunedin Study's estimates of lifetime prevalence and narrowing the prospective *versus* retrospective discrepancy we reported here.

A second limitation is that gaps between the Dunedin Study's four 12-month assessment windows did not allow us to count individuals who experienced an episode of disorder only between windows. We previously reported that our 'net' of 1-year DIS diagnoses at ages 18, 21, 26 and 32 has captured all but eight of the cohort members who reported treatment for mental-health or substance-use problems between assessment windows (Moffitt *et al*. 2007). Nevertheless, the number of cohort members we failed to count here because their only episodes of disorder occurred between study windows and went untreated is unknown. This limitation has the net effects of lowering our estimate of lifetime prevalence and narrowing the prospective *versus* retrospective discrepancy.

A third reason why Dunedin's lifetime rates are underestimates is that Dunedin data are right-hand censored at age 32. Retrospective surveys suggest that many new cases should be expected after age 32 (Kessler *et al*. 2003*a*). On the one hand, the number of new-onset cases after age 32 has probably been overestimated because survey respondents often recall their onset age as older than it was, and forget episodes from early life (Simon *et al*. 1995). On the other hand, new cases will be diagnosed as the Dunedin cohort ages. Our estimate to age 32 is an underestimate of lifetime prevalence for the full life course.

Implications

We compared retrospective *versus* prospective methods of ascertaining lifetime prevalence while holding constant the use of the DSM definitional approach to diagnosis in both types of studies. Therefore, this article is uninformative (and agnostic) about the validity of diagnoses of depression, anxiety and substance dependence as defined by DSM-IV. That is a separate debate (Horwitz & Wakefield, 2007). Our rather more modest aim was to point out that objections voiced to surveys' higher than expected lifetime prevalence of disorder are objections to prevalence that is only half what it could be in reality, because a very great deal of disorder has been lost to recall failure. Limitations of our research are such that we cannot here provide an estimate of the true prevalence of lifetime psychiatric disorders, but the findings can be taken as evidence that existing and oft-cited retrospective prevalence rates undercount not trivially, but substantially. This substantial undercounting is consequential because it can generate misleading findings in etiological research (Kendler *et al*. 1993; Foley *et al*. 1998) and misleading estimates of economic disease burden (Tang & Lopez, 1997). It is time for critical thinking about retrospective data.

It is not a new idea that retrospective surveys underdetect mental disorder (Kramer *et al*. 1980). However, despite repeated demonstrations that this underdetection is real, resources are still invested in collecting retrospective data and journals continue to publish them as epidemiological information. In addition, peer reviewers still recommend rejection of papers from longitudinal studies on the basis that their cumulative number of prospectively diagnosed cases is far too high, as compared to survey prevalence rates. Survey researchers have taken care to explain that their surveys' prevalence estimates are *lower* than they could be (Kessler *et al*. 1994, 2003*a*, 2005*a*) but paradoxically, much criticism continues to stem from the widespread belief that surveys' prevalence rates are *higher* than they should be. If the much-higher-than-expected lifetime prevalence now emerging from prospective studies is correct, then this widespread belief is a myth. If lifetime prevalence rates are as high as those we report here (or higher), debates may shift. It may be time to stop asking how surveys can achieve acceptably low rates of disorder. Instead, researchers might begin to ask why so many people experience a DSM-defined disorder at least once during their lifetimes, and what this prevalence means for etiological theory, the construct validity of the DSM approach to defining disorder, service-delivery policy, the economic burden of disease, and public perceptions of the stigma of mental disorder.

Acknowledgements

This research was supported by the New Zealand Health Research Council, the US National Institute of Health (grants MH45070, MH49414, MH077874, AG032282), and the UK Medical Research Council (grants G0100527, G0601483). A. Caspi is a Royal Society Wolfson Merit Award holder. We thank Dunedin Study founder P. Silva, study staff, the study members and their families.

Declaration of Interest

None.

Note

Supplementary material accompanies this paper on the Journal's website (http://journals.cambridge.org/psm).

References

Andrews G, Anstey K, Brodaty H, Issakidis C, Luscombe G (1999). Recall of depressive episodes 25 years previously. *Psychological Medicine* **29**, 787–791.

Andrews G, Poulton R, Skoog I (2005). Lifetime risk for depression: restricted to a minority or waiting for most? *British Journal of Psychiatry* **187**, 495–496.

APA (1987). *Diagnostic and Statistical Manual of Mental Disorders*, 3rd edn, revised. American Psychiatric Association: Washington, DC.

APA (1994). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. American Psychiatric Association: Washington, DC.

Boden JM, Fergusson DM, Horwood LJ (2006). Illicit drug use and dependence in a New Zealand birth cohort. *Australian and New Zealand Journal of Psychiatry* **40**, 156–163.

Brugha TS, Bebbington PE, Jenkins R (1999). A difference that matters: comparisons of structured and semi-structured psychiatric diagnostic interviews in the general population. *Psychological Medicine* **29**, 1013–1020.

Compton WM, Thomas YF, Stinson FS, Grant BF (2007). Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry* **64**, 566–576.

Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry* **60**, 837–844.

Cummings G, Finch S (2005). Inference by eye: confidence intervals and how to read pictures of data. *American Psychologist* **60**, 170–180.

Degenhardt L, Chiu WT, Sampson N, Kessler RC, Anthony JC, Angermeyer M, Bruffaerts R, de Girolamo G, Gureje O, Huang Y, Karam A, Kostyuchenko S, Lepine JP, Mora ME, Neumark Y, Ormel JH, Pinto-Meza A, Posada-Villa J, Stein DJ, Takeshima T, Wells JE (2008). Toward a global view of alcohol, tobacco, cannabis, and cocaine use: findings from the WHO World Mental Health Surveys. *PLoS Medicine* **5**, e141.

Feehan M, McGee R, Nada Raja S, Williams SM (1994). DSM-III-R disorders in New Zealand 18-year-olds. *Australian and New Zealand Journal of Psychiatry* **28**, 87–99.

Foley DL, Meale MC, Kendler KS (1998). Reliability of lifetime history of major depression: implications for heritability and comorbidity. *Psychological Medicine* **28**, 857–870.

Grant BF, Compton WM, Crowley TJ, Hasin DS, Helzer JE, Li TK, Rounsaville BJ, Volkow ND, Woody GE (2007). Errors in assessing DSM-IV substance use disorders. *Archives of General Psychiatry* **64**, 379–380.

Hasin DS, Grant BF (2004). The co-occurrence of DSM-IV alcohol abuse in DSM-IV alcohol dependence. *Archives of General Psychiatry* **61**, 891–896.

Hasin DS, Stinson FS, Grant BF (2007). Prevalence, correlates, disability and comorbidity of DSM-IV alcohol abuse and dependence in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry* **64**, 830–842.

Hickie I (2007). Is depression overdiagnosed? *British Medical Journal* **335**, 329.

Horwitz AV, Wakefield JC (2007). *The Loss of Sadness*. Oxford University Press: New York.

ICPSR (2006). Inter-university Consortium for Political and Social Research [producer and distributor] National Comorbidity Survey; Replication (NCS-R), 2001–2003 [Computer file]. Conducted by Harvard Medical School, Department of Health Care Policy/University of Michigan, Survey Research Center. ICPSR04438-v3. Ann Arbor, MI.

ICPSR (2007). Inter-university Consortium for Political and Social Research [producer and distributor], National Comorbidity Survey: Baseline (NCS-1), 1990–1992 [computer file]. Conducted by University of Michigan, Survey Research Center. ICPSR06693-v4. Ann Arbor, MI.

Insel TR, Fenton WS (2005). Psychiatric epidemiology, it's not just about counting anymore. *Archives of General Psychiatry* **62**, 590–592.

Jaffee SR, Harrington HL, Cohen P, Moffitt TE (2005). Cumulative prevalence of psychiatric disorder in youths. *Journal of the American Academy of Child and Adolescent Psychiatry* **44**, 406–407.

Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ (1993). The lifetime history of major depression in women: reliability of diagnosis and heritability. *Archives of General Psychiatry* **50**, 863–870.

Kessler RC, Andrade LH, Bijl RV, Offord DR, Demler OV, Stein DJ (2002). The effects of co-morbidity on the onset and persistence of generalized anxiety disorder in ICPE surveys. *Psychological Medicine* **32**, 1213–1225.

Kessler RC, Berglund P, Chiu W-T, Demler O, Heeringa S, Hiripi E, Jin R, Pennell B-E, Walters EE, Zaslavsky A, Zheng H (2004). The US National Comorbidity Survey Replication (NCS-R): design and field procedures. *International Journal of Methods in Psychiatric Research* **13**, 69–92.

Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, Rush AJ, Walters EE, Wang PS; National Comorbidity Survey Replication (2003*a*). The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *Journal of the American Medical Association* **289**, 3095–3105.

Kessler RC, Berglund PA, Demler O, Jin R, Merikangas KR, Walters EE (2005*a*). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication (NCS-R). *Archives of General Psychiatry* **62**, 593–602.

Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE (2005*b*). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry* **62**, 617–627.

Kessler RC, Gruber M, Hettema JM, Hwang I, Sampson N, Yonkers KA (2007). Co-morbid major depression and generalized anxiety disorders in the National Comorbidity Survey follow-up. *Psychological Medicine* **38**, 365–374.

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 Study. *Archives of General Psychiatry* **51**, 8–19.

Kessler RC, Merikangas KR (2007). Drug use disorders in the National Comorbidity Survey: have we come a long way? – Reply. *Archives of General Psychiatry* **64**, 381–382.

Kessler RC, Merikangas KR, Berglund P, Eaton WW, Koretz DS, Walters EE (2003*b*). Mild disorders should not be eliminated from the DSM-V. *Archives of General Psychiatry* **60**, 1117–1122.

Kim-Cohen J, Caspi A, Moffitt TE, Harrington HL, Milne BJ, Poulton R (2003). Prior juvenile diagnosis in adults with mental disorder: developmental follow-back of a prospective-longitudinal cohort. *Archives of General Psychiatry* **60**, 709–717.

Kramer M, von Korff M, Kessler L (1980). The lifetime prevalence of mental disorders: estimation, uses and limitations. *Psychological Medicine* **10**, 429–435.

Kruijshaar ME, Barendrecht J, Vos T, de Graaf R, Spijker J, Andrews G (2005). Lifetime prevalence estimates of major depression: an indirect estimation method and a quantification of recall bias. *European Journal of Epidemiology* **20**, 103–111.

Lewinsohn PM, Hops H, Roberts RE Seeley JR, Andrews JA (1993). Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. *Journal of Abnormal Psychology* **102**, 133–144.

Mechanic D (2003). Policy challenges in improving mental health services: some lessons from the past. *Psychiatric Services* **53**, 1227–1232.

Moffitt TE, Caspi A, Rutter M, Silva PA (2001). *Sex Differences in Antisocial Behaviour: Conduct Disorder, Delinquency, and Violence in the Dunedin Longitudinal Study*. Cambridge University Press: Cambridge, UK.

Moffitt TE, Harrington HL, Caspi A, Kim-Cohen J, Goldberg D, Gregory AM, Poulton R (2007). Depression and generalized anxiety disorder: cumulative and sequential comorbidity in a birth cohort followed to age 32. *Archives of General Psychiatry* **64**, 651–660.

Murray CJL, Lopez AD (1997). Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *Lancet* **349**, 1498–1504.

NAMHC (1993). Health care reform for Americans with severe mental illnesses: report of the National Advisory Mental Health Council. *American Journal of Psychiatry* **150**, 1447–1465.

Narrow WE, Rae DS, Robins LN, Regier DA (2002). Revised prevalence of mental disorders in the United States: using a clinical significance criterion to reconcile 2 surveys' estimates. *Archives of General Psychiatry* **59**, 115–123.

Newman DL, Moffitt TE, Caspi A, Magdol L, Silva PA, Stanton W (1996). Psychiatric disorder in a birth cohort of young adults: prevalence, comorbidity, clinical significance, and new cases incidence from age 11 to 21. *Journal of Consulting and Clinical Psychology* **64**, 552–562.

Newman SC, Bland RC (1998). Incidence of mental disorders in Edmonton. *Journal of Psychiatry Research* **32**, 273–282.

Oakley-Browne MA, Wells JE, Scott KM (2006). *Te Rau Hinengaro: The New Zealand Mental Health Survey*. NZ Ministry of Health: Wellington (www.moh.govt.nz).

Parker G (2007). Is depression overdiagnosed? *British Medical Journal* **335**, 328.

Patten SB (2003). Recall bias and major depression lifetime prevalence. *Social Psychiatry and Psychiatric Epidemiology* **38**, 290–296.

Pincus HA, Zarin DA, First M (1998). 'Clinical significance' and DSM-IV. *Archives of General Psychiatry* **55**, 1145.

Regier DA (2000). Community diagnosis counts. *Archives of General Psychiatry* **57**, 223–224.

Regier DA, Kaelber CT, Rae DS, Farmer ME, Knauper B, Kessler RC, Norquist GS (1998). Limitations of diagnostic criteria and assessment instruments for mental disorders: implications for research and policy. *Archives of General Psychiatry* **55**, 109–115.

Robins LN, Cottler L, Bucholz KK, Compton W (1995). *Diagnostic Interview Schedule for DSM-IV*. Washington University School of Medicine: St Louis, MO.

Robins LN, Helzer JE, Cottler L, Goldring E (1989). *Diagnostic Interview Schedule, Version III-R*. Washington University School of Medicine: St Louis, MO.

Robins LN, Regier DA (1991). *Psychiatric Disorders in America*. Free Press: New York.

SAS Institute (2004). SAS OnlineDoc® 9.1.3. SAS Institute Inc.: Cary, NC.

Simon GE, VonKorff M (1995). Recall of psychiatric history in cross-sectional surveys: implications for epidemiologic research. *Epidemiology Review* **17**, 211–227.

Simon GE, VonKorff M, Uston TB, Gater R, Gureje O, Sartorius N (1995). Is the lifetime risk of depression actually increasing? *Journal of Clinical Epidemiology* **48**, 1109–1118.

StataCorp (2005). *Stata Statistical Software: Release 9.1*. Stata Corporation: College Station, TX.

Wells JE, Horwood LJ (2004). How accurate is recall of key symptoms of depression? A comparison of recall and longitudinal reports. *Psychological Medicine* **34**, 1001–1011.

Wells JE, Oakley Browne MA, Scott KM, McGee MA, Baxter J, Kokaua J (2006). Te Rau Hinengaro: the New Zealand Mental Health Survey: overview of methods and findings. *Australian and New Zealand Journal of Psychiatry* **40**, 835–844.

Wilhelm K, Mitchell PB, Niven H, Finch A, Wedgwood L, Scimone A, Blair IP, Parker G, Schofield PR (2006). Life events, first depression onset and the serotonin transporter gene. *British Journal of Psychiatry* **188**, 210–215.