turning knowledge into practice

Investigating Alternative Ways of Estimating the Proportion of the Population with Serious Mental Illness Using the National Survey on Drug Use and Health

Prepared by

Phillip S. Kott (presenter), Dan Liao, and Jeremy Aldworth

RTI International, Rockville, MD, Washington, DC and Research Triangle Park, NC

Presented at

The 139th Annual Meeting of the American Public Health Association

Washington, DC • October 29–November 2, 2011

Phone 301-468-8281 • **e-mail** pkott@rti.org

Presenter Disclosure

Phillip S. Kott

The following personal financial relationships with commercial interests relevant to this presentation existed during the past 12 months:

Has been employed by RTI International,

an independent, nonprofit institute providing research, development, and technical services to government and commercial clients worldwide.

Acknowledgement

- This project is funded by the Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, under Contract no. 283-2008-00004 and Project no. 0211838
- The views expressed in this presentation do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government
- This presentation is sponsored by RTI International's Statistics and Epidemiology Division

Outline

- Background
- Defining Serious Mental Illness (SMI)
- Different Estimators
- Estimating Standard Errors
- Some Results
- Future Directions and Open Questions

Background

The National Survey on Drug Use and Health (**NSDUH**) is an annual household survey of the US civilian, non-institutionalized population, ages 12 or older conducted for the Substance Abuse and Mental Health Services Administration (SAMHSA)

- Contains approx. 67,500 respondents per year (since 1999).
- Produces national and state estimates of substance abuse and mental illness.

Background *(cont'd)*

- SAMHSA needs estimates of Serious Mental Illness (**SMI**). But no space on NSDUH for full a diagnostic assessment of SMI.
- Instead, full psychological assessment on a small stratified Bernoulli subsample of 18+ population using structured clinical interviews, the Mental Health Surveillance Study (**MHSS**).
- MHSS had roughly 750 relevant respondents in 2008 and 500 each in 2009 and 2010 (out of 46K adult NSDUH respondents).
- Plans are for roughly 1,500 each in 2011 and 2012.

RTI INTERNATIONAL

Serious Mental Illness (SMI)*

Mental Disorder:

- Diagnosable DSM-IV [**Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition**] mental disorder in past year (excluding substance use and developmental disorders)

Functional Impairment:

- Level of functional impairment substantially interferes with one or more major life activities

**Federal Register*: June 24, 1999.

Serious Mental Illness *(cont'd)*

A respondents to the MHSS is determined to have SMI if:

- Diagnosis of any applicable past year DSM-IV mental disorder
- Global Assessment of Functioning (GAF) score $\leq$ 50 (Level of impairment serious or worse)

This is viewed as the *Gold Standard* ("true") SMI determination based on which a logistic model for SMI will be estimated from the MHSS sample and then applied to all adult NSDUH respondents.

Scales in NSDUH MH Module

- **Psychological Distress:**
  - K6 scale (Kessler et al. 2003)
  - 6 items: item scores = 0–4; total score = 0–24
- **Functional Impairment:**
  - World Health Organization Disability Assessment Schedule (WHODAS, abbreviated)
  - 8 items: item scores = 0–3; total score = 0–24

(0 = no distress/impairment)

Other Potentially Useful NSDUH Predictors of SMI

- Past year or lifetime:
  - depression,
  - major depressive episode (DSM-IV criteria),
  - anxiety
- Suicidal thoughts in past year
- Received treatment for mental disorder in past year
  (not allowed because it would undercut a purpose)
- Demogrphics

RTI INTERNATIONAL

Estimators

- **Direct estimator:** Estimate of proportion of the population with SMI (in a domain) based purely on MHSS respondents:

$$Direct\ estimator = \frac{\sum_{\substack{MHSS \\ respondents \\ in\ \ domain}} w_k^{MHSS} y_k^{true}}{\sum_{\substack{MHSS \\ respondents \\ in\ \ domain}} w_k^{MHSS}}$$

← 1 if *k* is SMI, 0 otherwise

RTI INTERNATIONAL

Estimators *(cont'd)*

- **Cutpoint estimator:** Sample-weighted estimate of SMI proportion based on assigning each **NSDUH** respondent an SMI status (yes/no) and then estimating proportion with that status.

$$Cutpoint\ \ Estimator = \frac{\sum\limits_{\substack{NSDUH\\ respondents\\ in\ \ domain}} w_k^{NSDUH}\, y_k^{assigned}}{\sum\limits_{\substack{NSDUH\\ respondents\\ in\ \ domain}} w_k^{NSDUH}}$$

- This is what is currently used.

RTI INTERNATIONAL

Estimators *(cont'd)*

- Assignments are made from a fitted logistic model:

$$\sum_{\substack{MHSS \\ respondents \\ in\ full\, sample}} w_k^{MHSS}\left[y_k^{true} - \underset{\substack{\uparrow \\ p_k}}{p(\mathbf{x}_k{}'\mathbf{b})}\right]\mathbf{x}_k = \mathbf{0}.$$

$\mathbf{x}_k$ is a vector containing K6 and WHODAS measures and a constant.

- $y_k^{assigned}$ is set to 1 for respondents with p_k-values above a *cutpoint* and to 0 for the rest, so that the estimated number of incorrectly-assigned 1's (i.e., *false positives*) is as equal as possible to the estimated number of incorrectly-assigned 0's (i.e., *false negatives*).

Estimators *(cont'd)*

- **Probability estimator:** Sample-weighted estimate of SMI proportion based on assigning each **NSDUH** respondent a probability of being SMI.

$$Probability\ estimator = \frac{\sum_{\substack{NSDUH \\ respondents \\ in\ \ domain}} w_k^{NSDUH}\, p_k}{\sum_{\substack{NSDUH \\ respondents \\ in\ \ domain}} w_k^{NSDUH}}$$

- **Interpretation:** $w_k^{NSDUH}\, p_k$ are estimated with SMI and $w_k^{NSDUH}(1-p_k)$ without.

RTI INTERNATIONAL

Estimators *(cont'd)*

$$\text{Bias-adjusted probability estimator} = \frac{\displaystyle\sum_{\substack{NSDUH\\ respondents\\ in\ \ domain}} \left[w_k^{NSDUH}\, p_k + w_k^{MHSS}\,(y_k^{true} - p_k) \right]}{\displaystyle\sum_{\substack{NSDUH\\ respondents\\ in\ \ domain}} w_k^{NSDUH}}$$

$$\text{Bias-adjusted cutpoint estimator} = \frac{\displaystyle\sum_{\substack{NSDUH\\ respondents\\ in\ \ domain}} \left[w_k^{NSDUH}\, y_k^{assigned} + w_k^{MHSS}\,(y_k^{true} - y_k^{assigned}) \right]}{\displaystyle\sum_{\substack{NSDUH\\ respondents\\ in\ \ domain}} w_k^{NSDUH}}$$

$w_k^{MHSS} = 0$ when k is not in the MHSS sample.

RTI INTERNATIONAL

Estimators *(cont'd)*

- The bias-adjusted estimators are nearly unbiased under probability sampling theory. *They are model free*.
- When the domain is covered by the entire sample the bias-adjusted probability estimator and the (unadjusted) probability estimator coincide.
- The two cutpoint estimators are close to equal.

Estimating Standard Errors

- One can construct linearized standard errors for the direct and bias-adjusted estimators since each has the form:

$$Generic\,estimator = \frac{\displaystyle\sum_{\substack{NSDUH\\ respondents\\ in\ \ domain}} w_k^{NSDUH} z_k}{\displaystyle\sum_{\substack{NSDUH\\ respondents\\ in\ \ domain}} w_k^{NSDUH}},$$

where

$$z_k = \frac{w_k^{MHSS}}{w_k^{NSDUH}} y_k^{true}, \quad \text{or} \quad p_k + \frac{w_k^{MHSS}}{w_k^{NSDUH}} (y_k^{true} - p_k),$$

$$\text{or} \quad y_k^{assigned} + \frac{w_k^{MHSS}}{w_k^{NSDUH}} (y_k^{true} - y_k^{assigned})$$

RTI INTERNATIONAL

Assessing Biases

- We can test for the potential bias of a probability (or cutpoint) estimator within a domain by computing the “generic estimator” and its standard error by setting

$$z_k = \frac{w_k^{MHSS}}{w_k^{NSDUH}}(y_k^{true} - p_k)$$

$$\left(\text{or} \quad \frac{w_k^{MHSS}}{w_k^{NSDUH}}(y_k^{true} - y_k^{assigned}) \right).$$

- Is the resulting *bias measure* significantly different from zero?

Additional Standard-Error Comments

- We have also developed a linearized standard error for the unadjusted probability estimator under the assumption that the fitted logistic model is true.
- One cannot do that for the unadjusted cutpoint estimator.
- Our attempt to construct a replication-based standard-error measures using Fay's version of balanced repeated replication has not been supported by simulations.

RTI INTERNATIONAL

Some Results: Based on a Model Using Alternative K6 and WHODAS Measures with 2008/2009 Data

Domain	Estimator	Estimate	SE	Naïve SE
All	Direct estimator	4.99	0.98	—
All	Probability estimator, bias-adjusted	4.33	0.89	—
All	Cut-point estimator, bias-adjusted	3.82	1.16	—
All	Probability estimator, unadjusted	4.33	0.77	0.05
All	Cut-point estimator, unadjusted	3.79	?	0.11
All	Probability estimator bias	—	0.89	
All	Cut-point estimator bias	-0.03	1.17	

Some Results *(cont'd)*

Domain	Estimator	Estimate	SE	Naïve SE
Hispanic	Direct estimator	1.32	0.77	—
Hispanic	Probability estimator, bias-adjusted	1.32	1.06	—
Hispanic	Cut-point estimator, bias-adjusted	1.63	1.12	—
Hispanic	Probability estimator, unadjusted	3.93	0.71	0.13
Hispanic	Cut-point estimator, unadjusted	3.10	?	0.25
Hispanic	Probability estimator bias	2.61	1.06	
Hispanic	Cut-point estimator bias	1.47	1.11	

Some Results *(cont'd)*

- The probability estimator is also significantly biased for Nonhispanic blacks and for persons who sought treatment for mental disorders, but not for either sex or any investigated age group, region, level of education, or county level of urbanization.
- The cutpoint estimator was not significantly biased in any domain we investigated, but its standard error, when measurable, can be large.

Future Directions and Open Questions

- SAMHSA plans to recompute estimates based on a logistic model fit with multiple years of data using some variant of the cutpoint estimator, perhaps employing the whole-sample probabilty estimator to determine the cutpoint.
- What about other types of mental illness (e.g. moderate mental illness)?
- Should other covariates be added to the model and/or another response function be fit?
- Can a workable measure of the cutpoint estimator's standard error be developed?

RTI INTERNATIONAL