

Data Driven System Transformation

John A. Pandiani, Ph.D. and Joan A. Mongeon, M.S.
Vermont Department of Health

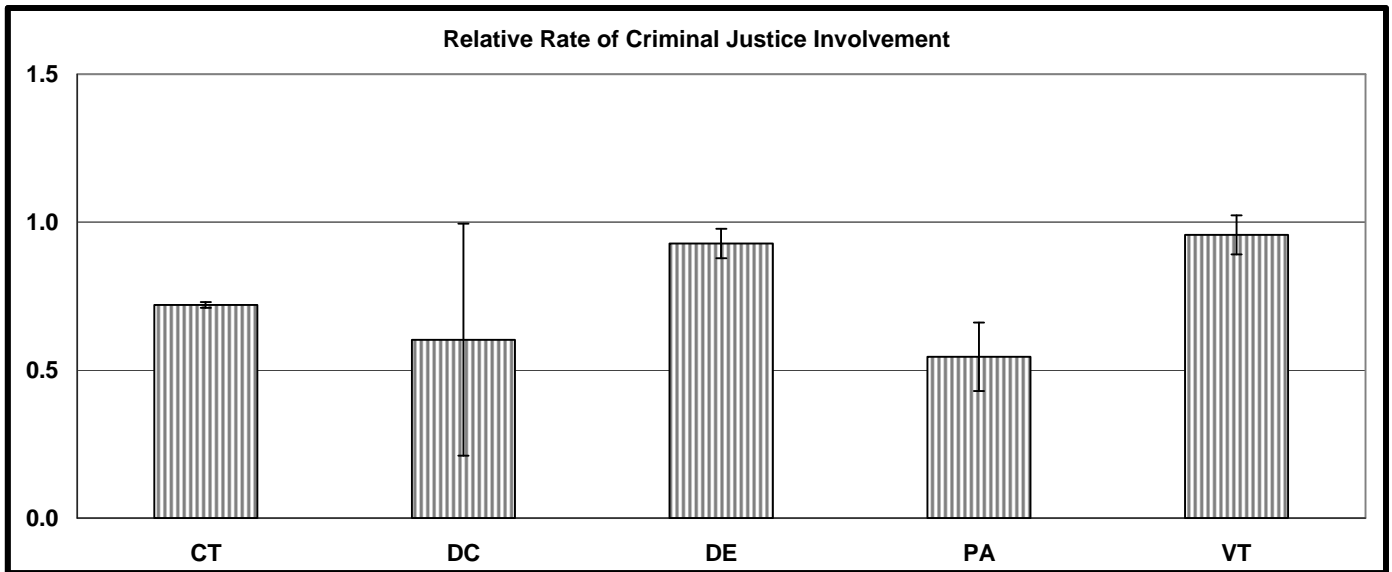
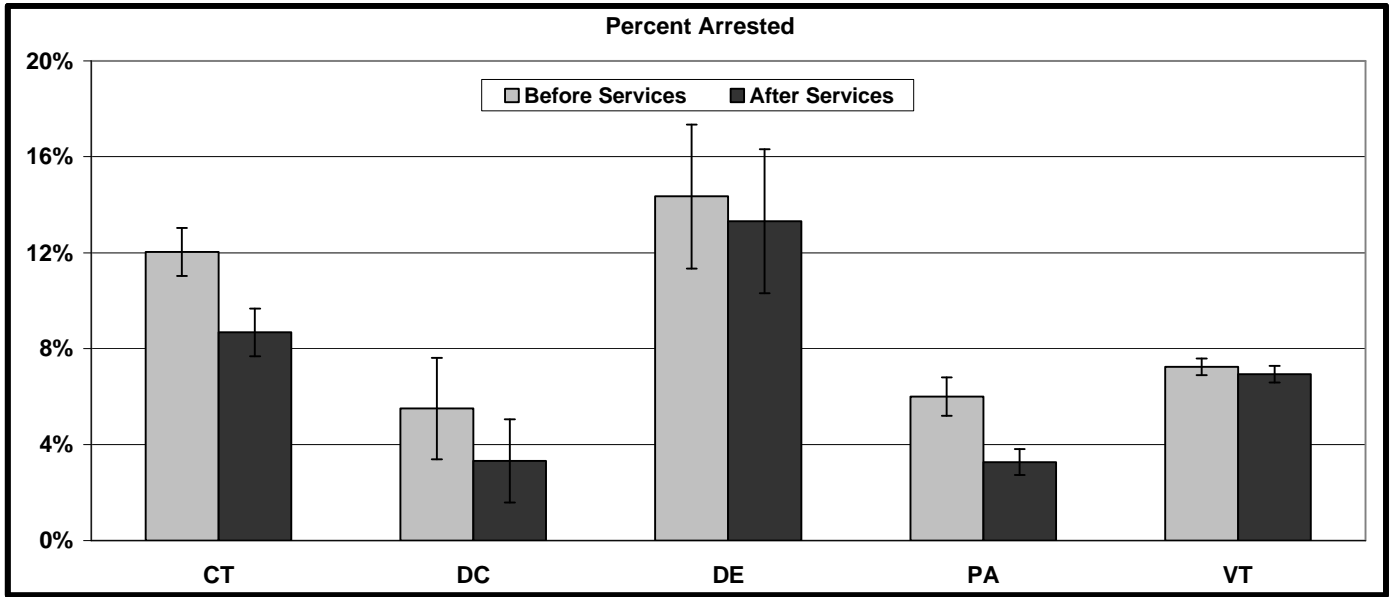
Steven M. Banks, Ph.D. The Bristol Observatory

Today, data are everywhere. This presentation discusses approaches to analyzing data and presenting findings in a way that supports system transformation. The SAMHSA National Outcome Measures (NOMs) require states to produce comparable measures of the performance of community based mental health and substance abuse programs. The data necessary to produce many of the NOMs has already been collected and is stored in electronic databases, awaiting analysis. Administrative databases are uniquely suited to informing the transformation process and monitoring the transformation of results of the care. The our systems of service systems challenge facing researchers today is to maximize the utility of available data. This presentation will use statistical techniques designed to produce valid and reliable information while protecting the confidentiality of medical records (HIPAA) and the personal privacy of individuals. Findings will be presented in a way that is accessible to a broad community of learners. Findings will be discussed in relation to the process of system transformation and in relation to underlying values.

**USE
THE
DATA**

www.thebristolobservatory.com

Criminal Justice Involvement Before and After Mental Health Services in 5 States

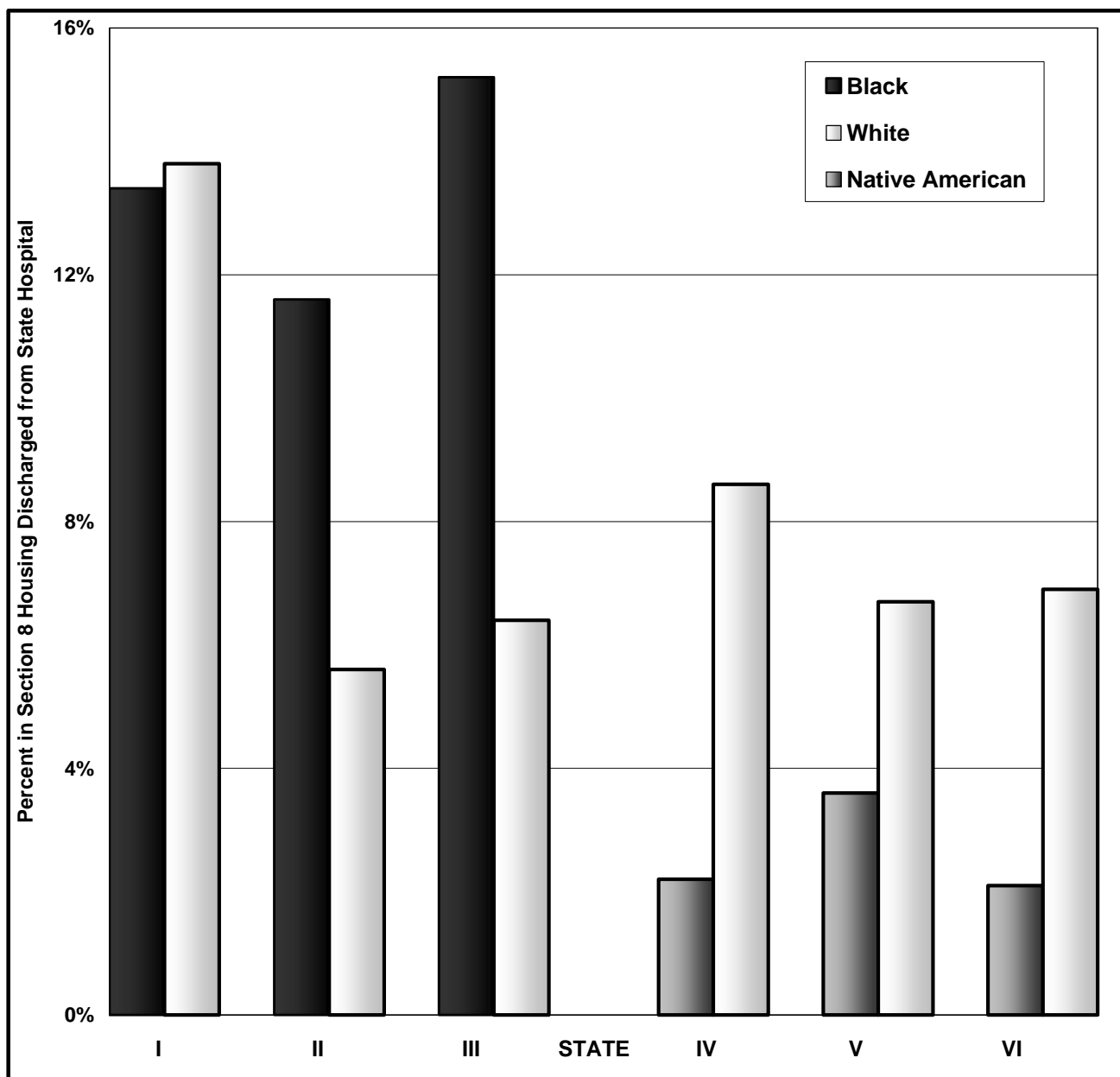


	STATE									
	Connecticut		Washington DC		Delaware		Pennsylvania		Vermont	
	Before	After	Before	After	Before	After	Before	After	Before	After
Total Served by MH Services	% 12.0%	8.7%	5.5%	3.3%	14.3%	13.3%	6.0%	3.3%	7.2%	6.9%
	± 1.00%	1.00%	± 2.11%	1.74%	± 3.00%	3.00%	± 0.80%	0.54%	± 0.35%	0.35%
Relative Rate	0.72	±0.01	0.60	±0.39	0.93	±0.05	0.55	±0.12	0.96	±0.07

Analysis is based on data provided by State Mental Health and Criminal Justice Agencies. Criminal Justice data sets describe rate of arrests in Connecticut, Washington DC, Delaware and Vermont. Criminal Justice data sets describe charges for Bedford County in Pennsylvania. Because these data sets do not include unique person identifiers, Probabilistic Population Estimation was used to determine overlap between Mental Health and Criminal Justice caseloads. Each estimate of caseload overlap is accompanied by a 95% confidence interval (±).

Relative Rate is the rate of criminal justice involvement before services divided by the rate of criminal justice involvement after services. A change ratio of 1.0 indicates no change in criminal justice involvement. A change ratio that is significantly greater than 1.0 indicates an increase in criminal justice involvement. A change ratio of 2.0, for instance, would indicate the after services rate was twice the before services rate. A change ratio that is significantly less than 1.0 indicates a decrease in criminal justice involvement. A change ratio of 0.5, for instance, would indicate the after services rate was half the before services rate.

Section 8 Utilization Rates After State Hospital Discharge in 6 States



	I		II		III		IV		V		VI	
	Black	White	Black	White	Black	White	Native American	White	Native American	White	Native American	White
Total Discharged												
#	66	177	80	142	305	470	302	401	138	691	45	472
Living in Section 8 Housing												
#	6	24	9	8	46	30	7	35	5	46	1	33
%	13%	14%	12%	6%	15%	6%	2%	9%	4%	7%	2%	7%

Findings for this report are based on analysis of HIPAA compliant extracts from the national Section 8 Housing database housed at The Bristol Observatory; and extracts from the state hospital database maintained by the National Association of State Mental Health Program Directors Research Institute. Because these data sets do not include unique person identifiers Probabilistic Population Estimation was used to determine overlap (with 95% confidence intervals).

PROBABILISTIC POPULATION ESTIMATION

Probabilistic Population Estimation is a statistical procedure that determines the number of people (with known confidence intervals) who are represented in data sets that do not contain unique person identifiers. Probabilistic Population Estimation uses information on the distribution of birth dates in a data set to determine the number of people represented in the data set. The number of people necessary to produce the number of birthdays observed in a single birth year cohort, for instance, would be calculated using the following formula:

$$P_j(l_j) = \sum_{i=1}^{l_j} \frac{365}{365-i}$$

where “P_j” is the number of people and “i” is the number of birth dates observed. Similar logic is used to determine the number of people who appear in more than one data set. The table below provides illustrative results of Probabilistic Population Estimation for populations of specified size.

Population Estimates for Specified Numbers of Birth Dates within a Year

Birth Dates	Number of People	Birth Dates	Number of People
1	1.003 ± 0.103	180	249 ± 20
10	10.15 ± 0.776	250	423 ± 38
20	20.6 ± 1.54	300	632 ± 64
50	54 ± 4	330	860 ± 101
100	117 ± 9	360	1603 ± 325

POPULATION OVERLAP

In order to probabilistically determine the number of people shared across data sets that do not include a common person identifier, the sizes of three populations are determined and the results are compared. The number of people in each of the original data sets are the first two populations. The number of people in a data set that is formed by combining the two original data sets is the third data set.

The number of people who are shared by the two data sets is the difference between the sum of the numbers of people represented in the two original data sets and the number of people represented in the combined data set. This occurs because the sum of the number of people represented in the two original data sets includes a double count of every person who is represented in both data sets. The number of people represented in the combined data set does not include this duplication. The difference between these two numbers is the size of the duplication between the two original data sets, the size of the caseload overlap. In terms of mathematical set theory, the intersection of two sets is the difference between the sum of the sizes of the two sets (A+B) and the union of the two sets (A∪B):

$$(A \cap B) = (A + B) - (A \cup B).$$

Related Readings

- Interpreting the Results of the CATIE Study. (Pandiani and Banks) *Psychiatric Services* 57: 140-141, January 2006
- Access to Child and Adolescent Mental Health Services. (Pandiani, Banks, Simon, Van Vleck, Pomeroy) *Journal of Child and Family Studies* 14: 3, 431-441, September 2005
- Involvement in the Criminal Justice System Among Recipients of Mental Health Services After September 11. (Pandiani JA, Knisley MB, Banks SM, Simon M, Blackburn P) *Psychiatric Services* 56: 80-84, January 2005
- Impact of Multi-Agency Employment Services on Employment Rates. (Pandiani, Tracy, Simon, Banks) *Community Mental Health Journal* 40: 333-345, August 2004
- Large Data Sets Are Powerful. (Pandiani and Banks) *Psychiatric Services* 54: 746, May 2003
- Terrorism and People with Mental Illness. (Pandiani and Banks) *Psychiatric Services* 53:1475, November 2002
- Measuring Access To Mental Health Care: A Multi-indicator Approach to Program Evaluation. (Pandiani, Banks, Bramley, Pomeroy, Simon) *Evaluation and Program Planning* 25: 271-285. August 2002
- Consumer Satisfaction and Treatment Outcomes: Dissatisfaction with Mental Health Services and Incarceration after Treatment. (Pandiani, Schacht, Banks) *Administration and Policy in Mental Health* 29: 145-155, November 2001
- Probabilistic Population Estimation of the Size and Overlap of Data Sets Based on Date of Birth. (Banks and Pandiani) *Statistics in Medicine* 20: 1421-1430, May 2001
- Caseload Segregation/Integration and Treatment Outcomes for Children and Adolescents. (Banks, Pandiani, and Schacht) *Journal of Emotional and Behavioral Disorders* 9: 232-238, Winter 2001
- Elevated Risk of Being Charged with a Crime for People with a Severe and Persistent Mental Illness. (Pandiani, Banks, Clements, and Schacht) *Justice Research and Policy* 2: 19-36, Fall 2000
- Utilization of Local Jails and General Hospitals by State Psychiatric Center Patients. (Banks, Stone, Pandiani, Cox, and Morchauser) *The Journal of Behavioral Health Services and Research* 27: 454-459, November 2000
- Caseload Segregation/Integration: A Measure of Shared Responsibility for Children and Adolescents. (Banks, Pandiani, and Schacht) *Journal of Emotional and Behavioral Disorders* 7: 66-71, Summer 1999
- Using Incarceration Rates to Measure Mental Health Program Performance. (Pandiani, Banks, and Schacht) *Journal of Behavioral Health Services and Research* 25: 300-311, August 1998
- Personal Privacy vs. Public Accountability: A Technological Solution to an Ethical Dilemma. (Pandiani, Banks, and Schacht) *Journal of Behavioral Health Services and Research* 25: 456-463, November 1998.
- The Utilization of State and General Hospitals for Inpatient Psychiatric Care. (Banks and Pandiani) *American Journal of Public Health* 88: 448-451, March 1998
- A Global Measure of Access to Mental Health Services for a Managed Care Environment. (Pandiani, Banks, and Gauvin) *Journal of Mental Health Administration* 24: 268-277, Summer 1997

The Bristol Observatory (TBO)

TBO is a contract research firm that specializes in human services program evaluation and service system research with an emphasis on treatment outcomes. Our measurement of treatment outcomes is frequently based on analysis of large administrative/operational databases from multiple public agencies using statistical tools that protect the personal privacy of the individuals represented in those databases. One of our primary tools in this work is Probabilistic Population Estimation, a statistical data mining tool that uses anonymous data sets to produce the information on caseload size and overlap in complex systems of care. Probabilistic Population Estimation allows our researchers to measure treatment outcomes, levels of access to care, and caseload overlap where the absence of unique person identifiers and/or concerns about personal privacy precludes direct linkage of records. For more information, visit: www.thebristolobservatory.com

The Vermont Mental Health Performance Indicator Project (PIP)

The PIP encourages rational data-based thinking and decision making within systems of care in Vermont by producing and widely distributing brief data reports on a weekly basis to service providers, consumers, administrators, advocates, and others. These reports rely heavily on analysis of existing administrative databases as a way of learning about the performance of programs and systems of care. All PIP reports are available on-line at <http://www.healthyvermonters.info/ddmhs/docs/res-eval/pip-reports.html>. The PIP was recently recognized by the Annapolis Coalition on Behavioral Workforce Education as one of the “Innovative Educational Practices” highlighted in the November, 2004 special issue of Administration and Policy in Mental Health.

