Adm Policy Ment Health (2015) 42:111–120
DOI 10.1007/s10488-014-0550-8

ORIGINAL ARTICLE

Neighborhood Sociodemographic Predictors of Serious Emotional Disturbance (SED) in Schools: Demonstrating a Small Area Estimation Method in the National Comorbidity Survey (NCS-A) Adolescent Supplement

Jennifer Greif Green · Margarita Alegría · Ronald C. Kessler · Katie A. McLaughlin · Michael J. Gruber · Nancy A. Sampson · Alan M. Zaslavsky

Published online: 17 April 2014

Abstract We evaluate the precision of a model estimating school prevalence of SED using a small area estimation method based on readily-available predictors from area-level census block data and school principal questionnaires. Adolescents at 314 schools participated in the National Comorbidity Supplement, a national survey of DSM-IV disorders among adolescents. A multilevel model indicated that predictors accounted for under half of the variance in school-level SED and even less when considering block-group predictors or principal report alone. While Census measures and principal questionnaires are significant predictors of individual-level SED, associations are too weak to generate precise school-level predictions of SED prevalence.

Keywords Small area estimation · Assessment · Serious emotional disturbance · National comorbidity survey adolescent supplement

J. G. Green (✉)
School of Education, Boston University, Two Silber Way, Boston, MA 02215, USA
e-mail: jggreen@bu.edu

M. Alegría
Center for Multicultural Mental Health Research, Cambridge Health Alliance, Somerville, MA, USA

M. Alegría
Department of Psychiatry, Harvard Medical School, Boston, MA, USA

R. C. Kessler · M. J. Gruber · N. A. Sampson · A. M. Zaslavsky
Department of Health Care Policy, Harvard Medical School, Boston, MA, USA

K. A. McLaughlin
Department of Psychology, University of Washington, Seattle, WA, USA

Introduction

To effectively address youth mental disorders in schools, researchers are increasingly encouraging a public health approach, which involves implementing school-wide programs, policies, and systems to monitor and respond to student emotional and behavioral problems (Doll and Cummings 2008; Dowdy et al. 2010; Horner et al. 2009; Stiffman et al. 2010). Monitoring the prevalence of mental disorders and adjusting service provision accordingly is a cornerstone of this approach, however, few schools systematically engage in screenings for mental disorders (Romer and McIntosh 2005). As a result, researchers have developed a variety of less expensive and less intrusive methods to generate estimates of school mental health need.

The "synthetic estimation" methodology was first applied to mental health epidemiology to estimate prevalence of serious mental illness (SMI) among adults and serious emotional disturbance (SED) among youths at the state and county levels. These analyses reweight national survey data to match the distribution of socio-demographic characteristics of smaller areas (Goldsmith et al. 1998; Holzer et al. 1981; Hudson 2009; Konrad et al. 2009). Decisions about resource allocation—for example, allotment of US Block Grant funds for community mental health services—in the public service sector frequently rely on synthetic estimation to identify areas of highest need. This method is gaining popularity as publically-available datasets with geographical information become more accessible and their utility for public health assessment and surveillance is recognized (Hudson and Abbott, 2013; Tranmer et al. 2005; Zhang et al. 2013). For example, a regression-synthetic estimation method was recently used in England and Wales to develop PsyMaptic, an online tool designed to predict the incidence of psychotic disorders in small regions, using area demographic information (Kirkbridge et al. 2013). Such online technology makes small area estimates of mental disorders available to a wide range of consumers and policymakers aiming to improve allocation of mental health resources.

The primary advantage of synthetic school-level estimates of SED from Census block-group (BG) data is that data are easily accessible and free. Screening individuals' mental health can be costly (Kuo et al. 2009) and have high non-response rates (Husky et al. 2011), possibly because of perceptions that data might be used to stigmatize or over-identify students. Further, a number of neighborhood characteristics have been found consistently to predict resident mental health in community epidemiological studies of child-adolescent mental disorders (Dupéré et al. 2009; Mair et al. 2008; Pickett and Pearl 2001; Xue et al. 2005). However, the accuracy of small-area estimates of mental illness based on small-area Census data has been questioned, as the strength of associations between aggregate BG variables and the mental illness of residents is often modest, resulting in imprecise small-area estimates of prevalence (Hudson 2009; Kessler et al. 1999; Li and Zaslavsky 2010). Furthermore, synthetic estimation relies on the assumption that the prevalence of disorders is constant for demographic strata, so area prevalence depends only on demographic composition; for example, low-income adolescent boys might be assumed to have the same rates of SED in low- and high-poverty areas. Violations of these assumptions might make synthetic estimates biased.

The current paper evaluates the feasibility of using Census BG data to estimate school-level SED among adolescents at 314 schools that participated in the National Comorbidity Survey Adolescent Supplement (NCS-A), a national survey of the prevalence and correlates of DSM-IV disorders among U.S. adolescents. We compare the utility of Census BG data to a brief survey that asked school principals to estimate the incidence of emotional/behavioral problems in their schools, representing another source of data on school-level SED entailing minimal cost and burden.

Methods

Sample

Between February 2001 and January 2004, adolescents aged 13–17 were interviewed face-to-face in dual-frame household and school samples selected to be representative of the US population (Kessler et al. 2009a, b; Merikangas et al. 2009). A representative sample of schools was drawn in the counties selected for the survey. Sample selection began by targeting 289 schools, however, initially only 81 agreed to participate. Schools primarily refused to participate because they were reluctant to release student information. As a result, recruitment efforts were expanded and replacement schools were targeted for schools that declined to participate and were matched in terms of school size, geographic area, and demographic characteristics. The expansion of these recruitment efforts resulted in a final sample of 314 schools. A comparison of household sample respondents who attended nonparticipating schools with school sample respondents from replacement schools found no evidence of bias in estimates of either the prevalence of either mental disorders or treatment (Kessler et al. 2009a).

The school sample included 9,244 adolescents at 314 schools, with a 74.7 % response rate (9,244/12,380) conditional on school participation. In addition, one parent or guardian was asked to complete a self-administered questionnaire (SAQ) about the participating adolescent's developmental history and mental health. The SAQ response rate, conditional on adolescent participation, was 83.7 % in the school sample (7,739/9,244). This report focuses on the 5,940 adolescent-parent pairs in the school sample for whom complete data were available from both adolescent interviews and parent SAQs. Incomplete parent data were taken into consideration by weighting procedures that are described elsewhere (Kessler et al. 2009a, b).

After a complete description of the study was provided to parents or guardians, written informed consent was obtained before adolescents were approached. Written assent was then obtained from adolescents before either adolescents or parents-guardians were surveyed. Each respondent received $50 for participation. Recruitment and consent procedures were approved by the Human Subjects Committees of both Harvard Medical School and the University of Michigan. The completed survey data were weighted for residual discrepancies between sample and population socio-demographic and geographic distributions, as detailed elsewhere (Kessler et al. 2009a). The weighted socio-demographic distributions of the sample closely approximate those of the Census population.

The principal of each participating school was asked to complete a survey about the school's resources and policies related to mental health services and its level of need for services. Data on need for services are used as predictors in our models along with small-area Census BG data on the assumption that it would be relatively easy to obtain comparable principal questionnaire data in any future effort to estimate school-level prevalence of SED.

Measures

Individual-level SED

Adolescents were administered a modified version of the Composite International Diagnostic Interview (CIDI), a fully-structured interview designed for use by trained lay interviewers (Merikangas et al. 2009). Diagnoses assessed include: mood disorders (major depressive disorder [MDD] or dysthymia, and bipolar I-II disorder), anxiety disorders (panic disorder with or without agoraphobia, agoraphobia without history of panic disorder, social phobia, specific phobia, generalized anxiety disorder, post-traumatic stress disorder, and separation anxiety disorder), behavior disorders (attention-deficit/hyperactivity disorder [ADHD], oppositional-defiant disorder [ODD], conduct disorder [CD], and intermittent explosive disorder), and substance disorders (alcohol and drug abuse, alcohol and drug dependence with abuse). In addition, parents completed a self-administered questionnaire (SAQ) to provide diagnostic information about MDD/dysthymia, ADHD, ODD, and CD, the disorders for which parent reports have previously been shown to play the largest part in diagnosis (Braaten et al. 2001; Grills and Ollendick 2002). All but two diagnoses were made using DSM-IV diagnostic hierarchy rules. The exceptions were ODD, which was defined with or without CD, and substance abuse, which was defined with or without dependence. The current report focuses on disorders that were present in the 12 months prior to the interview.

The NCS-A included a clinical reappraisal study ($n = 347$), which blindly re-interviewed a subsample of respondents with the Schedule for Affective Disorders and Schizophrenia for School-Age Children Lifetime Version (K-SADS) (Kaufman et al. 1997). As detailed elsewhere (Kessler et al. 2009c), the results of this study indicated that parent and adolescent report were best combined using the "or" rule, under which a symptom was considered present if it was endorsed by either respondent. One exception was for a diagnosis of ADHD, which was best estimated using parent-report only (Green et al. 2010a). Concordance between lifetime CIDI/SAQ and K-SADS diagnoses was good, with area under the receiver operating characteristic curve (AUC) of 0.86–1.0 for mood disorders, 0.79–0.94 for anxiety disorders, 0.78–0.98 for behavior disorders, and 0.87 for any disorder.

Consistent with the US substance abuse and mental health services administration (SAMHSA) definition of serious emotional disturbance (SED; Substance Abuse and Mental Health Services Administration 1993), the NCS-A estimated SED by requiring a DSM-IV diagnosis significantly interfering with children's functioning in family, school, or community activities. This SAMHSA definition differs from that provided in the Individuals with Disabilities in Education Act (IDEA), which specifically requires disorders to adversely affect educational performance (U.S. Department of Education 2004). In the NCS-A K-SADS clinical reappraisal study, clinician-assessed 12-month SED was defined as having one or more 12-month DSM-IV mental disorders and either a children's global assessment scale (CGAS) (Shaffer et al. 1983) score ≤ 50, bipolar I disorder (regardless of CGAS score), or a suicide attempt in the last 12 months (again, regardless of CGAS score). Previously, an approximation of this clinical assessment in the full NCS-A sample was developed using forward stepwise logistic regression analysis in the clinical reappraisal sample to predict SED from information available in the CIDI (Kessler et al. 2012). The predictors included 12-month DSM-IV/CIDI diagnoses of all Axis I disorders other than substance disorders, summary measures of total number of disorders, self-reports in the NCS-A interview about suicidality (ideation, plans, attempts), scores on the Sheehan Disability Scale, (Leon et al. 1997) responses to the K6 scale, (Kessler et al. 2002) responses to questions about the number of days out of 365 in the past year when the adolescent was completely unable to carry out his or her usual daily activities because of specific disorders, information about overnight hospitalization for emotional or behavior problems in the past 12-months, and information about intensity of outpatient treatment for emotional and behavior problem in the past 12-months. Concordance in the clinical reappraisal sample of the predicted probability of SED from this model with the clinician-defined assessment of SED was good, as indicated by an AUC of 0.85. The same equation was used to predict the probability of SED for adolescents in the NCS-A school sample (Kessler et al. 2009c). The dichotomous SED variable was then imputed for NCS-A respondents who did not participate in the clinical reappraisal study (Kessler et al. 2012). All analyses were conducted using these imputed dichotomous SED values.

Neighborhood Characteristics

Census data on neighborhood characteristics were extracted from the Geolytics 2000 Census BG dataset (www.geolytics.com) for each NCS-A participant. We used Census BG data for the neighborhoods in which the school's enrolled students reside, calculated as an average of BG statistics weighted by the proportion of students in each BG. We conducted a review of literature on the effects of neighborhood characteristics on health and mental health (Krieger et al. 2002; Morenoff et al. 2007; Pickett and Pearl 2001). Based on this review, we selected twenty-two Census BG predictors that were used in these prior studies and associated with key outcomes. General constructs assessed by these predictors were neighborhood income, residential employment, family structure, education, racial/ethnic composition, and age composition. Variables selected were the percentages of: families with income <\$10,000, families with income >\$50,000, families with incomes below 1.5 times the official federal poverty line, families receiving public assistance, residents unemployed, families with a female head of household, never married residents, adult residents completing <12 years of education, adult residents completing $\geq$ 16 years of education, residents in a professional or managerial occupation, residents who are non-Hispanic black, residents who are Hispanic, residents who are foreign-born, residents who are homeowners, residents who have been in the same residence for the past 5 years, and six population age categories (% 0–17, 18–29, 30–39, 40–49, 50–69, 70+). Census data were also used to define Census region (Northeast, Midwest, South, West) and urbanicity (major metropolitan area, other urban, rural) of the adolescents.

Principal Questionnaire

The school principal questionnaire asked "About how many times during a typical school year do you have each of the following types of problems in your school?" with respect to 6 internalizing problems (e.g., a student is reported because of depression) and 9 externalizing problems (e.g., physical attack or fight with a weapon). Open-ended numerical estimates were divided by the number of students in the school to estimate the proportion of students in the school with each problem. In some schools, these proportions summed to greater than 100 %, indicating that principals could have counted the same student(s) multiple times because they either presented with more than one of the problems listed or with a problem on more than one occasion. A factor analysis supported a two-factor structure that corresponded to internalizing and externalizing problems. We calculated internalizing and externalizing summary scores.

Analysis Methods

We first developed a model for predicting school SED rates from Census BG variables in three steps: (1) we conducted a factor analysis of Census BG variables to create factor-weighted composite predictors, (2) we used those composite predictors to predict individual-level SED, and (3) the best-fitting individual-level model was re-estimated to predict school-level SED. We then assessed and compared the predictive power of three models respectively using both Census BG data and principal report, Census BG data only, and principal report only. These steps are described in more detail below.

First, factor analysis of the Census BG variables was conducted (Appendix 1). Results identified factors consistent with those used in previous studies of health and mental health (Hull et al. 2008; Lantz and Pritchard 2010; Morenoff et al. 2007). Weighted factor scores derived from this analysis were used to create composite Census BG predictors. Second, we entered these Census BG factor scores into logistic regression models to predict individual-level SED with standard errors adjusted for clustering by school, using PROC SURVEYLOGISTIC (SAS 9.1.3). In addition to the census BG composites, candidate predictors in our models included squares of the census BG composites, responses from principal questionnaires, region of the country, urbanicity, and interactions of the BG factors with region and with principal reports. Models were estimated first with main effects of factor scores, region, and urbanicity. Next, we added the squared factor scores, as well as the remaining predictors. We finally added interactions of predictors. At each stage we eliminated non-significant predictors, with the final model including only statistically significant predictors (based on a Wald χ^2 test) of SED.

Third, to estimate the variance of school-level prevalence, the best-fitting model identified from the preliminary logistic regression was re-estimated with a multilevel random effects logistic regression model using PROC GLIMMIX (SAS 9.1.3). This step was needed because the small sample sizes per school (mean = 5,940/314 $\approx$ 19) caused sample estimates of prevalence to be over-dispersed relative to actual school-level prevalence. An estimate of the squared correlation between predicted and population prevalence on the logit scale was calculated from model parameter estimates (see Appendix 2).

Finally, to explore the implications of our fitted models for identifying high-SED-prevalence schools, we calculated the distribution of the rankings of NCS-A SED prevalence for schools at the 95th, 90th, and 80th percentiles of predicted SED prevalence, under the fitted normal random-effects model and assuming the approximately normal distribution of the fixed-effects linear predictor (see Appendix 2). We performed similar calculations for groups of schools with top-ranking predictions of prevalence, focusing on those in the top 5, 10 and 20 % of predicted values. These analyses were repeated using the full predictive model (Census BG and principal report), Census BG only, and principal report only.

Results

Factor Analysis of Census BG Data

Principal axis factor analysis yielded five unrotated factors with eigenvalues greater than 1.0 (unrotated eigenvalues = 6.3, 2.8, 2.5, 1.6, 1.2). Inspection of standardized partial regression coefficients of items on factor scores showed that the four factor solution closely represented dimensions found in previous research (Morenoff et al. 2007) for socioeconomic disadvantage, affluence-gentrification,

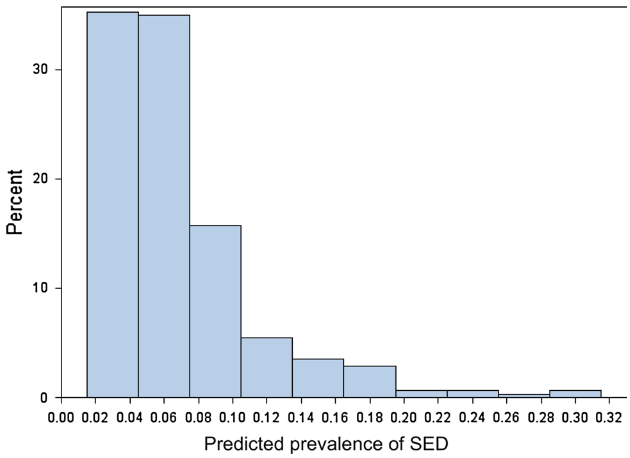

Fig. 1 Histogram of SED predicted prevalence across 314 NCS-A schools. Histogram excludes outliers (two schools with sample prevalences of SED > 40 %)

racial-ethnic composition, and older age composition. We therefore calculated four standardized factor-weighted scores based on the four-factor promax-rotated solution.

Associations of Census Data and Principal Ratings with Individual and School SED

The best-fitting multilevel model for the associations of Census BG and principal reports with student individual-level SED included the four Census BG factors, region, urbanicity, and principal reports of externalizing problems, as well as interactions of the four Census BG factors both with the four regions of the country and with principal reports of externalizing problems. Parameters from this best fitting model were used to estimate the predicted prevalence of SED in each school and its prediction interval with 80 % coverage probability.

The mean predicted prevalence of SED across schools was 7.3 %. A histogram of predicted prevalence had a positive skew, with most predicted prevalence below 10 % and almost all (95 %) below 26 % (Fig. 1). Two schools with substantially higher sample prevalence of SED (45–53 %) also had high predicted prevalence (over 55 %), with elevated principal reports of externalizing behaviors.

We estimated R^2 with and without these two outliers. When all schools were included, R^2 was 0.49, indicating that the predictive value of Census BG variables and principal ratings accounted for less than half of the variability in school-level SED. Excluding the two outliers, R^2 was 0.45, suggesting that these outliers had a minimal impact on the predictive accuracy of the model. We repeated this analysis with Census BG data alone (without principal ratings) and found that R^2 dropped to 0.40. When we considered principal ratings alone (without Census BG data), R^2 dropped further to 0.21.

We next examined the predictive median prevalence and 80 % prediction intervals for several values of SED predicted probability to determine the accuracy of model estimation at various points along the predicted probability continuum, again assuming our normal-logistic random effects model and estimates of model parameters. For the full model (Census BG and principal reports), a hypothetical school predicted to be in the 50th percentile of prevalence had a predicted median SED prevalence of 6.9 and 80 % prediction interval of 5.0–9.4 %, an interval much too large to provide useful estimates of school SED. At the 95th percentile, median SED prevalence was 10.5 % and the 80 % prediction interval was 7.7–14.2 %, illustrating the near-proportional widening of prediction intervals for prevalence at higher levels engendered by the inverse-logit transformation (Table 1). Prediction intervals based on Census BG data only or principal data only were slightly wider as expected given the lower R^2 of these models.

Finally, under the same model assumptions we estimated the percent of schools with predicted SED prevalence in the top percentile intervals whose actual SED prevalence would be in each of those intervals, to further illustrate predictive accuracy. Of the schools predicted to be in the top 10 % of SED prevalence by the full model, just under half (49.3 %, interpretable as positive predictive power and also equal to sensitivity, under our model) were actually in the top 10 % of observed SED prevalence and a little over two-thirds (71.9 %) were in the top 20 % of observed prevalence (Table 2). At the high end of the continuum, of those schools predicted to be in the top 5 %, approximately 41.9 % were actually in the top 5 %, over half (61.5 %) in the top 10 %, and almost all (98.1 %) in the top 50 %.

Discussion

Consistent with the results of several previous studies that documented associations of neighborhood socio-demographic characteristics with individual-level risk of SED/SMI, (Dupéré et al. 2009; Mair et al. 2008; Pickett and Pearl 2001; Xue et al. 2005) we found significant associations of both BG-level characteristics and principal reports with individual-level SED in the NCS-A sample. Nonetheless, predictions based on these models were too imprecise to be useful for policy planning or funding allocation purposes.

This result is consistent with the conclusions of two previous studies that evaluated the precision of estimates of SMI and/or SED based on associations between small-area Census BG data and individual-level prevalence estimates

Table 1 Prediction intervals (PI) illustrating precision of Census BG- and principal report-based predictors at different percentile points of predicted SED

Predicted percentile	Census BG & Principal		Census BG only		Principal only	
	Median	80 % PI	Median	80 % PI	Median	80 % PI
5	4.4	(3.2–6.1)	4.4	(3.0–6.3)	5.3	(3.5–7.8)
10	4.9	(3.5–6.7)	4.8	(3.3–7.0)	5.6	(3.7–8.3)
25	5.7	(4.1–7.9)	5.7	(3.9–8.2)	6.2	(4.1–9.1)
50	6.9	(5.0–9.4)	6.8	(4.7–9.7)	6.9	(4.6–10.2)
75	8.2	(6.0–11.2)	8.1	(5.6–11.6)	7.7	(5.2–11.3)
90	9.6	(7.0–13.0)	9.5	(6.6–13.4)	8.5	(5.7–12.4)
95	10.5	(7.7–14.2)	10.4	(7.3–14.7)	9.0	(6.1–13.1)

Table 2 Percent of schools with predicted SED prevalence in highest Q % range of predictions whose actual SED prevalence could be expected to be in top *P* % of schools

Predicted range Q	Observed interval											
	Census BG & Principal				Census BG only				Principal only			
	$P = 50$	20	10	5	$P = 50$	20	10	5	$P = 50$	20	10	5
50	76.0	36.5	19.2	9.8	73.8	35.4	18.8	9.7	65.5	30.6	16.4	8.6
20	91.2	58.6	35.9	20.4	88.5	55	33.6	19.1	76.4	41.8	24.5	13.9
10	95.9	71.9	49.3	30.7	93.8	67.1	45.2	28	81.8	49.0	30.5	18.1
5	98.1	81.6	61.5	41.9	96.6	76.5	56	37.5	85.8	55.4	36.2	22.5

(Kessler et al. 1999; Li et al. 2010). The proportion of variance explained by our best model was modest ($R^2 = 0.49$) and even lower for models relying only on Census BG data ($R^2 = 0.40$) or on principal data alone ($R^2 = 0.21$). Consequently, identification of schools with the highest prevalence could only be accomplished with moderate sensitivity and positive predictive power.

While it is possible that we could have found stronger associations using different Census variables, the ones we used include all those that have been found to be important in previous studies of the associations between Census-based neighborhood characteristics and mental health. Similarly, it is possible that we could have asked different questions of principals that would have yielded stronger predictions of individual-level SED, but we are aware of no data documenting such associations. The more plausible conclusion in light of these observations is that accurate estimation of school-level SED prevalence cannot be obtained from estimation using easily-accessible socio-demographic data of the sort we used in this exercise. The prediction intervals from our estimates alone indicate that estimates would be too imprecise to be valuable to district administrators and others allocating mental health resources to use for planning.

In contrast, a prior study examined the extent to which the K6 (Kessler et al., 2002, 2003), a 6-item measure of emotional distress, could be used to estimate school-level SED in this same NCS-A population (Green et al. 2010b; Li et al. 2010). Although the K6 had only fair concordance with SED at the individual level, K6 scores could be used to generate highly accurate estimates of school-level SED prevalence (Li and Zaslavsky 2010). This result suggests that it may be productive to continue research to refine very short measures, like the K6, to allow school-level SED to be estimated from information about the distribution of responses to very short screening scales (Green et al. 2010b). A number of student survey instruments have been developed for this purpose (Levitt et al. 2007). However, many of these instruments are quite long (Achenbach 1991) and shorter scales often focus on only a single disorder or class of disorders (Garrison et al. 1991; Levitt et al. 2007). There is consequently increasing interest in brief but broad measures that schools can feasibly use both to estimate the prevalence of SED and to identify individual students for referral to mental health services. The Strengths and Difficulties Questionnaire (Goodman 2001), BASC-2 Behavioral Emotional Screening System Student Form (Dowdy et al. 2011), and Brief Problem Checklist (Chorpita et al. 2010) are three relatively brief (12–30 item) scales that appear promising for this purpose.

The finding that reports from principals contributed to SED prediction, albeit only marginally, also raises questions about whether other data could be collected easily from individual schools to more effectively estimate SED. For example, many schools use administrative data (e.g., suspensions, absences, failures) to identify students requiring support services (Horner et al. 2009). The advantage of these data is their accessibility and direct relevance to the educational mission of schools. However, their contribution to predicting school-level SED is untested and, by their nature, they will be better at estimating disorders associated with externalizing rather than internalizing symptoms.

The current study has several limitations. First, the initial response rate of NCS-A schools was quite low. However, analyses comparing selected schools with matched replacements found no evidence of bias (Kessler et al. 2009a). In addition, adolescents whose parents did not complete the survey or had missing data were excluded from analyses. Data were re-weighted to account for missing parent data, but it is possible that there were systematic biases associated with these missing data. Second, the NCS-A excluded homeless adolescents and non-English speakers, limiting the generalizability of study findings and potentially weakening the associations of NCS-A SED with Census variables. Third, limitations of the NCS-A measure of SED may have introduced error in measurement. NCS-A disorder severity was imputed based on a clinical reappraisal study that relied on telephone (rather than in-person) clinical interviews (Kessler et al. 2009c). This and other limitations of these interviews suggest caution in interpreting results despite good concordance between gold standard clinician diagnoses of SED and imputed SED values. Fourth, principal reports of internalizing and externalizing problems were limited by possible reporting bias by principals (e.g., under-reporting if they were unaware of incidences, under-reporting of internalizing problems, over-reporting if the same student was counted in multiple categories or for the same behavior on multiple occasions). Fifth, Census block groups include many non-adolescents and many families whose adolescents likely attended schools that did not participate in the current study, potentially attenuating predictive associations. Sixth, we analyzed accuracy of identification of high-prevalence schools nationally, but our data did not enable us to evaluate accuracy of discrimination within a state or school district, which might be lower if schools tend to be more homogenous within a district than nationally. Finally, our calculation of the distribution of predicted SED prevalence for schools relied on statistical assumptions of normality of random effects and linear predictors, of which only the latter could be verified from our data.

Our analyses accounted for individual-level covariates, but did not include school-level covariates, including variations in school provision of mental health services. Results from a prior study of national trends in school-based services for students with SED suggest that students identified under IDEA as having an emotional or behavioral disorder generally attend larger schools that are comprised of a higher proportion of students receiving special education services than the national average (Wagner et al. 2006). This might be the result of districts "clustering" students with disabilities in a single school. Findings that students with emotional and behavioral disorders disproportionately attend schools outside of their neighborhood, and are more likely to change schools because of re-assignment, support this suggestion (Wagner et al. 2005; Wagner et al. 2006). Although we were unable to explicitly examine this trend in the current study, future research on school-level SED would benefit from including questions about school policies in this regard and comparing schools within a single district to identify trends in school-level SED relative to the local population.

Despite these limitations, the results reported here suggest that regression estimation methods are insufficient to obtain precise small-area estimation of school-level SED. For schools seeking to estimate SED prevalence, more promising methods are emerging that involve using very brief scales. This type of brief scale, if found to provide accurate estimates of school-level SED, could have considerable potential as an addition to ongoing health and mental health surveillance efforts, such as the Youth Risk Behavior Surveillance Survey conducted biennially by the Centers for Disease Control. This could be an important development given the fact that accurate estimation of school-level SED prevalence can improve the ability of schools to monitor the needs of students and inform allocation of limited resources for addressing unmet need for treatment of serious emotional disturbance.

Acknowledgements Preparation of this paper was supported by NIMH K01-MH085710 (Green) and NIMH K01-MH092526 (McLaughlin). This work is also supported by the National Institutes of Health/National Institute on Minority Health and Health Disparities (NIH/NIMHD) Recovery Act Project, which funded Challenge Grant 5RC1MD004588. The National Comorbidity Survey Replication Adolescent Supplement (NCS-A) is supported by the National Institute of Mental Health (NIMH; U01-MH60220 and R01-MH66627) with supplemental support from the National Institute on Drug Abuse (NIDA), the Substance Abuse and Mental Health Services Administration (SAMHSA), the Robert Wood Johnson Foundation (RWJF; Grant 044780), and the John W. Alden Trust. The views and opinions expressed in this report are those of the authors and should not be construed to represent the views of any of the sponsoring organizations, agencies, or U.S. Government. A complete list of NCS-A publications can be found at http://www.hcp.med.harvard.edu/ncs. Send correspondence to ncs@hcp.med.harvard.edu. The NCS-A is carried out in conjunction with the World Health Organization World Mental Health (WMH) Survey Initiative. We thank the staff of the WMH Data Collection and Data Analysis Coordination Centres for assistance with instrumentation, fieldwork, and consultation on data analysis. The WMH Data Coordination Centres have received support from NIMH (R01-MH070884, R13-MH066849, R01-MH069864, R01-MH077883), NIDA (R01-DA016558), the Fogarty International Center of the National Institutes of Health (FIRCA R03-TW006481), the John D. and Catherine T. MacArthur Foundation, the Pfizer Foundation, and the Pan American Health Organization. The WMH Data Coordination Centres have also received unrestricted educational grants from Astra Zeneca, BristolMyersSquibb, Eli Lilly and Company, GlaxoSmithKline, Ortho-

McNeil, Pfizer, Sanofi-Aventis, and Wyeth. A complete list of WMH publications can be found at http://www.hcp.med.harvard.edu/wmh/.

Conflict of interest Dr. Alegría has served as an expert presenter for Shire US, Inc. Dr. Kessler has been a consultant for AstraZeneca, Analysis Group, Bristol-Myers Squibb, Cerner-Galt Associates, Eli Lilly & Company, GlaxoSmithKline Inc., HealthCore Inc., Health Dialog, Hoffman-LaRoche, Inc., Integrated Benefits Institute, John Snow Inc., Kaiser Permanente, Matria Inc., Mensante, Merck & Co, Inc., Ortho-McNeil Janssen Scientific Affairs, Pfizer Inc., Primary Care Network, Research Triangle Institute, Sanofi-Aventis Groupe, Shire US Inc., SRA International, Inc., Takeda Global Research & Development, Transcept Pharmaceuticals Inc., and Wyeth-Ayerst. Dr. Kessler has served on advisory boards for Appliance Computing II, Eli Lilly & Company, Mindsite, Ortho-McNeil Janssen Scientific Affairs, Johnson & Johnson, Plus One Health Management and Wyeth-Ayerst. Dr. Kessler has had research support for his epidemiological studies from Analysis Group Inc., Bristol-Myers Squibb, Eli Lilly & Company, EPI-Q, GlaxoSmithKline, Johnson & Johnson Pharmaceuticals, Ortho-McNeil Janssen Scientific Affairs., Pfizer Inc., Sanofi-Aventis Groupe, Shire US, Inc., and Walgreens Co. Dr. Kessler owns 25 % share in DataStat, Inc. The remaining authors report no conflict of interest.

Appendices

Appendix 1

See Table 3.

Appendix 2: Technical Appendix

In our multilevel logistic model, the linear predictor including the random effect is logit P(SED) $= x'\beta + \delta$, where $\delta \sim N(0,\sigma^2)$ and σ^2 is the random-effects variance. We assume that the distribution of the fixed-effects linear predictor across schools is approximately normal, $x'\beta = z \sim N(\mu, S_x^2)$, where S_x^2 is the variance across schools of the mean linear predictor $x'\beta$, estimated by the sample variance of the school means. An estimate of the squared correlation between predicted and population prevalence on the logit scale was calculated from model parameter estimates as $R^2 = S_x^2/(S_x^2 + \sigma^2)$, the fraction of SED variance among schools explained by the model.

A school at the q quantile of the predictor distribution has $x'\beta = z_0 = \mu + S_x\,\Phi^{-1}(q)$. The predictive distribution of the linear predictor including the random effect δ is then $x'\beta + \delta \sim N(z_0, \sigma^2)$ with $(1 - p)$-level prediction interval $(z_0 - \Phi^{-1}(1 - p/2)\sigma, z_0 + \Phi^{-1}(1 - p/2)\sigma)$; these bounds can then be transformed to probabilities through the inverse logit transformation.To estimate probabilities of exceeding the p quantile of prevalence for the population of schools at or above the q quantile of predicted prevalence, we note that this is a conditional probability $P(z + \delta > \mu + \Phi^{-1}(p)\sqrt{(S_x^2 + \sigma^2)} \mid z > \mu + \Phi^{-1}(q)S_x)$where z and δ have the unconditional normal distributions given above. The

Table 3 Rotated (promax) tetrachoric factor analysis (standardized regression coefficients) of Census BG variables in NCS-A counties (n = 314 schools)

	Factor 1 Socioeconomic disadvantage	Factor 2 Affluence/ Gentrification	Factor 3 Older age composition	Factor 4 Ethnic/Racial composition
% ≥ 16 years education	*−0.89*	0.32	0.07	−0.10
% < 12 years education	*0.83*	−0.07	−0.05	0.32
% Professional/managerial occupation	*−0.82*	0.19	0.10	−0.12
% Families with income > $50 K	*−0.82*	−0.10	−0.27	0.01
% Families with income < $10 K	*0.73*	0.24	0.19	−0.19
% Families in poverty	*0.73*	0.34	0.00	−0.08
% Families with public assistance	*0.63*	0.18	−0.18	−0.04
% Unemployed in civilian labor force	*0.45*	0.28	−0.17	−0.13
% Never married	0.14	*0.81*	−0.05	−0.09
% 18–29 years old	−0.01	*0.77*	0.10	0.06
% In same residence in 1995	0.13	*−0.74*	−0.05	−0.19
% Homeowners	−0.25	*−0.72*	−0.23	−0.17
% 50–69 years old	−0.06	*−0.61*	0.32	−0.15
% 30–39 years old	−0.32	*0.37*	−0.30	0.22
% Families female householder	−0.05	0.28	*0.79*	−0.08
% 0–17 years old	0.30	−0.09	*−0.77*	0.01
% 70 + years old	0.15	−0.44	*0.76*	0.02
% 40–49 years old	−0.30	−0.19	*−0.35*	−0.21
% Hispanic	0.23	0.13	−0.12	*0.78*
% Foreign born	−0.05	0.29	0.00	*0.77*
% Non-hispanic black	0.47	0.23	−0.20	*−0.50*

Primary factor loadings are italicized

Correlations among factors: F1-F2: 0.27; F1-F3: −0.03; F1-F4: 0.10; F2-F3: 0.05; F2-F4: 0.04; F3-F4: −0.12

denominator is $1 - q$ and the numerator can be expressed as the probability of a quadrant of a bivariate normal distribution.

References

Achenbach, T. M. (1991). *Manual for the child behavior checklist/4-18 and 1991 profile*. Burlington: University of Vermont: Department of Psychiatry.

Braaten, E. B., Biederman, J., DiMauro, A., Mick, E., Monuteaux, M. C., Muehl, K., et al. (2001). Methodological complexities in the diagnosis of major depression in youth: An analysis of mother and youth self-reports. *Journal of Child and Adolescent Psychopharmacology, 11*(4), 395–407.

Chorpita, B. F., Reise, S., Weisz, J. R., Grubbs, K., Becker, K. D., Krull, J. L., et al. (2010). Evaluation of the brief problem checklist: Child and caregiver interviews to measure clinical progress. *Journal of Consulting and Clinical Psychology, 78*(4), 526–536.

Doll, B., & Cummings, J. A. (2008). *Population-based approaches to promoting the competency and wellness of children*. Thousand Oaks, CA: National Association of School Psychologists.

Dowdy, E., Ritchey, K., & Kamphaus, R. W. (2010). School-based screening: A population-based approach to inform and monitor children's mental health needs. *School Mental Health, 2*(4), 166–176.

Dowdy, E., Twyford, J. M., Chin, J. K., DiStefano, C. A., Kamphaus, R. W., & Mays, K. L. (2011). Factor structure of the BASC–2 behavioral and emotional screening system student form. *Psychological Assessment, 23*(2), 379–387.

Dupéré, V., Leventhal, T., & Lacourse, E. (2009). Neighborhood poverty and suicidal thoughts and attempts in late adolescence. *Psychological Medicine, 39*, 1295–1306.

Garrison, C. Z., Addy, C. L., Jackson, K. L., McKeown, R. E., & Waller, J. L. (1991). The CES-D as a screen for depression and other psychiatric disorders in adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 30*(4), 636–641.

Goldsmith, H. F., Holzer, C. E., & Manderscheid, R. W. (1998). Neighborhood characteristics and mental illness. *Evaluation and Program Planning, 21*(2), 211–225.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(11), 1337–1345.

Green, J. G., Avenevoli, S., Finkelman, M., Gruber, M. J., Kessler, R. C., Merikangas, K. R., et al. (2010a). Attention deficit hyperactivity disorder: Concordance of the adolescent version of the composite international diagnostic interview version 3.0 (GDI) with the K-SADS in the US national comorbidity survey replication adolescent (NCS-A) supplement. *International Journal of Methods in Psychiatric Research, 19*(1), 34–49.

Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., & Kessler, R. C. (2010b). Improving the K6 short scale to predict serious emotional disturbance in adolescents in the USA. *International Journal of Methods in Psychiatric Research, 19*, 23–35.

Grills, A. E., & Ollendick, T. H. (2002). Issues in parent-child agreement: The case of structured diagnostic interviews. *Clinical Child and Family Psychology Review, 5*(1), 57–83.

Holzer, C. E., Jackson, D. J., & Tweed, D. (1981). Horizontal synthetic estimation: A strategy for estimating small area health related characteristics. *Evaluation and Program Planning, 4*, 29–34.

Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A. W., et al. (2009). A randomized, wait-list controlled effectiveness trial assessing school-wide positive behavior support in elementary schools. *Journal of Positive Behavior Interventions, 11*(3), 133–144.

Hudson, C. G. (2009). Validation of a model for estimating state and local prevalence of serious mental illness. *International Journal of Methods in Psychiatric Research, 18*(4), 251–264.

Hudson, G. G., & Abbott, M. W. (2013). Modeling the geographic distribution of serious mental illness in New Zealand. *Social Psychiatry and Psychiatric Epidemiology, 48*, 25–36.

Hull, P., Kilbourne, B., Reece, M., & Husaini, B. (2008). Community involvement and adolescent mental health: Moderating effects of race/ethnicity and neighborhood disadvantage. *Journal of Community Psychology, 36*(4), 534–551.

Husky, M. M., Sheridan, M., McGuire, L., & Olfson, M. (2011). Mental health screening and follow-up care in public high schools. *Journal of the American Academy of Child and Adolescent Psychiatry, 50*(9), 881–891.

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., et al. (1997). Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*(7), 980–988.

Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S. T., et al. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine: A Journal of Research in Psychiatry and the Allied Sciences, 32*(6), 959–976.

Kessler, R. C., Avenevoli, S., Costello, E. J., Green, J. G., Gruber, M. J., Heeringa, S., et al. (2009a). Design and field procedures in the US national comorbidity survey replication adolescent supplement (NCS-A). *International Journal of Methods in Psychiatric Research, 18*(2), 69–83.

Kessler, R. C., Avenevoli, S., Costello, E. J., Green, J. G., Gruber, M. J., Heeringa, S., et al. (2009b). National comorbidity survey replication adolescent supplement (NCS-A): II. overview and design. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*(4), 380–385.

Kessler, R. C., Avenevoli, S., Costello, E. J., Green, J. G., Gruber, M. J., McLaughlin, K. A., et al. (2012). Severity of 12-month DSM-IV disorders in the NCS-R adolescent supplement (NCS-A). *Archives General Psychiatry, 69*, 381–389.

Kessler, R. C., Avenevoli, S., Green, J., Gruber, M. J., Guyer, M., He, Y., et al. (2009c). National comorbidity survey replication adolescent supplement (NCS-A): III. concordance of DSM-IV/CIDI diagnoses with clinical reassessments. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*(4), 386–399.

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., et al. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry, 60*(2), 184–189.

Kessler, R. C., Berglund, P. A., Walters, E. E., Leaf, P. J., Kouzis, A. C., Bruce, M. L., et al. (1999). A methodology for estimating the 12-month prevalence of serious mental illness. In R. W. Manderscheid & M. J. Henderson (Eds.), *Mental health, United States 1998* (pp. 99–109). Washington DC: U.S. Government Printing Office.

Kirkbridge, J. B., Jackson, D., Perez, J., Fowler, D., Winton, F., Coid, J. W., et al. (2013). A population-level prediction tool for the incidence of first-episode psychosis: Translational epidemiology based on cross-sectional data. *BMJ Open,*. doi:10.1136/bmjopen-2012-001998.

Konrad, T. R., Ellis, A. R., Thomas, K. C., Holzer, C. E., & Morrissey, J. P. (2009). County-level estimates of need for mental health professionals in the United States. *Psychiatric Services, 60*(10), 1307–1314.

Krieger, N., Chen, J., Waterman, P., Soobader, M., Subramanian, S., & Carson, R. (2002). Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: Does the choice of area-based measure and geographic level matter? *American Journal of Epidemiology, 156*, 471–482.

Kuo, E., Vander Stoep, A., McCauley, E., & Kernic, M. A. (2009). Cost-effectiveness of a school-based emotional health screening program. *Journal of School Health, 79*, 277–285.

Lantz, P. M., & Pritchard, A. (2010). Socioeconomic indicators that matter for population health. *Preventing Chronic Disease, 7*(4), A74.

Leon, A. C., Olfson, M., Portera, L., Farber, L., & Sheehan, D. V. (1997). Assessing psychiatric impairment in primary care with the sheehan disability scale. *International Journal of Psychiatry in Medicine, 27*(2), 93–105.

Levitt, J. M., Saka, N., Romanelli, L. H., & Hoagwood, K. (2007). Early identification of mental health problems in schools: The status of instrumentation. *Journal of School Psychology, 45*(2), 163–191.

Li, F., Green, J. G., Kessler, R. C., & Zaslavsky, A. M. (2010). Estimating prevalence of serious emotional disturbance in schools using a brief screening scale. *International Journal of Methods in Psychiatric Research, 19*, 88–98.

Li, F., & Zaslavsky, A. M. (2010). Using a short screening scale for small-area estimation of mental illness prevalence for schools. *Journal of the American Statistical Association, 105*(492), 1323–1332.

Mair, C., Diex Roux, A. V., & Galea, S. (2008). Are neighborhood characteristics associated with depressive symptoms? A review of evidence. *Journal of Epidemiology and Community Health, 62*, 940–946.

Merikangas, K. R., Avenevoli, S., Costello, E. J., Koretz, D., & Kessler, R. C. (2009). National comorbidity survey replication adolescent supplement (NCS-A): I. background and measures. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*(4), 367–379.

Morenoff, J. D., House, J. S., Hansen, B. B., Williams, D. R., Kaplan, G. A., & Hunte, H. E. (2007). Understanding social disparities in hypertension prevalence, awareness, treatment, and control: The role of neighborhood context. *Social Science and Medicine, 65*(9), 1853–1866.

Pickett, K. E., & Pearl, M. (2001). Multilevel analyses of neighborhood socioeconomic context and health outcomes: A critical review. *Journal of Epidemiology and Community Health, 55*(2), 111–122.

Romer, D., & McIntosh, M. (2005). The roles and perspectives of school mental health professionals in promoting adolescent mental health. In T. Walsh (Ed.), *Treating and preventing adolescent mental health disorders: What we know and what we don't know: A research agenda for improving the mental health of our youth* (pp. 597–615). New York: Oxford University Press.

Shaffer, D., Gould, M. S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., et al. (1983). A children's global assessment scale (CGAS). *Archives of General Psychiatry, 40*(11), 1228–1231.

Stiffman, A. R., Stelk, W., Horwitz, S. M., Evans, M. E., Outlaw, F. H., & Atkins, M. (2010). A public health approach to children's mental health services: Possible solutions to current service inadequacies. *Administration and Policy in Mental Health and Mental Health Services Research, 37*, 120–124.

Substance Abuse and Mental Health Services Administration. (1993). Final notice establishing definitions for (1) children with a serious emotional disturbance, and (2) adults with a serious mental illness. *Federal Register, 58*, 29422–29425.

Tranmer, M., Pickles, A., Fieldhouse, E., Elliot, M., Dale, A., Brown, M., et al. (2005). The case for small area microdata. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 168*, 29–49.

U.S. Department of Education (2004). Twenty-fourth annual report to congress on implementation of the Individuals with Disabilities Education Act. Washington, DC.

Wagner, M., Friend, M., Bursuck, W. D., Kutash, K., Duchnowski, A. J., Sumi, W. C., et al. (2006). Educating students with emotional disturbances: A national perspective on programs and services. *Journal of Emotional and Behavioral Disorders, 14*, 12–30.

Wagner, M., Kutash, K., Duchnowski, A. J., Epstein, M. H., & Sumi, W. C. (2005). The children and youth we serve: A national picture of the characteristics of students with emotional disturbances receiving special education. *Journal of Emotional and Behavioral Disorders, 13*, 79–96.

Xue, Y., Leventhal, T., Brooks-Gunn, J., & Earls, F. J. (2005). Neighborhood residence and mental health problems of 5- to 11-year-olds. *Archives of General Psychiatry, 62*(5), 554–563.

Zhang, X., Onufrak, S., Holt, J. B., & Croft, J. B. (2013). A multilevel approach to estimating small area childhood obesity prevalence at the census block-group level. *Preventing Chronic Disease, 10*, 120252.