Behavioral Health is Essential To Health
Prevention Works
Treatment is Effective
People Recover
Getting the Most out of a Limited Sample Size Field Test: Experiences from the National Survey on Drug Use and Health

Jonaki Bose, Dicy Painter

Center for Behavioral Health Statistics and Quality,
Substance Abuse and Mental Health Services Administration

Doug Currivan, Larry Kroutil, Kevin Wang

RTI International

FCSM Research Conference
5 November 2013
Washington, DC
Outline

– Introduction to the NSDUH
– Reasons for Field Test
– Ideal Situation
– Reality
– Our Approach
National Survey on Drug Use and Health (NSDUH)

- **Sponsor**: Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Behavioral Health Statistics and Quality (CBHSQ)
- **Data collection conducted under contract with RTI International in North Carolina**
- **Purpose**: Estimate prevalence, correlates and trends of substance use and mental illness in U.S.
- **History**: Conducted since 1971, annually since 1990
NSDUH Design

• Representative national samples since 1971
• State sampling and computer-assisted interviewing began in 1999
• Civilian, noninstitutional population, age 12+
• 0, 1, or 2 persons selected
• Face-to-face interview at respondents’ homes
• $30 respondent incentive since 2002
• ~68,000 respondents per year (oversample 12-25)
NSDUH Questionnaire

• Core (stays constant across years)
  – Demographics (e.g., age, gender, race/ethnicity)
  – Use of tobacco, alcohol, and illicit drugs

• Noncore (can change across years)
  – Mental illness and depression
  – Substance use and mental disorders (DSM-IV)
  – Treatment for substance use or mental health problems
  – Health conditions, service utilization
  – Additional demographics
NSDUH Questionnaire - Core Modules

- Demographics
- Initiation, Recency, Frequency of Use
  - Tobacco
  - Alcohol
  - Marijuana
  - Cocaine
  - Heroin
  - Hallucinogens
  - Inhalants
  - Pain Relievers*
  - Tranquilizers*
  - Stimulants*
  - Sedatives*

*Misuse (also known as nonmedical use).
Why Redesign?

In general, ongoing surveys should periodically be assessed, and updated if necessary. For NSDUH,

- Update questionnaire and sample design to meet current data needs and measure new phenomena
  - Are state and substate estimates still a priority?
  - Prescription drug questions outdated
  - Increased need for data on mental health, treatment, prescription drug use
- Implement better data collection and estimation methods
- Adjust cost to expected budget levels
Past Year Substance Use Estimates With and Without New Noncore Questions: 2006

Number in millions

<table>
<thead>
<tr>
<th>Substance Type</th>
<th>Core Data</th>
<th>Core + New Qs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methamphetamine</td>
<td>1.5</td>
<td>1.9</td>
</tr>
<tr>
<td>Stimulants</td>
<td>3.4</td>
<td>4.8</td>
</tr>
<tr>
<td>Sedatives</td>
<td>0.9</td>
<td>2.6</td>
</tr>
</tbody>
</table>

Note: New questions added in 2006 noncore section asked about the stimulants Adderall® and non-prescription methamphetamine, and the sedative Ambien®.
Redesign Approach

• Assess data needs (2008-2012)
  – Consult data users to determine what data are needed
• Methods research and development (2008-2012)
  – Explore various design alternatives in terms of cost and impact on data quality and analytic capability—cognitive testing, analysis of existing data, lit review, etc.
• Present plan to leadership to gain acceptance
• Field test new design (Sep-Oct 2012)
• Dress rehearsal (Sep-Oct 2013)
• January 2015 implement the redesign
Summary of Limited Redesign Plan

• A few new questions added in 2013

• Sample design adjustment in 2014
  – State, age allocation; increase cluster size; drop HOI

• Instrumentation update in 2015
  – New contact materials and equipment
  – Change in the sponsor presented to the respondent USPHS to HHS.
  – New Rx drug module; other smaller changes
  – No major changes to core drugs, to maintain trends
  – Maintain $30 incentive payment, to avoid break in trend
New 2013 Questions

These were included in the field test

• Serving in military reserves
• Medical marijuana
• Height/weight
• Discussions one has had with a doctor about substance use and abuse in the past year.
Changes to Prescription Drug Module

- Update list of specific drugs, and maintain current drugs every year
- Ask about past year use of specific drugs, instead of lifetime
- On-screen pictures of pills
- Medical use and misuse
- Clearer definition, and separate components of and motivation for misuse
- Methamphetamine moved to separate module (out of prescription stimulants module)
Other Questionnaire Changes

- On-screen calendar for more accurate reporting by reference periods
- Some items moved to ACASI
- New items for sexual orientation, military families, and disability
- Improved measure of education status
- Improved health condition questions; age first diagnosed
- Combine snuff and chew into “smokeless”
- “New” definition of female binge, heavy drinking (4+ drinks)
- Update hallucinogens, inhalants lists
  - Ketamine, DMT/AMT, Salvia divinorum
  - Felt-tip pens, computer keyboard cleaner
Questions Needing Answers

• With the new design, can trends still be maintained for tobacco, alcohol, and marijuana?
• How did moving items from CAPI to ACASI affect estimates of health insurance coverage, sources of income, personal and family income, employment, etc.?
• What did estimates from new questions look like?
• What did the new prescription drug data look like?
  How was the new module performing?
Questions Needing Answers (continued)

• How long was the redesigned survey?
  – Overall
  – For specific modules (e.g., prescription drugs)
  – For particular subgroups of respondents
• How did the new equipment perform in the field?
• How did the updated contact materials perform in the field?
• What was the interviewer experience like?
Field Test Goals

• Opportunity to test changes under field conditions
  – Major changes being considered (e.g., prescription drugs)
  – Multiple changes being considered

• Opportunity to anticipate effects on estimates, trends, data quality, burden, etc. in advance of full-scale implementation

• Opportunity to reconsider any proposed changes

*However*, if the field test design has limitations, then making inferences and decisions becomes challenging.
Ideal Field Test Conditions

- Mimics all essential survey conditions
  - Field staff
  - Response rates
  - Sample design and field conditions
- Only differences are elements being tested
- Split-sample designs
- Sufficiently large sample sizes for statistical power
- Sufficient time to analyze the data and make informed decisions
NSDUH 2012 Field Test Design

- Sample size of 2,000 cases
- No split sample; part of main survey field work
- Excluded Alaska and Hawaii
- No Spanish interviews
- Used experienced NSDUH field staff
- Used the new age allocation but did not use other elements of the new sample design that will be implemented in 2014.
- Modified editing, imputation and weighting procedures
Limitations of Field Test Design

• The field test sample size constrained the kinds of inferences and decisions that realistically could be made.
• However, decisions still needed to be made regardless of the results of field testing.
• Therefore we attempted to maximize all the resources that were available to us.
Identifying Comparison Data Within the NSDUH

• Split samples may be ideal
  – Allow for the isolation of the effects
  – Need larger sample size

• Used comparison data sets (with relevant exclusions)
  – Main study data collected during similar field period
  – Main study data from the year before (all 4 quarters)
Past Year Use of Alcohol, Cigarettes, Marijuana, Cocaine, and Heroin among Persons Aged 12 or Older: 2012 Comparison and 2012 QFT

Percent Using in Past Year

- **Alcohol**: 67.6% (2012 Comparison), 66.8% (2012 QFT)
- **Cigarettes**: 26.1% (2012 Comparison), 28.0% (2012 QFT)
- **Marijuana**: 12.1% (2012 Comparison), 12.4% (2012 QFT)
- **Cocaine**: 1.7% (2012 Comparison), 1.5% (2012 QFT)
- **Heroin**: 0.2% (2012 Comparison), 0.2% (2012 QFT)

QFT = Questionnaire Field Test.

† Difference between this estimate and the 2012 QFT estimate is statistically significant at the .05 level.

Note: Error bars denote the bounds of asymmetric 95 percent confidence intervals.

Note: Samples do not include Alaska or Hawaii and do not include Spanish-language interviews.

Note: The survey data for the 2012 comparison estimate were collected in quarter 3 and quarter 4, 2012, through December 2, 2012.

Note: The survey data for the 2012 QFT estimate were collected from September 1 through November 3, 2012.
Other Sources of Information

- Prior work in the field
- Prior work within the survey
- External data sources
- Timing data
- Interviewer debrief items
- Rates of refusals and don’t knows
- ‘Other, specify’ responses
Comparison with External Data Sources

• New height-weight questions
  – Used NHIS and NHANES (reported/measured)
  – Evaluated effects of different upper and lower bounds
  – Changed the upper and lower bounds in the programming based on the results
Using Timing and Debriefing Data

- How do estimates of prescription drug misuse compare in the field test and main survey?
- Even if results suggest that the revised questions are capturing more complete/accurate reports of misuse, are the changes justifiable from a timing (i.e., respondent burden) perspective?
Comparison of Selected Prescription Drug Estimates

**Percentage, Aged 12 or Older**

- **Any Psychotherapeutic:**
  - Field Test (n = 2,044) - 17.9%
  - 2011 Comparison (n = 65,928) - 20.5%
  - 2012 Comparison (n = 31,213) - 21.0%
  - Difference is statistically significant at the .05 level.

- **Pain Relievers:**
  - Field Test (n = 2,044) - 12.0%
  - 2011 Comparison (n = 65,928) - 13.6%
  - 2012 Comparison (n = 31,213) - 14.4%
  - Difference is statistically significant at the .05 level.

- **Any Psychotherapeutic:**
  - Field Test (n = 2,044) - 8.1%
  - 2011 Comparison (n = 65,928) - 5.7%
  - 2012 Comparison (n = 31,213) - 5.9%
  - Difference is statistically significant at the .05 level.

- **Pain Relievers:**
  - Field Test (n = 2,044) - 6.0%
  - 2011 Comparison (n = 65,928) - 4.3%
  - 2012 Comparison (n = 31,213) - 4.4%
  - Difference is statistically significant at the .05 level.

* + Difference between estimate and the field test estimate is statistically significant at the .05 level.*
Total Interview Timing, by Age Group

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Field Test (n = 2,044)</th>
<th>2011 Comparison (n = 65,928)</th>
<th>2012 Comparison (n = 31,213)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 or Older</td>
<td>55.99</td>
<td>58.30</td>
<td></td>
</tr>
<tr>
<td>12 to 17</td>
<td>58.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 to 25</td>
<td>68.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26 to 49</td>
<td>69.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 to 64</td>
<td>74.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65 or Older</td>
<td>68.43</td>
<td></td>
<td>69.39</td>
</tr>
</tbody>
</table>
Interview Timing for Prescription Drug Module, by Age Group

Field Test (n = 2,044)
2011 Comparison (n = 65,928)
2012 Comparison (n = 31,213)

Median Time, in Minutes

- 12 or Older: 4.92, 4.77, 4.77
- 12 to 17: 4.92, 4.77, 4.77
- 18 to 25: 4.92, 4.77, 4.77
- 26 to 49: 5.92, 5.77, 5.77
- 50 to 64: 6.68, 6.75, 6.75
- 65 or Older: 8.28, 6.68, 6.75
Interviewer Debriefing Items

• Did the respondent make any comments about the interview being too long?

• Please describe the respondent's [R's] comments about the prescription drug questions.

• What comments, if any, did the respondent [R] make about the lead letter or in response to the lead letter?

• What comments, if any, did the respondent [R] (or parent/guardian) make about the Q&A [question and answer] brochure?
CAPI or ACASI?

• Notable differences in estimates for some sources of income and health insurance

• Because of need to measure changes related to insurance due to the Affordable Care Act, the only way to justify the move to ACASI is to demonstrate that ACASI estimates remain the same or are better.

• How do you establish that estimates are better?
  – Closer to “gold-standard” estimates
  – Theoretical basis
New question on *family members serving in the military* had 8.9% weighted missing data rate in the field test. For the dress rehearsal:

- Definition of immediate family was added to the question vs. being offered only as help text
- A new response category for “another member of the immediate family” was added
- An “other, specify” follow-up question was added for when the new response category is selected
Limitations

• Be completely confident about absence of impact on the early parts of the questionnaire
• Determine what the prescription drug estimates will be in 2015
• Evaluate the effect on substance abuse and dependence estimates
• Isolate the effects of the different changes
• Evaluate context effects
Final Challenge: Time

• A field test becomes less valuable if all the information that can be gleaned from it is not gleaned in time for the next level of decision making.
  – For example we were not able to make the decision to move items back into CAPI in time for the dress rehearsal.

• Be prepared for snags:
  – Potential nonresponse bias that cannot be fully accounted for by the simplified nonresponse adjustments needed to meet the field test analysis schedule
Next Steps

- Repeat some field test analyses with dress rehearsal data
- Combine field test and dress rehearsal data for some analyses
- Finalize questionnaire and submit OMB package on May 2014
- Analyze 2014 data to see if there is any evidence of effects on estimates of the sample design change
  - Interviewer effects in states with major changes in field staff (e.g., California)