Published in final edited form as:
J Psychiatr Res. 2012 September ; 46(9): 1217–1222. doi:10.1016/j.jpsychires.2012.05.017.

Lifetime rates of psychopathology in single versus multiple diagnostic assessments: Comparison in a community sample of probands and siblings

Thomas M. Olino[1,‡], **Stewart A. Shankman**[2], **Daniel N. Klein**[3], **John R. Seeley**[4], **Jeremy W. Pettit**[5], **Richard F. Farmer**[4], and **Peter M. Lewinsohn**[4]

[1]University of Pittsburgh, Pittsburgh, PA, USA

[2]University of Illinois-Chicago, Chicago, IL, USA

[3]Stony Brook University, Stony Brook, NY, USA

[4]Oregon Research Institute, Eugene, OR, USA

[5]Florida International University, Miami, FL, USA

Abstract

Lifetime prevalence rates of psychopathology vary a great deal depending on whether they are estimated from cross-sectional or prospective longitudinal studies, with the former yielding significantly lower rates. Such findings, however, come from comparisons of separate studies from different countries and cohorts. Here, we compare lifetime rates of psychopathology between a community sample of individuals assessed on multiple occasions to their siblings who completed only a single diagnostic evaluation. Data come from the Oregon Adolescent Depression Project. We included 442 original participants who completed four prospective diagnostic assessments over the course of fifteen years, and 657 of their siblings who completed a single lifetime assessment. Comparisons of rates of depressive, bipolar, anxiety, and substance use disorders were made using survival analysis. We found that rates of depressive disorders, specifically major depressive disorder, were elevated among individuals who completed multiple diagnostic assessments relative to individuals who completed a single lifetime assessment. We did not find significant differences in rates of aggregate anxiety, bipolar, or substance use disorders. Within a single cohort, cross-sectional surveys appear to underestimate the lifetime rates of major depression relative to prospective, longitudinal designs. This suggests that disorders with an episodic course may be under-reported in cross-sectional surveys. Rates of anxiety, bipolar, and substance use disorders did not differ across assessment methods. To further evaluate method effects on lifetime estimates of psychopathology, future work may benefit from comparing rates of

‡Please direct correspondence to Thomas M. Olino, Western Psychiatric Institute and Clinic, 3811 O'Hara St, Pittsburgh, PA, 15213; olinotm@upmc.edu; Phone: (412) 383-5434; Fax: (412) 383-5426..

Thomas M. Olino conceptualized the research question, planned and conducted the data analysis, and wrote the first draft of the manuscript.
Stewart A. Shankman and Daniel N. Klein provided critical feedback on the conceptualization of the data analysis.
John R. Seeley, Jeremy W. Pettit, Richard F. Farmer, and Peter M. Lewinsohn provided access to the data and provided critical feedback on the manuscript.
All authors provided valuable feedback contributing to the final submitted manuscript.

retrospectively- and prospectively-derived diagnoses in individuals who are repeatedly assessed over a lengthy follow-up period.

Keywords

assessment; psychopathology; prevalence; epidemiology

Research on prevalence rates of psychopathology have important implications for justifying service allocation and policy decisions and estimating the economic cost and public health burden of psychiatric illness (Insel and Fenton, 2005). Such work may also provide guidance about defining cases in a diagnostic system, such as distinguishing non-disorder from disorder or for determining symptom thresholds required for a positive diagnosis. To estimate lifetime disorder prevalence rates, epidemiological work has generally relied on retrospective, cross-sectional, studies of individuals with a wide range of ages (Andrade et al., 2002; Compton et al., 2007; Fergusson and Horwood, 2001; Kessler et al., 2005; Kessler et al., 1994; Robins and Regier, 1991; Vollebergh et al., 2001). Although it is efficient, using retrospective recall of past disorders to derive lifetime prevalence rates, has several potential biases. Memory effects, for example, might differentially affect the recall of lifetime disorders between the young and old, producing the appearance of cohort effects (Fombonne, 1994). Similarly, when lifetime disorders are assessed on a single occasion, recall of earlier disorders may be biased depending on whether individuals are presently in an episode or disorder-free (Aneshensel et al., 1987). Alternative assessment strategies, such as those that involve multiple assessments with the same cohort over an extended time interval, may be required to minimize these limitations.

Recently, researchers have adopted a developmental epidemiological approach to estimate the lifetime prevalence of psychopathology (e.g., Costello et al., 2003; Moffitt et al., 2001). Diagnostic assessments in such studies were prospective and assessed the onset and course of disorders over multiple time intervals with the same age-based cohort (e.g., disorders over the previous 2 years rather than `ever' in life). Such an approach has several advantages, such as a reducing recall burden and the impact of confounding age with the recall period evaluated. This approach also presents a number of challenges. As participants develop, for example, diagnostic interviews may be modified so as to be more developmentally appropriate, or there may be revisions of the diagnostic criteria. In longitudinal studies, participants may also discontinue their participation, with attrition being non-random and thus affecting some groups of participants more than others (e.g., those with more severe forms of psychopathology).

Although research is limited, rates of specific forms of psychopathology differ between prospective observational studies and cross-sectional studies, with rates from the former type of study usually higher than that of the latter. However, no study has directly compared the lifetime prevalence rates across prospective, longitudinal assessments and a single cross-sectional lifetime assessment in the same dataset. Based on existing datasets, Moffitt et al. (2010) compared lifetime rates of psychopathology, specifically depression, anxiety, and alcohol and cannabis use disorders, estimated from both one prospective longitudinal study (Moffitt *et al.*, 2001) and several retrospective cross-sectional studies (i.e., the National Comorbidity Study [NCS; Kessler *et al.*, 1994], NCS-R National Comorbidity Study Replication [NCS-R; Kessler et al., 2005], and New Zealand Mental Health Survey [NZMHS; Oakley-Browne *et al.*, 2006]). When comparing across studies, Moffitt et al. (2010) found that the prevalence of past year disorders were similar across prospective, longitudinal, and retrospective, cross-sectional studies, suggesting that recall biases were not functioning differently across designs in the short-term. However, lifetime prevalence rates

of depression, aggregate and individual anxiety disorders, and alcohol and cannabis dependence disorders in prospective studies were approximately double the rates observed in the retrospective, cross-sectional studies. This suggests that prospective assessments may provide more accurate estimates of psychopathology and that cross-sectional studies may underestimate lifetime rates of depressive, anxiety, and substance use disorders (SUDs).

In Moffitt et al (2010), studies selected for comparison were those conducted (a) within the same country but with different sampling methodologies (i.e., Dunedin longitudinal study, NZMHS cross-sectional study), and (b) in different countries with the same cross-sectional methodology (i.e., NZMHS with the NCS and NCS-R). Although the authors were able to control for sampling method and assessment region, other uncontrolled factors may have affected their results, including cohort effects. The cross-sectional studies included participants who were assessed when aged 18–32, which includes a mixture of individuals who were assessed both within and outside of the peak age range for psychiatric disorder risk. Analyses nonetheless treated lifetime diagnoses as being present or absent, which implies that all individuals were assessed through the full period of risk. While the authors reported findings for similar-aged participants with lifetime and past-year disorders, the birth cohorts that they represented differed across samples. Some emerging evidence demonstrates significant increases in rates of emotional problems and psychopathology in more recent birth cohorts (Collishaw et al., 2010); consequently, prevalence comparisons across different cohorts may be confounded by generational effects. Ideally, to reduce such potential confounding effects, studies should directly compare lifetime prevalence rates derived from prospective, longitudinal assessments with those from a single cross-sectional lifetime assessment within the same sample. To our knowledge no such studies presently exist. A next best approach is to directly compare lifetime prevalence rates across prospective, longitudinal assessments and a single cross-sectional assessment among two highly similar cohorts, such as participants and their siblings, from the same study.

This paper examines lifetime prevalence rates of psychopathology between two sets of individuals from the same region, cohort, and family of origin: probands prospectively assessed for psychiatric disorders on four separate occasions and their siblings who were similarly assessed on a single occasion. Data come from the Oregon Adolescent Depression Project (OADP; Lewinsohn et al., 1993), which began as a longitudinal, epidemiological study of adolescent psychopathology. Probands completed up to four diagnostic interviews over the course of 15 years. A family study component was later added to the study that assessed lifetime psychopathology in siblings based on a single measurement occasion. An advantage of comparing probands to their siblings is that the two samples come from the same birth cohort. Additionally, probands and siblings share genetic and environmental variance. While the age range of participants reported on here is wide (i.e., age 13–45), survival analyses are implemented to model heterogeneity in the length of surveillance of disorder onset.

Consistent with Moffitt et al. (2010), we anticipate that higher lifetime rates of psychopathology will be found in prospective, longitudinal assessments relative to a single, lifetime assessment. However, we speculate that this result will be most pronounced for disorders with more episodic than chronic courses. Thus, we anticipate that rates of major depressive disorder (MDD) will be higher when based on prospective assessments than when derived from a single assessment. However, because dysthymic disorder, anxiety disorders (Penninx *et al.*, 2011), and substance use disorders (SUDs) (Rohde et al., 2001; Sher et al., 2005) tend to be more chronic conditions, we expect that lifetime prevalence rates will not differ as markedly according to assessment method. That is, individuals will likely have a longer history with these problems that will, in turn, reduce the possible influence of recall biases related to temporal factors. We also examined lifetime rates of

bipolar disorder, which may present in an episodic course. However, management of bipolar disorder generally requires long-term attention. Thus, we make no *a priori* hypotheses for rates of bipolar disorder.

Materials and Methods

Participants

The present study uses data from the OADP (Lewinsohn et al., 1993), which included multiple diagnostic assessments of probands and a single lifetime assessment of their siblings. All procedures were approved by the local institutional review board and all participants provided written informed consent before procedures were implemented.

Probands—Probands came from a large cohort of high school students who were assessed twice during adolescence (~ ages 16 and 17), a third time at approximately age 24, and a fourth time at approximately age 30. Participants were randomly selected from nine high schools in western Oregon. A total of 1,709 adolescents (ages 14–18; mean age 16.6, SD = 1.2; 52.1% female) completed the initial (T_1) assessments between 1987 and 1989. The participation rate at T_1 was 61%. Approximately one year later, 1,507 of the adolescents (88%) returned for a second evaluation (T_2). Diagnostic information for all assessments were the result of direct interviews with the identified proband (i.e., parental interviews were not conducted). Differences between the sample and the larger population from which it was selected, and between participants and those who declined to participate or dropped out of the study before T_2, were small (Lewinsohn et al., 1993). However, individuals with a history of a disruptive behavior disorder at T_1 were more likely to drop-out of the study (16.8% vs. 6.0%, $\chi^2[1, N=1{,}709] = 31.22$, $p < .001$).

All adolescents with a history of a psychopathology by T_2 ($n = 644$) and a random sample of adolescents with no history of psychopathology by T_2 ($n = 457$) were invited to participate in a third (T_3) evaluation. All non-white T_2 participants were retained in the T_3 sample to maximize racial and ethnic diversity. Such a stratification strategy reduces study costs by maximizing the representativeness of the study population with fewer participants. Of the 1,101 T_2 participants selected for a T_3 interview, 941 (85%) completed the age 24 evaluation. The diagnostic histories of T_3 participants were not significantly different at T_2 when compared with those who dropped out of the study between T_2 and T_3 or not selected as a result of the stratification procedure. At age 30, all T_3 participants were asked to complete another interview assessment (mean age = 30.45, SD = .70, range = 28–34). Of the 941 who participated in the T_3 assessment, 816 (87%) completed the T_4 assessment. Differences were small between those who participated in T_3 but not T_4 compared to those who participated in both T_3 and T_4 (Olino et al., 2008).

Siblings—At the time of the T_3 proband assessment, all first degree family relatives (i.e., parents and full siblings) were asked to participate in lifetime diagnostic assessments. We focus on siblings to minimize cohort effects that may be observed when comparing parents and offspring and to increase similarity in age and environment of individuals with multiple and single diagnostic assessments. We only included data from siblings who participated in direct interviews (versus diagnostic information derived from informant reports and best-estimate diagnoses) to maximize comparability between assessment methods across probands and siblings. Diagnostic information was collected on 1,094 siblings, and direct interview data were available for 736.

To maximize the observation period of these participants, data from probands were included in the present analyses provided that diagnostic information was available for the T_4

assessment *and* if lifetime diagnostic data were available from direct interviews for at least one sibling. Using these criteria, the final sample included 442 probands and 657 siblings.

Proband diagnostic measures—At T_1 and T_2, participants were interviewed with a version of the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS; Orvaschel et al., 1982) that combined features of the Epidemiologic and Present Episode versions and included additional items to derive Diagnostic and Statistical Manual of Mental Disorders, 3rd edition revised (DSM-III-R; American Psychiatric Association, 1987) diagnoses. Follow-up assessments at T_2 and T_3 were jointly administered with the Longitudinal Interval Follow-Up Evaluation (LIFE; Keller *et al.*, 1987). The K-SADS/LIFE procedure provided information regarding the onset and course of disorders since the previous interview. The T_4 interview consisted of a joint administration of the LIFE and the Structured Clinical Interview for DSM-IV (SCID; First et al., 1996) to probe for new or continuing episodes since T_3. Diagnoses were based on DSM-III-R criteria for T_1 and T_2 and Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV; American Psychiatric Association, 1994) criteria for T_3 and T_4. For this report, lifetime diagnoses were used as the indicators of psychopathology, such that a proband was considered to have the diagnosis if it was identified as a past or current disorder at any of the four assessments.

A subset of interviews from each wave was rated from audio or videotapes by a second interviewer for reliability purposes: $T_1 = 263$, $T_2 = 162$, $T_3 = 190$, and $T_4 = 124$ interviews. Diagnostic agreement among raters was indexed by kappa. To avoid potential inflation, deflation, and/or unreliability of the kappa statistic, inter-rater reliability was calculated only for categories diagnosed 10 or more times by both raters combined. Fleiss (1981) provides guidelines for the interpretation of kappa, whereby values ≥ 0.75 denote excellent agreement beyond chance, those between 0.75 and 0.40 are indicative of good to fair agreement, and coefficients < 0.40 reflect poor agreement. Across the four assessment waves, inter-rater diagnostic reliability was good to excellent for disorders that occurred with sufficient frequency. The reliability of one disorder, major depressive disorder (MDD), could be determined for each of the 4 waves (*M* kappa = .84; *range*: .81 – .86). Alcohol abuse/dependence and cannabis abuse/dependence were diagnosed with sufficient frequency among raters during 3 of the 4 waves, and the mean kappas were, respectively, .77 (range: .74 to .82) and .79 (range: .72 to .83). Hard drug abuse/dependence was diagnosed with sufficient frequency among raters during 2 of the 4 waves, with the mean kappa for this disorder being .76 (range: .69 to .83). Kappa coefficients for dysthymia (.56), specific phobias (.66), and panic disorder (.81) could only be determined for 1 of the 4 assessment waves. Social phobia, bipolar spectrum disorders, and obsessive-compulsive disorder (OCD) were not diagnosed with sufficient frequency during any assessment wave to allow an evaluation of diagnostic reliability.

Sibling diagnostic measures—Siblings of OADP participants were interviewed using the Structured Clinical Interview for DSM-IV, non-patient version (SCID; First et al., 1996) at the time of the T_3 assessment of the proband. Interviewers conducted sibling assessments without knowledge of proband diagnoses. The interrater reliability of lifetime diagnoses (based on 184 randomly selected interviews) was excellent for depression (including both MDD and DYS; $\kappa = .94$), any anxiety disorder ($\kappa = .90$), AUD ($\kappa = .86$), and SUD, inclusive of alcohol, cannabis, hard drug abuse/dependence ($\kappa = .89$).

Proband interviews at T_3 and T_4 and sibling interviews were conducted by telephone, which generally yields comparable results to face-to-face interviews (Rohde et al., 1997; Sobin et al., 1993). Most interviewers had advanced degrees in a mental health field and several years of clinical experience.

Data analysis—Comparisons of incidence of psychopathology were estimated using survival analysis (Cox proportional hazards models) implemented in Mplus version 6.11 (Muthén and Muthén, 1998–2010). To accommodate non-independence of observations due to multiple individuals coming from the same family, analyses were conducted using the `TYPE IS COMPLEX' analysis option. This implemented Taylor series linearization for computing standard errors to make them appropriate for non-independence. All analyses controlled for age at the final assessment (for probands, age at T_4; for siblings, age at their only assessment) and participant sex. Analyses were also conducted implementing sample weights to adjust for the stratified random sampling used at the T_3 assessment to provide generalizable prevalence estimates. The predictor of interest was whether participants were original probands who completed four diagnostic assessments or were siblings assessed on a single occasion.

Results

Chi-square tests compared the lifetime rates of disorders examined here between probands included in the present analyses to probands not included in the present report. No significant differences in lifetime rates of disorders were found (all *p*s > .20).

Table 1 displays demographic information for probands and siblings. Overall, probands and siblings did not differ on gender distribution; however, probands were older than their siblings at the time of the last assessment. There was nonetheless substantial overlap in ages for both groups.

Table 2 displays the lifetime rates of broad and narrow diagnostic categories for probands and siblings with and without adjustments for stratified random sampling at the T_3 assessment. For both probands and siblings, MDD and SUDs had substantial lifetime rates while anxiety disorders had more modest rates.

To formally compare the rates of disorders between probands and siblings, we estimated survival models. After adjusting for participant age and sex, we found that probands who completed four assessments had significantly higher rates of non-bipolar depressive disorders, including both MDD and dysthymic disorder, than siblings who completed only one lifetime diagnostic interview (Table 2). This difference appeared to be primarily accounted for by MDD, where probands demonstrated significantly higher rates than siblings, as opposed to dysthymic disorder where the difference was non-significant.

Due to low prevalence rates of bipolar disorders, we combined bipolar I, bipolar II, and cyclothymia into a single bipolar spectrum category. We did not find a significant difference in rates of bipolar spectrum disorders between probands and siblings. No significant difference was observed between probands and siblings on history of any anxiety disorder. The only significant difference in rates of individual anxiety disorders between probands and siblings was for social phobia, where siblings demonstrated *higher* rates than probands. Lastly, we examined rates of aggregate SUDs and individual categories of AUD and drug use disorders (DUD). No significant differences in rates of aggregate SUD, AUD, or DUD were found.

To examine whether age differences between probands and siblings may have unduly influenced the results, the analyses were repeated using only sib-pairs where the difference in age at the time of the final (proband) and only (sibling) assessment was less than five years. This reduced the sample to 219 probands and 278 siblings. These analyses were largely consistent with the previously reported results. Rates of depression and MDD were significantly higher for probands relative to siblings and non-significant differences were

observed for all other disorders. The previously reported difference between probands and siblings for social phobia was non-significant in the restricted sample.

Discussion

The present study compared rates of lifetime psychopathology among probands who participated in four prospective diagnostic assessments across 15 years to their siblings who completed a single lifetime retrospective assessment. Previous work (Moffitt et al., 2010) found that prospective assessments yielded rates of depressive, anxiety, and substance use disorders that were twice as high as rates from a single lifetime assessment, findings that were interpreted as indicating that cross-sectional assessments underestimate lifetime rates of disorders.

We expected to replicate the findings reported by Moffitt et al. (2010), particularly for disorders that follow episodic courses. Indeed, we found that individuals who completed prospective, longitudinal assessments had higher rates of depressive disorders, specifically MDD, than individuals who completed a single, lifetime assessment. This is consistent with previous studies, using a fully within-person design, that found that a longer duration of episode was associated with greater recall of previous episodes (Andrews et al., 1999; Wells and Horwood, 2004). A potential alternative explanation for these findings is that the increased number of assessments may have yielded increased rates of false positive results. Thus, it could be of interest to examine whether there are differences in the severity of individual episodes from prospective and lifetime retrospective methods. However, these data are not available for the siblings reported on here. An additional possibility involves the assessment methods employed. While both of the diagnostic interviews administered to the participants were semi-structured, it is possible that differences in the screening questions for MDD may have influenced the results. As the K-SADS, but not the SCID, includes irritability and suicidality as screening items, it is possible that the K-SADS is more sensitive to MDD episodes in youth.

We did not have a hypothesis regarding bipolar spectrum disorders, as these conditions are often episodic, but require chronic management. We found that lifetime rates of bipolar spectrum disorders did not significantly differ based on whether the assessment was prospective or cross-sectional. However, the HR comparing probands and siblings on lifetime prevalence rates was the largest of all comparisons. As the prevalence of bipolar disorder, was quite low in both probands and siblings, we had limited power to detect differences. Thus, further work is needed to clarify whether this finding is due to low power or a true null result.

In contrast to Moffitt et al. (2010), we did not find evidence that rates of dysthymic disorder, anxiety disorders, or SUDs differed between multiple longitudinal assessments and a lifetime cross-sectional assessment. Earlier research has suggested that anxiety disorders (Penninx *et al.*, 2011) and SUDs (Rohde et al., 2001; Sher et al., 2005) tend to be chronic. Thus, there may be greater recall of these conditions due to the longer duration of the episode. Similarly, rates of dysthymic disorder did not significantly differ between multiple and single assessments, further suggesting that the chronicity of a disorder may be a key feature of accurate disorder recall.

Surprisingly, we found that rates of social phobia were *higher* for individuals who completed a single lifetime assessment relative to those who participated in multiple prospective assessments. This may be due to a variety of factors. One possibility is that adolescents' may discount or underrate social anxiety compared to those who are older. Thus, adolescents may not recognize the impairment associated with social anxiety until

later in development. Alternatively, retrospective recall of social anxiety may be influenced by the mislabeling of developmentally normative heightened social anxiety as being pathological. It is also possible that the assessment of social phobia using the K-SADS may be less sensitive than that of the SCID. Finally, it is important to note that rates of social phobia did not differ between probands and siblings when the analysis was limited to pairs that were close in age. Although these latter analyses had less power due to the smaller N, they raise the possibility that the difference in the main analyses is an artifact of the age difference between the two groups.

Differences in findings between the present study and those from Moffitt et al. (2010) may be explained by an important methodological difference. In the current research, we utilized diagnostic information collected with semi-structured diagnostic interviews, the K-SADS and SCID-IV, whereas all of the studies examined by Moffitt et al. (2010) used fully structured diagnostic interviews, the Diagnostic Interview Schedule – III (Robins et al., 1989) and – IV (Robins et al., 1995), Composite International Diagnostic Interview 1.1 (Robins *et al.*, 1988), and the Composite International Diagnostic Interview 3.0 (Kessler and Ustün, 2004). One important difference between the administration of fully structured and semi-structured diagnostic interviews is the ability of the interviewer to probe responses. Thus, interviewers may elicit more information, and in the process cue respondents' memories. This probing may reduce the differences in rates between cross-sectional and prospective studies, particularly for chronic disorders. As episodic disorders are subject to greater recall difficulty (Andrews *et al.*, 1999), semi-structured interviews cannot make up for having to rely on a single retrospective assessment. To address some of these possibilities, novel analytic approaches (e.g., normal language item response theory) could be employed to directly compare information obtained from responses for disorder criteria across semi-structured and fully structured diagnostic interviews (Markon, 2008).

The present study has a number of strengths, including a reasonably large sample of individuals from a similar birth cohort and the use of direct reports and semi-structured diagnostic interviews. Findings, however, should be interpreted in light of a number of limitations. First, we relied on a between-subjects design to address the research question. It would be important to examine these methodological questions in future studies using within-subject methods (Wells and Horwood, 2004). Second, we only considered disorders that were assessed on all four diagnostic assessments for probands and assessed on a single occasion for siblings. Some diagnoses were not included because they did not fulfill this criterion (e.g., post-traumatic stress disorder, generalized anxiety disorder, separation anxiety disorder, oppositional defiant disorder, conduct disorder, attention deficit hyperactivity disorder). Third, despite substantial overlap, the mean ages of the probands and siblings at the final assessment significantly differed. However, we estimated models with survival analytic methods that take into account different observation periods and included age as a covariate. In addition, we repeated our analyses, restricting the sample to probands and siblings who were within five years of age of each other and the major conclusions were unchanged. Fourth, our data only speak to differences in rates of diagnoses. We were unable to examine differences in psychiatric morbidity, including number of episodes, episode duration, or particular symptoms.

In summary, and in contrast to Moffitt et al. (2010), we found that only rates of MDD are underestimated when psychopathology is assessed on a single occasion relative to multiple longitudinal assessments. Rates of anxiety disorders, dysthymia, bipolar disorders, and SUDs, were not observed to significantly differ. Future work would benefit from examining these questions using within-person prospective/retrospective designs and comparing the results of fully and semi-structured diagnostic interviews.

Acknowledgments

We would like to thank the OADP participants for their continued participation. We also thank the support of the numerous interviewers who participated in data collection.

This work was supported by National Institute of Mental Health Grants RO1 MH40501, RO1 MH50522, and RO1 MH52858 (Dr. Lewinsohn) and K01 MH092603 (Dr. Olino).

Role of Funding Source The project described was supported by National Institute Of Mental Health grants R01MH66023 (Dr. Klein), R01 MH40501, R01 MH50522, R01 MH52858, R01 DA012951 (Dr. Lewinsohn), and K01 MH092603 (Dr. Olino). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute Of Mental Health or the National Institutes of Health.

References

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 3rd edition, rev.. Author; Washington, DC: 1987.

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th edition, rev.. Author; Washington, DC: 1994.

Andrade L, Walters EE, Gentil V, Laurenti R. Prevalence of ICD-10 mental disorders in a catchment area in the city of São Paulo, Brazil. Soc Psychiatry Psychiatr Epidemiol. 2002; 37:316–25. [PubMed: 12111023]

Andrews G, Anstey K, Brodaty H, Issakidis C, Luscombe G. Recall of depressive episode 25 years previously. Psychol Med. 1999; 29:787–91. [PubMed: 10473305]

Aneshensel CS, Estrada AL, Hansell MJ, Clark VA. Social psychological aspects of reporting behavior: lifetime depressive episode reports. J Health Soc Behav. 1987:232–46. [PubMed: 3680917]

Collishaw S, Maughan B, Natarajan L, Pickles A. Trends in adolescent emotional problems in England: a comparison of two national cohorts twenty years apart. Journal of Child Psychology and Psychiatry. 2010; 51:885–94. [PubMed: 20497281]

Compton WM, Thomas YF, Stinson FS, Grant BF. 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. Arch Gen Psychiatry. 2007; 64:566–76. [PubMed: 17485608]

Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. Arch Gen Psychiatry. 2003; 60:837–44. [PubMed: 12912767]

Fergusson DM, Horwood JL. The Christchurch Health and Development Study: Review of findings on child and adolescent mental health. Australas Psychiatry. 2001; 35:287–96.

First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JBW. The Structured Clinical Interview for DSM-IV Axis I Disorders – Non-patient edition. Biometrics Research Department, New York State Psychiatric Institute; New York: 1996.

Fleiss JL. The measurement of interrater agreement. Statistical methods for rates and proportions. 1981:212–36.

Fombonne E. Increased rates of depression: update of epidemiological findings and analytical problems. Acta Psychiatr Scand. 1994; 90:145–56. [PubMed: 7810337]

Insel TR, Fenton WS. Psychiatric epidemiology: It's not just about counting anymore. Arch Gen Psychiatry. 2005; 62:590–2. [PubMed: 15939836]

Keller MB, Lavori PW, Friedman B, Nielsen E, Endicott J, McDonald-Scott P, Andreasen NC. The Longitudinal Interval Follow-up Evaluation - a comprehensive method for assessing outcome in prospective longitudinal-studies. Arch Gen Psychiatry. 1987; 44:540–8. [PubMed: 3579500]

Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry. 2005; 62:593–602. [PubMed: 15939837]

Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, Wittchen HU, Kendler KS. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Arch Gen Psychiatry. 1994; 51:8–19. [PubMed: 8279933]

Kessler RC, Ustün TB. The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). International Journal of Methods in Psychiatric Research. 2004; 13:93–121. [PubMed: 15297906]

Lewinsohn PM, Hops H, Roberts RE, Seeley JR, Andrews JA. Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III--R disorders in high school students. J Abnorm Psychol. 1993; 102:133–44. [PubMed: 8436689]

Markon, KE. Natural language item response theory and the assessment of psychopathology. 22nd Annual Meeting of the Society for Research in Psychopathology; Pittsburgh, PA. 2008.

Moffitt, TE.; Caspi, A.; Rutter, M.; Silva, PA. Sex differences in antisocial behaviour: Conduct disorder, delinquency, and violence in the Dunedin Longitudinal Study. Cambridge University Press; Cambridge, UK: 2001.

Moffitt TE, Caspi A, Taylor A, Kokaua J, Milne BJ, Polanczyk G, Poulton R. How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. Psychol Med. 2010:899–909. [PubMed: 19719899]

Muthén, LK.; Muthén, BO. Mplus User's Guide. Sixth Edition. Muthén & Muthén; Los Angeles, CA: 1998–2010.

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

Olino TM, Klein DN, Lewinsohn PM, Rohde P, Seeley JR. Longitudinal associations between depressive and anxiety disorders: A comparison of two trait models. Psychol Med. 2008; 38:353–63. [PubMed: 17803836]

Orvaschel H, Puig-Antich J, Chambers WJ, Tabrizi MA, Johnson R. Retrospective assessment of prepubertal major depression with the Kiddie-SADS-E. J Am Acad Child Adolesc Psychiatry. 1982; 21:392–7.

Penninx BWJH, Nolen WA, Lamers F, Zitman FG, Smit JH, Spinhoven P, Cuijpers P. Two-year course of depressive and anxiety disorders: Results from the Netherlands Study of Depression and Anxiety (NESDA). J Affect Disord. 2011; 133:76–85. [PubMed: 21496929]

Robins, LN.; Cottler, L.; Bucholz, KK.; Compton, W. Diagnostic interview schedule for DSM-IV (DIS-IV). Washington University School of Medicine; St Louis, MO: 1995.

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

Robins, LN.; Regier, DA. Psychiatric disorders in America: The Epidemiologic Catchment Area study. Free Press; New York: 1991.

Robins LN, Wing J, Wittchen HU, Helzer JE, Babor TF, Burke J, Farmer A, Jablenski A, Pickens R, Regier DA. The Composite International Diagnostic Interview: An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. Arch Gen Psychiatry. 1988; 45:1069–77. [PubMed: 2848472]

Rohde P, Lewinsohn PM, Kahler CW, Seeley JR, Brown RA. Natural course of alcohol use disorders from adolescence to young adulthood. J Am Acad Child Adolesc Psychiatry. 2001; 40:83–90. [PubMed: 11195569]

Rohde P, Lewinsohn PM, Seeley JR. Comparability of telephone and face-to-face interviews in assessing axis I and II disorders. Am J Psychiatry. 1997; 154:1593–8. [PubMed: 9356570]

Sher KJ, Grekin ER, Williams NA. The development of alcohol use disorders. Annual Review of Clinical Psychology. 2005; 1:493–523.

Sobin C, Weissman MM, Goldstein RB, Adams P, Wickramaratne P, Warner V, Lish JD. Diagnostic Interviewing for Family Studies - Comparing Telephone and Face-to-Face Methods for the Diagnosis of Lifetime Psychiatric-Disorders. Psychiatr Genet. 1993; 3:227–33.

Vollebergh WAM, Iedema J, Bijl RV, de Graaf R, Smit F, Ormel J. The structure and stability of common mental disorders: the NEMESIS study. Arch Gen Psychiatry. 2001; 58:597–603. [PubMed: 11386990]

Wells JE, Horwood LJ. How accurate is recall of key symptoms of depression? A comparison of recall and longitudinal reports. Psychol Med. 2004; 34:1001–11. [PubMed: 15554571]

Table 1

Demographic Characteristics of Probands and Siblings

	Probands (*n* = 442)	**Siblings (*n* = 657)**	**F/χ^2**
Female (%)	259 (58.6)	358 (54.5)	1.81
Age (years)	30.07 (.71)	24.72 (5.69)	385.55 ***
Age Range	28–34	13–45	

*** $p < .001$.

Table 2

Rates of Disorders Hazard Ratios (and 95% Confidence Intervals) for Comparisons Between Multiple (Probands) and Single Lifetime (Siblings) Assessments

	Proband		Siblings		
	Unweighted % (*n*)	**Weighted %**	**Unweighted % (*n*)**	**Weighted %**	**HR (95% CI)**
Depressive Disorders	55.0 (243)	47.3	27.5 (181)	25.9	1.86 (1.46–2.36)***
MDD	54.8 (242)	47.1	25.1 (165)	23.8	2.08 (1.61–2.68)***
Dysthymia	4.3 (19)	3.2	4.4 (29)	3.8	0.62 (0.33–1.15)
Bipolar Spectrum	3.2 (14)	2.9	1.1 (7)	1.0	2.44 (0.81–7.42)
Anxiety Disorders	13.8 (61)	11.1	9.7 (64)	9.4	0.76 (0.52–1.13)
Panic Disorder	7.0 (31)	5.7	2.4 (16)	2.2	1.69 (0.89–3.24)
Agoraphobia	0.7 (3)	0.5	0.5 (3)	0.3	0.95 (0.19–4.84)
Social Phobia	4.1 (18)	3.5	4.6 (30)	4.8	0.48 (0.26–0.89)*
Specific Phobia	5.0 (22)	3.6	3.0 (20)	3.0	1.18 (0.51–2.72)
OCD	1.4 (6)	0.9	1.4 (9)	1.3	0.61 (0.23–1.65)
SUD	41.0 (181)	36.9	39.0 (256)	36.6	0.92 (0.72–1.18)
AUD	36.0 (159)	32.4	31.2 (205)	29.3	0.94 (0.72–1.22)
DUD	20.6 (91)	17.4	24.5 (161)	22.2	0.77 (0.56–1.07)

HR = Hazard ratio; CI = confidence interval; MDD = major depressive disorder; OCD = obsessive-compulsive disorder; SUD = substance use disorders; AUD = alcohol use disorders; DUD = drug use disorders. Associations are adjusted for sex and age at assessment.

* $p < .05$;

*** $p < .001$.