Small-area estimation of mental illness prevalence for schools

Fan Li ¹   Alan Zaslavsky ²

¹Department of Statistical Science
Duke University

²Department of Health Care Policy
Harvard Medical School

March 6, 2012
Introduction

- Mental disorders account for 15% of overall burden of diseases in U.S.

- Early onset leads research focus on children and adolescents.

- Significant variation exists in prevalence of mental illness across schools and geographical regions.

- Information about prevalence of serious emotional disturbance (SED) among youth in small areas (states, counties) are valuable for mental health policy planning.

- Carrying out surveys to obtain direct estimates of SED in small areas are prohibitively expensive.

- Most of current literature is based on "synthetic estimation" (reweight national surveys).
Introduction

- Mental disorders account for 15% of overall burden of diseases in U.S.
- Early onset leads research focus on children and adolescents.
Introduction

- Mental disorders account for 15% of overall burden of diseases in U.S.
- Early onset leads research focus on children and adolescents.
- Significant variation exists in prevalence of mental illness across schools and geographical regions.
Introduction

- Mental disorders account for 15% of overall burden of diseases in U.S.
- Early onset leads research focus on children and adolescents.
- Significant variation exists in prevalence of mental illness across schools and geographical regions.
- Information about prevalence of serious emotional disturbance (SED) among youth in small areas (states, counties) are valuable for mental health policy planning.
Introduction

- Mental disorders account for 15% of overall burden of diseases in U.S.
- Early onset leads research focus on children and adolescents.
- Significant variation exists in prevalence of mental illness across schools and geographical regions.
- Information about prevalence of serious emotional disturbance (SED) among youth in small areas (states, counties) are valuable for mental health policy planning.
- Carrying out surveys to obtain direct estimates of SED in small areas are prohibitively expensive.
Introduction

- Mental disorders account for 15% of overall burden of diseases in U.S.
- Early onset leads research focus on children and adolescents.
- Significant variation exists in prevalence of mental illness across schools and geographical regions.
- Information about prevalence of serious emotional disturbance (SED) among youth in small areas (states, counties) are valuable for mental health policy planning.
- Carrying out surveys to obtain direct estimates of SED in small areas are prohibitively expensive.
- Most of current literature is based on “synthetic estimation” (reweight national surveys).
Motivation

- Data on short screening scales for SED, e.g., the K6, are collected in three major national health tracking surveys -

1. National Health Interview Survey (NHIS)
2. CDC's Behavioral Risk Factors Surveillance Survey (BRFSS)
3. SAMHSA's National Household Survey on Drug Use and Health (NSDUH)

- Motivated by the National Comorbidity Survey - Adolescent (NCS-A): 1st nationally representative population survey to evaluate mental health of adolescents.

- Data on both clinical diagnosis and short screening scale are available.
Data on short screening scales for SED, e.g., the K6, are collected in three major national health tracking surveys - (1) National Health Interview Survey (NHIS) (2) CDC’s Behavioral Risk Factors Surveillance Survey (BRFSS) (3) SAMHSA’s National Household Survey on Drug Use and Health (NSDUH)
Motivation

- Data on short screening scales for SED, e.g., the K6, are collected in three major national health tracking surveys -
  (1) National Health Interview Survey (NHIS)
  (2) CDC’s Behavioral Risk Factors Surveillance Survey (BRFSS)
  (3) SAMHSA’s National Household Survey on Drug Use and Health (NSDUH)

- Short screening scales are known to be correlated with clinical diagnosis. Can we utilize such information?
Motivation

- Data on short screening scales for SED, e.g., the K6, are collected in three major national health tracking surveys -
  1. National Health Interview Survey (NHIS)
  2. CDC’s Behavioral Risk Factors Surveillance Survey (BRFSS)
  3. SAMHSA’s National Household Survey on Drug Use and Health (NSDUH)
- Short screening scales are known to be correlated with clinical diagnosis. Can we utilize such information?
- Motivated by the National Comorbidity Survey - Adolescent (NCS-A): 1st nationally representative population survey to evaluate mental health of adolescents.
Motivation

- Data on short screening scales for SED, e.g., the K6, are collected in three major national health tracking surveys -
  1. National Health Interview Survey (NHIS)
  2. CDC’s Behavioral Risk Factors Surveillance Survey (BRFSS)
  3. SAMHSA’s National Household Survey on Drug Use and Health (NSDUH)

- Short screening scales are known to be correlated with clinical diagnosis. Can we utilize such information?

- Motivated by the National Comorbidity Survey - Adolescent (NCS-A): 1st nationally representative population survey to evaluate mental health of adolescents.

- Data on both clinical diagnosis and short screening scale are available.
NCS-A

- NCS-A school sample: 9244 adolescents from 318 schools, 42 sampling strata.
NCS-A

- NCS-A school sample: 9244 adolescents from 318 schools, 42 sampling strata.
NCS-A

- NCS-A school sample: 9244 adolescents from 318 schools, 42 sampling strata.
- Each individual is measured by
  1. Short screening scale - K6: 6 questions about mental health during last 30 days, measured on 0-4 scale.
NCS-A

- NCS-A school sample: 9244 adolescents from 318 schools, 42 sampling strata.
- Each individual is measured by
  1. Short screening scale - K6: 6 questions about mental health during last 30 days, measured on 0-4 scale.
  2. Clinical diagnostic score of SED: CIDI-A.
NCS-A

- NCS-A school sample: 9244 adolescents from 318 schools, 42 sampling strata.
- Each individual is measured by
  1. Short screening scale - K6: 6 questions about mental health during last 30 days, measured on 0-4 scale.
  2. Clinical diagnostic score of SED: CIDI-A.
- Individual socio-demographic and school information is also collected.
NCS-A school sample: 9244 adolescents from 318 schools, 42 sampling strata.

Each individual is measured by

1. Short screening scale - K6: 6 questions about mental health during last 30 days, measured on 0-4 scale.
2. Clinical diagnostic score of SED: CIDI-A.

Individual socio-demographic and school information is also collected.

Goal: develop a methodology to provide small area estimates of a gold-standard measure of SED from short screening scale, together with socio-demographic variables.
Small-area estimation

- “Small-area”: domains with inadequate sample sizes for precise direct estimates.
Small-area estimation

- “Small-area”: domains with inadequate sample sizes for precise direct estimates.
- Small-area estimation (SAE): borrow strength across domains.
Small-area estimation

- “Small-area”: domains with inadequate sample sizes for precise direct estimates.
- Small-area estimation (SAE): borrow strength across domains.
- Multivariate models improve small-area estimation for each outcome (De Souza, 1992).
Small-area estimation

- “Small-area”: domains with inadequate sample sizes for precise direct estimates.
- Small-area estimation (SAE): borrow strength across domains.
- Multivariate models improve small-area estimation for each outcome (De Souza, 1992).
- In most existing multivariate SAE, all outcomes are observed in each domain.
Small-area estimation

- “Small-area”: domains with inadequate sample sizes for precise direct estimates.
- Small-area estimation (SAE): borrow strength across domains.
- Multivariate models improve small-area estimation for each outcome (De Souza, 1992).
- In most existing multivariate SAE, all outcomes are observed in each domain.
- Our study is different: except for a relatively small calibration sample, only one outcome (the K6) is observed.
Small-area estimation

- “Small-area”: domains with inadequate sample sizes for precise direct estimates.
- Small-area estimation (SAE): borrow strength across domains.
- Multivariate models improve small-area estimation for each outcome (De Souza, 1992).
- In most existing multivariate SAE, all outcomes are observed in each domain.
- Our study is different: except for a relatively small calibration sample, only one outcome (the K6) is observed.
- Our goal: using a model estimated from the calibration sample, we predict the small area quantities of another (missing) outcome (SED) from the observed one.
Outline for small-area prediction

1. Build a model on the bivariate (both K6 and SED) NCS-A data and estimate model parameters.
Outline for small-area prediction

1. Build a model on the bivariate (both K6 and SED) NCS-A data and estimate model parameters.
2. Derive formulas for predicting the small-area means of SED from K6 given the estimated parameters.
Outline for small-area prediction

1. Build a model on the bivariate (both K6 and SED) NCS-A data and estimate model parameters.

2. Derive formulas for predicting the small-area means of SED from K6 given the estimated parameters.

3. For a new data set with only K6, collect auxiliary information (e.g., socio-demographic) for each individual.
Outline for small-area prediction

1. Build a model on the bivariate (both K6 and SED) NCS-A data and estimate model parameters.
2. Derive formulas for predicting the small-area means of SED from K6 given the estimated parameters.
3. For a new data set with only K6, collect auxiliary information (e.g., socio-demographic) for each individual.
4. Plug in parameter estimates from the NCS-A and the auxiliary information of the new sample into small area prediction formulas derived in Step 2.
Data and notations

- We focus on two-level hierarchical structure with individual $(i, j)$ for $j = 1, \ldots, J_i$ belonging to cluster $i$ for $i = 1, \ldots, I$.

- Each individual has two continuous outcomes: $Y_1, Y_2$. For example, in the NCS-A: $Y_1$ is sum of the K6 scores; $Y_2$ is the linear clinical diagnostic score of SED.

- Auxiliary information (covariates): individual-level $X_{ij1}, X_{ij2}$; cluster-level $Z_{i1}, Z_{i2}$.

- Objective: provide small area prediction of second-level mean of $Y_2$ from $Y_1$ for a distinct new sample where only $Y_1$ and $X, Z$ are observed.
Data and notations

- We focus on two-level hierarchical structure with individual $(i,j)$ for $j = 1, \ldots, J_i$ belonging to cluster $i$ for $i = 1, \ldots, I$.
- Each individual has two continuous outcomes: $Y_1, Y_2$. For example, in the NCS-A: $Y_1$ is sum of the K6 scores; $Y_2$ is the linear clinical diagnostic score of SED.
Data and notations

- We focus on two-level hierarchical structure with individual \((i, j)\) for \(j = 1, \ldots, J_i\) belonging to cluster \(i\) for \(i = 1, \ldots, I\).
- Each individual has two continuous outcomes: \(Y_1, Y_2\). For example, in the NCS-A: \(Y_1\) is sum of the K6 scores; \(Y_2\) is the linear clinical diagnostic score of SED.
Data and notations

- We focus on two-level hierarchical structure with individual \((i, j)\) for \(j = 1, \ldots, J_i\) belonging to cluster \(i\) for \(i = 1, \ldots, I\).
- Each individual has two continuous outcomes: \(Y_1, Y_2\). For example, in the NCS-A: \(Y_1\) is sum of the K6 scores; \(Y_2\) is the linear clinical diagnostic score of SED.
- Auxiliary information (covariates): individual-level \(X_{ij1}, X_{ij2}\); cluster-level \(Z_{i1}, Z_{i2}\).
We focus on two-level hierarchical structure with individual \((i, j)\) for \(j = 1, ..., J_i\) belonging to cluster \(i\) for \(i = 1, ..., I\).

Each individual has two continuous outcomes: \(Y_1, Y_2\). For example, in the NCS-A: \(Y_1\) is sum of the K6 scores; \(Y_2\) is the linear clinical diagnostic score of SED.

Auxiliary information (covariates): individual-level \(X_{ij1}, X_{ij2}\); cluster-level \(Z_{i1}, Z_{i2}\).

Objective: provide small area prediction of second-level mean of \(Y_2\) from \(Y_1\) for a distinct new sample where only \(Y_1\) and \(X, Z\) are observed.
Bivariate model for NCS-A: two continuous outcomes

- Two-level bivariate random effects model:

\[
Y_{ij1} = X_{ij1}\beta_1 + Z_{i1}\alpha_1 + v_{i1} + e_{ij1},
\]
\[
Y_{ij2} = X_{ij2}\beta_2 + Z_{i2}\alpha_2 + v_{i2} + e_{ij2},
\]

(1)
Bivariate model for NCS-A: two continuous outcomes

- Two-level bivariate random effects model:

\[
Y_{ij1} = X_{ij1}\beta_1 + Z_{i1}\alpha_1 + v_{i1} + e_{ij1},
\]

\[
Y_{ij2} = X_{ij2}\beta_2 + Z_{i2}\alpha_2 + v_{i2} + e_{ij2},
\]

with \((v_{i1}, v_{i2})' \sim N(0, \Sigma_v)\), \((e_{ij1}, e_{ij2})' \sim N(0, \Sigma_e)\), and

\[
\Sigma_e = \begin{pmatrix}
\sigma^2_{e1} & \rho_e\sigma_{e1}\sigma_{e2} \\
\rho_e\sigma_{e1}\sigma_{e2} & \sigma^2_{e2}
\end{pmatrix}, \quad \Sigma_v = \begin{pmatrix}
\sigma^2_v & \rho_v\sigma_v\sigma_{v2} \\
\rho_v\sigma_v\sigma_{v2} & \sigma^2_{v2}
\end{pmatrix}.
\]
Bivariate model for NCS-A: two continuous outcomes

- Two-level bivariate random effects model:
  \[ Y_{ij1} = X_{ij1} \beta_1 + Z_{i1} \alpha_1 + v_{i1} + e_{ij1}, \]
  \[ Y_{ij2} = X_{ij2} \beta_2 + Z_{i2} \alpha_2 + v_{i2} + e_{ij2}, \]
  (1)

with \((v_{i1}, v_{i2})' \sim N(0, \Sigma_v), (e_{ij1}, e_{ij2})' \sim N(0, \Sigma_e)\), and

\[ \Sigma_e = \begin{pmatrix} \sigma_{e1}^2 & \rho_e \sigma_{e1} \sigma_{e2} \\ \rho_e \sigma_{e1} \sigma_{e2} & \sigma_{e2}^2 \end{pmatrix}, \quad \Sigma_v = \begin{pmatrix} \sigma_{v1}^2 & \rho_v \sigma_{v1} \sigma_{v2} \\ \rho_v \sigma_{v1} \sigma_{v2} & \sigma_{v2}^2 \end{pmatrix}. \]

- For a continuous and a binary outcome, we can assume a probit model for the binary, implementation is similar.
We adopt the hierarchical Bayes approach to fit model (1).

Assume uninformative uniform priors for $\beta, \alpha$:

$$\beta \propto 1,$$

$$\alpha \propto 1.$$

Assume inverse Wishart prior for $\Sigma$:

$$\Sigma^{-1} \sim \text{Wishart}(b_0, b_0 \Sigma^{-1})$$

$b_0$: prior sample size; $\Sigma_0$: prior covariance matrix.

Prior ignorance: set $b_0 = 3$ and $\Sigma_0$ diagonal.

Sensitivity analysis show the results are robust to the above prior settings, with a slightly conservative estimation of correlation.

Model is fitted by the software MLwiN 2.02 (Browne, 2005).
We adopt the hierarchical Bayes approach to fit model (1).
Assume uninformative uniform priors for $\beta, \alpha$: $\beta \propto 1$, $\alpha \propto 1$.

Assume inverse Wishart prior for $\Sigma$:

$\Sigma^{-1} \sim \text{Wishart}(b_0, b_0 \Sigma^{-1}_0)$,

$b_0$: prior sample size; $\Sigma_0$: prior covariance matrix.

Prior ignorance: set $b_0 = 3$ and $\Sigma_0$ diagonal.

Sensitivity analysis show the results are robust to the above prior settings, with a slightly conservative estimation of correlation.

Model is fitted by the software MLwiN 2.02 (Browne, 2005).
We adopt the hierarchical Bayes approach to fit model (1).

Assume uninformative uniform priors for $\beta, \alpha$: $\beta \propto 1$, $\alpha \propto 1$.

Assume inverse Wishart prior for $\Sigma$:

$$
\Sigma^{-1} \sim \text{Wishart}(b_0, b_0 \Sigma_0^{-1}),
$$

$b_0$: prior sample size; $\Sigma_0$: prior covariance matrix.
We adopt the hierarchical Bayes approach to fit model (1). Assume uninformative uniform priors for $\beta, \alpha$: $\beta \propto 1, \alpha \propto 1$. Assume inverse Wishart prior for $\Sigma$:

$$\Sigma^{-1} \sim \text{Wishart}(b_0, b_0\Sigma_0^{-1}),$$

$b_0$: prior sample size; $\Sigma_0$: prior covariance matrix. Prior ignorance: set $b_0 = 3$ and $\Sigma_0$ diagonal.
We adopt the hierarchical Bayes approach to fit model (1).
Assume uninformative uniform priors for $\beta, \alpha$: $\beta \propto 1$, $\alpha \propto 1$.
Assume inverse Wishart prior for $\Sigma$:

$$\Sigma^{-1} \sim \text{Wishart}(b_0, b_0 \Sigma_0^{-1}),$$

$b_0$: prior sample size; $\Sigma_0$: prior covariance matrix.
Prior ignorance: set $b_0 = 3$ and $\Sigma_0$ diagonal.
Sensitivity analysis show the results are robust to the above prior settings, with a slightly conservative estimation of correlation.

Model is fitted by the software MLwiN 2.02 (Browne, 2005)
Model Fitting: Hierarchical Bayes

- We adopt the hierarchical Bayes approach to fit model (1).
- Assume uninformative uniform priors for $\beta, \alpha$: $\beta \propto 1, \alpha \propto 1$.
- Assume inverse Wishart prior for $\Sigma$:

$$\Sigma^{-1} \sim \text{Wishart}(b_0, b_0\Sigma_0^{-1})$$

$b_0$: prior sample size; $\Sigma_0$: prior covariance matrix.
- Prior ignorance: set $b_0 = 3$ and $\Sigma_0$ diagonal.
- Sensitivity analysis show the results are robust to the above prior settings, with a slightly conservative estimation of correlation.
- Model is fitted by the software MLwiN 2.02 (Browne, 2005)
Small-area prediction: general settings

- **Calibration sample**: both $Y_1$, $Y_2$ are observed, where the parameters $(\alpha, \beta, \Sigma_e, \Sigma_v)$ in model (1) are estimated.
Small-area prediction: general settings

- **Calibration sample**: both $Y_1$, $Y_2$ are observed, where the parameters $\left(\alpha, \beta, \Sigma_e, \Sigma_v\right)$ in model (1) are estimated.

- **Target population**: a distinct new population of clusters that we want to estimate cluster-mean $Y_2$, but with only data on
Small-area prediction: general settings

- **Calibration sample**: both $Y_1$, $Y_2$ are observed, where the parameters $(\alpha, \beta, \Sigma_e, \Sigma_v)$ in model (1) are estimated.
- **Target population**: a distinct new population of clusters that we want to estimate cluster-mean $Y_2$, but with only data on
- **Survey sample**: outcome $Y_1$ and auxiliary variables $X_1, X_2$ are available.
Small-area prediction: general settings

- **Calibration sample**: both $Y_1, Y_2$ are observed, where the parameters $(\alpha, \beta, \Sigma_e, \Sigma_v)$ in model (1) are estimated.
- **Target population**: a distinct new population of clusters that we want to estimate cluster-mean $Y_2$, but with only data on
- **Survey sample**: outcome $Y_1$ and auxiliary variables $X_1, X_2$ are available.
- Both target population and survey sample are within the same cluster, but there are two different scenarios:
Small-area prediction: general settings

- **Calibration sample**: both $Y_1$, $Y_2$ are observed, where the parameters $(\alpha, \beta, \Sigma_e, \Sigma_v)$ in model (1) are estimated.
- **Target population**: a distinct new population of clusters that we want to estimate cluster-mean $Y_2$, but with only data on
- **Survey sample**: outcome $Y_1$ and auxiliary variables $X_1, X_2$ are available.
- Both target population and survey sample are within the same cluster, but there are two different scenarios:
  1. In-sample: the target population is equal to or a subset of the survey sample.
Small-area prediction: general settings

- **Calibration sample**: both $Y_1, Y_2$ are observed, where the parameters $(\alpha, \beta, \Sigma_e, \Sigma_v)$ in model (1) are estimated.
- **Target population**: a distinct new population of clusters that we want to estimate cluster-mean $Y_2$, but with only data on
- **Survey sample**: outcome $Y_1$ and auxiliary variables $X_1, X_2$ are available.
- Both target population and survey sample are within the same cluster, but there are two different scenarios:
  1. In-sample: the target population is equal to or a subset of the survey sample.
  2. Out-of-sample: the target population is distinct from the survey sample.
Small-area prediction: general settings

- **Calibration sample**: both $Y_1$, $Y_2$ are observed, where the parameters $(\alpha, \beta, \Sigma_e, \Sigma_v)$ in model (1) are estimated.
- **Target population**: a distinct new population of clusters that we want to estimate cluster-mean $Y_2$, but with only data on
- **Survey sample**: outcome $Y_1$ and auxiliary variables $X_1, X_2$ are available.
- Both target population and survey sample are within the same cluster, but there are two different scenarios:
  1. In-sample: the target population is equal to or a subset of the survey sample.
  2. Out-of-sample: the target population is distinct from the survey sample.

Lead to different small-area prediction formulas.
Small-area prediction formulas

- For an in-sample case, the key is to find the distribution of 
  \( s_i = v_{i2} + \bar{e}_{i2} \) conditional on \( Y_1, X_1, Z_1 \).
Small-area prediction formulas

- For an in-sample case, the key is to find the distribution of $s_i = \nu_{i2} + \bar{e}_{i2}$ conditional on $Y_1, X_1, Z_1$.
- After some algebra, we can show

$$s_i \mid Y_{i.1}, X_{i.1}, Z_{i1}, \theta \sim N(\tilde{\mu}_{si}, \tilde{\sigma}_{si}^2),$$

where $\tilde{\mu}_{si} = c_{si}(\bar{Y}_{i1} - \bar{X}_{i1}\beta_1)$, and

$$\tilde{\sigma}_{si}^2 = \sigma_{v2}^2 + \frac{\sigma_{e2}^2}{J_{i2}} - \frac{(\rho_v\sigma_{v1}\sigma_{v2} + \rho_e\sigma_{e1}\sigma_{e2}/\sqrt{J_{i1}J_{i2}})^2}{\sigma_{v1}^2 + \sigma_{e1}^2/J_{i1}},$$

with prediction coefficient $c_{si} = \rho_v\sigma_{v1}\sigma_{v2} + \rho_e\sigma_{e1}\sigma_{e2}/\sqrt{J_{i1}J_{i2}}$. 
Small-area prediction formulas

- For a in-sample case, the key is to find the distribution of $s_i = v_{i2} + \bar{e}_{i2}$ conditional on $Y_1, X_1, Z_1$.
- After some algebra, we can show

$$s_i | Y_{i.1}, X_{i.1}, Z_{i1}, \theta \sim N(\tilde{\mu}_{si}, \tilde{\sigma}^2_{si}),$$

where $\tilde{\mu}_{si} = c_{si}(\bar{Y}_{i1} - \bar{X}_{i1}\beta_1)$, and

$$\tilde{\sigma}^2_{si} = \sigma^2_{v2} + \frac{\sigma^2_{e2}}{J_{i2}} - \frac{(\rho_v \sigma_{v1} \sigma_{v2} + \rho_e \sigma_{e1} \sigma_{e2}/\sqrt{J_{i1}J_{i2}})^2}{\sigma^2_{v1} + \sigma^2_{e1}/J_{i1}},$$

with prediction coefficient

$$c_{si} = \frac{\rho_v \sigma_{v1} \sigma_{v2} + \rho_e \sigma_{e1} \sigma_{e2}/\sqrt{J_{i1}J_{i2}}}{\sigma^2_{v1} + \sigma^2_{e1}/J_{i1}}.$$
Small-area prediction

Prediction for the cluster-level mean of a continuous outcome:

\[ E(\bar{Y}_{i,2}|Y_{i1}, x_{i1}, z_{i1}, x_{i2}, z_{i2}, \theta) \approx \bar{X}_{i2}\beta_2 + Z_{i2}\alpha_2 + \tilde{\mu}_{si} \]

where parameters \( \theta \) as given (the estimates from model (1) and the NCS-A data).
Small-area prediction

- Prediction for the cluster-level mean of a continuous outcome:

\[ E(\bar{Y}_{i,2} \mid Y_{i1}, X_{i1}, Z_{i1}, X_{i2}, Z_{i2}, \theta) \approx \bar{X}_{i2}\beta_2 + Z_{i2}\alpha_2 + \tilde{\mu}_{si} \]

where parameters \( \theta \) as given (the estimates from model (1) and the NCS-A data).

- Variance of the estimates can be obtained numerically.
Small-area prediction

- Prediction for the cluster-level mean of a continuous outcome:

  \[
  E(\bar{Y}_{i,2} | Y_{i1}, X_{i1}, Z_{i1}, X_{i2}, Z_{i2}, \theta) \approx \bar{X}_{i2}\beta_2 + Z_{i2}\alpha_2 + \tilde{\mu}_{si}
  \]

  where parameters \( \theta \) as given (the estimates from model (1) and the NCS-A data).

- Variance of the estimates can be obtained numerically.

- The key to the out-of-sample case is to obtain conditional distribution of the cluster-level random effects \( v_{i2} \) - the formulas have the same form but simpler.
Small-area prediction

- Prediction for the cluster-level mean of a continuous outcome:

\[ E(\bar{Y}_{i,2} | Y_{i1}, X_{i1}, Z_{i1}, X_{i2}, Z_{i2}, \theta) \approx \bar{X}_{i2}\beta_2 + Z_{i2}\alpha_2 + \tilde{\mu}_{si} \]

where parameters \( \theta \) as given (the estimates from model (1) and the NCS-A data).

- Variance of the estimates can be obtained numerically.

- The key to the out-of-sample case is to obtain conditional distribution of the cluster-level random effects \( \nu_{i2} \) - the formulas have the same form but simpler.

- Formulas for binary outcome prediction are also developed.
Measure of reliability

We measure accuracy of the prediction by ratio of the posterior to prior variance of the cluster-average random effects/erros:

$$ r^2 = \frac{\text{Var}(v_{i2} + \bar{e}_{i2} | Y_{i1}, \theta)}{\text{Var}(v_{i2} + \bar{e}_{i2} | \theta)}. $$

Reliability parameter $\zeta$ (variance reduction):

$$ \zeta = 1 - r^2. $$

$\zeta = 1$: perfect prediction; $\zeta = 0$: SAE data for $Y_1$ completely uninformative.
Measure of reliability

- We measure accuracy of the prediction by ratio of the posterior to prior variance of the cluster-average random effects/erros:

\[
r^2 = \frac{\text{Var}(v_{i2} + \bar{e}_{i2} | Y_{i1}, \theta)}{\text{Var}(v_{i2} + \bar{e}_{i2} | \theta)}.
\]

- Reliability parameter \( \zeta \) (variance reduction): \( \zeta = 1 - r^2 \).
Measure of reliability

- We measure accuracy of the prediction by ratio of the posterior to prior variance of the cluster-average random effects/erros:

\[ r^2 = \frac{\text{Var}(v_{i2} + \bar{e}_{i2}|Y_{i..1}, \theta)}{\text{Var}(v_{i2} + \bar{e}_{i2}|\theta)}. \]

- Reliability parameter \( \zeta \) (variance reduction): \( \zeta = 1 - r^2 \).
- \( \zeta = 1 \): perfect prediction;
  \( \zeta = 0 \): SAE data for \( Y_1 \) completely uninformative.
Measure of reliability

- Reliability of in-sample case:

\[ \zeta_i = \frac{[\rho_v + \rho_e \sigma_{e1} \sigma_{e2} / (J_{i1} \sigma_{v1} \sigma_{v2})]^2}{[1 + \sigma_{e1}^2 / (J_{i1} \sigma_{v1}^2)][1 + \sigma_{e2}^2 / (J_{i2} \sigma_{v2}^2)]}. \]
Measure of reliability

- Reliability of in-sample case:

\[
\zeta_i = \frac{\left[\rho_v + \rho_e \sigma_{e1} \sigma_{e2} / (J_{i1} \sigma_{v1} \sigma_{v2})\right]^2}{\left[1 + \sigma_{e1}^2 / (J_{i1} \sigma_{v1}^2)\right] \left[1 + \sigma_{e2}^2 / (J_{i2} \sigma_{v2}^2)\right]}.
\]

- High correlation $\rho_v$, $\rho_e$, large observed cluster sample size $J_i$, large ratio $\sigma_{v}^2 / \sigma_{e}^2$ all contribute to increased reliability of the results.
Measure of reliability

- Reliability of in-sample case:

\[
\zeta_i = \frac{[\rho_v + \rho_e \sigma_{e1} \sigma_{e2} / (J_i \sigma_v \sigma_e)]^2}{[1 + \sigma_{e1}^2 / (J_i \sigma_v^2)] [1 + \sigma_{e2}^2 / (J_i \sigma_e^2)]}.
\]

- High correlation \( \rho_v, \rho_e \), large observed cluster sample size \( J_i \), large ratio \( \sigma_v^2 / \sigma_e^2 \) all contribute to increased reliability of the results.

- As \( J_i \) increases, reliability converges to its upper bound \( \rho_v^2 \).
Measure of reliability

- Reliability of in-sample case:
  \[
  \zeta_i = \frac{[\rho_v + \rho_e \sigma_{e1} \sigma_{e2}/(J_i \sigma_{v1} \sigma_{v2})]^2}{[1 + \sigma_{e1}^2/(J_i \sigma_{v1})][1 + \sigma_{e2}^2/(J_i \sigma_{v2})]}. 
  \]

- High correlation $\rho_v$, $\rho_e$, large observed cluster sample size $J_i$, large ratio $\sigma_{v}^2/\sigma_{e}^2$ all contribute to increased reliability of the results.

- As $J_i$ increases, reliability converges to its upper bound $\rho_V^2$.

- For the same cluster sample size, in-sample cases have higher reliability than out-of-sample cases.
Socio-demographic variables: age, sex, race/ethnicity, and age at entrance into primary school. School-level predictors include school size and public/private school status.

The diagnostic instrument is a modification of the WHO's Composite International Diagnostic Interview (CIDI) appropriate for adolescents (CIDI-A).
Application to NCS-A

- Schools are leading providers of mental health services to children and adolescents in the US.
- We focus on the NCS-A school sample of the NCS-A to predict school-level SED prevalence from K6.
Application to NCS-A

- Schools are leading providers of mental health services to children and adolescents in the US.
- We focus on the NCS-A school sample of the NCS-A to predict school-level SED prevalence from K6.
- Schools with less than 10 students are dropped, resulting in 9022 students from 282 schools.
Application to NCS-A

- Schools are leading providers of mental health services to children and adolescents in the US.
- We focus on the NCS-A school sample of the NCS-A to predict school-level SED prevalence from K6.
- Schools with less than 10 students are dropped, resulting in 9022 students from 282 schools.
- Socio-demographic variables: age, sex, race/ethnicity, and age at entrance into primary school. School-level predictors include school size and public/private school status.

The diagnostic instrument is a modification of the WHO’s Composite International Diagnostic Interview (CIDI) appropriate for adolescents (CIDI-A).
Application to NCS-A

- Schools are leading providers of mental health services to children and adolescents in the US.
- We focus on the NCS-A school sample of the NCS-A to predict school-level SED prevalence from K6.
- Schools with less than 10 students are dropped, resulting in 9022 students from 282 schools.
- Socio-demographic variables: age, sex, race/ethnicity, and age at entrance into primary school. School-level predictors include school size and public/private school status.
- The diagnostic instrument is a modification of the WHO’s Composite International Diagnostic Interview (CIDI) appropriate for adolescents (CIDI-A).
The outcomes in NCS-A

- $Y_2$ - **Probability of SED**: calculated as a function of variables in the CIDI-A, from a probit model estimated from a small validation example.
The outcomes in NCS-A

- $Y_2$ - Probability of SED: calculated as a function of variables in the CIDI-A, from a probit model estimated from a small validation example.

- $Y_1$ - The K6 (Kessler and Mroczec, 1994): six short questions on mental health during the past 30 days - “how often did you feel (1) nervous; (2) hopeless; (3) restless or fidgety; (4) so depressed that nothing could cheer you up; (5) everything was an effort; (6) worthless?”
The outcomes in NCS-A

- $Y_2$ - **Probability of SED**: calculated as a function of variables in the CIDI-A, from a probit model estimated from a small validation example.

- $Y_1$ - **The K6** (Kessler and Mroczec, 1994): six short questions on mental health during the past 30 days - “how often did you feel (1) nervous; (2) hopeless; (3) restless or fidgety; (4) so depressed that nothing could cheer you up; (5) everything was an effort; (6) worthless?” Each item uses a frequency response scale coded from 0 (never) to 4 (all the time). Sum of all six items.
The outcomes in NCS-A

- $Y_2$ - Probability of SED: calculated as a function of variables in the CIDI-A, from a probit model estimated from a small validation example.

- $Y_1$ - The K6 (Kessler and Mroczec, 1994): six short questions on mental health during the past 30 days - “how often did you feel (1) nervous; (2) hopeless; (3) restless or fidgety; (4) so depressed that nothing could cheer you up; (5) everything was an effort; (6) worthless?” Each item uses a frequency response scale coded from 0 (never) to 4 (all the time). Sum of all six items.

- $Y_1$ - Augmented K6 (Green et al., 2010): K6 supplemented by five additional CIDI items that specifically assessed behavior disorders.
## Results of model fitting

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Augmented K6</th>
<th>SED</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est</td>
<td>SE</td>
</tr>
<tr>
<td>intercept</td>
<td>2.635*</td>
<td>(0.165)</td>
</tr>
<tr>
<td>sex(male)</td>
<td>-0.643*</td>
<td>(0.068)</td>
</tr>
<tr>
<td>age 14-year</td>
<td>0.266*</td>
<td>(0.115)</td>
</tr>
<tr>
<td>age 15-year</td>
<td>0.513*</td>
<td>(0.124)</td>
</tr>
<tr>
<td>age 16-year</td>
<td>0.373*</td>
<td>(0.125)</td>
</tr>
<tr>
<td>age 17-year</td>
<td>0.550*</td>
<td>(0.127)</td>
</tr>
<tr>
<td>age 18-year</td>
<td>0.496*</td>
<td>(0.171)</td>
</tr>
<tr>
<td>black</td>
<td>0.597*</td>
<td>(0.101)</td>
</tr>
<tr>
<td>hispanic</td>
<td>0.440*</td>
<td>(0.104)</td>
</tr>
<tr>
<td>other race</td>
<td>0.920*</td>
<td>(0.151)</td>
</tr>
<tr>
<td>start schl at 7</td>
<td>0.252*</td>
<td>(0.075)</td>
</tr>
<tr>
<td>start schl &gt; 7</td>
<td>0.765*</td>
<td>(0.151)</td>
</tr>
<tr>
<td>schl size</td>
<td>0.153</td>
<td>(0.091)</td>
</tr>
<tr>
<td>public schl</td>
<td>0.288*</td>
<td>(0.135)</td>
</tr>
</tbody>
</table>

**Table:** Estimates of coefficients of two-level models with posterior standard deviation (*Significant at 0.05 level two sided test).
### Variance Components

<table>
<thead>
<tr>
<th></th>
<th>K6 and SED</th>
<th>Aug. K6 and SED</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>school</td>
<td>individual</td>
</tr>
<tr>
<td>Variance (K6)</td>
<td>0.316</td>
<td>10.531</td>
</tr>
<tr>
<td>Variance (SED)</td>
<td>0.037</td>
<td>0.596</td>
</tr>
<tr>
<td>K6-SED correlation</td>
<td>0.695</td>
<td>0.516</td>
</tr>
</tbody>
</table>

**Table:** Estimated variance components of two-level models.

School-level correlation is strong despite modest individual-level correlation. Augmented K6 further improves school-level correlation.
Variance Components

<table>
<thead>
<tr>
<th></th>
<th>K6 and SED</th>
<th>Aug. K6 and SED</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>school</td>
<td>individual</td>
</tr>
<tr>
<td>Variance (K6)</td>
<td>0.316</td>
<td>10.531</td>
</tr>
<tr>
<td>Variance (SED)</td>
<td>0.037</td>
<td>0.596</td>
</tr>
<tr>
<td>K6-SED correlation</td>
<td>0.695</td>
<td>0.516</td>
</tr>
</tbody>
</table>

**Table:** Estimated variance components of two-level models.

- School-level correlation is strong despite modest individual-level correlation.
Variance Components

Table: Estimated variance components of two-level models.

- School-level correlation is strong despite modest individual-level correlation.
- Augmented K6 further improves school-level correlation.
Figure: Scatterplot of school-level K6 and the augmented K6 versus predicted SED for schools with more than 25 screened students.
Predictive models

- In a school-wide screening, the target population is exactly the survey sample: in-sample scenario with $J_{i1} = J_{i2} = J_i$, $X_1 = X_2 = X$. 

Table: Reliability of small-area prediction

<table>
<thead>
<tr>
<th>$J_i$</th>
<th>1</th>
<th>30</th>
<th>200</th>
<th>$\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\zeta_{out}$</td>
<td>0.00</td>
<td>0.28</td>
<td>0.60</td>
<td>0.72</td>
</tr>
<tr>
<td>$\zeta_{in}$</td>
<td>0.31</td>
<td>0.54</td>
<td>0.67</td>
<td>0.72</td>
</tr>
</tbody>
</table>

- Bivariate analysis enables $c_{si}$ to be extrapolated in practical cases with large school, but univariate analysis cannot.
Predictive models

- In a school-wide screening, the target population is exactly the survey sample: in-sample scenario with $J_{i1} = J_{i2} = J_i$, $X_1 = X_2 = X$.

- The prediction coefficients $c_{si}$ and reliability depend on the cluster sample size $J_i$.

<table>
<thead>
<tr>
<th></th>
<th>$1$</th>
<th>$30$</th>
<th>$200$</th>
<th>$\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\zeta^{\text{out}}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\zeta^{\text{in}}$</td>
<td>0.31</td>
<td>0.54</td>
<td>0.67</td>
<td>0.72</td>
</tr>
<tr>
<td>$\zeta^{\text{out}}$</td>
<td>0.00</td>
<td>0.28</td>
<td>0.60</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Table: Reliability of small-area prediction
Predictive models

▶ In a school-wide screening, the target population is exactly the survey sample: in-sample scenario with $J_i = J_i = J_i$, $X_1 = X_2 = X$.

▶ The prediction coefficients $c_{si}$ and reliability depend on the cluster sample size $J_i$.

**Table: Reliability of small-area prediction**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>30</th>
<th>200</th>
<th>$\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\zeta^{\text{out}}$</td>
<td>0.00</td>
<td>0.28</td>
<td>0.60</td>
<td>0.72</td>
</tr>
<tr>
<td>$\zeta^{\text{in}}$</td>
<td>0.31</td>
<td>0.54</td>
<td>0.67</td>
<td>0.72</td>
</tr>
</tbody>
</table>

▶ Bivariate analysis enables $c_{si}$s to be extrapolated in practical cases with large school, but univariate analysis cannot.
Model Validation

- Use posterior predictive checks (Gelman, Meng and Stern, 1996) to check modeling fitting.
- Generate copies of the NCS-A data using 1000 posterior draws of the parameters from model (1) and calculate posterior predictive p-values.

Table: Summary statistics from observed and simulated data. Posterior predictive p-values are shown in parenthesis.

<table>
<thead>
<tr>
<th>summary statistics</th>
<th>K6</th>
<th>Φ−1(SED)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>obs</td>
<td>sim</td>
</tr>
<tr>
<td>Individual mean</td>
<td>-1.72</td>
<td>-1.73 (0.41)</td>
</tr>
<tr>
<td>Individual s.d.</td>
<td>0.63</td>
<td>0.63 (0.47)</td>
</tr>
<tr>
<td>Mean of school means</td>
<td>-1.72</td>
<td>-1.72 (0.50)</td>
</tr>
<tr>
<td>S.d. of school means</td>
<td>0.20</td>
<td>0.21 (0.58)</td>
</tr>
<tr>
<td>Individual correlation with K6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>School correlation with K6</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Comparison with the observed school means

- The NCS-A have both SED and K6, we can compare the model-based predictions with the observed SED.

  - We predict school SED prevalence from K6 in the NCS-A as if only K6 had been measured.
  - Average observed and predicted school prevalence is 6.1% and 5.9% respectively.
  - The predicted and observed distributions of school means are well matched except in the upper tail (observed prevalence > 0.10).
  - The observed SED prevalence of 254 out of 282 (90.7%) schools falls into the 95% in-sample predictive intervals.
  - Despite the under-estimation, our method identifies 72.4% (21 out of 29) schools with the top 10% observed prevalence.
Comparison with the observed school means

- The NCS-A have both SED and K6, we can compare the model-based predictions with the observed SED.
- We predict school SED prevalence from K6 in the NCS-A as if only K6 had been measured.

Average observed and predicted school prevalence is 6.1% and 5.9% respectively.

The predicted and observed distributions of school means are well matched except in the upper tail (observed prevalence > 0.10).

The observed SED prevalence of 254 out of 282 (90.7%) schools falls into the 95% in-sample predictive intervals.

Despite the under-estimation, our method identifies 72.4% (21 out of 29) schools with the top 10% observed prevalence.
Comparison with the observed school means

- The NCS-A have both SED and K6, we can compare the model-based predictions with the observed SED.
- We predict school SED prevalence from K6 in the NCS-A as if only K6 had been measured.
- Average observed and predicted school prevalence is 6.1% and 5.9% respectively.
Comparison with the observed school means

- The NCS-A have both SED and K6, we can compare the model-based predictions with the observed SED.
- We predict school SED prevalence from K6 in the NCS-A as if only K6 had been measured.
- Average observed and predicted school prevalence is 6.1% and 5.9% respectively.
- The predicted and observed distributions of school means are well matched except in the upper tail (observed prevalence > 0.10)
Comparison with the observed school means

- The NCS-A have both SED and K6, we can compare the model-based predictions with the observed SED.
- We predict school SED prevalence from K6 in the NCS-A as if only K6 had been measured.
- Average observed and predicted school prevalence is 6.1% and 5.9% respectively.
- The predicted and observed distributions of school means are well matched except in the upper tail (observed prevalence > 0.10)
- The observed SED prevalence of 254 out of 282 (90.7%) schools falls into the 95% in-sample predictive intervals.
Comparison with the observed school means

- The NCS-A have both SED and K6, we can compare the model-based predictions with the observed SED.
- We predict school SED prevalence from K6 in the NCS-A as if only K6 had been measured.
- Average observed and predicted school prevalence is 6.1% and 5.9% respectively.
- The predicted and observed distributions of school means are well matched except in the upper tail (observed prevalence > 0.10).
- The observed SED prevalence of 254 out of 282 (90.7%) schools falls into the 95% in-sample predictive intervals.
- Despite the under-estimation, our method identifies 72.4% (21 out of 29) schools with the top 10% observed prevalence.
Comparison with the direct estimates

- A second check: compare direct and model-based estimates for aggregates of schools.
  - Collapse the 42 geographical strata into 14 larger strata.
  - Direct estimates: averages of the observed school means of SED probability within each strata weighted by the strata size.
  - In 10 out of 14 strata, the bivariate predictions fall within the 95% confidence intervals of the direct estimates.
  - The same strata were identified as having the two highest observed and predicted prevalences.
A second check: compare direct and model-based estimates for aggregates of schools.

Collapse the 42 geographical strata into 14 larger strata.
Comparison with the direct estimates

- A second check: compare direct and model-based estimates for aggregates of schools.
- Collapse the 42 geographical strata into 14 larger strata.
- Direct estimates: averages of the observed school means of SED probability within each strata weighted by the strata size.

In 10 out of 14 strata, the bivariate predictions fall within the 95% confidence intervals of the direct estimates. The same strata were identified as having the two highest observed and predicted prevalences.
Comparison with the direct estimates

- A second check: compare direct and model-based estimates for aggregates of schools.
- Collapse the 42 geographical strata into 14 larger strata.
- Direct estimates: averages of the observed school means of SED probability within each strata weighted by the strata size.
- In 10 out 14 strata, the bivariate predictions fall within the 95% confidence intervals of the direct estimates.
Comparison with the direct estimates

- A second check: compare direct and model-based estimates for aggregates of schools.
- Collapse the 42 geographical strata into 14 larger strata.
- Direct estimates: averages of the observed school means of SED probability within each strata weighted by the strata size.
- In 10 out 14 strata, the bivariate predictions fall within the 95% confidence intervals of the direct estimates.
- The same strata were identified as having the two highest observed and predicted prevalences.
Comparison with alternative methods

(A) A standard regression-synthetic model: first regress individual-level SED score on covariates without K6, and then calculate individual predictions and thus predict school means (e.g., Hudson, 2010).

\[ Y_{ij} = X_{ij} \beta + v_i + e_{ij}, \quad v_i \sim N(0, \sigma^2_v), \quad e_{ij} \sim N(0, \sigma^2_e). \]
Comparison with alternative methods

(A) A standard regression-synthetic model: first regress individual-level SED score on covariates without K6, and then calculate individual predictions and thus predict school means (e.g., Hudson, 2010).

(B) A similar regression-synthetic model including K6 as a predictor.

Y_{ij} = X_{ij} \beta + v_i + e_{ij},
\quad v_i \sim N(0, \sigma^2_v),
\quad e_{ij} \sim N(0, \sigma^2_e).

(D) The same as (C) but including K6 as a predictor.

Y_{ij} = \alpha + X_{ij} \beta + v_i + e_{ij},
\quad v_i \sim N(0, \sigma^2_v),
\quad e_{ij} \sim N(0, \sigma^2_e).
Comparison with alternative methods

(A) A standard regression-synthetic model: first regress individual-level SED score on covariates without K6, and then calculate individual predictions and thus predict school means (e.g., Hudson, 2010).

(B) A similar regression-synthetic model including K6 as a predictor.

(C) A univariate random-effects logistic model without K6 as a predictor.

\[ Y_{ij2} = X_{ij}\beta + v_i + e_{ij}, \quad v_i \sim N(0, \sigma_v^2), \quad e_{ij} \sim N(0, \sigma_e^2). \]
Comparison with alternative methods

(A) A standard regression-synthetic model: first regress individual-level SED score on covariates without K6, and then calculate individual predictions and thus predict school means (e.g., Hudson, 2010).

(B) A similar regression-synthetic model including K6 as a predictor.

(C) A univariate random-effects logistic model without K6 as a predictor.

\[ Y_{ij2} = X_{ij} \beta + v_i + e_{ij}, \quad v_i \sim N(0, \sigma^2_v), \quad e_{ij} \sim N(0, \sigma^2_e). \]

(D) The same as (C) but including K6 as a predictor.

\[ Y_{ij2} = Y_{ij1} \alpha + X_{ij} \beta + v_i + e_{ij}, \quad v_i \sim N(0, \sigma^2_v), \quad e_{ij} \sim N(0, \sigma^2_e). \]
Comparison with alternative methods

Table: Comparison of errors of prediction of school-level SED prevalences in NCS-A from different SAE models.

<table>
<thead>
<tr>
<th>Model</th>
<th>MSE ($\times 10^3$)</th>
<th>MAE ($\times 10^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) Synthetic without K6</td>
<td>2.35</td>
<td>2.94</td>
</tr>
<tr>
<td>(B) Synthetic with K6</td>
<td>1.90</td>
<td>2.55</td>
</tr>
<tr>
<td>(C) Univariate without K6</td>
<td>2.44</td>
<td>2.95</td>
</tr>
<tr>
<td>(D) Univariate with K6</td>
<td>1.67</td>
<td>2.38</td>
</tr>
<tr>
<td>Bivariate multilevel</td>
<td>1.54</td>
<td>2.30</td>
</tr>
</tbody>
</table>

MSE = mean squared error, MAE = mean absolute error.
Discussion

- Develop a methodology for bivariate small-area prediction based on multilevel models.
Discussion

- Develop a methodology for bivariate small-area prediction based on multilevel models.
- Apply to NCS-A to provide small-area estimation of SED prevalence of schools from the short screening scale K6.
Discussion

- Develop a methodology for bivariate small-area prediction based on multilevel models.
- Apply to NCS-A to provide small-area estimation of SED prevalence of schools from the short screening scale K6.
- Provide evidence: K6 feasible alternative to clinical diagnosis at school level.
Discussion

- Develop a methodology for bivariate small-area prediction based on multilevel models.
- Apply to NCS-A to provide small-area estimation of SED prevalence of schools from the short screening scale K6.
- Provide evidence: K6 feasible alternative to clinical diagnosis at school level.
- Ongoing work:
  1. Explore more flexible models (e.g., copula models with T distribution) to improve prediction in the upper tail of the prevalence distribution.
Discussion

- Develop a methodology for bivariate small-area prediction based on multilevel models.
- Apply to NCS-A to provide small-area estimation of SED prevalence of schools from the short screening scale K6.
- Provide evidence: K6 feasible alternative to clinical diagnosis at school level.
- Ongoing work:
  1. Explore more flexible models (e.g., copula models with T distribution) to improve prediction in the upper tail of the prevalence distribution.
  2. Develop open-source software packages for public usage.
Acknowledgements

- Harvard University HCP: Ron Kessler, Michael Gruber, Nancy Sampson
- Boston University: Jennifer Green
- The National Institute for Mental Health (NIMH), the National Institute on Drug Abuse (NIDA), the Substance Abuse and Mental Health Services Administration (SAMHSA), the Robert Wood Johnson Foundation, and the John W. Alden Trust.
Key References


