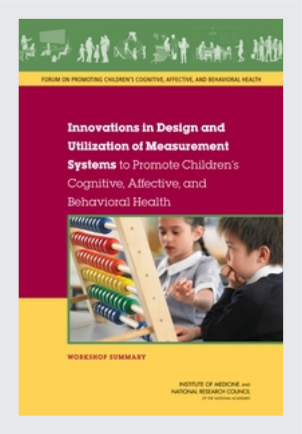

Innovations in Design and Utilization of Measurement Systems to Promote Children's Cognitive, Affective, and Behavioral Health: Workshop Summary

DETAILS

106 pages | 6 x 9 | PAPERBACK
ISBN 978-0-309-36748-6 | DOI 10.17226/21661

AUTHORS

Steve Olson and Noam I. Keren, Rapporteurs; Forum on Promoting Children's Cognitive, Affective, and Behavioral Health; Board on Children, Youth, and Families; Institute of Medicine; National Research Council

Innovations in Design and Utilization of Measurement Systems to Promote Children's Cognitive, Affective, and Behavioral Health

WORKSHOP SUMMARY

Steve Olson and Noam I. Keren, *Rapporteurs*

Forum on Promoting Children's Cognitive, Affective, and Behavioral Health

Board on Children, Youth, and Families

INSTITUTE OF MEDICINE *AND*
NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

NOTICE: The workshop that is the subject of this workshop summary was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

This activity was supported by contracts between the National Academy of Sciences and the American Academy of Pediatrics (Unnumbered Award); the American Board of Pediatrics (Unnumbered Award); the Annie E. Casey Foundation (213.0427); Autism Speaks (Unnumbered Award); the Centers for Disease Control and Prevention (200-2011-38807, TO #16); the Department of Justice Office of Juvenile Justice and Delinquency Prevention (2013-MU-MU-0002); the National Institutes of Health (HHSN26300035); the Robert Wood Johnson Foundation (71071); the Substance Abuse and Mental Health Services Administration (HHSP233201300244P); and the William T. Grant Foundation (182528). Additional support came from the American Orthopsychiatric Association, the American Psychological Association, the Hogg Foundation for Mental Health, the Society for Child and Family Policy and Practice, the Society of Clinical Child and Adolescent Psychology, and the Society of Pediatric Psychology. The views presented in this publication do not necessarily reflect the views of the organizations or agencies that provided support for the activity.

International Standard Book Number-13: 978-0-309-36748-6
International Standard Book Number-10: 0-309-36748-4

Additional copies of this workshop summary are available for sale from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

For more information about the Institute of Medicine, visit the IOM home page at: **www.iom.edu.**

Printed in the United States of America

The serpent has been a symbol of long life, healing, and knowledge among almost all cultures and religions since the beginning of recorded history. The serpent adopted as a logotype by the Institute of Medicine is a relief carving from ancient Greece, now held by the Staatliche Museen in Berlin.

Suggested citation: IOM (Institute of Medicine) and NRC (National Research Council). 2015. *Innovations in design and utilization of measurement systems to promote children's cognitive, affective, and behavioral health: Workshop summary.* Washington, DC: The National Academies Press.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. C. D. Mote, Jr., is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Victor J. Dzau is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. C. D. Mote, Jr., are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

PLANNING COMMITTEE FOR THE WORKSHOP ON INNOVATIONS IN DESIGN AND UTILIZATION OF MEASUREMENT SYSTEMS TO PROMOTE CHILDREN'S COGNITIVE, AFFECTIVE, AND BEHAVIORAL HEALTH[1]

C. HENDRICKS BROWN (*Chair*), Professor, Departments of Psychiatry, Behavioral Sciences, and Preventive Medicine, Northwestern University
WILLIAM R. BEARDSLEE, Director, Baer Prevention Initiatives; Chairman Emeritus, Department of Psychiatry, Boston Children's Hospital; and Gardner/Monks Professor of Child Psychiatry, Harvard Medical School
DARA BLACHMAN-DEMNER, Social Science Analyst, Crime, Violence, and Victimization Research Division, National Institute of Justice
DAVID A. BRENT, Academic Chief, Child and Adolescent Psychiatry, Western Psychiatric Institute and Clinic and Professor of Psychiatry, Pediatrics, and Epidemiology, School of Medicine, University of Pittsburgh
KIMBERLY E. HOAGWOOD, Society of Clinical Child and Adolescent Psychology, American Psychological Association and Professor and Vice Chair for Research, Department of Child and Adolescent Psychiatry, New York University School of Medicine
JENNIFER KAMINSKI, Acting Team Lead, Child Development Studies Team, Child Development and Disabilities Branch, Centers for Disease Control and Prevention
LAUREL K. LESLIE, Board of Directors, American Board of Pediatrics; Professor of Medicine and Pediatrics, Tufts University School of Medicine; and Director, Center for Aligning Researchers and Communities for Health, Tufts Clinical and Translational Science Institute
RUTH PEROU, Acting Mental Health Coordinator, Program Performance and Evaluation Office, Centers for Disease Control and Prevention
EVE E. REIDER, Health Scientist Administrator, Prevention Research Branch, National Institute on Drug Abuse
ANDY SHIH, Senior Vice President of Scientific Affairs, Autism Speaks
BELINDA E. SIMS, Health Scientist Administrator, Prevention Research Branch, National Institute on Drug Abuse

---

[1] Institute of Medicine planning committees are solely responsible for organizing the workshop, identifying topics, and choosing speakers. The responsibility for the published workshop summary rests with the workshop rapporteurs and the institution.

FORUM ON PROMOTING CHILDREN'S COGNITIVE, AFFECTIVE, AND BEHAVIORAL HEALTH[1]

WILLIAM R. BEARDSLEE (*Co-Chair*), Director, Baer Prevention Initiatives; Chairman Emeritus, Department of Psychiatry, Boston Children's Hospital; and Gardner/Monks Professor of Child Psychiatry, Harvard Medical School

C. HENDRICKS BROWN (*Co-Chair*), Professor, Departments of Psychiatry, Behavioral Sciences, and Preventive Medicine, Northwestern University

DARA BLACHMAN-DEMNER, Social Science Analyst, Crime, Violence, and Victimization Research Division, National Institute of Justice

THOMAS F. BOAT, Professor of Pediatrics and Dean Emeritus, University of Cincinnati College of Medicine, Cincinnati Children's Hospital Medical Center

FELESIA R. BOWEN, Assistant Professor and Director, Center for Urban Youth, Rutgers College of Nursing

DAVID A. BRENT, Academic Chief, Child and Adolescent Psychiatry, Western Psychiatric Institute and Clinic and Professor of Psychiatry, Pediatrics, and Epidemiology, School of Medicine, University of Pittsburgh

LAUREN CALDWELL, Director, Children, Youth, and Families Office and Public Interest Directorate, American Psychological Association

DAVID A. CHAMBERS, Associate Director, Dissemination and Implementation Research and Chief, Services Research and Clinical Epidemiology Branch, Division of Services and Intervention Research, National Institute of Mental Health

WILMA PETERMAN CROSS, Senior Public Health Advisor, Office of Disease Prevention, National Institutes of Health

COSTELLA GREEN, Branch Chief, Division of Community Programs, Center for Substance Abuse Prevention, Substance Abuse and Mental Health Services Administration

J. DAVID HAWKINS, Social Work Endowed Professor of Prevention, University of Washington School of Social Work

KIMBERLY E. HOAGWOOD, Society of Clinical Child and Adolescent Psychology, American Psychological Association and Professor and Vice Chair for Research, Department of Child and Adolescent Psychiatry, New York University School of Medicine

---

[1] Institute of Medicine forums and roundtables do not issue, review, or approve individual documents. The responsibility for the published workshop summary rests with the workshop rapporteurs and the institution.

COLLEEN HORTON, Policy Program Officer, Hogg Foundation for Mental Health, University of Texas, Austin
UMA KOTAGAL, Senior Vice President for Quality, Safety and Transformation, and Executive Director, James M. Anderson Center for Health Systems Excellence, Cincinnati Children's Hospital Medical Center
LAUREL K. LESLIE, Board of Directors, American Board of Pediatrics; Professor of Medicine and Pediatrics, Tufts University School of Medicine; and Director, Center for Aligning Researchers and Communities for Health, Tufts Clinical and Translational Science Institute
MARY ANN McCABE, Society for Child and Family Policy and Practice, Society of Pediatric Psychology, Associate Clinical Professor of Pediatrics, George Washington University School of Medicine and Affiliate Faculty in Psychology, George Mason University
JUSTIN MILNER, Senior Research Associate, Urban Institute
WILLIAM MODZELESKI, Senior Consultant, SIGMA Threat Management Associates
JENNIFER NG'ANDU, Program Officer, Robert Wood Johnson Foundation
JENNIFER OPPENHEIM, Public Health Advisor and Director, Project LAUNCH, Center for Mental Health Services, Substance Abuse and Mental Health Services Administration
LAWRENCE A. PALINKAS, Albert G. and Frances Lomas Feldman Professor of Social Policy and Health, and Director, Behavior, Health, and Society Research Cluster, University of Southern California School of Social Work
RUTH PEROU, Acting Mental Health Coordinator, Program Performance and Evaluation Office, Centers for Disease Control and Prevention
EVE E. REIDER, Health Scientist Administrator, Prevention Research Branch, National Institute on Drug Abuse
MARY JANE ROTHERAM-BORUS, Bat-Yaacov Professor of Child Psychiatry and Biobehavioral Sciences, Director, Global Center for Children and Families; and Director, Center for HIV Identification, Prevention & Treatment Services, Department of Psychiatry, University of California, Los Angeles
PAT SHEA, Deputy Director, Technical Assistance and Prevention, National Association of State Mental Health Program Directors
ANDY SHIH, Senior Vice President of Scientific Affairs, Autism Speaks
JOSÉ SZAPOCZNIK, Professor and Chair, Department of Public Health Sciences and Director, Miami Clinical Translational Science Institute and Center for Family Studies, University of Miami

VERA FRANCES TAIT, Associate Executive Director, and Director, Department of Child Health and Wellness, American Academy of Pediatrics
JENNIFER TYSON, Social Science Analyst, Office of Juvenile Justice and Delinquency Prevention, Department of Justice
DONALD WERTLIEB, Professor Emeritus, Eliot-Pearson Department of Child Development, Tufts University

Project Staff

MORGAN A. FORD, Senior Program Officer
WENDY KEENAN, Program Associate
TARA MAINERO, Associate Program Officer
NOAM I. KEREN, Research Associate
AMANDA PASCAVIS, Senior Program Assistant
ANTHONY JANIFER, Program Assistant

Board on Children, Youth, and Families Staff

PAMELLA ATAYI, Administrative Assistant
FAYE HILLMAN, Financial Associate
KIMBER BOGARD, Board Director

Reviewers

This workshop summary has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published workshop summary as sound as possible and to ensure that the workshop summary meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the process. We wish to thank the following individuals for their review of this workshop summary:

Denise Dougherty, Agency for Healthcare Research and Quality
Vetisha L. McClair, Centers for Medicare & Medicaid Services
Sarah Hudson Scholle, National Committee for Quality Assurance
Bradley D. Stein, RAND Corporation

Although the reviewers listed above have provided many constructive comments and suggestions, they did not see the final draft of the workshop summary before its release. The review of this workshop summary was overseen by **Hugh H. Tilson.** Appointed by the National Research Council and the Institute of Medicine, he was responsible for making certain that an independent examination of this workshop summary was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this workshop summary rests entirely with the rapporteurs and the institution.

Contents

ABBREVIATIONS AND ACRONYMS xvii

1 INTRODUCTION 1
Highlights from the Workshop, 3
Organization of the Report, 4
References, 4

2 MAXIMIZING THE VALUE OF NATIONAL, STATE, AND LOCAL MEASUREMENT SYSTEMS 5
Data on Children's Health and Well-Being in the State of Maryland, 5
Health Care Expenditures for Children with Disruptive Behavior Disorders in Medicaid/CHIP, 9
Multiple Service Systems Use Among Illinois Families, 11
Using Data to Improve Patient-Centered Outcomes: A Perspective from the Medicaid Medical Directors, 17
References, 20

3 MEASUREMENT SYSTEMS TO ASSESS INDIVIDUAL- AND POPULATION-LEVEL CHANGE 21
Evaluating Substance Abuse Prevention on a Large Scale Across State Populations, 21
Closing Research Data Gaps to Prevent Youth Suicides, 23

How Data Repositories Are Opening Access to Research Data on Autism and Mental Health Illnesses, 27
Using Data to Inform Decision Making in Maryland Public Schools, 29
Measuring Population-Level Progress in the Fight for Drug-Free Communities, 32
References, 35

4 **USING QUALITY MEASURES TO FACILITATE SYSTEM CHANGE** **37**
Ensuring High Quality in Measures of Child Care and Preschool: A Cautionary Tale, 37
The Juvenile Justice Reform and Reinvestment Initiative, 40
Can Child Mental Health Cross the Quality Chasm?, 43
Quality Measures and Improvement in Health Care, 47
References, 49

5 **TOWARD EFFICIENT AND SUSTAINABLE DELIVERY OF INTERVENTIONS** **51**
Measuring Implementation of Evidence-Based Prevention: Lessons from Communities That Care, 51
The Stages of Implementation Completion, 55
Implementing Programs with Fidelity, 57
Measuring Implementation Fidelity Using Computational Methods, 59
References, 61

6 **BREAKOUT GROUP DISCUSSIONS** **63**
Measurement at the Individual Level, 63
Measurement at the Community Level, 65
Measurement at the State and Systems Level, 66
Final Comments, 67

APPENDIXES

A WORKSHOP AGENDA 71
B BIOGRAPHIES OF WORKSHOP SPEAKERS 77

Figures and Table

FIGURES

2-1 The state of Maryland's indicators and results framework, 7
2-2 The Integrated Database on Child and Family Programs in Illinois, 12
2-3 Costs of services for multi-system families (MSFs), 14
2-4 Types of problems experienced by families, 15

3-1 Suicide rates among persons aged 15 to 19 years, 25
3-2 Emergency department self-inflicted injuries by age and sex, 25
3-3 Suicidal ideation and behavior among high school students, 26
3-4 The U.S. Department of Education's School Climate Model, 31
3-5 Framework for community change of the National Coalition Institute, 33

4-1 Matching risks in an evidence-based operating platform for juvenile justice (JJ), 41
4-2 All youth 12-month recidivism by matrix adherence level, 42
4-3 The "six P" conceptual framework, 45

5-1 The Communities That Care timeline, 53
5-2 The eight Stages of Implementation Completion, 56
5-3 Implementation fidelity over generations of trainers, 60

TABLE

2-1 Advantages and Disadvantages of Surveys and Administrative Data, 13

Abbreviations and Acronyms

ACA	Affordable Care Act
ADHD	attention deficit hyperactivity disorder
CADCA	Community Anti-Drug Coalitions of America
CDC	Centers for Disease Control and Prevention
CHIP	Children's Health Insurance Program
CLASS	Classroom Assessment Scoring System
CMMI	Center for Medicare & Medicaid Intervention
CMS	Centers for Medicare & Medicaid Services
CTC	Communities That Care
DBD	disruptive behavior disorder
DFC	Drug-Free Communities
ECERS	Early Childhood Environment Rating Scale
ED-STARS	Emergency Department Screen for Teens at Risk for Suicide
FERPA	Family Educational Rights and Privacy Act
FIMP	fidelity of implementation
GLS	Garrett Lee Smith
GUID	global unique identifier
HIPAA	Health Insurance Portability and Accountability Act

IOM	Institute of Medicine
JJRRI	Juvenile Justice Reform and Reinvestment Initiative
LGBTQ	lesbian, gay, bisexual, transgender, and queer
MAP	Measures Applications Partnership
MAX	Medicaid Analytic eXtract
MDS3	Maryland Safe and Supportive Schools
MSF	multi-system family
MTFC	Multidimensional Treatment Foster Care
NAASP	National Action Alliance for Suicide Prevention
NCHS	National Center for Health Statistics
NCQA	National Committee for Quality Assurance
NDAR	National Database for Autism Research
NIH	National Institutes of Health
NIMH	National Institute of Mental Health
NQF	National Quality Forum
NSF	National Science Foundation
NVDRS	National Violent Death Reporting System
PBIS	Positive Behavioral Interventions and Supports
PCORI	Patient-Centered Outcomes Research Institute
PMTO	Parent Management Training Oregon
SAMHSA	Substance Abuse and Mental Health Services Administration
SIC	Stages of Implementation Completion
SPEP	Standardized Program Evaluation Protocol
SPF-SIG	Strategic Prevention Framework State Incentive Grant
WISQARS	Web-Based Injury Statistics Query and Reporting System

1

Introduction

Many measurement systems to monitor the well-being of children and guide services are implemented across the community, state, and national levels in the United States. While great progress has been made in recent years in developing interventions that have been shown to improve the cognitive, affective, and behavioral health of children, many of these tested and effective interventions have yet to be widely implemented. One potential reason for this lag in implementation is a need to further develop and better utilize measures that gauge the success of evidence-based programs as part of a broad effort to prevent negative outcomes and foster children's health and well-being.

To address this issue, the Forum on Promoting Children's Cognitive, Affective, and Behavioral Health held a workshop in Washington, DC, on November 5–6, 2014, titled Innovations in Design and Utilization of Measurement Systems to Promote Children's Cognitive, Affective, and Behavioral Health. The forum was established as an outgrowth of the 2009 National Research Council and Institute of Medicine report *Preventing Mental, Emotional, and Behavioral Disorders Among Young People: Progress and Possibilities*, which called on the nation to make the prevention of such disorders and the promotion of mental health of young people a high priority (NRC and IOM, 2009). The forum engages in dialogue and discussion to connect the prevention, treatment, and implementation sciences with settings where children are seen and cared for, including health care, schools, social service and child welfare agencies, and the juvenile justice system, and to create systems that are effective and affordable in addressing children's needs.

The workshop on measurement systems was the third held by the forum. The first, on April 1–2, 2014, examined strategies for scaling tested and effective family-focused preventive interventions so they can have widespread economic and social benefits (IOM and NRC, 2014a). The second, on June 16–17, 2014, was on harvesting the scientific investment in prevention science to promote children's cognitive, affective, and behavioral health (IOM and NRC, 2014b). Both workshops pointed to the key role of measurement systems in guiding program implementation, evaluation, and adaptation and in building support for preventive interventions, which heightened the forum's interest in holding a workshop devoted specifically to measurement issues.

Many different measurement systems track the health and well-being of children in the United States. In part, these systems are designed to monitor and provide actionable feedback to improve the delivery of health care, social, and educational services for children at the local, regional, and national levels. They can provide information on such factors as:

- The selection of target communities;
- The targeting of risk and protective factors;
- Strategies to change health systems, policies, and practices; and
- Strategies to improve the overall delivery and quality of care.

However, many of these systems could be improved through such tools as data linkage and integration, quality improvement methods, and feedback mechanisms. For example, as forum co-chair Hendricks Brown emphasized in his introductory comments at the workshop, data generated and used in one system rarely cross over into other systems, so information from the health care system are rarely used in schools, and information from the juvenile justice or social welfare systems are rarely used in health.[1] In addition, significant gaps exist in the ability to measure factors that have a significant effect on children's health and well-being, Brown said. New measurement types and uses could provide innovative and efficient ways to assess, assemble, and integrate data to drive actions.

The workshop featured presentations on the use of data linkage and integration to inform research and practice related to children's cognitive, affective, and behavioral health; the use of quality measures to facilitate system change in health care, classroom, and juvenile justice settings; and tools developed to measure implementation of evidence-based prevention programs at scale to support sustainable program delivery, among other topics. Workshop presenters and participants discussed examples of inno-

---

[1] Please see Chapters 2, 4, and 6 for additional discussion of data usage and linkage across systems.

vative design and utilization of measurement systems, new approaches to build on existing data systems, and new data systems that could support the cognitive, affective, and behavioral health and well-being of children. As the other co-chair of the forum, William Beardslee, stated, the ability to improve the health of children through effective interventions "depends on the collection of data and the use of information for continuous quality improvement, so this is a crucially important forum."

HIGHLIGHTS FROM THE WORKSHOP

At the conclusion of both days of the workshop, the forum co-chairs, the moderators of the four panels that took place in the workshop (summarized in Chapters 2–5 of this report), and the leaders of the breakout discussions (summarized in Chapter 6) identified the major issues that emerged from the presentations and discussions. **This section draws on those comments to introduce broad topics that emerged from speakers' presentations at the workshop and should not be seen as the conclusions of the workshop as a whole. Each is relevant to the mission of the Forum on Promoting Children's Cognitive, Affective, and Behavioral Health:**

- Measurement of children's cognitive, affective, and behavioral health and well-being can be a powerful force for accountability and for the promotion of preventive interventions.
- Measurement can improve understanding of the factors that influence children, including the broad systems that both establish context and affect their lives directly.
- Large gaps still exist in the kinds of data that are being gathered and the ways information is being linked to help improve children's lives.
- A strong infrastructure can enable the collection, analysis, and dissemination of data that are accurate, timely, relevant, sustainable, and acceptable to providers and users of the information.
- Better and more accessible measures can contribute to better informed and more educated policy makers and members of the public, which can further the public health mission of detecting, responding to, and preventing threats to health.
- An increased emphasis on measures, better use of existing measures, and new kinds of measures could help encourage communities to adopt evidence-based programs.
- New technologies and platforms are enabling not just new and less expensive ways of gathering, linking, analyzing, and disseminating data but the collection and use of entirely new kinds of data.

- Data systems that are simultaneously embedded and transportable can grow to meet the needs that exist while providing useful information today.
- Communication and collaboration across sectors and across levels of government could greatly improve the development, implementation, and refinement of such measures.
- Short-term, medium-term, and long-term steps could be identified to make such improvements, which could help prioritize the actions of the public and private sectors.

ORGANIZATION OF THE REPORT

The broad topics identified by the forum co-chairs, the panel moderators, and the breakout group leaders emerged both from the presentations at the workshop and from subsequent discussions. In this summary report, important points raised by workshop members during discussion sessions have been incorporated into the summaries of speakers' presentations.

Chapter 2 describes several examples of the successful use of data integration to inform research and policy pertinent to children's cognitive, affective, and behavioral health. It also discusses how these and other systems might be improved. Chapter 3 looks at innovations in the use of measurement systems to monitor changes in children's health at the individual and population levels. Chapter 4 considers examples of how quality measures have been used to facilitate system change in early childhood, juvenile justice, and health care settings. Chapter 5 discusses tools developed to measure the implementation of evidence-based preventive programs at scale and offers suggestions about how such tools can be used to improve the sustainable delivery of programs. Finally, Chapter 6 summarizes the observations of the breakout groups and the comments of workshop participants on the major topics discussed at the workshop.

REFERENCES

IOM (Institute of Medicine) and NRC (National Research Council). 2014a. *Strategies for scaling effective family-focused preventive interventions to promote children's cognitive, affective, and behavioral health: Workshop summary*. Washington, DC: The National Academies Press.

IOM and NRC. 2014b. *Harvesting the scientific investment in prevention science to promote children's cognitive, affective, and behavioral health: Workshop summary*. Washington, DC: The National Academies Press.

NRC and IOM. 2009. *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. Washington, DC: The National Academies Press.

2

Maximizing the Value of National, State, and Local Measurement Systems

The workshop began with four presentations that featured examples of the successful use of data integration to inform research and policy pertinent to children's cognitive, affective, and behavioral health. These examples extended from the local level to the state level to the national level, but all had several features in common. They relied on data gathered, combined, and disseminated in unusual ways to better understand the problems facing children and ways of overcoming those problems.

DATA ON CHILDREN'S HEALTH AND WELL-BEING IN THE STATE OF MARYLAND

Maryland is a state of about 5.8 million people, explained Anne Sheridan, executive director of the Maryland Governor's Office for Children. Of the 857,000 children in the state's K–12 public education system, 43 percent are eligible for free and reduced price meals. With two large urban areas—Baltimore and the suburbs of Washington, DC—the state is rich demographically, with about 55 percent whites, 29 percent African Americans, 8 percent Latinos, and 5 percent Asians. The state also has a strong county government system, with 23 counties plus the City of Baltimore, "so everything that you do, you need to automatically multiply by 24," said Sheridan. The state's economy is grounded in its academic and educational institutions, its proximity to Washington, DC, its medical institutions, and its thriving hospitality and tourism industry—which Sheridan referred to as "eds, beds, feds, and meds."

The state now has more jobs than before the recession that began in 2008, and it has a relatively high median income. But it also has persistent pockets of poverty and challenges for many families, which has a great impact on children's health and well-being.

Maryland has a Governor's Office for Children and a Children's Cabinet consisting of the heads of the departments responsible for child welfare, education, health, disabilities, juvenile services, and the budget. Moreover, unlike children's cabinets in other states, the one in Maryland has a dedicated line item in the state budget devoted to early intervention and prevention programming in each of its jurisdictions. The cabinet functions as a grant maker. It asks jurisdictions to think about the highest needs in their areas, and funds are distributed to a wide variety of programs according to those needs. Historically, it has focused on children with the highest needs, who are also at risk of involvement in multiple state systems. "Not surprisingly," said Sheridan, these are "the children and families who also cost the state the most money."

The state uses two major data structures to measure improvements in children's health and well-being. The first is a results and indicators framework that has been used for many years to guide Children's Cabinet programming (see Figure 2-1). This approach focuses on planning, decision making, and budgeting around desired results or outcomes. It has 25 indicators that are associated with particular results, which enable data to be gathered that can be used to determine what is driving particular outcomes. The framework, which is publicly available, establishes shared goals, aligns grant making, and makes children a priority. The system was "a game changer when it was introduced 14 years ago," said Sheridan. It also is fairly simple and accessible. "We could train you up on it in an afternoon and feel like you'd have the basic principles." It has limitations in being produced only on an annual basis, not being a powerful tool for day-to-day management, and not being comprehensive. But "even with those limitations, it's proved to be valuable over time."

The second data structure is a statewide process called State Stat that was introduced by Governor Martin O'Malley after he was elected in 2006. Based on a similar process developed in New York City to track crime, State Stat sets clear expectations about statewide goals. Four areas fall under the purview of the Children's Cabinet:

- End childhood hunger in Maryland by 2015.
- Improve student achievement and school, college, and career readiness by 25 percent by the end of 2015.
- Reduce violent crimes committed against women and children by 25 percent by the end of 2018.
- Reduce infant mortality in Maryland by 10 percent by 2017.

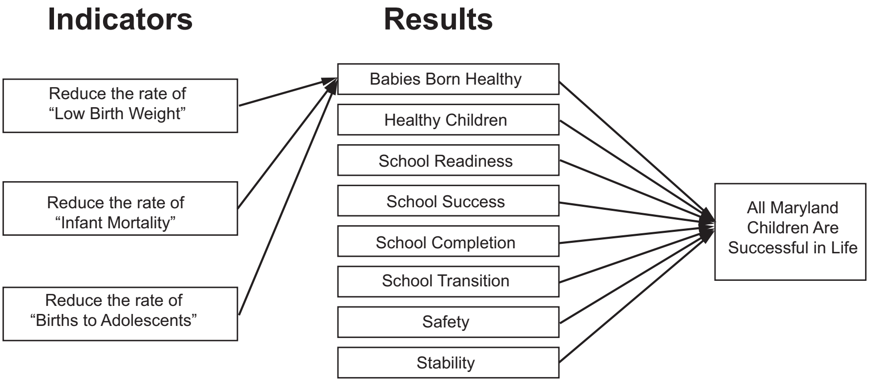

FIGURE 2-1 The state of Maryland's indicators and results framework. The framework has indicators associated with a variety of positive outcomes.
SOURCE: Sheridan, 2014, adapted from Children's Cabinet and Governor's Office for Children, 2012.

The State Stat process is a rigorous, data-driven procedure in which the state officials responsible for these goals convene, review the information, and take a "deep dive" into the things that are contributing to or detracting from success, said Sheridan. The process is also designed to foster collaboration and transparency so problems can be identified and solved.

As an example of how information is used, Sheridan discussed the children's hunger goal. Maryland has had a number of long-running child nutrition programs, but for a variety of reasons the programs remained underutilized. For example, every child who receives a free and reduced-price lunch should get a school breakfast, which contributes greatly to their ability to pay attention, to do well in school, and to remain healthy. To increase the use of the school breakfast program, the state convened a number of players at the state level, foundations, and corporate partners, and this group together mounted an outreach program that has encouraged schools to make breakfast part of the school day. From 2007 to 2008 and from to 2014 to 2015, the number of children eligible for a free or reduced-price lunch who ate a free or reduced-price breakfast at school every day rose from 44.6 percent to 60.9 percent. About 400 schools have introduced some sort of in-classroom grab-and-go breakfast, providing greater flexibility for students to eat breakfast, and about 75,000 more children are getting free and reduced-price breakfasts than was the case before.

A major obstacle in this effort, Sheridan said, was getting the county-wide and school-level data about breakfasts from the Maryland State

Department of Education. To some extent, the needed data were not being collected, and to some extent attitudes limited the release of data, noted Sheridan. Improvements on both fronts have resulted in school-level data mapped across the state. “You can go to the State Stat website, if you’re interested in how we’re doing on this goal, and pull your school down,” said Sheridan. “You can see how it’s doing, [which] is also a factor in communicating success, attracting more supporters for this enterprise, and hopefully showing what were the longer-term impacts of us doing a better job connecting kids to school breakfast.”

Analysis of the data showed that an increase in school breakfasts was linked to an improvement in test scores and attendance, with attendance linked to longer-term success. “We are able to make pretty definitive statements about the value proposition for making these investments,” said Sheridan. The result has been a “virtuous cycle of improvement, progress, and success . . . that’s really helped us.”

The ability to engage outside partners was “a big part of what drove the success of the school breakfast campaign,” said Sheridan. “Every big issue that the Governor’s Office for Children has worked on, its progress has been made possible by having [a] national partner. On our hunger goal, it’s been Share Our Strength. On our out of school time, it’s been largely the Charles Stewart Mott Foundation, which is building a statewide network.”

States will continue to be innovators and incubators of successful programs in the future, said Sheridan. “I don’t think the United States Congress is going to be offering a lot on that front for a number of years.” A major challenge at the state level is to make it safer and more attractive to invest in prevention. “It costs less to stay at the Four Seasons every night than it does for us to place you in a hardware-secured juvenile detention facility, but I am having a hard time getting folks to say that $2,000 a year for high-quality afterschool programming is a good investment. Need help here.” Policy makers are deluged with data, some of which is helpful and some of which is not. Data systems that are sustainable across administrations, so they do not change when a new administration is elected, can help align budgeting, messaging, and outcomes, Sheridan concluded. Also, she noted, “The difference between a dream and a goal is a deadline.”

“To improve outcomes for kids is an investment, and you have to look at it that way,” said Sheridan. “The more we can make data-driven arguments that *X* investment here leads to this outcome, . . . the better off we’re going to be.” At the local levels of government, people especially care about their own neighborhoods and communities, and this drive can be carried through to more broadly based levels of government as well.

HEALTH CARE EXPENDITURES FOR CHILDREN WITH DISRUPTIVE BEHAVIOR DISORDERS IN MEDICAID/CHIP

The Center for Medicare & Medicaid Innovation (CMMI) within the Centers for Medicare & Medicaid Services (CMS) conducts many of the innovative projects undertaken by CMS throughout the United States. Vetisha McClair, a research analyst at CMMI, described two intramural research projects conducted by members of this team as examples of the kinds of data on children's health and well-being that are available and how they can be used.

Medicaid and the Children's Health Insurance Program (CHIP) are together the largest source of health care for children in the country. Medicaid, which covers primarily low-income children, is a partnership between the federal government and states, and the benefit packages differ from state to state. CHIP is available in all states and it covers children from higher income families who are not eligible for care under their parent's private insurance or whose parents lack insurance.

Few studies have examined the expenditures for children on Medicaid/CHIP with disruptive behavior disorders (DBDs), said McClair. DBDs consists largely of conduct disorder, which is characterized by disregard for societal norms and the rights of others; oppositional defiant disorder, which is characterized by persistent patterns of anger and breaking rules; and are often comorbid with attention deficit hyperactivity disorder (ADHD). These conditions can become a large burden on both familial and health care resources and can significantly limit the educational experiences of children. Also, some children diagnosed with DBDs experience impairment that persists into adulthood, including the development of antisocial personality disorder.

A group of researchers within CMMI conducted a study, using health care claims data, to assess disease prevalence and expenditures for the treatment for DBDs among Medicaid/CHIP beneficiaries. The data came from Medicaid Analytic eXtract (MAX), which is a set of person-level and claim-level data files compiled annually that contain Medicaid eligibility, service utilization, and program payment information derived from state reporting. Eligibility files were combined with health care claims from in-patient, long-term care, prescription drug treatment, and other claims to develop an extensive dataset of information about the health care that children receive throughout the United States.

Medicaid claims data are very useful in examining national trends and intra-state health care policies. However, comparisons between individual states are not generally recommended, said McClair, due to differences in benefits and data quality. Individual states determine the benefit packages and services covered, therefore one cannot be ensured of uniformity in services between individual states.

The study was able to extract and examine claims data for 46 states and the District of Columbia, though the California and New York analyses were still under way at the time of the workshop because of the difficulties of analyzing such large samples. In addition, Maine and Alabama were excluded because of data inconsistencies. For children and teens under 20 with at least one diagnosis in a health service claim of conduct disorder, oppositional defiant disorder, and/or ADHD, outcomes were analyzed by sex, age, race and ethnicity, state, and delivery type (fee for service or managed care). Outcome measures included both prevalence in the study sample and expenditures.

The study found that the prevalence of DBDs among Medicaid beneficiaries increased slightly over the period studied, from 4.7 percent in 2006 to 5.3 percent in 2009. Within racial and ethnic groups, the prevalence of DBDs remained relatively stable, with whites and Hispanics/Latinos having a substantially higher prevalence than blacks (McClair et al., 2014). Over the 4-year period, the number of children eligible for Medicaid/CHIP increased from 24.8 million to 27.7 million. As a result, the total Medicaid payment for beneficiaries with a DBD increased. However, the total Medicaid payment for DBD treatment remained more or less flat, and the proportion of expenditures devoted to mental health among all services addressing mental health issues decreased slightly over this period, McClair said.

She also described a related study on the use of prescription drugs among Medicaid beneficiaries. Individuals with a DBD have high rates of prescription drug use, and the number of these individuals who use prescription drugs went up 25.6 percent between 2006 and 2009. The total number of prescriptions also increased by 27.2 percent from 2006 to 2009, and the Medicaid payments increased by 28.8 percent, with the payments per beneficiary staying largely the same (Zhao et al., 2014). A large portion of drugs prescribed to children with DBDs were psychotherapeutic drugs, and this proportion has remained relatively stable for all beneficiaries and within racial subgroups.

McClair drew several conclusions from these data. First, these results could be explained by a number of exogenous factors, including delivery system reform and state-level changes to Medicaid/CHIP benefits.

Second, further research is needed to address not only the costs associated with the treatment of children with DBDs but how expenditures may vary by state and how those monies are allocated.

Finally, Medicaid data are a viable source of information about the state of health care for vulnerable populations, such as low-income children. Understanding the most current information available on medical expenditures for children with these conditions is crucial to instigating changes in the health care system that promote increased access to care while containing costs, she said. "We need to continue research in this area and continue to use Medicaid data for both research and policy purposes."

In response to a question, McClair observed that Medicare has always attracted much of the attention within CMS, resulting in less focus on children's issues. But the Innovation Center has been tasked to look at specific populations and problems that drive costs, which has laid the groundwork for policy changes and new programs. For example, one potential area of focus in the future could include how CHIP may cover a child in a family but not cover a parent who has a substance abuse problem that is affecting the child. Innovative cross-agency programs could address problems within families in a more coordinated way, she said.

MULTIPLE SERVICE SYSTEMS USE AMONG ILLINOIS FAMILIES

Chapin Hall, which was an organization created to take care of orphans of the Civil War, was brought to the University of Chicago 30 years ago to have an impact on the well-being of children through better research and analysis. It has particularly sought to achieve that goal by building relationships with policy makers and by conducting research that meets the needs of the public sector. Known for its work with government social program administrative data, it employs a full range of research methods and addresses many issues relevant to children, families, and their neighborhoods.

Robert Goerge, senior research fellow at Chapin Hall, pointed out that subgroups of families are at highest risk of their children experiencing adverse childhood experiences. These families tend to be characterized by a combination of attributes:

- Chronically unemployed parents
- Very low socioeconomic status
- Long-term welfare program participants
- Single-parent families
- Mothers who had their first child as an adolescent
- Families without grandparents living with them
- Alcohol, drug use, mental illness
- Inadequate parenting skills, family breakdown, parental stress and mental illness, domestic violence, and parental history of maltreatment during childhood
- Community-level low socioeconomic status, living in an impoverished community, family size, and sibling spacing

To learn more about these families, Chapin Hall has created an Integrated Database on Child and Family Programs in Illinois that combines data from law enforcement, foster care and child maltreatment investiga-

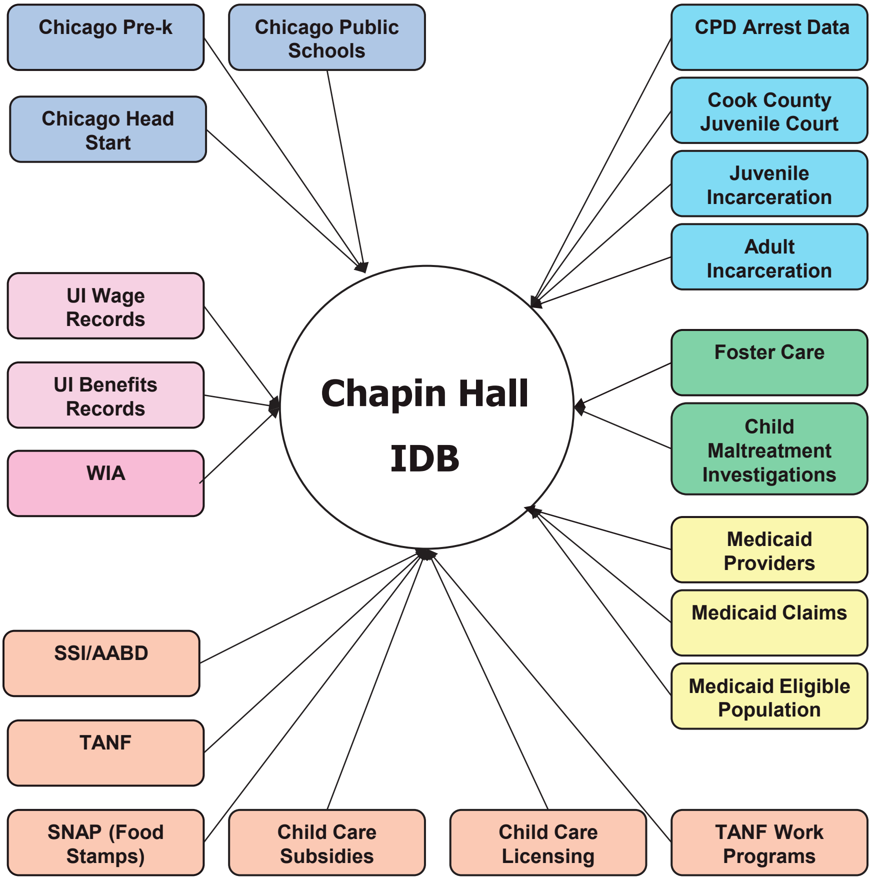

FIGURE 2-2 The Integrated Database on Child and Family Programs in Illinois. The database combines data from a number of social service agencies.
NOTE: AABD = Aid to the Aged, Blind, and Disabled; CPD = Chicago Police Department; IDB = Integrated Database; SNAP = Supplemental Nutrition Assistance Program; SSI = Social Security Income; TANF = Temporary Assistance for Needy Families; UI = Unemployment Insurance; WIA = Workforce Investment Act.
SOURCE: Goerge, 2014.

tions, Medicaid programs, welfare programs, employment records, and education (see Figure 2-2). The data span the period from 1977 to 2014, but most are from 1990 forward.

The process of making administrative data suitable for research use includes three important steps[1]:

---

[1] For discussion on privacy and use of personal data, see Chapter 6.

- De-duplication, or identifying records within a dataset that belong to the same individual;
- Record linkage, to identify and link individuals across datasets, which may require probabilistic record linkage; and
- Identifying relationships among individuals.

Administrative data have a variety of advantages and disadvantages compared with survey data, Goerge pointed out (see Table 2-1).

The data collected by Chapin Hall demonstrate that a small number of families use a large portion of the state's highest-cost services, including expensive and intensive health, human service, and corrections programs. Families were defined as individuals who were "linked" through membership in the Supplemental Nutritional Assistance Program or Temporary

TABLE 2-1 Advantages and Disadvantages of Surveys and Administrative Data

	Advantages	Disadvantages
Surveys based on data collection: sample surveys and censuses	• Can choose which questions to ask across multiple domains • Can be up to date (depending on how big of an effort it is)	Some respondents … … do not understand the question … have forgotten how it was … do not repond (nonresponse) … respond carelessly Burden on respondents can be high Expensive Low quality for estimates for small study domains (for sample surveys)
Register-based Surveys (administrative data and other non-traditional datasets)	• No further burden on the respondent for the statistics • Low costs • Almost complete coverage of population • Complete coverage of time • Respondents answer carefully to important administrative questions • Good possibilities for reporting for small areas, regional statistics, and longitudinal studies • Link records across datasets to take advantage of the relatively small amount of high-quality data in each data source	• Cannot ask questions • Dependent on the administrative system's population, object, and variable definitions • The reporting of administrative data can be slow; the time between the reference period and when data are available for statistical purposes can be long • Changes in the administrative systems make comparisons difficult • Variables that are less important for administrative work can be of lower quality

SOURCE: Goerge, 2014, adapted from Wallgren and Wallgren, 2007.

Assistance for Needy Families cases, or were involved in a Department of Children and Family Services case; problems were defined as an instance of mental health service, substance abuse treatment, adult incarceration, juvenile incarceration, or foster care placement. Multisystem families whose members had at least two different types of problems, which represent 23 percent of families, account for 63 percent of system involvement and 86 percent of costs of foster care, Medicaid, and adult and juvenile incarceration (see Figure 2-3). If the state could develop a deeper understanding of the circumstances in which these families live, their histories, their geographic location, and their trajectories, along with what services they have used, it could provide more adequate and efficient services, Goerge observed.

Analysis of the data also revealed the percentage of multisystem families by county across the state, with the southern part of the state having the highest prevalence of multisystem families. Within the city of Chicago, some of the 77 Chicago area communities that sociologists have been studying for many years had up to 40 percent of the study population in multisystem families. At an even more detailed geographic level, some census tracts have more than 60 percent of children living in multisystem families, with some tracts above 90 percent. Furthermore, the percentage of children living in multisystem families is tightly correlated with education failure, Goerge noted.

The most common problem experienced by families was mental health service, accounting for about 221,000 problems among the total of 465,000

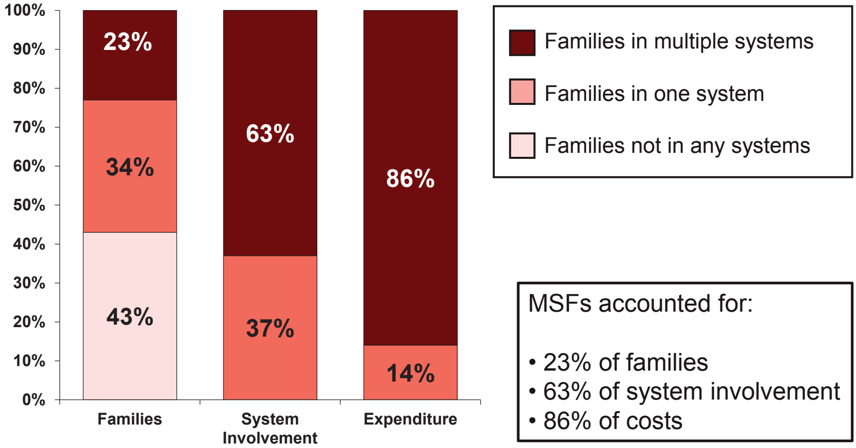

FIGURE 2-3 Costs of services for multi-system families (MSFs). Less than one-quarter of the families account for 86 percent of the costs of services.
SOURCE: Goerge, 2014.

problems experienced (see Figure 2-4). Among families with multiple problems, 94 percent had received mental health services. More than half of these families (54 percent) received both in-patient and out-patient services, which surprised the researchers, who expected more of the services to be out-patient. For substance abuse treatment, 42 percent of those who received any type of services were receiving in-patient services. "For about half of these families, it's a very intensive and costly service," noted Goerge.

The biggest surprise involved child abuse or neglect and intentional injury. Of the multisystem families, 81 percent had reported child abuse or neglect or a report of an intentional injury from the Medicaid claims data. "You wonder where the violence in Chicago comes from," said Goerge. "Its source is these children experiencing tremendous amounts of violence in their homes. That's why they're on the streets."

Future research directions for multisystem families include identifying unique characteristics of such families in specific places, discovering the problems that may have the greatest impact on the current state of the family, revealing whether single individuals account for multiple problems within a

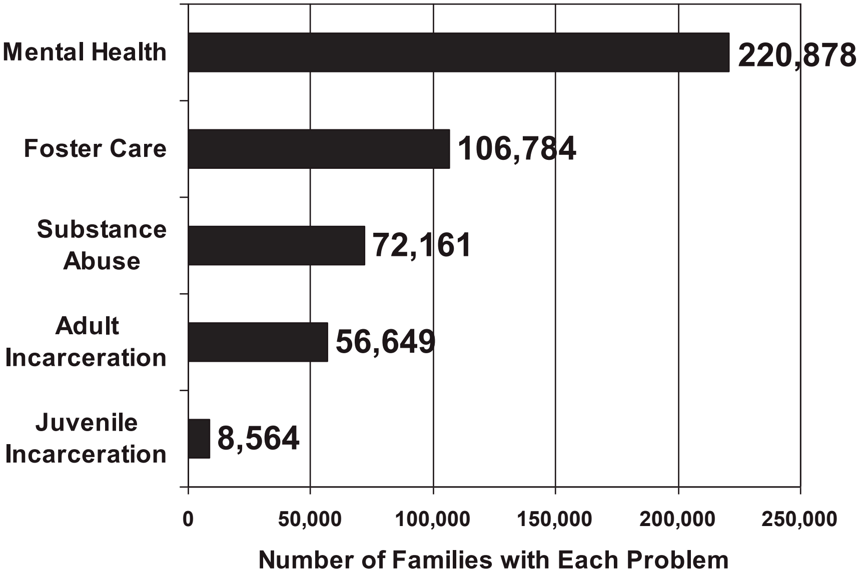

FIGURE 2-4 Types of problems experienced by families. Mental health was the most common problem for families.
NOTE: The total number of problems (465,036) exceeds the total number of families with problems (285,722) because some families have multiple problems.
SOURCE: Goerge, 2014.

family or whether several family members encounter problems, delineating between families that have had few service spells versus families with multiple service spells, estimating the cost of providing state services to multiple-problem families, looking at asthma and other chronic conditions, examining assets such as employment and education, and conducting longitudinal analysis to determine when these families became multisystem families.

Turning from this example to broader issues, Goerge observed that states and cities are developing their administrative data sources faster than ever, and they are using the data for many different purposes. They also are making the data public, enabling data entrepreneurs to create apps that inform the public and policy makers. In addition, a number of federal initiatives are promoting the development of administrative data.

However, this is happening to different degrees in different cities, counties, and states, and there is a wide variation in who has access to the data and in the quality of the data. Best practices have not been disseminated to a sufficient extent, Goerge said, and state and cities are taking years to develop these efforts. States often rely on large corporate vendors, and these vendors typically go only so far—for example, they generally do not give states the ability to track individuals across systems. Meanwhile, government agencies do not have the skilled staff necessary to take full advantage of the efforts. In addition, state bureaucracies often want to hold on to their data, said Goerge. “They are more likely to share their data with us than with each other.”

Special interests promote the perspective that problems can be addressed one by one, Goerge observed, but everyone knows that early nutrition and good parenting is associated with learning, learning is associated with getting a job, a parent having a job is associated with child well-being, and lack of school success is associated with criminal behavior. This is why breaking down silos is necessary to make policy and development programs to improve the well-being of individuals and families, he said. Goerge reiterated a point made by McClair—that no one agency or group is responsible for all of the problems facing individual families. “Our policy in the United States is not family focused. Our policy in the United States is we do something when a child or adult is ready to hurt somebody else or hurt themselves.”

Thirty years ago, when there were less data, most public-sector agencies had handfuls of analysts, but now they are fortunate to have one. At the same time, the federal government is requiring evidence-based practice and evaluation in many areas of social programs, which is a major challenge, given the lack of research expertise in these agencies. Importantly, data-sharing agreements are becoming more complicated as identity theft becomes more prevalent, as federal regulations proliferate, and to the extent that sometimes leaders and legal teams can view information as power and perceive that data sharing could potentially lead to negative media. Goerge noted that much change happens incrementally through negotiations between government, not-for-profit agencies, and health care providers over the contracting and budgeting processes. The more that people have information about what is happening to children and families, the more this behind-the-scenes work yields progress. Contracts are the easiest way to work with government, even though universities are generally concerned that they limit academic freedom. Furthermore, it is hard to justify using state dollars to fund evaluations, so the few evaluations that are done tend to be federally funded.

The greatest need, Goerge concluded, is to better connect academics and government. Government currently does not have enough human capital to link to the researchers who can help. In addition, barriers keep data from flowing to those who need the information. The nature of these barriers varies from sector to sector and place to place, but there are common themes. Incentives have to be put into place for all jurisdictions to use their data to get smarter about what they are doing, Goerge emphasized, and all federal research projects should be reviewed so that they are effectively using administrative data.

USING DATA TO IMPROVE PATIENT-CENTERED OUTCOMES: A PERSPECTIVE FROM THE MEDICAID MEDICAL DIRECTORS

The Medicaid Medical Directors Network, which started as a knowledge transfer project of the Agency for Healthcare Research and Quality, has been meeting since 2005 and is now part of the National Association of Medicaid Directors. Situated at the intersection of practice, policy, and payment, it engages in collaborative work to achieve its vision of "improving the lives of Medicaid patients through excellent health stewardship."

In the keynote address of the workshop, the director of the network, Jeff Schiff, who is medical director for the Minnesota Health Care Programs at the Minnesota Department of Health Services, described three projects of national significance that the network has undertaken. The first was on the use of psychotropic medicine among children. In 2007, pooled results of state analyses indicate that far more children in the foster care population received a psychotropic prescription than in the general population.[2] This finding has led to state policies and programs directed toward the use of such medications among this population, Schiff said.

The second project was on hospital readmissions. Trudnak et al. (2014) found that 9.4 percent of all hospital admissions among the Medicaid population in 19 states were 30-day readmissions. They also found a 10.9 percent mental and behavioral health readmission rate among children ages

---

[2] More information is available at http://rci.rutgers.edu/~cseap/MMDLNAPKIDS.html (accessed March 25, 2015).

12 and younger, and a 11.7 percent rate among those ages 13 to 20. “In Medicaid, that’s obviously real dollars,” said Schiff.

The third project, which he described in greater detail, involved early elective delivery. Slightly less than half of all births in America are paid for by the Medicaid program, and early elective deliveries, which have been associated with various negative cognitive outcomes later in life, account for nearly 9 percent of births paid for by Medicaid (Fowler et al., 2014). Early elective deliveries have become a concern for the American College of Obstetrics and Gynecology, the Collaborative Improvement and Innovation Network of the Maternal and Child Health Bureau, and the National Center for Health Statistics of the Centers for Disease Control and Prevention (CDC). The Medicaid Medical Directors Network set about getting other organizations involved in the issue, first to improve birth outcomes and second, as a proof of concept that this kind of quality work can be done as a state-led national project.

The research process was to agree on research questions based on input from the state Medicaid medical directors and staff, to define data parameters with the state data staff, to set up data submissions, to analyze the resulting data, and to report the results and make the link to policy. One interesting thing about this process is that the states are motivated by comparing their results with other states, Schiff noted. Similar to quality improvement processes at the provide level, reporting confidentially to states offers opportunities for state level discussions and subsequent policy actions. Additionally, by using an existing data stream for improvement creates an atmosphere where states strive to improve the quality of their data, Schiff added. Though the results of the study were embargoed at the time of the workshop, Schiff noted that the research did find significant numbers of early elective deliveries.

Schiff drew several conclusions from the studies he described:

- Sharing information so that states can compare their results to the national aggregate tends to spur quality improvements at the state level.
- Involvement in the research at the state and local levels creates ownership of the data and quality improvement.
- Measurement and quality improvement occur best when the results are relevant to families and frontline providers.
- State-level quality improvement can be linked to state policies.
- Data and quality improve together. “The data will never be perfect,” said Schiff. “We have to make decisions based on the data we have.”
- Resources and a resourced infrastructure enable such projects to be undertaken.

Schiff also pointed to several criteria for the development of measures:

- Importance
- Scientific soundness, including clinical soundness (process to outcome relationship) and measurement soundness (reliability and validity)
- Usability
- Feasibility of data collection

In addition to these criteria, he emphasized the importance of involving the people who are getting measured. Such involvement can avoid "measurement fatigue" that is felt by many providers and can help measures be adopted. Schiff also noted that the involvement of data experts can bring energy, excitement, and expertise to a project.

Data linkage can be a challenge with Medicaid and vital records data. Medicaid records include claims, service data, and data that come in through managed care. They also include enrollment data on such factors as race and level of poverty. Timeliness can be an issue, because old information is not always actionable, and data are not always of high quality. Deciding on identifiers, matching records, establishing common definitions, and linking data involve a variety of processes and actors, and all of these issues need to be addressed, said Schiff.

Data linkage also provides an opportunity to improve data, Schiff pointed out. For example, linkages could make it possible to incorporate patient-reported outcomes into the data. Electronic health records, even knowing one small piece such as body mass index is available, would help with monitoring data.

In conclusion, Schiff listed several possible topics for the forum to consider:

- Foster care and educational outcomes
- Autism treatment and outcomes registry
- ADHD diagnosis and educational outcomes
- Psychotropic medications and race/ethnicity
- Behavioral health screening and mental health services
- Social complexity and improved well-being

All of these issues are vertically relevant at all levels of the system. But pursuing any such set of issues will require an infrastructure for project development and process steps, the time of state data and policy staff members, and integrated support, including network maintenance.

REFERENCES

Children's Cabinet and Governor's Office for Children. 2012. *Maryland's results for child well-being*. Baltimore, MD: Maryland Children's Cabinet, Governor's Office for Children.

Fowler, T. T., J. Schiff, M. S. Applegate, K. Griffith, and G. L. Fairbrother. 2014. Early elective deliveries accounted for nearly 9 percent of births paid for by Medicaid. *Health Affairs (Millwood)* 33(12):2170-2178.

Goerge, R. 2014. *Multiple service systems use among Illinois families*. Presented at Institute of Medicine and National Research Council Workshop on Innovations in Design and Utilization of Measurement Systems to Promote Children's Cognitive, Affective, and Behavioral Health, Washington, DC.

McClair, V.L., C. Cross-Barnet, S. Chan, W.D. Clark, and L. Zhao. August, 2014. *Healthcare Expenditures for Children with Disruptive Behavior Disorders in Medicaid/CHIP*. Poster presented at the Annual Convention of the American Psychological Association, Washington, DC.

Medicaid Medical Directors Learning Network (MMDLN) and Rutgers Center for Education and Research on Mental Health Therapeutics. 2010. *Antipsychotic Medication Use in Medicaid Children and Adolescents: Report and Resource Guide from a 16-State Study*. MMDLN/Rutgers CERTs. Publication #1. Distributed by Rutgers CERTs at http://rci.rutgers.edu/~cseap/MMDLNAPKIDS.html (accessed March 25, 2015).

Sheridan, A. 2014. *Maximizing the value of national, state, and local measurement systems for research and policy*. Presented at Institute of Medicine and National Research Council Workshop on Innovations in Design and Utilization of Measurement Systems to Promote Children's Cognitive, Affective, and Behavioral Health, Washington, DC.

Trudnak, T., D. Kelley, J. Zerzan, K. Griffith, H. J. Jiang, and G. L. Fairbrother. 2014. Medicaid admissions and readmissions: Understanding the prevalence, payment, and most common diagnoses. *Health Affairs (Millwood)* 33(8):1337-1344.

Wallgren, A., and B. Wallgren. 2007. *Register-based statistics: Administrative data for statistical purposes*. Chichester, UK: Wiley.

Zhao, L., V.L. McClair, C. Cross-Barnet, S. Chan, and W. D. Clark. June, 2014. *Disruptive Behavior Disorders in Medicaid-enrolled Children: Disease Prevalence and Medication Expenditures*. Poster presented at AcademyHealth's Annual Research Meeting, San Diego, CA.

3

Measurement Systems to Assess Individual- and Population-Level Change

Many data systems already exist that can be used to monitor changes in children's health and well-being. Experience with these systems has revealed how innovations can improve health at the individual and population levels. Five presenters described examples of these systems and pointed toward how they could be replicated and expanded.

EVALUATING SUBSTANCE ABUSE PREVENTION ON A LARGE SCALE ACROSS STATE POPULATIONS

Robert Orwin, a senior study director in the Behavioral Health Group at Westat, led off the session by describing results from a national public health initiative to counter substance abuse: the Strategic Prevention Framework State Incentive Grant (SPF-SIG), an ongoing program of the Substance Abuse and Mental Health Services Administration (SAMHSA). Its goals are to prevent the onset of substance abuse problems and to reduce their progression; to reduce substance abuse-related problems and their consequences in communities, such as alcohol-related motor vehicle incidents; and to build capacity and infrastructure for prevention work at state and local levels. Orwin and his colleagues conducted an evaluation of the first two cohorts of SPF with funding from the National Institute on Drug Abuse in partnership with SAMHSA (Orwin et al., 2014). The study, which linked state, community, and intervention factors with population changes relating to substance abuse, is a major advance because of its use of large-scale population data, he said.

The SPF model features an interative process involving:

- Assessment of assets and needs at the local and state levels;
- Building, assessing, and increasing capacity;
- Strategic planning driven by data analysis;
- Funding of communities to implement and carry out intervention programs at the local level based on the assessment, capacity building, and planning process;
- Ongoing evaluation throughout the process, resulting in additional assessment and capacity building; and
- Sustainability and cultural competence that are woven throughout all steps of the program.

The SPF program differs from previous federal prevention efforts in two novel ways that are related to the workshop theme, Orwin said. One is that the initiative requires using population-based information to set state priorities and justify how resources are allocated to communities. This approach, known as "data-driven planning," had not yet been applied in a major way to substance abuse prevention at the time SPF began in 2004, he said. A second distinctive aspect is that SPF measures its effectiveness by relying on population-based outcomes, unlike the traditional approach of examining effects on individuals at the program level.

Twenty-four states and 2 territories participated in cohorts I and II of SPF, which funded 450 communities that initiated 2,534 interventions with goals such as reducing alcohol and marijuana use, underage drinking, binge drinking, and driving after drinking. Orwin summarized the results as generally being "very positive." For example, out of 174 communities that targeted 30-day alcohol use, 132 showed improvement; 79 performed significantly better, while 15 did significantly worse. Similarly, efforts to reduce driving after drinking in high school students led to improvements in 56 out of 78 communities (Diana et al., 2014). Aggregating those community-level findings to look at statewide effects revealed even stronger outcomes, he said.

The time period of the analysis—around 2006 to 2012—coincided with a secular downward trend in substance abuse across the United States. However, some states in the SPF project collected data from communities that did not offer intervention programs, which allowed for a comparison against the nationwide shift. "Even when the secular trend was taken into account, the results were generally quite impressive over a large scale," Orwin said. About two-thirds of the 450 communities and states targeting substance abuse improved relative to their comparison communities. Seven states that ran intervention programs reaching more than 50 percent of their populations achieved improvements relative to the national downward trend on most outcome measures (Diana et al., 2014).

The researchers examined which predictors at the state, community, and intervention levels might explain success in reducing substance abuse and its consequences. Generally, state-level factors relating to implementation, infrastructure development, and population were not all that predictive of the variations in outcome performance. “The communities were really where the action was,” he said, referring to community-level factors such as funding and organizational support, coalition capacity, SPF step scores, and intervention variables.

Communities that used their SPF grants to leverage additional prevention funding from other sources were more likely to achieve significant reductions in substance abuse measures. However, the results depended on the sources of that extra money. Block grants and community or municipal funds appeared to predict significant favorable changes in outcomes, whereas financing from foundations, corporations, or private donors did not. With few exceptions, factors relating to organizational support, such as state-provided technical assistance, had no effect on the variations in outcomes.

Community partners that were well-structured coalitions with good processes in place, paid leadership, membership diversity, and supportive communities achieved greater reductions in underage drinking outcomes. As far as intervention variables, an interesting question was which kinds of prevention strategies can promote population-level changes. Whereas traditional programming approaches try to change individuals’ behaviors, currently there is a major emphasis on so-called environmental strategies for health prevention, such as large-scale education campaigns and community-level actions such as changing zoning laws for liquor stores. The researchers found that the number of environmental strategies that a community implemented predicted more reductions in substance abuse outcomes. Tailoring interventions to the needs of the target population also predicted decreases in underage drinking outcomes (Diana et al., 2014).

The take-home messages from SPF are that researchers should do more evaluations with population data, do it better, and explain the results more simply, Orwin said. Protecting and expanding the data systems are also important goals, he added.

CLOSING RESEARCH DATA GAPS TO PREVENT YOUTH SUICIDES

According to the National Center for Health Statistics (NCHS), as of 2012, suicide accounted for 40,600 deaths in the United States—and 5,178 were young people from 10 to 24 years old, making it the second leading cause of death in that age group. For the past 3 years, Jane Pearson, associate director for preventive interventions in the Division of Services and Intervention Research at the National Institute of Mental Health (NIMH) and chair of the NIMH Suicide Research Consortium, has been working on a prioritized research agenda for suicide prevention, a project being implemented by a task force of the National Action Alliance for Suicide Prevention (NAASP). Experts have been looking at youth suicide numbers for decades, Pearson told the workshop audience, and "It was time to try to do something about it."

The task force set out to prioritize research objectives that, if implemented, could reduce all suicides — including youth suicides — by 20 percent in 5 years (NAASP, 2014). While the task force of course understands how long it takes to do research, it wanted to push to "find something that looked tractable, something that we should be able to do." The task force identified several gaps in research data around key questions: What causes suicide? How does one detect risk? What are adequate interventions? What kinds of prevention could work?

Pearson walked through several examples of data that are currently available. She started with the 2012 youth suicide numbers from NCHS, which were the most recent data available at the time of the workshop (see Figure 3-1). The suicide statistics are available online (and via downloadable app) through the CDC's Web-Based Injury Statistics Query and Reporting System (WISQARS).[1] WISQARS makes it possible to examine data at the state level and filter by gender, age, race, and ethnicity. The system also includes information on injury morbidity, which includes the numbers of suicide attempts as captured by a survey of hospital emergency departments (see Figure 3-2). Another source of suicide data is the recently expanded National Violent Death Reporting System (NVDRS), which offers an opportunity for state health departments to put together rich data around a death, such as linking to police or hospital records that may indicate stressors in an individual's life.

Looking at a graph of suicide statistics for the U.S. population, Pearson pointed out that there were upticks in the numbers in 2008 and 2009, perhaps reflecting the effects of the recession. What is disturbing is that the numbers for 2012 continue to show an uptick. "This is troublesome," she said. For 15- to 19-year-olds in particular, suicide rates declined in the early 2000s but have shown an uptick since then.

In the teenage and young adult years (ages 10 to 24), many more females attempt suicide than males do, but then that discrepancy evens out over time. Remarkably, Youth Risk Behavior surveys show that as many as 20 percent of youths seriously consider suicide. "What does that mean?" Pearson asked. "We just don't have a good sense of why there are some kids who think about it and nothing happens, and some kids think about it and do something." Experts talk about suicide in terms of a continuum

---

[1] Available at www.cdc.gov/ncipc/wisqars/default.htm (accessed March 25, 2015).

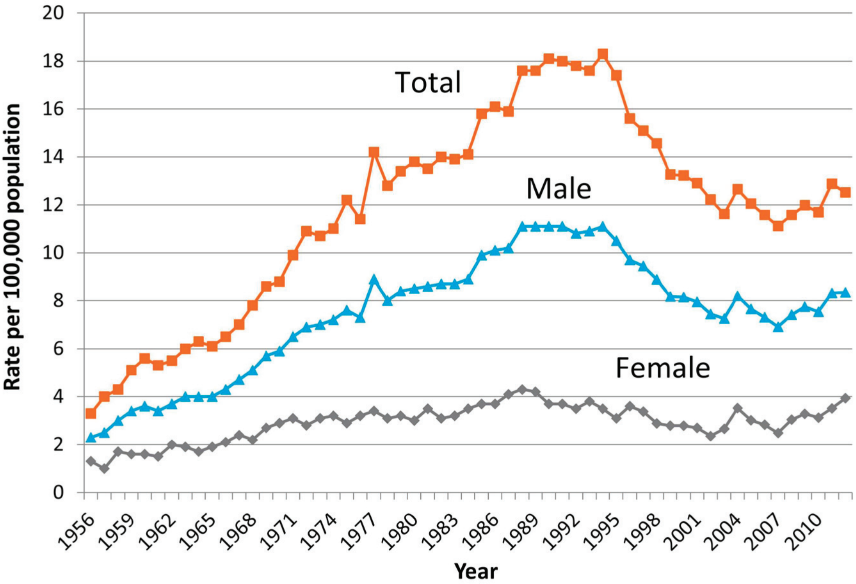

FIGURE 3-1 Suicide rates among persons aged 15 to 19 years. Suicide in the United States has risen slightly in recent years after a substantial decline.
SOURCE: CDC vital statistics; courtesy of Alex Crosby.

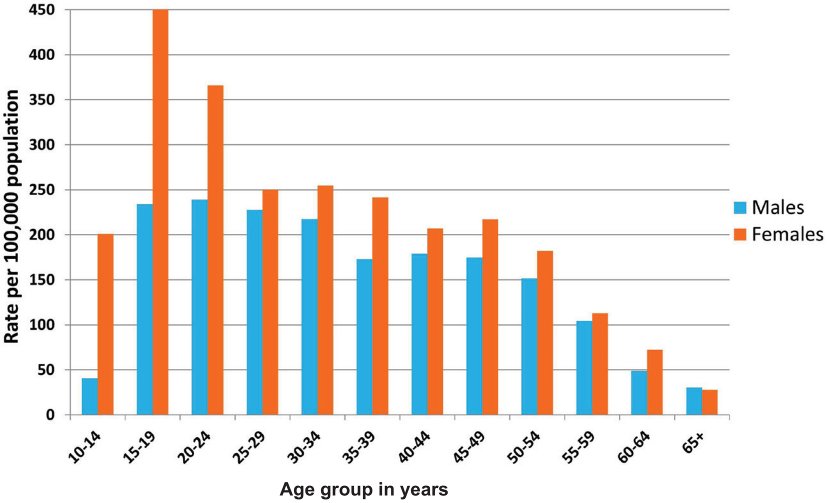

FIGURE 3-2 Emergency department self-inflicted injuries by age and sex. Data are shown for 2012 in the United States.
SOURCE: CDC WISQARS, 2012; courtesy of Alex Crosby.

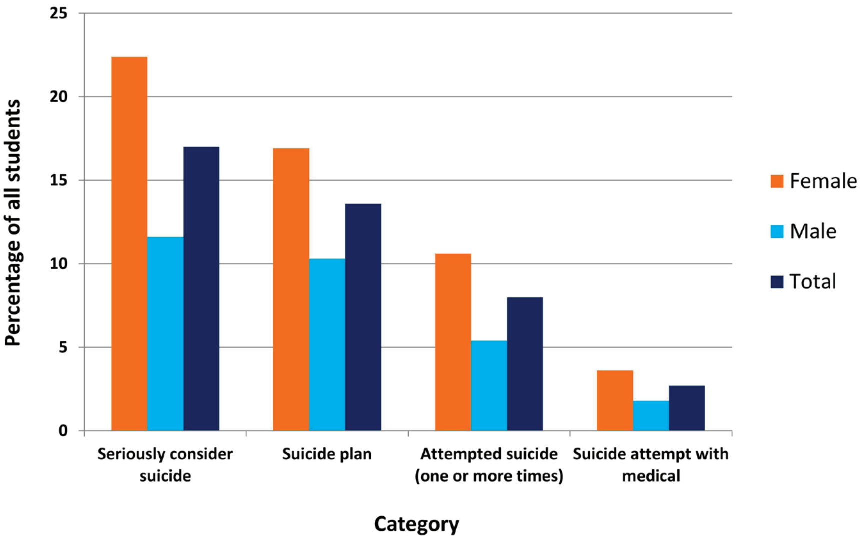

FIGURE 3-3 Suicidal ideation and behavior among high school students.
NOTE: Data from the Youth Risk Behavior Survey (YRBS) 2013, for the 12 months preceding the survey.
SOURCE: CDC Youth Risk Behavior Survey: http://www.cdc.gov/HealthyYouth/yrbs/index.htm (accessed March 25, 2015); courtesy of Alex Crosby.

that starts with people thinking about killing themselves, then planning an act, and attempting it (see Figure 3-3), but more longitudinal and phenomenological data are needed to understand those attempts, she said.

What can be done to reach youth at risk for suicide? The emergency room is a promising place for intervention.[2] A recently launched program called Emergency Department Screen for Teens at Risk for Suicide (ED-STARS)—funded by NIMH—is studying high-risk youth who receive care at 14 participating emergency departments.[3] Researchers will develop innovative new approaches for predicting suicide attempts, including a computerized adaptive screening tool similar to one created for identifying depression, and an implicit association task that has been adapted to screen for suicidal thinking (Cha et al., 2010; Gibbons et al., 2013).

The justice system is another route for addressing youth suicides. In

---

[2] For details, visit http://www.hcup-us.ahrq.gov/nedsoverview.jsp (accessed March 25, 2015).

[3] For more information, see http://www.nimh.nih.gov/news/science-news/2014/personalized-screen-to-id-suicidal-teens-in-14-ers.shtml (accessed March 25, 2015).

Utah, a 2002 study found that 63 percent of youths who died by suicide had had run-ins with the juvenile justice system (Gray et al., 2002). In a subsequent pilot program, University of Utah researchers screened youths going through the juvenile court system and found high rates of mental illness. Providing preventative interventions for the juvenile offenders improved their mental health and reduced new offenses (Moskos et al., 2007). Funding for that project was dropped, but the Utah researchers then applied for money from the Garrett Lee Smith (GLS) grants administered by SAMHSA.

The GLS grants support programs that train gatekeepers such as teachers and juvenile justice personnel on how to help a young person who is suicidal. A 2013 SAMHSA report to Congress found that in counties that implemented GLS training programs, the suicide rate went down in 10- to 24-year-olds, with 237 deaths prevented between 2007 and 2010 (SAMHSA, 2013). “It gives us hope that some of these programs that SAMHSA is implementing are making a difference,” Pearson said.

An important issue for prevention is measuring suicidal youth’s help seeking and getting them to the right kind of assistance. Studies indicate that “suicidal youth have pretty bad approaches to their own coping and getting help,” she said. They are more likely to reach out to their peers than adults (Gould et al., 2004; Pisani et al., 2012). “You have a lot of kids who get told about somebody who is suicidal, but they are asked to keep it a secret: Don’t tell anybody. This is a big issue.”

The National Science Foundation (NSF) and National Institutes of Health (NIH) have started funding research to explore how social media might be used to help youths with mental health issues such as substance abuse or depression. For example, a recent NSF-funded study of an online social network for lesbian, gay, bisexual, transgender, and queer (LGBTQ) youth suggested that users with fewer social connections were at greater risk for depression (Homan et al., 2014).

HOW DATA REPOSITORIES ARE OPENING ACCESS TO RESEARCH DATA ON AUTISM AND MENTAL HEALTH ILLNESSES

Greg Farber, director of NIMH’s Office of Technology Development and Coordination, described an innovative infrastructure, called the NIMH Data Archives, that collects information about research on human subjects. The infrastructure started with the National Database for Autism Research (NDAR), which Farber’s office oversees, and which recently expanded to include all data supported by NIMH from clinical trials as well as data related to the Research Domain Criteria initiative. In addition, NDAR has developed a “deep federation” in linking with other data repositories ranging from the Autism Tissue Program to the Simons Foundation Autism Research Initiative. “Qualified researchers can access the data," he said. "You can launch a query from the NIMH website"—or from several other sites—"and cover all of these other data repositories simultaneously."

NDAR was originally created in late 2006 as a joint initiative of the NIMH, the National Institute of Child Health and Human Development, the National Institute of Neurological Disorders and Stroke, and the National Institute of Environmental Health Sciences. "Putting data into a repository is like filling a pool with a garden hose. It takes a while before you have enough data to be interesting. At this point, the pool is pretty full. We have data from over 77,000 subjects. We are holding around 500 terabytes, a variety of different types of data."

Two principal building blocks shape how the data are organized: data dictionaries (which define the language characterizing research on autism and other illnesses) and the global unique identifier (GUID). The NDAR data dictionary provides a flexible and extensible framework for data definition by the research community and makes more than 500 instruments freely available for download from ndar.nih.gov. As data aggregates in the federal repositories, the community can begin to compare various data dictionaries and "start to pick winners," Farber said.

Another benefit of aggregating data is that if a researcher tries to send in an answer that falls outside the value range defined by a particular data dictionary, "We send it back and say, 'No, that is not quite right.'" At first, researchers hate being told their data are incorrect, he said, but they appreciate discovering errors sooner rather than later, and fix them. "The quality of data across the field is improved." The data repositories' query tool also allows investigators to quickly perform quality control checks on their data by comparing their results to the large samples of data in the repository.

Meanwhile, the GUID (rhymes with *fluid*) is a key building block designed to avoid including within the data repositories any information—such as Social Security numbers—that could personally identify individual study subjects. Using the GUID software, any researcher can enter information from a person's birth certificate (such as first name, last name, date, and place of birth) to generate a unique identifier number for that individual. If the same information is entered by other researchers in different laboratories, the same GUID number will be generated.

The GUID thus allows data from multiple research labs to be aggregated on the same research participants without having to share personally identifiable information about them. "This is a useful tool that we are happy to make available to a wide array of research communities," Farber said.[4]

The querying systems on the NDAR website let investigators easily

---

[4] For a video about informed consent issues, see http://www.youtube.com/watch?v=Tb6euCVoous (accessed March 25, 2015).

search the data repositories in various ways: by laboratory sources, by published scientific papers (linked to citations in PubMed), by data dictionary definitions, or by research concepts. Being able to access the shared datasets can enable real science. By running queries, scientists also can observe that many research participants are seen in many different laboratories, which was surprising, Farber said. Researchers usually assume "that we are drawing from a random sample when, in fact, we are drawing from a much smaller sample. That has real possibilities for biasing the sorts of results that we are getting."

How much is NDAR being used so far? "One question about these databases always is, if you build it, do they really come?" Farber said. But "Once [the pool] is full, people really do come." More than 270 users have been granted access to NDAR, and Farber's office has started seeing papers published based on data that came from the data registry (Richman et al., 2013; Sansone et al., 2012; Supekar et al., 2013). All in all, NDAR has made autism data useful and accessible, said Farber, and his office is happy to work with researchers on taking in data they have collected and making the information accessible and searchable through the NIMH Data Repositories.

USING DATA TO INFORM DECISION MAKING IN MARYLAND PUBLIC SCHOOLS

Catherine Bradshaw, professor and associate dean for research and faculty development at the University of Virginia's Curry School of Education, next presented a glimpse of how end-users are using datasets in real-world practice within the school setting. She described two federal initiatives in education: the Maryland Safe and Supportive Schools (MDS3) project and the Race to the Top. They offer examples of how data are being collected on children's academic performance and behavioral and mental health and applied to decisions about individual children or the adoption of evidence-based practices.

This work is in collaboration with the Maryland State Department of Education and is guided by a conceptual framework for prevention that builds upon multitiered systems of support. Maryland follows a well-known noncurricular, schoolwide tiered prevention model called Positive Behavioral Interventions and Supports (PBIS), which focuses on improving systems and practices through data-based decision making (Sugai and Horner, 2006).[5] The PBIS model applies a public health approach, wherein it is generally understood that about 80 percent of children will respond to universal interventions, while 20 percent need additional services. As the

---

[5] For more details, see http://www.pbis.org (accessed March 25, 2015).

magnitude of a problem rises, the level of intensity of support, resources, and data collection needed to address it increases as well.

Many schools do a good job in getting a universal level of supports into place, such as implementing schoolwide behavior management strategies, Bradshaw said. But when schools find that a student is not responding to that universal program and they want to understand why and which kinds of services might help, data-based decision making is necessary. Response to intervention is a data-based decision making framework used in education to guide that program selection process.

MDS3 is a collaboration among Johns Hopkins University, the Maryland State Department of Education, and Sheppard Pratt Health System, a large nonprofit mental health provider. Funded with $13.7 million from the U.S. Department of Education and around $1 million from the William T. Grant Foundation, MDS3 has three broad aims. One is to improve the school environment by cutting down on violence, bullying, and substance use and by improving connections among youth. The second is to develop a sustainable Web-based survey system for assessing school climate to guide the decision-making process. The third is implementing a continuum of evidence-based programs or practices to meet students' needs. In a pilot project, 58 schools implemented the school climate assessment system, and then half of the schools were randomized to a condition in which they selected from a menu of evidence-based programs. The model was tested as part of a 4-year randomized controlled trial. "We were very interested in how to improve the school climate in these school settings," Bradshaw said. Schools in the group that received interventions adopted the PBIS model and also chose to implement evidence-based programs such as Botvin's Life Skills Program for substance abuse prevention, or Check-In/Check-Out to boost student engagement and attendance. Schools received training and coaching by an implementation support provider, referred to as a school climate specialist. The overall framework for helping schools select and implement evidence-based programs is akin to the Communities That Care community-wide implementation model.

The MDS3 project uses the U.S. Department of Education's model of school climate, which outlines the three key areas of safety, engagement, and environment (see Figure 3-4). Many experts on school climate focus on students' perceptions of the school environment, but Bradshaw and her colleagues also wanted to examine behavioral issues, such as bullying or substance abuse, which can influence—or be influenced by—those perceptions.

The trial is currently in its final year of data collection, which includes gathering data on fidelity of implementation; observations of the school environment; and measures of school climate assessed via a Web-based surveillance system, called the MDS3 School Climate Survey, that allows students, parents, and school staff to complete a voluntary, anonymous

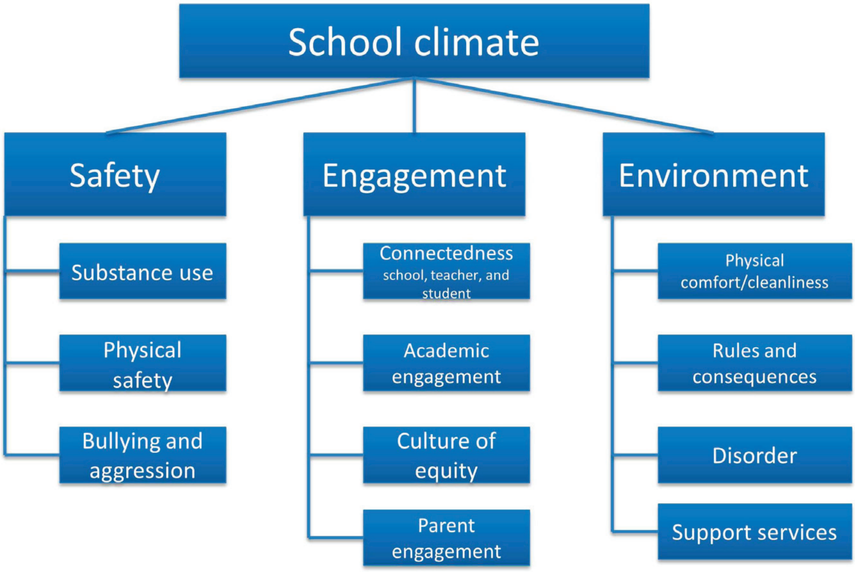

FIGURE 3-4 The U.S. Department of Education's School Climate Model. The model considers safety, engagement, and the school climate.
SOURCE: Adapted from Bradshaw et al., 2014b.

online questionnaire (Bradshaw et al., 2014a). School administrators can access the survey data in real time and instantly generate status reports through the password-protected website. The MDS3 initiative has expanded into middle and elementary schools, and more than 200 Maryland schools are now participating. Preliminary results from the randomized trial indicate significant improvements in several behavioral outcomes and aspects of school climate (Bradshaw et al., 2014a).

Meanwhile, Bradshaw and her colleagues worked with the Maryland State Department of Education to develop a data dashboard for Maryland's participation in the Race to the Top initiative. Goals included identifying which of 19 different indicators or risk factors could predict key outcomes, such as whether students graduated from or dropped out of high school, and whether students progressed to fifth or eighth grades on time or were held back. The researchers analyzed state data on three cohorts—each with more than 60,000 students—across elementary, middle, and high school. The 19 indicators included measures such as proficiency on standardized tests, yearly retention data, and yearly absences; demographic or racial information were not included in the initial analysis because the state did not want the resulting algorithm to be driven largely by demographic-based risk factors.

The data dashboard made it easy to visually stratify each student cohort into categories by the number of risk factors they had, and to see what the "cutpoints" were for the different outcomes. For instance, a high school student with zero risk factors is likely to graduate, but a teen with three or more risk factors is not. Faced with such a high-risk student, a school principal might respond by saying, "We have got to get to this kid and start thinking about ways that we can support [him or her]," Bradshaw said. The researchers found their model fit best for predicting the outcome of whether students did not graduate; academic achievement and retention were the best predictors (Pas and Bradshaw, 2014).

Wrapping up, Bradshaw highlighted several common themes from the two data-based initiatives in Maryland schools: The focus and framing of the data dashboard varies by user need, which may include different interests in school climate, dropout rates, particular types of data (e.g., academic performance or behavior), or decision making at the school level versus for individual cases. Predictive modeling can be helpful for guiding decision making. Incentives for data collection and use are important. Lastly, it is important to provide training and a framework for using that information to support decision making.

MEASURING POPULATION-LEVEL PROGRESS IN THE FIGHT FOR DRUG-FREE COMMUNITIES

In the last presentation of the panel, Kareemah Abdullah, director of the National Community Anti-Drug Coalition Institute and vice president of training operations at Community Anti-Drug Coalitions of America (CADCA), shared a perspective on measurement systems that assess population-level change through the context of coalitions. Based in Alexandria, Virginia, the nonprofit CADCA supports data-driven strategic planning and prevention of illicit drug use, underage drinking, and youth tobacco use for coalitions across the United States and its territories.[6] CADCA represents more than 5,000 community coalitions and affiliates nationwide and has helped build coalitions in 29 countries. Its goal is to unite these partners in bringing about population-level reductions in substance abuse rates.

The National Coalition Institute is the arm of CADCA that provides high-level training and technical assistance, evaluation and research, and capacity building for coalitions that receive funding through the Drug Free Communities (DFC) Support Act passed by Congress in 2001. The institute is charged with "making coalitions smarter faster," Abdullah said. CADCA's Community Problem-Solving Model aims to help coalitions achieve population-level change. Realizing programs are necessary but not

---

[6] For more information, see http://www.cadca.org (accessed March 25, 2015).

sufficient, coalitions focus on environmental strategies to achieve significant youth behavioral outcomes.

Its research shows that "one of the most important indicators for achieving population-level outcomes is engaging in essential processes and having the capacity to do so." The institute operates from a conceptual "framework for community change" in which training and technical assistance trigger coalitions to improve their capacity, which in turn enables them to implement essential processes that lead to a set of comprehensive intervention strategies (see Figure 3-5). Such strategies create community change that resonates at the population level, driving population-level improvements. This is the model that CADCA uses in providing the training that all DFC-funded coalitions are required to undertake.

CADCA has found that communities need a problem-solving process to be able to ask critical questions and identify the "local conditions" and risk factors for which interventions can be developed to reduce problem behaviors. "We are identifying the problem. We are then asking why that problem exists, what are the causal factors about the problem behavior among youth in your particular neighborhood. . . . Then, your comprehensive intervention strategies must be mapped, measured, and monitored at that local condition level." One example of a local condition is when youths are gaining access to alcohol because neighborhood merchants are not carding them.

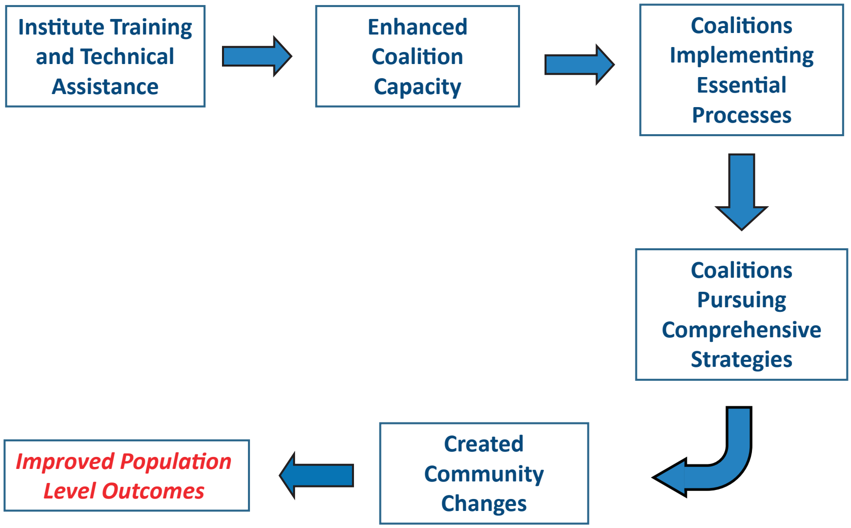

FIGURE 3-5 Framework for community change of the National Coalition Institute.
SOURCE: Abdullah, 2014.

CADCA focuses on seven strategies for behavioral change: providing information, enhancing skills, providing support, enhancing access or reducing barriers, changing consequences or incentives, changing physical design, and modifying or changing policies. The last four areas target changes in environment and are necessary for achieving the greatest improvements in youth behavioral outcomes, Abdullah explained. All seven strategies must be applied to each local condition that is fueling a substance abuse problem.

CADCA's Institute has been independently evaluated since its inception in 2002. An independent evaluation led by Pennie Foster-Fishman of Michigan State University examined the impacts of CADCA's institute training and technical assistance.[7] This longitudinal 4-year analysis tracked coalitions that had received DFC funding in 2008–2009. Overall, the percentage of coalitions that engaged in creating policy changes grew significantly in the 3 years following their CADCA training. Policy change was a major indicator for positive outcomes in youth behaviors, Abdullah said.

However, the analysis also revealed that the increase in policy-changing work dropped off at 4 years' posttraining (Foster-Fishman, 2014), indicating that coalitions "at this stage needed additional training and technical assistance," she said. Many coalitions needed to revisit their collective work and begin to focus on new local conditions because, in many cases, their strategies for addressing previously identified local conditions had succeeded.

Another evaluation conducted by ICF International studied the extent to which pursuit of CADCA's problem-solving approach had an impact on youth substance abuse. In a cross-site analysis of middle school and high school students, this study measured how many of the teens reported perceptions of parental disapproval of alcohol, marijuana, and tobacco use and perceptions of risk associated with substance use; and how many had used substances in the past month. Communities with DFC-funded coalitions that engaged in the problem-solving approach showed better outcomes in youth behaviors than did communities without DFC coalitions (DFC National Evaluation, 2013).

Overall, of all the elements in the CADCA's framework for change, the amount of "community change"—that is, program and policy changes—that coalitions produce has the strongest impact on population outcomes. Given the complex, messy nature of communities and coalition work, Abdullah recommended that measurement systems for assessing coalitions' effect on youth behavioral outcomes have three components: They should be simple, with a linear logic model. They should be able to capture varied

---

[7] For more information, see http://www.cadca.org/resources/detail/coalitions-trained-cadca-institute-are-more-effective-community-problem-solvers (accessed March 25, 2015).

interactions and outcomes, she added, because "dynamics need to be considered where multiple coordinated pathways are part of the outcomes and very complex." Lastly, they should look at complexity-based theories for both action and change, Abdullah said.

DFC coalitions have a real impact through their work in communities, she concluded, because the prevalence of substance abuse among young people increases or decreases based on their perceptions of harm and use.

REFERENCES

Abdullah, K. 2014. *Measurement systems to assess individual- and population-level change*. Presented at Institute of Medicine and National Research Council Workshop on Innovations in Design and Utilization of Measurement Systems to Promote Children's Cognitive, Affective, and Behavioral Health, Washington, DC.

Bradshaw, C. P., K. J. Debnam, S. Lindstrom Johnson, E. Pas, P. Hershfeldt, A. Alexander, S. Barrett, and P. J. Leaf. 2014a. Maryland's evolving system of social, emotional, and behavioral interventions in public schools: The Maryland Safe and Supportive Schools Project. *Adolescent Psychiatry* 4(3):194-206.

Bradshaw, C. P., T. E. Waasdorp, K. J. Debnam, and S. L. Johnson. 2014b. Measuring school climate in high schools: A focus on safety, engagement, and the environment. *Journal of School Health* 84(9):593-604.

Cha, C. B., S. Najmi, J. M. Park, C. T. Finn, and M. K. Nock. 2010. Attentional bias toward suicide-related stimuli predicts suicidal behavior. *Journal of Abnormal Psychology* 119(3):616-622.

Diana, A., R. G. Orwin, and J. Park. 2014 (February). *Outcomes from the SPF SIG cross-site evaluation, Cohorts I and II*. Presented at the Community Anti-Drug Coalitions of America Annual Meeting, National Harbor, MD.

Drug-Free Communities Support Program 2012 National Evaluation Report. 2013 (June). Fairfax, VA: ICF International. http://www.whitehouse.gov/sites/default/files/dfc_2012_interim_report_annual_report_-_final.pdf (accessed March 25, 2015).

Foster-Fishman, P., and Y. Mei. 2014 (September 15). *Longitudinal evaluation of the impact of CADCA's institute's training and TA on coalition effectiveness*. East Lansing, MI: Michigan State University. http://www.cadca.org/sites/default/files/pdfs/institute/lastestfindingsinstituteeval.pdf (accessed March 25, 2015).

Gibbons, R. D., G. Hooker, M. D. Finkelman, D. J. Weiss, P. A. Pilkonis, E. Frank, T. Moore, and D. J. Kupfer. 2013. The computerized adaptive diagnostic test for major depressive disorder (CAD-MDD): A screening tool for depression. *Journal of Clinical Psychiatry* 74(7):669-674.

Gould, M. S., D. Velting, M. Kleinman, C. Lucas, J. G. Thomas, and M. Chung. 2004. Teenagers' attitudes about coping strategies and help-seeking behavior for suicidality. *Journal of the American Academy of Child and Adolescent Psychiatry* 43(9):1124-1133.

Gray, D., J. Achilles, T. Keller, D. Tate, L. Haggard, R. Rolfs, C. Cazier, J. Workman, and W. M. McMahon. 2002. Utah youth suicide study, phase I: Government agency contact before death. *Journal of the American Academy of Child and Adolescent Psychiatry* 41(4):427-434.

Homan, C. M., N. Lu, X. Tu, M. C. Lytle, and V. M. B. Silenzio. 2014. Social structure and depression in TrevorSpace. Proceedings of the 17th ACM Conference on Computer-Supported Cooperative Work & Social Computing. February 15, 2014. Baltimore, MD.

Moskos, M. A., S. R. Halbern, S. Adler S, H. Kim, and D. Gray. 2007. Utah Youth Suicide Study: Evidence-based suicide prevention for juvenile offenders. *Journal of Law and Family Studies* 10:127-145.

NAASP (National Action Alliance for Suicide Prevention Research Prioritization Task Force). 2014. *A prioritized research agenda for suicide prevention: An action plan to save lives.* Rockville, MD: National Institute of Mental Health.

Orwin, R. G., A. Stein-Seroussi, J. M. Edwards, A. L. Landy, and R. L. Flewelling. 2014. Effects of the Strategic Prevention Framework State Incentives Grant (SPF-SIG) on state prevention infrastructure in 26 states. *Journal of Primary Prevention* 35(3):163-180.

Pas, E., and C. Bradshaw. 2014. *Student Risk Trends Dashboard Project*. A report prepared for the Maryland State Department of Education, Baltimore, MD.

Pisani, A. R., K. Schmeelk-Cone, D. Gunzler, M. Petrova, D. B. Goldston, X. Tu, and P. A. Wyman. 2012. Associations between suicidal high school students' help-seeking and their attitudes and perceptions of social environment. *Journal of Youth and Adolescence* 41(10):1312-1324.

Richman, D. M., L. Barnard-Brak, A. Bosch, S. Thompson, L. Grubb, and L. Abby. 2013. Predictors of self-injurious behaviour exhibited by individuals with autism spectrum disorder. *Journal of Intellectual Disabilities Research* 57(5):429-439.

SAMHSA (Substance Abuse and Mental Health Services Administration). 2013. Report to Congress: Garrett Lee Smith Youth Suicide Prevention Program. Rockville, MD: SAMHSA.

Sansone, S. M., K. F. Widaman, S. S. Hall, A. L. Reiss, A. Lightbody, W. E. Kaufmann, E. Berry-Kravis, A. Lachiewicz, E. C. Brown, and D. Hessl. 2012. Psychometric study of the Aberrant Behavior Checklist in Fragile X Syndrome and implications for targeted treatment. *Journal of Autism and Developmental Disorders* 42(7):1377-1392.

Sugai, G., and R. H. Horner. 2006. A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review* 35:245-259.

Supekar, K., L. Q. Uddin, A. Khouzam, J. Phillips, W. D. Gaillard, L. E. Kenworthy, B. E. Yerys, C. J. Vaidya, and V. Menon. 2013. Brain hyperconnectivity in children with autism and its links to social deficits. *Cellular Reproduction* 5(3):738-747.

4

Using Quality Measures to Facilitate System Change

Whether developed purely for research or to be applied in program improvement, measures have the potential to fundamentally change social service systems, including child care, juvenile justice, education, and health care. However, these changes are almost invariably complex and can have harmful as well as beneficial effects. Four speakers at the workshop examined the systemic changes that can occur as the result of the development and implementation of new measures and drew lessons on how to optimize the effects of such changes.

ENSURING HIGH QUALITY IN MEASURES OF CHILD CARE AND PRESCHOOL: A CAUTIONARY TALE

Quality rating and improvement systems, which link child care subsidy levels to quality ratings, emerged in the late 1990s and now operate in about three-quarters of the states (Child Trends, 2014). More recently, the Race to the Top Early Learning Challenge encouraged states to integrate quality-monitoring systems across funding streams, and the Improving Head Start for School Readiness Act of 2007 required lower-quality Head Start grantees to recompete for funding (though none were actually required to do so until 2011).[1] These policy initiatives have accelerated a trend of adopting measures designed for other purposes for high-stakes

---

[1] Additional information can be found at the *Federal Register* 75, no. 183 [September 22, 2010]: 57717, and at http://www2.ed.gov/programs/racetothetop-earlylearningchallenge/index.html (accessed March 25, 2015).

uses, noted Rachel Gordon, professor in the Department of Sociology and associate director of the Institute of Government and Public Affairs at the University of Illinois at Chicago.

These initiatives also have directed attention to the use of measures that are valid and reliable. For example, the Race to the Top Early Learning Challenge required states to use "valid and reliable" indicators of the overall quality of the early learning environment and of the quality of adult–child interactions.[2] Such use of the terms *reliable* and *valid* suggest that these are static properties of a measure for all time, all purposes, and all populations, Gordon observed, but noting "This isn't consistent with our contemporary thinking about measurement." Instead, the developers and users of measures need to consider the intents of each research and policy use and weigh the body of reliability and validity evidence against each use, she said, which is consistent with the latest *Standards for Educational and Psychological Testing* (AERA/APA/NCME, 2014). For instance, the body of evidence desired to demonstrate reliability and validity for program self-assessment may be different than that for teacher professional development, and both of these could differ for policy decision making and accountability. Similarly, the developers and users of measures need to build in continuous and local validation of measures selected for various uses and allow for the refinement of measures over time and place, Gordon said.

She used two examples to talk about these issues in greater detail. The first is the Early Childhood Environment Rating Scale (ECERS),[3] and the second is the Classroom Assessment Scoring System (CLASS).[4] These are intensive observational measures where independent observers spend at least several hours in classrooms. Together, they are used in about 90 percent of states (Child Trends, 2014).

As the high stakes use of quality rating and improvement systems came into being, states aimed to assure that publicly funded programs were of high quality and to incentivize advancement. In response to a question, Gordon pointed out that at first it was hard to get child care centers to participate voluntarily. Some states then moved to include all centers in the program and started them out at the bottom level, after which they could move up in the rankings. "The rating and improvement is meant to partly be marketing" for the centers, she said, but "it is meant to also give information to consumers about making choices."

---

[2] Race to the Top Early Learning Challenge requirements can be found at https://www.federalregister.gov/articles/2011/08/26/2011-21756/applications-for-new-awards-race-to-the-top-early-learning-challenge (accessed March 25, 2015).

[3] Additional information on ECERS can be found at http://ers.fpg.unc.edu/early-childhood-environment-rating-scale-ecers-r (accessed March 25, 2015).

[4] Additional information on CLASS can be found at http://teachstone.com/the-class-system (accessed March 25, 2015).

Earlier interpretations concluded that a strong association existed between these quality measures and student readiness outcomes. However, this evidence often focused on statistical significance and not the size of associations, did not rigorously adjust for selection (confounds), and may have reflected measures of quality levels typical several decades ago, prior to contemporary licensing and programmatic standards, Gordon noted. The emerging consensus is that the associations with readiness are not always significant and are generally small—typically 0.1 or smaller in effect size (Abner et al., 2013; Burchinal et al., 2011; Gordon et al., 2013; Keys et al., 2013).

How these measures were designed may relate to their limitations for high-stakes uses, Gordon explained. The ECERS-R emerged in the 1970s from a checklist to help child care practitioners improve the quality of their settings. It reflects developmentally appropriate practices, including a predominance of child-initiated activities selected from a wide array of options and a "whole child" approach that integrates physical, emotional, social, and cognitive components. The ECERS-R has more than 400 indicators across 43 items, grouped in a way that makes sense in the context of practice. These features of the measure may be valid from a philosophical perspective, Gordon said, but they do not necessarily focus on the kinds of intentional teaching that increasing evidence indicates is best for school readiness.

The CLASS was developed more recently based on research suggesting that interactions between students and adults are the primary mechanism of student development and learning (Pianta et al., 2008). Its predecessor was part of a research study, and it was later aimed at professional development and coaching before being adopted in high-stakes policy contexts. It has a very different structure than ECERS-R, because it requires observers to assimilate what they see in order to assign scores to just a few items in the categories of emotional support, classroom organization, and instructional support.

In a high-stakes context, a measure should provide very high agreement if it is being used for specific cutoffs, Gordon noted. However, a recent publication from the CLASS developers reveals that inter-rater reliability is low (Cash et al., 2012). The CLASS developers also recently found a bi-factor structure (Hamre et al., 2014), with one general dimension (responsive teaching) and two specific dimensions (proactive management and routines, and cognitive facilitation). These differ from the subscales written into policy, and domains may align differently than originally thought with aspects of quality specific to readiness for school. CLASS scores also tend to cluster in just a few scale categories, and this limited variation may make it difficult to reveal changes in quality over time and may attenuate associations between quality measures and cognitive outcomes. Because public investments expect high-quality preschool promoting academic school readiness, these results could have the unintended consequence of discouraging investments in early child care, Gordon said.

The bottom line is that information about reliability and validity needs to be independently collected, Gordon concluded. For example, she briefly mentioned a pilot study at the University of Illinois known as the Early Investments Initiative that uses new technology to take a careful look at variation in quality within and across the school day and across quality definitions, measures, and standards. She also pointed out that technology can be leveraged to gather evidence, both for greater understanding and for feedback for teachers and parents. A major objective is to provide continuous feedback and learn from teachers about the reliability and validity of measures, Gordon said. "Quality is the right goal, but we want to be sure we have the tools aligned with our current high-stakes policy use."

THE JUVENILE JUSTICE REFORM AND REINVESTMENT INITIATIVE

The Juvenile Justice Reform and Reinvestment Initiative (JJRRI) is a comprehensive approach to reforming the juvenile justice system using a research-based, data-driven, decision-making platform to inform system improvements and service delivery. The ultimate goal of the initiative, said Kristen Kracke, a social science specialist at the U.S. Department of Justice's Office of Juvenile Justice and Delinquency Prevention, is to improve both outcomes and cost effectiveness.

Research in juvenile justice has demonstrated that early intervention is less expensive and more effective than involvement with the juvenile justice system. It also has shown that, for youth involved in the juvenile justice system, interventions focusing on control (such as detention) are less effective than therapeutic approaches for reducing recidivism, and that the deeper youth go into the system, the more likely they are to reoffend (Lipsey, 2009; Lipsey and Cullen, 2007; Loughran et al., 2009).

JJRRI,[5] which is an innovation pilot funded by the Office of Management and Budget, involves three jurisdictions over a 3-year project period. Its short-term goals are improved program delivery, better matching of youth to services, system improvements, and reinvestment of cost savings to the front end of youth services in the community rather than in confinement. Its long-term goals are decreased recidivism rates and improved outcomes for youth, improved cost effectiveness of juvenile justice services, and a reduction in public cost and reinvestment in community services.

---

[5] For more information on JJRRI, see http://www.ojjdp.gov/grants/solicitations/FY2012/JJRRI.pdf (accessed March 25, 2015).

A major objective of the initiative is to create an evidence-based operating platform, which would increase the match between the risk assessment and supervision levels, and between needs assessments and program options (see Figure 4-1). Implementation of the platform has involved the installation of a program rating instrument known as the Standardized Program Evaluation Protocol (SPEP). Developed by Mark Lipsey at Vanderbilt University's Peabody Research Institute based on a meta-analysis of more than 600 intervention studies, SPEP assesses how well current program practice matches the profile of programs with research evidence for effectiveness. Pilot programs are working with SPEP to derive an initial set of scores, which guide the program improvement process, and a second set of scores designed to reveal changes in program quality.

SPEP covers four key areas:

1. Program or service type
2. Quality of service delivery
3. Amount of service, including duration and intensity of contact (face-to-face, group, etc.)
4. Risk level of youth served

It is a holistic process, Kracke emphasized, designed to align all parts of the system, including risk and needs assessments, service selection, ongoing case management, and reinvestment in front-end outcome-driven community-based services. "It is a continuous improvement process for all of the partners involved—the courts, the agencies, the providers, and the youth themselves."

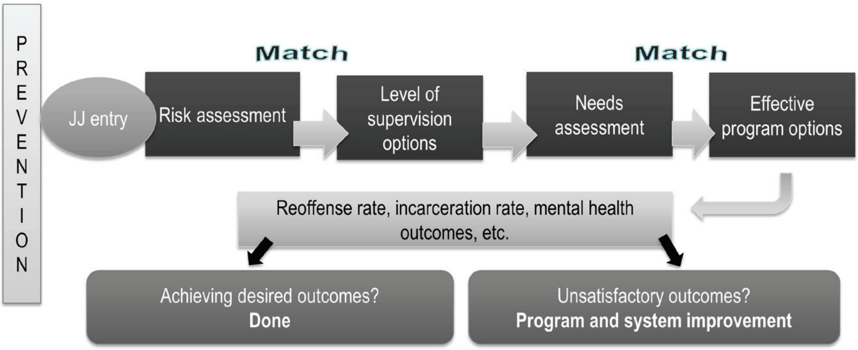

FIGURE 4-1 Matching risks in an evidence-based operating platform for juvenile justice (JJ). An evidence-based operating platform for JJ matches risk to supervision and needs to effective programs.
SOURCE: Kracke, 2014.

As an example of how the alignment process works, Kracke described in greater detail a part of the system known as the dispositional matrix. The dispositional matrix is a structured decision-making tool for courts on dispositions for youth that matches risk levels and offense types to recommend a supervision level. In Florida's implementation of the matrix, for example, low-risk offenders remain in the community with minimal supervision, moderate-risk offenders are placed in more structured community programs with intensive probation for higher risk youth, and residential placement is reserved for the highest risk offenders after community-based alternatives have been exhausted. The underlying principle is to place youth in an optimal placement while trying to meet the youth's needs in the most cost-effective way.

In Florida, when youth are placed in the least restrictive environment, their 12-month recidivism is lower than for placements that are either below or above the guidelines (see Figure 4-2) (Baglivio and Mark, 2014). These are "really powerful data," said Kracke. The information also makes it possible to monitor the results of differential placement by jurisdictions while not taking away judicial discretion. "If a jurisdiction or a district is making discretionary decisions always above the guidelines or always below the guidelines, this gives them the data to show them the direct impact of that—and the state knows."

"It is very exciting work," Kracke concluded. "We think it is one of the next best things in terms of helping us meet our mission, which is to have the juvenile justice system be rare, fair, and beneficial." Evidence-

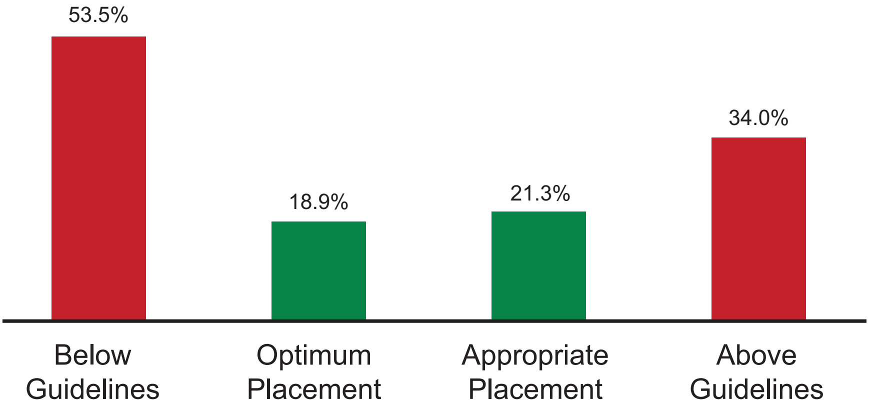

FIGURE 4-2 All youth 12-month recidivism by matrix adherence level. Optimum placement in the least restrictive environment reduces recidivism.
SOURCE: Kracke, 2014, from Baglivio and Russell, 2014.

based practice takes money and time to cultivate, she noted. But evidence developed in one area also can have applications across the social sciences.

CAN CHILD MENTAL HEALTH CROSS THE QUALITY CHASM?

As with the educational system and juvenile justice system, the health care system is in the process of building an infrastructure of structural measures with associated process and outcome measures. In particular, the Affordable Care Act (ACA) contains a large number of provisions focused on quality measurement and accountability, observed Harold Pincus, professor and vice-chair of the Department of Psychiatry and codirector of the Irving Institute for Clinical and Translational Research at Columbia University, and also director of quality and outcomes research at New York-Presbyterian Hospital. To take just one example, within the area of mental health care, the value-based inpatient psychiatry quality reporting program is a pay-for-reporting, rather than a pay-for-performance system, that is having an impact on the mental health care provided by hospitals, Pincus said.

Part of the concern with quality in the health care system dates to the Institute of Medicine (IOM) reports *To Err Is Human* (IOM, 1999) and *Crossing the Quality Chasm* (IOM, 2001). The latter report established six domains of quality health care:

1. **Safe—**Avoids injuries of care
2. **Effective—**Provides care based on scientific knowledge and avoids services not likely to help
3. **Patient-centered—**Respects and responds to patient preferences, needs, and values
4. **Timely—**Reduces waits and sometimes harmful delays for those receiving and giving care
5. **Efficient—**Avoids waste, including waste of equipment, supplies, ideas, and energy
6. **Equitable—**Care does not vary in quality due to personal characteristics (gender, ethnicity, geographic location, or socioeconomic status)

A subsequent IOM report, *Improving the Quality of Health Care for Mental and Substance-Use Conditions*, looked specifically at the quality chasm for mental health and substance abuse, including care for children (IOM, 2005). The report identified a series of obstacles for mental health and substance abuse conditions, including obstacles to patient-centered care, poor linkages across the health care system, and an insufficient workforce capacity for quality improvement.

Pincus directed his attention to another obstacle—a weak measurement and improvement infrastructure—and cited the following six conclusions from the report:

1. Clinical assessment and treatment practices are not well standardized and classified for use with administrative datasets. Data about the kinds of care being provided, the kind of conditions being treated, and outcomes are not readily available.
2. Outcomes measurement is not widely applied despite the availability of reliable and valid instruments and the demonstrated value of measurement-based care (Harding et al., 2011).
3. Not enough attention has been given to the development or implementation of performance measures for mental health and substance abuse care, especially for children.
4. Quality improvement measures have not yet permeated the day-to-day operations of mental health services.
5. The workforce is not trained in quality measures and improvement.
6. Policies have not effectively incentivized quality and efficiency.

In addition, Pincus pointed out that with the profusion of new measures being developed, the number particular to the mental health of children is small in proportion to the need. Efforts on behalf of children therefore are needed to develop reliable, valid, and feasible quality measures.

The ACA requires the establishment of a National Quality Strategy to implement better measurement and improvement of quality, and the Centers for Medicare & Medicaid Services (CMS) also has a number of quality programs for hospitals, physicians, payers, and health care settings. Pincus called attention to several key features of these quality measurement systems.

First, measures can improve performance by teaching people how to improve. Furthermore, this can be done across domains of safety, effectiveness, equity, efficiency, patient-centeredness, and timeliness.

Measurements also can be used for accountability. Such measurements require a higher threshold, said Pincus, when they are used for such purposes as public reporting and payment.

Measures can be used across the different silos of the health care system, including mental health, and at multiple levels. For example, Pincus outlined a "six P" conceptual framework at the level of patients, providers, practices or delivery systems, health plans, public and private purchasers, and populations and policies (see Figure 4-3). These levels can be analyzed in terms of five types of measures:

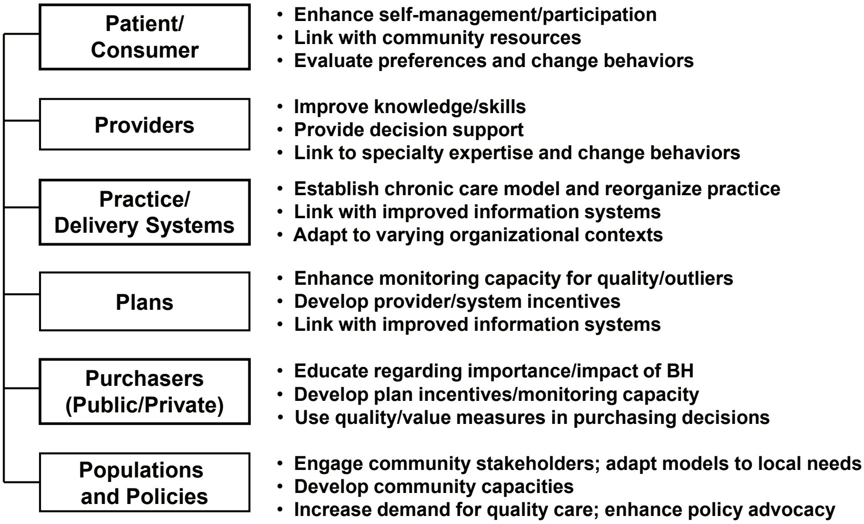

FIGURE 4-3 The "six P" conceptual framework.
NOTE: BH = behavioral health.
SOURCE: Pincus, 2014.

1. **Structure**—Are adequate personnel, training, facilities, security, quality improvement infrastructure, information technology resources, and policies available for providing care?
2. **Process**—Are evidence-based processes of care accessible? Are they delivered with fidelity?
3. **Outcome**—Does care improve clinical outcomes?
4. **Patient experience**—What do users and other stakeholders think about the system's structure, the care they have received, and their outcomes?
5. **Resource use**—What resources are expended for the structure, processes of care, and outcomes?

Developing indicators in turn involves a series of steps:

- Establishing an evidence base
- Translating evidence to guidelines
- Translating guidelines to measure concepts
- Operationalizing concepts to measure specifications (which includes determining a numerator and denominator)
- Testing for reliability, validity, feasibility
- Aligning measures across multiple programs
- Stewardship, including updating measures over time

The data for quality measures can come from multiple sources, including administrative data, chart reviews, electronic health records, registries, and patient surveys. In addition, benchmarks need to be set, said Pincus. "What rate is right? If you are looking at adherence to certain types of medications, should everybody be 100 percent adherent to all their medications? Or are there some elements of patient preference, adverse effects and other kinds of issues that need to be taken into account?"

Multiple players are involved in the measurement process, including evidence developers (such as researchers, the National Institutes of Health [NIH], and Patient-Centered Outcomes Research Institute [PCORI]), guideline developers (such as professional associations), measure developers and stewards (such as the National Committee for Quality Assurance [NCQA] and CMS), measure endorsers (such as the National Quality Forum [NQF]), measure selectors and advisers (such as NQF, Measures Applications Partnership [MAP], and CMS), and measure users, including CMS, health plans, organizations, the media, and the public. Pincus focused on the endorsement criteria of the NQF, which include the following sets of standardized criteria[6]:

- **Importance of measure—**Extent to which the specific measure focus is evidence-based, important to making significant gains in health care quality, and improving health outcomes for a specific high-priority (high-impact) aspect of health care where there is variation in or overall less-than-optimal performance
- **Scientific acceptability—**Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented
- **Usability—**Extent to which potential audiences are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient health care for individuals or populations
- **Feasibility—**Extent to which the required data are readily available or could be captured without undue burden and can be implemented for performance evaluation

Finally, Pincus noted that indicators can be used to improve quality at the clinical level, the organizational level, and the policy level. Measurement needs to be built into the processes of care, he said, because successful interventions require that people be followed over time and that outcomes be

---

[6] Additional information about the NQF criteria for evaluation of measures can be found at http://www.qualityforum.org/docs/measure_evaluation_criteria.aspx (accessed March 25, 2015).

measured. For example, an accreditation process could incorporate within it an expectation of longitudinal measurement, which gets reported as a measure of outcomes. But a key question remains—accountability. “You have a lot of players,” Pincus noted, and then asked “How do you hold them accountable? What sort of shared accountability is necessary and how can it be operationalized?” Some innovations supported by the ACA, such as accountable care organizations, can specify accountability, but these provisions may not be operationalized in a way that has an impact on care for children.

QUALITY MEASURES AND IMPROVEMENT IN HEALTH CARE

As part of its accreditation process for health plans, NCQA looks at performance-based measures of health care quality. For example, a measure it has been using in recent years is the percentage of children who receive appropriate follow-up care when they are prescribed medications for attention deficit hyperactivity disorder (ADHD). “It is not very impressive performance,” said Sarah Scholle, vice president of research and analysis for NCQA, noting that only about 45 percent of children get appropriate follow-ups. Similarly, the number of people who receive a follow-up within 30 days of being hospitalized for a mental illness is worryingly low—only about 55 percent to 75 percent for different kinds of health plans, despite the seriousness of mental health hospitalization.

Having a quality measure does not necessarily focus attention on improvement in the area, Scholle pointed out. “Just because you have a quality measure, it doesn’t mean that it is going to galvanize the kind of attention, even when there is some accountability attached to it.”

In a recent review of behavioral health quality measures, NCQA found 496 such measures, though many are variations on a much smaller number of themes. But only 12 percent of these measures are nationally endorsed, and only 10 percent focus on children and adolescents. Most of the measures used data that can be captured from claims data, because it is much easier to capture that information than it is to derive data from medical records.

A good quality measure needs to be reliable and valid for the particular use of that measure, said Scholle, and it needs a good evidence base. The evidence is the foundation for a measure’s development, along with a review of the environment, guidelines, and important concepts surrounding that measure. Draft measure specifications need to be tested with feedback from the people who are going to be measured and the people who are going to use the measure.

As an example of this process, Scholle described the work that has been done to improve a measure for adolescent depression.[7] First, the measure developers articulated a vision of what high-quality care for adolescent depression would look like. A screening process would lead to a good assessment to determine a diagnosis. Depending on the severity of illness, a patient might undergo brief supportive counseling or different kinds of treatment. Ideally, symptoms and functionality would be followed and assessed over time as they either respond to treatment or go into remission.

Using an existing measure in adults that is based on a patient health questionnaire, the group developed measures for monitoring, remission, and treatment adjustment for patients aged 12 to 17. Preliminary data indicate that a major challenge with the measure was the low rates of symptom monitoring for depressed adolescents, with 83 percent of a sample of 684 adolescents not being followed up 4 to 8 months after a diagnosis of major depression. Among those who were followed up, 5 percent of the total sample were in remission 4 to 8 months later, 4 percent had responded without remission, and 8 percent did not respond. Among those who were not followed up, many may have gotten better, but "you have to go out and find those kids and make sure they are okay," said Scholle.

The logic model for quality measurement starts with structure, said Scholle, including training and ongoing supervision in evidence-based therapy, an infrastructure for collection of patient-reported data, and systems for sharing information across care teams (Lewandowski et al., 2013). It then progresses through various processes, including access to and use of behavioral health services, receipt of evidence-based therapy, and monitoring of symptoms and functioning using standardized tools. Outcomes include fewer harmful events, diminished symptoms, improved functioning, and school attendance.

According to this logic model, important issues for evidence-based therapies are which therapies to use, which populations to target, criteria for determining whether the evidence-based treatment is carried out, the data sources for capturing treatment, and access to confidential records.

This work on depression is following a model being used by other federal agencies to develop outcome measures for adults, both in the general medical sector and in behavioral health, Scholle stated. Key questions are whether tools are available, which tools to use or develop, methods for data collection, the expectations for improvement over time, and accountability. "For every measure, we would have to think about 'what do you expect?'" said Scholle. "Who is accountable? Is it the individual clinician?

---

[7] This project was supported by grant number U18HS020503 (PI: Scholle) from the Agency for Healthcare Research and Quality (AHRQ) and the Centers for Medicare & Medicaid Services (CMS).

Is it the practice? Is it the delivery system? Is it the health plan?" Often, the responsibility belongs to the system as a whole, not just to the clinician or the patient, Scholle said. For example, a clinician may provide the care while a care coordinator makes sure that a patient returns for a scheduled appointment.

Scholle also pointed out, in response to a comment about the difficulties posed by electronic health record systems for clinicians, that much work needs to be done to make such systems more usable, both for clinicians and the care team in clinical care as well as for quality measurement. In part, a major challenge is rethinking who should be doing what within that system. For example, some pieces of the documentation can be done by other members of the care team. In addition, the model of the electronic health record is now based on paper records rather than thinking about the system as an electronic interface that includes clinical decision support and an interface to families and patients.

The ACA is changing the incentives for primary care providers and health care institutions. It is helping to create joint accountability through models like shared savings programs, health homes, patient-centered medical homes, and incentives for states and health plans, so improving mental health and substance use outcomes becomes a community responsibility. However, existing quality measures for mental health and substance use show only limited improvement, and measures assessing psychosocial interventions are lacking. Efforts to develop outcome measures for children and adolescents are under way but face challenges. New efforts to develop quality measures should focus on demonstrating how measures can inform clinical care and provide opportunities to monitor meaningful aspects of quality, Scholle concluded.

REFERENCES

Abner, K., R. A. Gordon, K. Kaestner, and S. Korenman. 2013. Does domain-specific quality of child care mediate associations between the type of care and child development. *Journal of Marriage and Family* 75:1203-1217.

AERA (American Educational Research Association)/APA (American Psychological Association)/NCME (National Council on Measurement in Education). 2014. Standards for Educational and Psychological Testing. Washington, DC: AERA. http://www.apa.org/science/programs/testing/standards.aspx (accessed March 25, 2015).

Baglivio, M., and R. Mark. 2014. *The Florida Department of Juvenile Justice Dispositional Matrix: A Validation Study.* Bureau of Research and Planning, Florida Department of Juvenile Justice. February 2014. http://www.djj.state.fl.us/docs/research2/the-fdjj-disposition-matrix-validation-study.pdf?sfvrsn=0 (accessed March 25, 2015).

Burchinal, M., K. Kainz, and Y. Cai. 2011. How well do our measures of quality predict child outcomes? A meta-analysis and coordinated analysis of data from large-scale studies of early childhood settings. In *Quality Measurement in Early Childhood Settings*, edited by M. Zaslow, I. Martinez-Beck, K. Tout, and T. Halle. Baltimore, MD: Brookes Publishing.

Cash, A. H., B. K. Hamre, R. C. Pianta, and S. S. Myers. 2012. Rater calibration when observational assessment occurs at large scale: Degree of calibration and characteristics of raters associated with calibration. *Early Childhood Research Quarterly* 27:529-542.

Child Trends. 2014. QRIS compendium. http://qriscompendium.org (accessed March 25, 2015).

Gordon, R. A., K. Fujimoto, R. Kaestner, S. Korenman, and K. Abner. 2013. An assessment of the validity of the ECERS-R with implications for assessments of child care quality and its relation to child development. *Developmental Psychology* 49:146-160.

Hamre, B., B. Hatfield, R. Pianta, and F. Jamil. 2014. Evidence for general and domain-specific elements of teacher-child interactions: Associations with preschool children's development. *Child Development* 85:1257-1274.

Harding, K. J., A. J. Rush, M. Arbuckle, M. H. Trivedi, and H. A. Pincus. 2011. Measurement-based care in psychiatric practice: A policy framework for implementation. *Journal of Clinical Psychiatry* 72(8):1136-1143.

IOM (Institute of Medicine). 1999. *To err is human: Building a safer health system*. Washington, DC: National Academy Press.

IOM. 2001. *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: National Academy Press.

IOM. 2005. *Improving the quality of health care for mental and substance-use conditions*. Washington, DC: The National Academies Press.

Keys, T. D., G. Farkas, M. R. Burchinal, G. J. Duncan, D. L. Vandell, W. Li, E. A. Ruzek, and C. Howes. 2013. Preschool center quality and school readiness: Quality effects and variation by demographic and child characteristics. *Child Development* 84(4):1171-1190.

Kracke, K. 2014. *Juvenile Justice Reform and Reinvestment Initiative (JJRRI)*. Presented at Institute of Medicine and National Research Council Workshop on Innovations in Design and Utilization of Measurement Systems to Promote Children's Cognitive, Affective, and Behavioral Health, Washington, DC.

Lewandowski, R. E., M. C. Acri, K. E. Hoagwood, M. Olfson, G. Clarke, W. Gardner, S. H. Scholle, S. Byron, K. Kelleher, H. A. Pincus, S. Frank, and S. M. Horwitz. 2013. Evidence for the management of adolescent depression. *Pediatrics* 32(4):e996-e1009.

Lipsey, M. W. 2009. The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims and Offenders* 4:124-147.

Lipsey, M. W., and F. T. Cullen. 2007. The effectiveness of correctional rehabilitation: A review of systematic reviews. *Annual Review of Law and Social Science* 3:297-320.

Loughran, T. A., E. P. Mulvey, C. A. Schubert, J. Fagan, A. R. Piquero, and S. H. Losoya. 2009. Estimating a dose-response relationship between length of stay and future recidivism in serious juvenile offenders. *Criminology* 47(3):699-740.

Pianta, R. C., K. M. La Paro, and B. K. Hamre. 2008. *Classroom assessment scoring system manual, PreK*. Baltimore, MD: Brookes Publishing.

Pincus, H. 2014. *Can child mental health cross the quality chasm? Children's behavioral health, healthcare reform, and the "quality measurement industrial complex."* Presented at Institute of Medicine and National Research Council Workshop on Innovations in Design and Utilization of Measurement Systems to Promote Children's Cognitive, Affective, and Behavioral Health, Washington, DC.

5

Toward Efficient and Sustainable Delivery of Interventions

Delivery of evidence-based prevention on a large scale has required the development of measures to monitor implementation, including selection and adoption of evidence-based programs, training and technical assistance, fidelity monitoring, and other factors. Measurement of implementation factors can provide information related to the decision-making process for evidence-based prevention, quality of program or service delivery, sustainability of implementation, and influence of implementation factors on intervention outcomes. In the final panel of the workshop, four presenters talked about lessons learned from implementation monitoring and recommendations for moving the field forward given the need for sustained program quality to improve outcomes.

MEASURING IMPLEMENTATION OF EVIDENCE-BASED PREVENTION: LESSONS FROM COMMUNITIES THAT CARE

The National Research Council (NRC) and Institute of Medicine (IOM) report *Preventing Mental, Emotional, and Behavioral Disorders Among Young People* (2009) summarized the burgeoning knowledge base for prevention science. It noted that 40 years of prevention science research advances have produced a strong understanding of the epidemiology and etiology of problem behaviors, a wealth of efficacy trials that have tested preventive interventions, and research findings on how to build an effective infrastructure to use prevention science to achieve community impact.

"If only it would be so easy," said Richard Catalano, the Bartley Dobb Professor for the Study and Prevention of Violence and co-founder of the

Social Development Research Group in the School of Social Work at the University of Washington. Despite the research advances that have occurred, prevention approaches that do not work or have not been evaluated are much more widely used than are evidence-based programs (Ringwalt et al., 2009). "The challenge for the 21st century is how can we build a prevention infrastructure to increase the use of tested and effective prevention policies and programs with fidelity and impact at scale," Catalano said.

A major difficulty in overcoming this challenge is that communities are different from one another, and communities need to decide locally what policies and programs they use. Overcoming this difficulty requires building the capacity of local coalitions to reduce common risk factors for multiple negative outcomes, according to Catalano, which in turn requires several actions:

- Assessing and prioritizing epidemiological levels of risk, protection, and problems
- Choosing proven programs that match local priorities
- Implementing chosen programs with fidelity to those targeted

Catalano used the Communities That Care (CTC) program as an example of this approach. CTC is a proven method to build community commitment and capacity to prevent underage drinking, tobacco use, and delinquent behavior, including violence. Developed in 1988, the program underwent 15 years of implementation and improvement through community input prior to being tested in a randomized controlled trial involving 12 pairs of matched communities across 7 states from Maine to Washington. The positive effects of the program have been independently replicated in a statewide test in Pennsylvania.

CTC has succeeded by building a prevention infrastructure, said Catalano. It first creates citizen coalitions of diverse stakeholders, including key leaders, elected officials, judges, faith leaders, parents, educators, and business leaders. These groups assess and prioritize risk, protection, and behavior problems with a student survey. They address locally prioritized risk with tested and effective prevention programs that are matched to those priorities. Then they support and sustain high-fidelity implementation of these chosen programs.

The trial of CTC provided funding for the selected programs, training in the selected programs, and fidelity and reach monitoring of the selected programs, along with funding for a community coordinator and training. It also provided weekly phone technical assistance and two site visits per year. Assessment surveys are done every 2 years.

The timeline for implementation encompasses the process, evaluation, and measurable outcomes (see Figure 5-1). Milestones and benchmarks

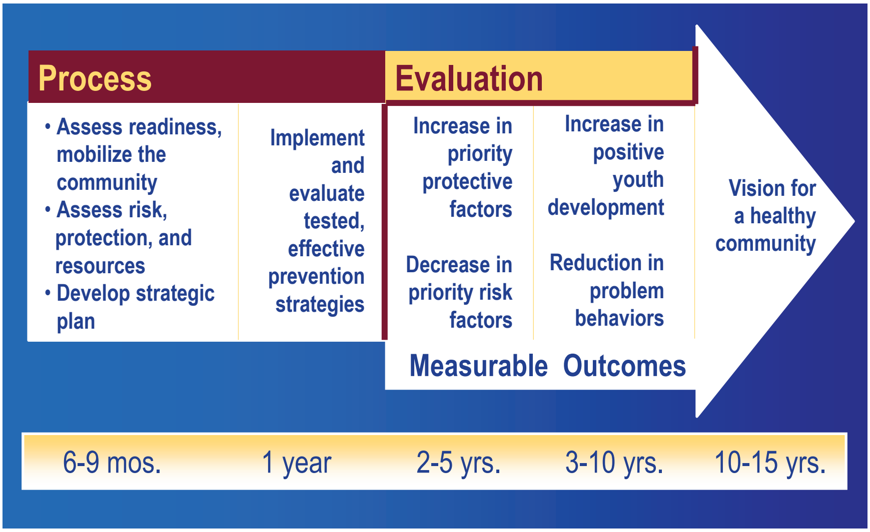

FIGURE 5-1 The Communities That Care timeline. The timeline calls for measurable outcomes 2 to 5 years after the process begins.
SOURCE: Catalano, 2014.

assess the key components of CTC's strategy, including goals, steps, actions, and conditions needed for CTC implementation to build prevention infrastructure. The milestones and benchmarks are listed in CTC training manuals and discussed in training workshops, incorporated into the community coordinator's job performance objectives, and reviewed by technical assistance providers and coordinators during weekly phone calls.

As an example of phase 1, readiness for CTC, Catalano cited the milestone of "the community is ready to begin CTC," with the more specific benchmark "A key leader 'champion' has been identified to guide the CTC process." For phase 5, implementing the Community Action Plan, a milestone is "implementers of evidence-based programs, policies, or practices have the necessary skills, expertise, and resources to implement with fidelity," with the benchmark "implementers have received needed training and technical assistance."

The CTC implementation has been maintained with fidelity over time, Catalano noted. The percentage of milestones completed across communities and raters ranged between 83 percent and 91 percent in the fifth year (Fagan et al., 2009).

A second measurement instrument is the CTC youth survey, which assesses young people's experiences and perspectives in the sixth, eighth, tenth, and twelfth grades. It identifies levels of risk and protective factors for substance use, crime, violence, and depression at the state, district, city, school, or neighborhood levels and provides a foundation for the selection of tested and effective actions.

The selection of programs based on the survey results presents a challenge to measurement strategies, noted Catalano, because it generates the need to evaluate fidelity across a range of programs while continuing to encourage local ownership, high fidelity, and the sustainability of prevention programs. All CTC sites are expected to achieve a high level of fidelity, including

- **Adherence—**Implementing the core content and components
- **Delivery of sessions—**Implementing the specified number, length, and frequency of sessions
- **Quality of delivery—**Ensuring that implementers are prepared, enthusiastic, and skilled
- **Participant responsiveness—**Ensuring that participants are engaged and retaining material

To measure fidelity, assessment checklists were obtained from developers or created by research staff. The checklists were completed by program staff, coalition members, and reviewed locally as well as analyzed at the University of Washington. These checklists showed high levels of implementation adherence and participant responsiveness, said Catalano.

Catalano concluded with the following recommended actions:

- Build capacity and provide tools (such as the CTC milestones and benchmarks) to achieve effective prevention infrastructure.
- Build capacity and provide tools to assess and prioritize local risk, protection and youth outcomes, match priorities to evidence-based programs, and repeat assessment periodically.
- Build capacity and provide tools to ensure program fidelity and engagement of target population.
- Create citizen–advocate–scientists to affect risk, protection, substance use, delinquency, and violence community wide.

CTC has been able to get communities to choose the right programs, implement them with fidelity, and achieve positive outcomes, Catalano said. Furthermore, creating citizen-scientists has built advocacy for evidence-based prevention at the local level. These citizens are in the best position to support evidence-based practices in their communities, the best choices for children and families, and taking evidence-based prevention to scale.

In response to a question, Catalano pointed to the interplay that can occur between states and local communities, with each interested in what is happening at the other level. For example, communities can be the impetus to get a state involved, or a two-way feedback loop can result in modifications to programs at both levels.

THE STAGES OF IMPLEMENTATION COMPLETION

A growing body of measures target key aspects of implementation, including organizational culture and climate, organizational readiness, leadership, attitudes toward evidence-based practices, and the use of research evidence. However, there remains a gap in the measure of the implementation process itself, said Lisa Saldana, senior research scientist at Oregon Social Learning Center. This gap encompasses measures of the implementation process, including the rate of implementation, implementation activities, and patterns of implementation behavior, along with measures of implementation outcomes, including milestones and penetration.

Implementation of evidence-based practices entails extensive planning, training, and quality assurance, Saldana noted. It also requires a complex set of interactions among developers, system leaders, frontline staff, and consumers, who ultimately have to buy in to whether the interventions meet their needs. This is a recursive process of well-defined stages or steps that are not necessarily linear, said Saldana.

Little is known about which methods and interactions are most important for successful implementation. In addition, little is known about how and if the process influences successful outcomes. "Just because this is the way we have been doing it does not necessarily imply that that is what is necessary," said Saldana.

The Stages of Implementation Completion (SIC) was developed as part of an implementation trial focused on the scale-up of the Multidimensional Treatment Foster Care (MTFC) intervention (Chamberlain et al., 2011). Fifty-three sites in California and Ohio involved in youth foster care were observed over the entire implementation process. Some sites dropped off along the way, providing an opportunity to observe the factors that influenced how far they got in the implementation process. A tool was developed to measure the rate and thoroughness of implementation. Eight stages of implementation were measured, from engagement to competency. The tool was not only data driven but *date* driven, said Saldana, in that activities were tracked along with the dates on which they occurred. The SIC involves assessment of implementation behavior of different levels of agents. The eight stages of implementation were broken down into the three phases of pre-implementation, implementation, and sustainment, involved different levels of agents (see Figure 5-2).

The SIC yields three overall scores. The first is duration, or the amount of time that it takes to complete a stage. The second is proportion, or the

	8 Stages:	**Involvement:**
Pre	1. Engagement	System leader
	2. Consideration of feasibility	System leader, agency
	3. Readiness planning	System leader, agency
Imp	4. Staff hired and trained	Agency, practitioner
	5. Adherence mointoring established	Practitioner, client
	6. Services and consultation	Practitioner, client
	7. Ongoing services consultation, fidelity, feedback	Practitioner, client
Sus	8. Competency (certification)	System leader, agency, practitioner, client

FIGURE 5-2 The eight Stages of Implementation Completion. The stages are divided into pre-implementation, implementation, and sustainment phases.
SOURCE: Saldana et al., 2011.

number of recommended activities that are completed. The third is the stage score, which is how far along within the implementation process an agency or organization has progressed.

The recursive nature of implementation makes it impossible to look at each stage across the entire SIC, said Saldana. Agencies might be in more than one stage at a time, or they might also be going backward because of feedback from stakeholders. Rasch-based modeling helps account for these difficulties, said Saldana. This modeling has demonstrated reliability across all eight stages. It also has demonstrated face validity while identifying three clusters of sites based on their pre-implementation behavior: those that complete activities relatively quickly, those that are relatively slow, and noncompleters. Finally, it has demonstrated predictive validity in that sites that both took longer to complete each stage and completed fewer activities had a significantly lower hazard of successful program start-up during the study period (Saldana et al., 2011). “You have to hit a little bit of a sweet spot,” said Saldana. “You don’t want to go too fast. You also don’t want to go too slow. You don’t necessarily need to complete all of the activities, but there are some critical key activities that, if you skip, your chances of success are going to be decreased.”

These initial results have been successfully replicated in other sites. In a sample of 75 recent MTFC interventions, sites were successfully clustered, failed sites spent significantly longer in pre-implementation than successful sites, and sites that took longer to complete stages one through three had a significantly lower hazard of successful program start-up (Saldana, 2014).

In response to a question, Saldana noted that the pre-implementation process is also key, in that it predicts competencies that are needed for the sustainment phase. And during the sustainment phase, all levels of agents need to be involved for sustainability within a community, with ongoing dialogue and feedback between programs and policy makers.

Saldana and her colleagues are now trying to adapt the SIC to other programs and service sectors. Will similar utility be found, she asked. Is there a universality in implementation? One finding from a much broader study of interventions in schools, child welfare, juvenile justice, and substance abuse is that there are more similarities than differences in the types of activities they are including in their implementation strategies (Saldana et al., 2014). As a result, a universal SIC was being developed at the time of the workshop, and "preliminary studies of the universal SIC psychometrics are very promising," according to Saldana. This could lead to a standardized method of measuring implementation processes and milestones across different practices and service sectors, which could make it possible to detect sites that are at risk for implementation failure and evaluate alternative implementation strategies.

SIC can reliably distinguish poor versus good performers, Saldana concluded. It can reliably distinguish between different implementation strategies and provides meaningful prediction of implementation milestones. Results from the SIC also can provide data-driven evidence that can be used to engage in conversations with high-level systems leaders, who also have a powerful influence on implementation. For example, Saldana and her colleagues have been developing an interactive website that can provide high-level systems leaders with information on the value of key elements in the implementation process.

IMPLEMENTING PROGRAMS WITH FIDELITY

When programs that were developed in academia under well-controlled conditions are transferred into the community, they are not necessarily implemented the same way they were while being developed. Marion Forgatch, senior scientist emerita at the Oregon Social Learning Center and executive director of Implementation Sciences International Inc., illustrated this process by discussing implementations of the Parent Management Training Oregon (PMTO) model, which provides interventions to parents to help protect children and enhance their development (Forgatch and Patterson, 2010). Randomized controlled trials in a variety of sample populations have linked parenting practices not only with positive outcomes for children but with improved outcomes for the parents.

The program has been implemented in a number of locations through a process known as full transfer. Control of the program is placed into the hands of the community. As part of a broader infrastructure, a governing authority is tasked with sustaining model fidelity and effective treatment outcomes. This process starts with a visionary leader or group committed to affecting lasting change, said Forgatch. This person or group needs strong social and political capital and the resources to support the necessary structure. A leader also needs longevity; "They are not here today and gone tomorrow," said Forgatch. "It starts at the top, but it is sustained at the bottom by the families and the practitioners who are satisfied with the methods and the outcomes."

The full transfer approach has tremendous capacity to improve reach, because each new generation of practitioners who are trained can in turn train other practitioners. For example, from a first generation of 29 practitioners trained in Norway, more than 1,000 have now been trained in an intervention provided to families in clinics. "We spend a lot of time training the first generation," said Forgatch. "These are the people who carry the program forward, so they have to know the model, and they have to know the procedures really well." The program has been implemented throughout the United States, in Europe, and in locations in Africa and Latin America.

The implementation process is evaluated through an instrument called the fidelity of implementation (FIMP). This measure is based on direct observation of therapy sessions, with ratings by observers of practitioner adherence to the model and competence. Ratings are in five categories:

1. **Knowledge**—Proficiency in understanding and applying core components
2. **Structure**—Session management, pacing/timing, responsiveness
3. **Teaching**—Promotes mastery, use of role play, problem solving
4. **Process**—Clinical and strategic skills, supportive context for learning
5. **Overall**—Growth, satisfaction, likely return, adjust context, difficulty

Sessions are scored to evaluate training and certification, drift across and within generations of practitioners, and to assess mechanisms of change. The only people able to score FIMP are certified therapists to ensure that competent adherence to the model is being properly assessed.

Technology is a central feature of the evaluation. A database and web portal known as FIMP Central provide videotapes of sessions where coders can access the video and make their ratings, with the ratings being retained on the portal so supervisors of the coding teams can assess reliability. Online and in-person meetings further enhance the reliability of assessments.

FIMP has several uses, including

- Teaching tool for training and coaching
- Evaluation of training and certification
- Evaluation of drift across generations
- Evaluation of drift within a generation
- Assessment of mechanisms, such as whether fidelity predicts improved parenting and whether fidelity predicts improved parent outcomes

In the last category, for instance, Forgatch cited a recent study showing that fidelity assessed during treatment predicts change in observed parenting before and after treatment and change in child outcomes (Forgatch and Domenech Rodríguez, in press). "We are very pleased with those findings. It said, 'Yes, our fidelity measure is indeed predicting the mechanism of change in our model.'"

FIMP also has been used to study the fidelity of implementation over successive generations of training. These results showed that fidelity does drop in the second generation of training, but it then recovers as successive generations of trainers improve their training skills (see Figure 5-3).

MEASURING IMPLEMENTATION FIDELITY USING COMPUTATIONAL METHODS

In the past, measuring the fidelity of implementation has generally involved direct observation or the observation of recordings, which requires highly trained individuals and can be time consuming and costly. But the fidelity of implementation also can be measured using computational methods, said Carlos Gallo, a research assistant professor in the Department of Psychiatry and Behavioral Sciences at Northwestern University. He provided two examples as proofs of concept: one from a program known as *Familias Unidas*, and the other from audio recordings of an intervention known as the Good Behavior Game.

Familias Unidas is an evidence-based parent training intervention for Hispanic youth. It is delivered in family visits at home by a school counselor, generally in a bilingual context, where an adolescent or young adult is speaking English, a parent is speaking Spanish, and a school counselor is switching between the two. The goal of the computational analysis is to recognize the type of questions or statements the facilitator is making to the parents to engage them and communicate acceptance, trust, and respect. The evaluation focuses on an aspect of fidelity known as joining. Linguistic structures linked to joining include statements (such as "You like school."), yes/no questions (such as "Do you like school?"), and open-ended questions

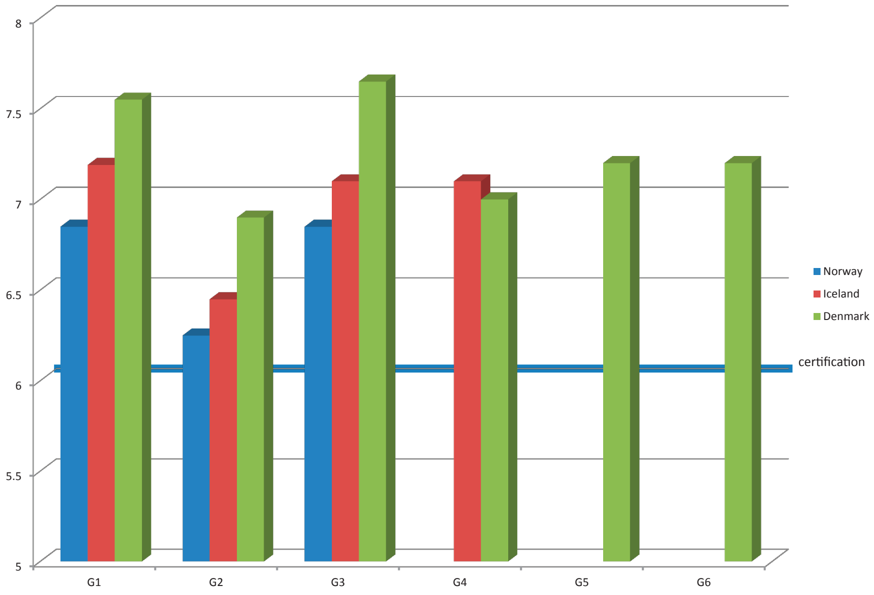

FIGURE 5-3 Implementation fidelity over generations of trainers. Though implementation fidelity declined in the second generation of trainers, it did not drop below the certification level, and fidelity recovered in subsequent generations.
NOTE: X-axis shows trainer generation number, Y-axis depicts implementation fidelity level.
SOURCE: Forgatch, 2014.

(such as “Why do you like school?”). The more open-ended the question, the better the degree of joining.

Gallo and his colleagues wrote a program that scans for these linguistic patterns and then rates each sentence (Gallo et al., 2014). The program can identify different kinds of statements with a high degree of reliability compared with a human rater. But the cost of a human rater is approximately $800 per session, which, said Gallo, “becomes prohibitively expensive as local agencies want to pick up on these interventions and carry on the same work carefully.”

The second example he described is based on nonlinguistic cues. The Good Behavior Game is a universal classroom behavior management strategy for first grade teachers that has been shown to influence adolescent and young adult drug abuse, sexual risk behavior, delinquency, and suicidal behavior (Kellam et al., 2008). It calls for the intervention to be delivered by the teacher in a neutral tone, not an angry, frustrated, or sad tone. “You can imagine how challenging this can be when you are giving instructions to 40 kids that are running around in a classroom. You need to keep sane and speak neutrally." Furthermore, speaking neutrally is a key competence in this intervention, though coaches for the intervention find that doing so is challenging.

In a laboratory test, Gallo and his colleagues have taken audio clips spoken by an angry person or a person speaking neutrally and have enabled a computer to learn about features of the frequency domain for these samples. By analyzing patterns in voice frequency, the computer can determine whether something is said in an angry or neutral way. The program is 98 percent accurate in distinguishing a neutral from an angry tone, and 87 percent successful in distinguishing a neutral from an emotional tone. "This is something that we can do in a split second throughout an hour session," Gallo noted. Teachers then can be asked in the same session when and why an emotion started and how it progressed, rather than trying to remember days after a session was recorded.

Computational linguistics has several other advantages over human raters, Gallo noted. It can scan 100 percent of a session, not just a sample of that session, and it can do so on a much more detailed level. Eventually, such programs could be adapted for tablets or phones, where they could display summary statistics of what happened in a session. They can flag particular moments in a session for review and analysis. Facilitators can receive global assessments of their delivery of an intervention and compare themselves with peers. Such programs also could point to places where teachers or other facilitators are adapting an intervention for a particular circumstance, which could provide valuable lessons. "These facilitators know their community a lot better. They know what are the needs of the target audience."

These techniques are developing very rapidly, Gallo noted, driven partly by the very wide range of applications for computational linguistics. For example, the programs they have developed could be applied to many other interventions because the linguistic processes they have analyzed occur in many contexts. Furthermore, as several workshop participants observed in the discussion session, resistance to such applications is diminishing as people become more accustomed to being recorded.

REFERENCES

Catalano, R. F. 2014. *Measuring implementation of evidence-based prevention to improve impact and sustainability: Lessons from Communities That Care*. Presented at Institute of Medicine and National Research Council Workshop on Innovations in Design and Utilization of Measurement Systems to Promote Children's Cognitive, Affective, and Behavioral Health, Washington, DC.

Chamberlain, P., C. H. Brown, and L. Saldana. 2011. Observational measure of implementation progress in community based settings: The Stages of Implementation Completion (SIC). *Implementation Science* 6:116.

Fagan, A. A., K. Hanson, J. D. Hawkins, and M. Arthur. 2009. Translational research in action: Implementation of the communities that care prevention system in 12 communities. *Journal of Community Psychology* 37(7):809-829.

Forgatch, M. S. 2014. *Fidelity: Preventing drift*. Presented at Institute of Medicine and National Research Council Workshop on Innovations in Design and Utilization of Measurement Systems to Promote Children's Cognitive, Affective, and Behavioral Health, Washington, DC.

Forgatch, M. S., and M. M. Domenech Rodríguez. in press. Interrupting coercion: The iterative loops among theory, science, and practice. In *Oxford Handbook of Coercive Relationship Dynamics* edited by T. J. Dishion and J. J. Snyder. New York: Oxford University Press.

Forgatch, M. S., and G. R. Patterson. 2010. Parent Management Training–Oregon Model: An intervention for antisocial behavior in children and adolescents. In *Evidence-Based Psychotherapies for Children and Adolescents*, 2nd edition, edited by J. R. Weisz and A. E. Kazdin. New York: Guilford.

Gallo, C. G., H. Pantin, J. Villamar, G. Prado, M. Tapia, M. Ogihara, G. Cruden, and C. H. Brown. 2014. Blending qualitative and computational linguistics methods for fidelity assessment: Experience with the Familias Unidas preventive intervention. *Administration and Policy in Mental Health*. E-pub ahead of print. http://dx.doi.org/10.1007/s10488-014-0538-4 (accessed August 4, 2015).

Kellam, S. G., C. H. Brown, J. M. Poduska, N. S. Ialongo, W. Wang, P. Toyinbo, and H. C. Wilcox. 2008. Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. *Drug and Alcohol Dependence* 95:S5-S28.

NRC (National Research Council) and IOM (Institute of Medicine). 2009. *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. Washington, DC: The National Academies Press.

Ringwalt, C., A. A. Vincus, S. Hanley, S. T. Ennett, J. M. Bowling, and L. A. Rohrbach. 2009. The prevalence of evidence-based drug use prevention curricula in U.S. middle schools in 2005. *Prevention Science* 10(1):33-40.

Saldana, L. 2014. The stages of implementation completion for evidence-based practice: Protocol for a mixed methods study. *Implementation Science* 9(43):1-11.

Saldana, L., P. Chamberlain, W. Wang, and H. Brown. 2011. Predicting program start-up using the Stages of Implementation Measure. *Administration and Policy in Mental Health and Mental Health Services Research* 39(6):419-425.

Saldana, L., J. Chapman, H. Schaper, and M. Campbell. 2014. *Standardized measurement of implementation: The universal SIC*. Presented at the 7th Annual Conference on the Science of Dissemination and Implementation, Bethesda, MD.

6

Breakout Group Discussions

One session of the workshop was devoted to a set of three breakout groups that examined strategies to improve measurement and the integration of data at the (1) individual, (2) community, and (3) state and system levels. This final chapter of the workshop summary describes the major points discussed by those three groups. It also includes observations made by workshop participants in general discussions of the major issues raised at the workshop.

MEASUREMENT AT THE INDIVIDUAL LEVEL

As children age and develop, measurement issues at the individual level change. In addition, children are situated within families and broader social systems that both influence and respond to their development. These complexities create challenges in measuring the strengths and the needs of children as well as services, processes, and resulting outcomes.

The breakout group began by discussing several gaps in measurement. Several important features of families are not currently well measured, such as family structure, family functioning, parental mental health, and child and family strengths. Some children may be more biologically susceptible to adversity, but individual-level markers are not currently measured or incorporated into decision making or datasets. Better integration of preschool and school data into measurements from other sectors could yield indicators of how a child is doing socially, academically, cognitively, and emotionally.

In the measurement of services and processes, a team approach could fill existing gaps, especially with colocation and coordination of care. With regard to outcomes, brief and straightforward outcome measures could reflect key principles. For example, is a child on a trajectory to graduate from high school? Other issues involving measurement gaps include the information that can feasibly be derived from primary care and other service systems, the ability of data and measures to cross-talk among systems, the gradual transition from the parent or caregiver to the child or adolescent as the source of information, and whether data are readily available or are difficult to access.

The breakout group also discussed obstacles to data integration. Regulations deriving from the Health Insurance Portability and Accountability Act (HIPAA) and the Family Educational Rights and Privacy Act (FERPA) can reduce integration and cross-talk, despite opportunities to use data covered by these regulations productively. Integration is also an issue between research studies and the implementation of interventions in practice. Differing languages and measures point to a need for more systematic and standardized measures in research, policy, and practice, including measures of implementation fidelity.

Developmental changes also can be a factor in data integration. For example, young people who are aging out of care systems can lose services; they also can gain legal rights and greater control over data. Such transitions complicate the challenge of using data productively. Similarly, differing cultures and languages across professions can have a strong influence on the ability to share data and set measurement priorities. For instance, in the context of adult health care, behavioral health may refer to chronic disease self-management, self-efficacy, management of substance abuse, or other personal attributes.

Recent data breaches have increased concerns about the loss of control over data and how data might be used. For example, if a parent mentions in one context that he or she has mental health issues, will that information be disseminated in ways that cannot be controlled and could have harmful consequences? In general, ethical issues in measurement and data use require greater emphasis and study, members of the breakout group observed.

Data sharing requires infrastructure, time, energy, and a willingness to overcome a natural reluctance to share potentially sensitive information, but no system now exists to finance such sharing. Basic questions such as who will be entering data into a system, who cleans up the data, and who makes the data accessible remain unanswered. Families may be able to enter some data into a system, but who will work with families to determine what they want, need, and trust?

Determining data collection approaches, terms, and measures will require having different systems talking with each other to develop shared understandings. Workforce development and training also will be required to understand cross-sector approaches and overcome cultural differences, breakout some group participants observed. The U.S. Department of Health and Human Services could advance this agenda by prioritizing measurement and assessment, by creating an office focused on child health, and by organizing a White House conference on children and youth to reinvigorate public support for improving youth services.

Measures could be bundled to provide greater insights into system issues. For example, the Affordable Care Act (ACA) includes provisions that could advance needed steps, though the ACA has few provisions directed specifically toward children in this area.

Families need to feel that they own and trust measurement processes if they are to participate fully. Could service providers work with families so they are involved in entering and using data? This would allow the building of a culture in which measurement tools contribute to learning and to the building of relationships between service providers and family members. Data collection could also be structured in such a way as to counter the fragmentation of health care if data were made available not just to single providers but to teams of service providers in ways that maintain the trust of individuals.

Individuals and families often do not have ready access to the services they need. For example, would it be possible for them to get better access to providers, and especially specialists, during the hours and days of the week that are convenient to working families? Measures could be developed to study and improve this issue.

Finally, the breakout group on measurement issues at the individual level discussed the need for public use measures. Assessment tools that are copyrighted or otherwise restricted can limit measurement capabilities. With greater availability and numbers of tools, the most appropriate assessment for a particular need could be identified and used. Similarly, common measures that are usable across research, policy, and practice could increase the scope, influence, and power of those measures. One approach, for example, could be to extend the lexicon developed by the Agency for Healthcare Research and Quality (AHRQ) around behavioral and mental health and use it across domains and sectors.

MEASUREMENT AT THE COMMUNITY LEVEL

The breakout group on issues at the community level talked about the definition of the term *community*. Most often, communities are defined geographically, but they could also be defined socially, economically, culturally, technologically, or along many other dimensions, and different forms of community call for different measurement tools.

A major issue discussed by the breakout group was that a great deal of data are already being collected. These data originate in a large number of programs, jurisdictions, systems, and levels of government. They include economic data, public health data, education data, health data, census data, and data from the foster care, juvenile justice, and other social service systems. Data also can come from unconventional sources; for example, the website Zillow provides a wealth of community-level data.

However, this great wealth of data is highly fragmented, which results in an inability to form a holistic view of children's health and well-being. Data collected at the federal, state, and local levels are not well connected, and public- and private-sector data are isolated from each other. Greater integration, dissemination, and use of these disparate data sources could help achieve many of the objectives sought by workshop participants while reducing duplicated effort.

Some members of the breakout group noted that funding could be directed explicitly toward the analysis and use of data and not just the collection of data could capitalize on the data that exist. Greater emphasis on the use of data in funding streams could incentivize such a change; another possibility would be to make future funding dependent at least in part on the use of the data generated by current funding. One interesting proposal discussed by the breakout group was to integrate all of the existing forms of data for a single community and determine both the potential and remaining gaps in such a dataset.

Greater integration of data will require attention to privacy issues, some members of the breakout group noted. Privacy issues may evolve as people begin to create their own personal data vaults and as a result of new policies. Greater public understanding of data issues also could change public attitudes toward privacy.

Considerable work is being done on the development of new and better assessment tools at the community level. However, some forms of data, such as positive indicators of children's health and well-being or various forms of contextual and environmental data, are still notably lacking. As with existing measurement tools, the data from new measures could be used more efficiently and effectively to solve existing problems and identify strengths.

The community has a vested interest in the health and well-being of its children, even more so than do state or federal governments. This interest in children could be the foundation for greater trust and collaboration in building community-oriented data systems.

MEASUREMENT AT THE STATE AND SYSTEMS LEVEL

Integration of data is a particularly pressing concern at the state level, noted many participants in the third breakout group. State-level departments that deal with issues affecting children's health and well-being may not want to share information with each other, recognizing only the potential downsides rather than the benefits of such sharing. However, some states have forged quite successful data-sharing systems.

An interesting project would be to survey state systems to determine how they are sharing data, how they are disseminating data, and whether they have developed systems that could be disseminated more widely across the states. A variety of issues could be examined in such a project, from general ethical concerns to practical issues such as developing an adequate data-sharing infrastructure or building relationships between researchers and state agencies to collect and use data.

States currently collect and have access to a wide variety of data, but gaps remain. Examples mentioned by breakout group members include predictor variables, like risk and protective factors; the specific services individual children are receiving; longitudinal information on children; and the early identification of problems. Children need safe, stable, and nurturing relationships, one breakout group member mentioned, yet little information is collected directly about these relationships.

States have an opportunity as the health care system adopts electronic health records to use data from these records to further children's health and well-being. In particular, young children often have been overlooked in the past and could be a point of emphasis in collaborative efforts. States would need to work with software developers to understand what types of data need to be collected and how those data could be collected rather than grafting data systems onto electronic health records once they are up and running.

The breakout group discussed the interactions among state-level personnel and the frontline personnel who are often collecting data at the local level. Frontline personnel need to understand the purpose of the data being collected to build support for the data collection system and to provide input on how data can be used. They also need technical assistance from the states if they are to gather information accurately and reliably. Finally, communities need to receive information back from the state to maintain and extend their data collection efforts, according to some members of the breakout group.

FINAL COMMENTS

Finally, several presenters and participants in the workshop made comments that elaborated on or extended the observations of the breakout groups.

Felisia Bowen, Rutgers School of Nursing, emphasized the need to develop culturally appropriate instruments, given the diversity of the U.S. population. The best way to do so, she added, is to invite the members of culturally distinct groups to share in the development of instruments, which also educates them in the process and increases the likelihood of sharing.

The children in U.S. schools speak many different languages, observed Laurel Leslie, Tufts University School of Medicine, and many live in poverty, which can be expected to have effects on their health literacy. As Hendricks Brown of Northwestern University pointed out, a tool is not adequate if it leaves out large portions of the population that have less access to health care.

Andy Shih of Autism Speaks pointed to the need to conduct field trials of new instruments in a range of cultural settings, as was done with a recent diagnostic screening tool for autism that can be administered by nonspecialists. Conducting field trials in different cultures makes it possible to identify elements that may be transcultural so that a tool can be generalized.

Robert Goerge, University of Chicago pointed out that many of the barriers to better measures are not technical but social and cultural. Laws and regulations in the United States protect privacy and independence, which limit what can be done with data after it is collected. As Mary Ann McCabe, George Washington University, added, these considerations differ among data types. For example, deidentified administrative data tends to be viewed differently than personally identifiable health care data, particularly in sensitive areas such as mental health.

David Keller observed that one way to develop new tools is to require the use of such tools in payment systems. For example, value-based payment formulas have been tied to tool development in adult health care, though the same measures have not been applied in pediatrics. In general, added Brown, actionable information from a more comprehensive and integrated data system could support the investments in infrastructure needed to build and sustain such a system. Brown also pointed to the opportunities to use qualitative information more strategically, particularly in monitoring the implementation and adaptation of programs.

Harold Pincus observed that, in health care, one challenge is to redesign the workforce to establish base measurements of care, and an electronic health record (EHR) can be built around that. In that case, an EHR would be usable not just by clinicians but by patients and parents. The Office of the National Coordinator for Health Information Technology is working on developing standards for behavioral health EHRs, and other groups are thinking about how patient portals can be used more effectively.

Jeff Schiff, Minnesota Department of Human Services, pointed more generally to the potential of new technologies to gather patient-reported outcomes. “All of us will, at some time in the next 15 minutes, look at our cell phones,” he said. “We have to start thinking about different modalities of patient-reported outcomes. And then not just look at patient-reported satisfaction or patient-reported well-being but how people are actually functioning in their families and in their communities."

Finally, Schiff emphasized the power of forceful advocacy. "If we can be a little impatient and a little persnickety about this and get to be a thorn in a bunch of people's shoes, we may be able to move forward," he said. Legislators tend to react more forcefully to stories than to data, Schiff said, so he always tries to talk about data in the form of stories. At the same time, a little bit of data can go a long way, especially if it drives home an important point. "If we can get the data to be used, we can make a difference."

A

Workshop Agenda

Innovations in Design and Utilization of Measurement Systems to Promote Children's Cognitive, Affective, and Behavioral Health

November 5–6, 2014

National Academy of Sciences Building
2101 Constitution Avenue, NW, Washington, DC

AGENDA

There are a number of measurement systems to monitor the well-being of children in the United States and to guide improved delivery of health care, social, and educational services for children at the local, regional, and national levels. The usefulness of these systems may be improved by the use of tools, such as data linkage and integration; quality improvement methods; and feedback mechanisms to enhance sustainable delivery of interventions. This workshop aims to discuss examples of innovative design and utilization of measurement systems to collect and analyze data relevant to promoting children's cognitive, affective, and behavioral health, and to identify common features of these systems to guide future research and practice.

Day 1, Wednesday, November 5, 2014 (LECTURE ROOM)

8:00 a.m. **BREAKFAST**

8:45 a.m. **Welcome and overview of meeting agenda**
C. Hendricks Brown and William Beardslee, Forum Co-Chairs

8:50 a.m. **Panel 1: Maximizing the Value of National, State, and Local Measurement Systems for Research and Policy**

Speakers will present examples of the successful use of data integration to inform research and policy pertinent to promoting children's cognitive, affective, and behavioral health, and discuss how measurement systems might be improved.

Moderator: Ruth Perou, Centers for Disease Control and Prevention, Planning Committee

Speakers:
- Anne Sheridan, Executive Director, Governor's Office for Children, Maryland
- Vetisha McClair, Division of Research on Vulnerable Populations and Care Transformation, Center for Medicare & Medicaid Innovation
- Robert Goerge, Chapin Hall at the University of Chicago

9:40 a.m. **Questions and discussion**

10:00 a.m. **Keynote Presentation: Using Data to Improve Patient-Centered Outcomes: A View from the Medicaid Medical Directors**
Jeff Schiff, Chair, Medicaid Medical Directors Network

10:20 a.m. **Questions and discussion**

10:30 am. **BREAK** (**15 minutes**)

10:45 a.m. **Panel 2: Measurement Systems to Assess Individual- and Population-Level Change**

Speakers will describe innovations in the use of measurement systems to monitor changes in children's health at the individual and population levels.

Moderator: William Beardslee, Harvard Medical School, Planning Committee

Speakers:

- Robert Orwin, Westat
- Jane Pearson, Division of Services and Intervention Research, National Institute of Mental Health
- Greg Farber, Office of Technology Development and Coordination, National Institute of Mental Health
- Catherine Bradshaw, University of Virginia
- Kareemah Abdullah, Community Anti-Drug Coalitions of America

12:05 p.m. **Questions and discussion**

12:25 p.m. **Preview of afternoon session**
C. Hendricks Brown and William Beardslee, Forum Co-Chairs

12:30 p.m. **LUNCH**

1:30 p.m. **Breakout group discussions**

Forum members and workshop attendees will divide into three groups to discuss how to improve measurement and integration of data at the (1) individual-, (2) community-, and (3) state-level. (Note: Group assignments and specific questions for discussion will be available at the meeting.)

3:15 p.m. **BREAK (15 minutes)**

3:30 p.m. **Breakout group reports**

4:00 p.m. **Full group discussion**

4:50 p.m. **Closing remarks and preview of next day**
C. Hendricks Brown and William Beardslee, Forum Co-Chairs

5:00 p.m. **Adjourn for day**

Day 2, Thursday, November 6, 2014 (NAS 120)

8:30 a.m. **BREAKFAST**

9:00 a.m. **Welcome and reflections from Day 1**
C. Hendricks Brown and William Beardslee, Forum Co-Chairs

9:05 a.m. **Panel 3: Utilizing Quality Measures to Facilitate System Change**

Panelists will present examples of how quality measures have been used to facilitate system change in early childhood, juvenile justice, and health care settings.

Moderator: Dara Blachman-Demner, National Institute of Justice, Planning Committee

Speakers:
- Rachel Gordon, University of Illinois at Chicago
- Kristen Kracke, Office of Juvenile Justice and Delinquency Prevention
- Harold Pincus, Columbia University
- Sarah Scholle, National Committee for Quality Assurance

10:10 a.m. **Questions and discussion**

10:30 a.m. **BREAK (15 minutes)**

10:45 a.m. **Panel 4: Toward Efficient and Sustainable Delivery of Interventions**

Speakers will discuss tools developed to measure implementation of evidence-based preventive programs at scale and offer suggestions about how such tools can be used to improve the sustainable delivery of programs.

Moderator: Eve Reider, National Institute on Drug Abuse, Planning Committee

Speakers:
- Richard Catalano, University of Washington
- Lisa Saldana, Oregon Social Learning Center
- Marion Forgatch, Oregon Social Learning Center
- Carlos Gallo, Northwestern University

11:50 a.m.	**Questions and discussion**
12:10 p.m.	**Closing remarks** C. Hendricks Brown and William Beardslee, Forum Co-Chairs
12:15 p.m.	**Adjourn workshop**

WORKSHOP PLANNING COMMITTEE

C. Hendricks Brown, Ph.D., Northwestern University (*Chair*)
William R. Beardslee, M.D., Boston Children's Hospital and Harvard Medical School
Dara Blachman-Demner, Ph.D., National Institute of Justice
David A. Brent, M.D., M.S.Hyg., University of Pittsburgh
Kimberly E. Hoagwood, Ph.D., New York University
Jennifer Kaminski, Ph.D., Centers for Disease Control and Prevention
Laurel K. Leslie, M.D., M.P.H., Tufts University
Ruth Perou, Ph.D., Centers for Disease Control and Prevention
Eve E. Reider, Ph.D., National Institute on Drug Abuse
Andy Shih, Ph.D., Autism Speaks
Belinda E. Sims, Ph.D., National Institute on Drug Abuse

B

Biographies of Workshop Speakers

Kareemah Abdullah is the Director of the National Community Anti-Drug Coalition Institute (Institute) and the Vice President of Training Operations for Community Anti-Drug Coalitions of America (CADCA) in Alexandria, Virginia. The Institute established by Congress under the Drug-Free Communities Support Act is administered by the Executive Office of the President, Whitehouse Office of National Drug Control Policy (ONDCP) and managed by the Substance Abuse Mental Health Services Administration (SAMHSA). Ms. Abdullah is responsible for executive and strategic direction of the Institute with an emphasis on the development of coalitions serving economically disadvantaged areas throughout the United States, territories, and tribal communities. As the Vice President for Training Operations, Ms. Abdullah is the chief architect for CADCA's youth programs, featuring the National Youth Leadership Initiative (NYLI), military services including CADCA's VetCorps Program, as well as fee-for-service operations. Ms. Abdullah also served as the Vice President for Development accountable for CADCA's strategic partnership and resource procurement. For 7 years, she was the Institute's Deputy Director for Training and Technical Assistance. Under her leadership, the National Coalition Academy was created in partnership with the National Guard Bureau and its Regional Counter-drug Training School Network. The Academy and other components of the Institute's comprehensive national training delivery system, established by Ms. Abdullah, was designed to provide substantial support to communities throughout the nation. This drug demand reduction system represents the largest sustained training and technical assistance undertaking in CADCA's, ONDCP's and SAMHSA's history. The Institute's systems change training model is used in local, state, regional and international settings. Previously, Ms. Abdullah, a Certified Prevention Specialist Level 4, served as Vice-President of the Board of Directors for the Prevention Credentialing Consortium for the State of Georgia and Chief Executive Officer of Genesis Prevention Coalition, Inc., based in Atlanta, Georgia. Additionally, she had a successful corporate career that spanned more than 20 years. Ms. Abdullah has presented papers internationally as a U.S. delegate on the Institute's independently evaluated population-level change model and the highly acclaimed NYLI. She has acquired and leveraged millions of federal, state, local, and private dollars and other resources via contracts, grants, cooperative agreements, workforce development, and interagency collaborations. Ms. Abdullah is a distinguished administrator, master trainer, facilitator and coach, effective interviewer, broadcast moderator, talk show host, and keynote speaker.

Catherine Bradshaw, Ph.D., M.Ed., is a Professor and the Associate Dean for Research and Faculty Development at the Curry School of Education at the University of Virginia (UVA). Prior to her current appointment at UVA, she was an Associate Professor and the Associate Chair of the Department of Mental Health at the Johns Hopkins Bloomberg School of Public Health. She maintains an affiliation with Johns Hopkins as the Deputy Director of the Centers for Disease Control and Prevention (CDC)-funded Johns Hopkins Center for the Prevention of Youth Violence and co-director of the National Institute of Mental Health (NIMH)-funded Johns Hopkins Center for Prevention and Early Intervention. She holds a doctorate in developmental psychology from Cornell University and a master's of education in counseling and guidance. She collaborates on research projects examining bullying and school climate; the development of aggressive and problem behaviors; effects of exposure to violence, peer victimization, and environmental stress on children; and the design, evaluation, and implementation of evidence-based prevention programs in schools. She received an award from President Obama in 2009 for her research on the use of evidence-based prevention programs in schools. She has co-authored more than 150 articles and chapters and presently collaborates on federally supported randomized trials of school-based prevention programs, including Positive Behavioral Interventions and Supports and social-emotional learning curricula. She also has expertise in implementation science and coaching models. Dr. Bradshaw works with the Maryland State Department of Education and several school districts to support the development and implementation of programs and policies to prevent bullying and school violence, and to foster safe and supportive learning environments. She collaborates on federally funded research grants supported by NIMH, National Institute on Drug Abuse (NIDA), CDC, U.S. Department of Education, National Institute of

Justice, and the Institute of Education Sciences. She has also consulted with the National Education Association, the United Nations, and the World Bank on issues related to bullying, mental health, and school-based prevention. She is an Associate Editor for the *Journal of Research on Adolescence* and the editor of *Prevention Science*, and co-editor of the *Handbook of School Mental Health*.

Richard Catalano, Ph.D., is the Bartley Dobb Professor for the Study and Prevention of Violence and the co-founder of the Social Development Research Group in the School of Social Work at the University of Washington. For more than 30 years, he has led research and program development to promote positive youth development and prevent problem behavior. His work has focused on discovering risk and protective factors for positive and problem behavior, designing and evaluating programs to address these factors, and using this knowledge to understand and improve prevention service systems in states and communities. He has served on expert panels for the National Academy of Sciences, federal and state governments, and foundations. He has published more than 350 articles and book chapters and his work is highly cited (H factor of 50). His work has been recognized by practitioners (1996 National Prevention Network's Award of Excellence); criminologists (Fellow of the Academy of Experimental Criminology, 2007 August Vollmer Award from the American Society of Criminology, and 2003 Paul Tappan Award from the Western Society of Criminology); prevention scientists (2001 Prevention Science Award, 2012 Presidential Award from the Society for Prevention Research, President-elect of the Society for Prevention Research), and social workers (Fellow in the American Academy of Social Work and Social Welfare). Dr. Catalano is a member of the Washington State Academy of Sciences. He is the codeveloper of the Social Development Model; the parenting programs Guiding Good Choices, Supporting School Success, Staying Connected with Your Teen, and Focus on Families; the school-based program, Raising Healthy Children; and the community prevention approach, Communities That Care.

Gregory Farber, Ph.D., has a B.S. in chemistry from Penn State University (1984) and received his Ph.D. in physical chemistry from the Massachusetts Institute of Technology (1988). Dr. Farber's research in graduate school involved determining the three dimensional structure and mechanism of the enzyme xylose isomerase in the laboratory of Dr. Gregory A. Petsko. After graduate school, Dr. Farber received a Life Sciences Research Fellowship to work on mechanistic enzymology with Dr. W. W. Cleland at the University of Wisconsin. Following his postdoctoral fellowship, Dr. Farber returned to Penn State as an Assistant Professor of Biochemistry and rose to the rank of Associate Professor with tenure by 1998. His research included work on structural movies of enzyme action, molecular evolution, and mechanistic enzymology. Following a sabbatical year in the Division of Biological Infrastructure at the National Science Foundation, Dr. Farber decided to stay in government service. He moved to the National Center for Research Resources (NCRR), part of the National Institutes of Health (NIH), in 2000. At NCRR, he managed a number centers and individual investigator awards in technology development and bioinformatics, as well as a cohort of interdisciplinary research centers. Dr. Farber concluded his service at NCRR as the Director of the Office of Extramural Activities and the Director of the Office of Construction Grants. In June 2011, Dr. Farber became the Director of the Office of Technology Development and Coordination at NIMH. That office is responsible for coordinating all technology development and bioinformatics activities at NIMH, including common data element activities, overseeing the National Database for Autism Research, managing the NIMH component of the BRAIN Initiative, managing the Human Connectome project on behalf of the NIH Neuroscience Blueprint, and also overseeing the NIMH small business program.

Marion S. Forgatch, Ph.D., is Senior Scientist Emerita at the Oregon Social Learning Center (OSLC), where she developed and tested programs for families with children at-risk or referred for child adjustment problems and substance abuse. In 2001, Dr. Forgatch founded Implementation Sciences International Inc. (ISII), a nonprofit affiliate of OSLC, to conduct implementation projects based on Parent Management Training, the Oregon Model (PMTO), a set of evidence-based practices developed and tested by the OSLC group. At ISII she serves as Executive Director and Director of Research. She and her team conduct large-scale implementations for systems providing services to families in child mental health and child welfare. Implementations include nationwide programs (Denmark, Iceland, The Netherlands, Norway), statewide programs (Kansas, Michigan), and citywide programs (Detroit, Mexico City, New York City). Since 2000, Dr. Forgatch and her team have adapted and applied the *Parenting through Change (PTC)* program for use with diverse populations, including non-English speaking Latinos, mothers living in shelters to escape homelessness or violence, parents whose children have been removed for reasons of abuse/neglect, and most recently for members of the National Guards reintegrating home following service in the Iraq and Afghanistan wars. Dr. Forgatch is co-author with Dr. Gerald R. Patterson of *Parents and Adolescents*, a set of books for parents and clinicians. She has also co-authored numerous journal articles, book chapters, and audio and video tapes for parents. A Fellow of the Association for Psychological Science, Dr. Forgatch has received awards from the Society of Prevention Research (Friend of the Early Career Prevention Network in 2003 and Award for

International Collaborative Prevention Research in 2008) and the American Family Therapy Academy (Distinguished Contribution to Family Systems Research in 2012).

Carlos Gallo, Ph.D., is currently a Research Assistant Professor at the department of Psychiatry and Behavioral Sciences at Northwestern University. He obtained a Ph.D. in Computational Psycholinguistics from the University of Rochester followed by postdoctoral studies at Harvard University and the University of Miami. Dr. Gallo applies his expertise in bilingualism and natural language processing to developing computational methods for measuring fidelity in parent training prevention interventions.

Robert M. Goerge, Ph.D., is a Senior Research Fellow at Chapin Hall at the University of Chicago with more than 30 years of research focused on improving the available data and information on children and families, particularly those who require specialized services related to maltreatment, disability, poverty, or violence. Dr. Goerge developed Chapin Hall's Integrated Database on Child and Family Programs in Illinois, which links the administrative data on social service receipt, education, criminal and juvenile justice, employment, health care, and early childhood programs to provide a comprehensive picture of child and family use of publicly provided or financed service programs. His work provides high-quality information to policy makers to improve the programs serving children and their families. He is also the Principal Investigator of the National Survey of Early Care and Education. In addition to his Chapin Hall work, he is a Senior Fellow at the Computation Institute, where he is co-Principal Investigator on a National Science Foundation (NSF)-funded grant entitled, "An Urban Sciences Research Coordination Network for Data-Driven Urban Design and Analysis." He is the Executive Director of the Master's Degree in Computational Analysis and Public Policy and a Senior Fellow at the Harris School for Public Policy Studies. He is a member of the Panel on Modernizing Crime Statistics of the National Academy of Sciences, and a Technical Work Group member of the National Study of Child and Adolescent Well-Being. Dr. Goerge received his Ph.D. from the School of Social Service Administration of the University of Chicago. He is also co-founder of the International Society for Child Indicators.

Rachel A. Gordon, Ph.D., is a Professor in the Department of Sociology and Associate Director of the Institute of Government and Public Affairs at the University of Illinois at Chicago. Dr. Gordon's research broadly examines contextual, social, and policy factors that affect children and families. She has studied how child care and preschool quality affect child development, the relationships between youth gang participation and delinquency, the association between community context and child well-being, the causes and consequences of grandmother coresidential support for young mothers, and an evaluation of an innovative job program for young couples. She is the author of two textbooks (*Regression Analysis for the Social Sciences* and *Applied Statistics for the Social and Health Sciences*) and has published her research in leading academic journals, including the *American Journal of Evaluation*, *Child Development*, *Criminology*, *Demography*, *Developmental Psychology*, *Journal of Marriage and Family*, and *Journal of Research on Adolescence*. Throughout her career, Dr. Gordon has worked at the intersection of academic research and social policy, including through directing the Illinois Family Impact Seminars. Her latest research has been funded by NIH and Institute of Education Sciences (IES) to examine the psychometric properties of widely used measures of preschool and child care quality, including the Early Childhood Environment Rating Scale-Revised Edition (ECERS-R) and Classroom Assessment Scoring System (CLASS). She is also collaborating with the Collaborative for Academic, Social and Emotional Learning and the Washoe County School District (Reno, Nevada) to improve measurement of students' social and emotional competencies using item response theory methods.

Kristen Kracke, M.S.W., is a Social Science Specialist for the U.S. Department of Justice's Office of Juvenile Justice and Delinquency Prevention (OJJDP), and serves as a senior technical expert in the area of juvenile justice systems improvement, evidence-based practice, research, and resource development specifically in the areas of youth justice, trauma, and children's exposure to violence. In this role, she has worked in a range of capacities including research, performance management, strategic planning, training, and program development. Ms. Kracke has developed and currently manages a number of national reform efforts using data-informed practice. She has been leading the development of promising approaches in the field of juvenile justice and children's exposure to violence for more than 20 years and has directed collaborative systems change efforts to improve policy and practice for children and families at both the federal and local levels.

Vetisha L. McClair, Ph.D., joined the Center for Medicare & Medicaid Innovation (CMMI) at the Centers for Medicare & Medicaid Services (CMS) in 2012. She is a Research Analyst in the Research and Rapid Cycle Evaluation Group, Division of Research on Vulnerable Populations and Care Transformation. Dr. McClair leads the evaluation of many of CMMI's mental health-related demonstrations, including a group of Healthcare Innovation Awards (HCIA) focused on behavioral health and substance abuse. The comprehensive evaluation of this diverse set of 10 awardees from the initial round of the HCIA initiative will yield findings with important implications for public policies and programs that affect individuals with behavioral and substance use disorders. Dr. McClair also co-leads an interagency group organized by CMMI concentrated on the integration of behavioral health and primary care. Additionally, she provides consultation to CMMI and other CMS components in the development and evaluation of new mental health related projects. Her research interests relate to mental health service use and the effect of psychosocial/demographic factors on treatment outcomes. Previously, she completed a postdoctoral fellowship in the Section on Developmental Genetic Epidemiology in NIMH after earning her Ph.D. and M.S. in Counseling Psychology from the University of Illinois at Urbana-Champaign and her B.S. in Psychology from Howard University. She has also provided psychological services in a variety of settings (Veterans Administration Medical Centers, disability services, and out-patient hospitals) and is a licensed psychologist in the state of Maryland.

Robert Orwin, Ph.D., is a Senior Study Director in the Behavioral Health Group at Westat, with 30 years of experience in program evaluation and 20 years of experience in the issues of homelessness, mental health, substance abuse prevention, treatment, and policy. His expertise includes evaluation and survey research, research design, and data analysis. Dr. Orwin has directed research projects for SAMHSA; the National Institute on Alcohol Abuse and Alcoholism; NIDA; CDC; the National Heart, Lung, and Blood Institute; the U.S. Food and Drug Administration; and various other agencies. He has directed or co-directed several multisite evaluations of services for individuals and families and was recently director of the cross-site evaluation of the Center for Substance Abuse Prevention's Strategic Prevention Framework. Dr. Orwin's other research has included evaluations of the National Youth Anti-Drug Media Campaign and the effects of gender-sensitive substance abuse treatment on increasing employment and decreasing criminal justice involvement among women.

Jane Pearson, Ph.D., chairs the NIMH's Suicide Research Consortium. She is the Associate Director for Preventive Interventions in the Division of Services and Intervention Research, and she currently leads the staffing for the National Action Alliance for Suicide Prevention Research Prioritization Task Force, which includes the implementation of the first prioritized research agenda for suicide prevention (www.suicide-research-agenda.org). Dr. Pearson serves as the NIH representative to the U.S. Department of Health and Human Services Federal Steering Group on Suicide Prevention.

Harold Alan Pincus, M.D., is Professor and Vice-Chair of the Department of Psychiatry and co-director of the Irving Institute for Clinical and Translational Research at Columbia University and Director of Quality and Outcomes Research at New York-Presbyterian Hospital. Dr. Pincus also serves as a Senior Scientist at the RAND Corporation. Previously he was Director of the RAND-University of Pittsburgh Health Institute and Executive Vice Chairman of the Department of Psychiatry at the University of Pittsburgh, where he still maintains an adjunct professorship. He is the National Director of the Atlantic Philanthropies' Health and Aging Policy Fellowship and previously directed the Robert Wood Johnson Foundation's national program on Depression in Primary Care: Linking Clinical and Systems Strategies and the Hartford Foundation's national program on Building Interdisciplinary Geriatric Research Centers. Dr. Pincus has also served as the Deputy Medical Director of the American Psychiatric Association (APA) and the founding director of APA's Office of Research. Prior to joining APA, he was the Special Assistant to the Director of NIMH. Dr. Pincus graduated from the University of Pennsylvania and received his medical degree from Albert Einstein College of Medicine in New York. Following completion of residency at George Washington University Medical Center, Dr. Pincus was named a Robert Wood Johnson Foundation Clinical Scholar. As a Clinical Scholar, Dr. Pincus served as a professional staff member of the President's Commission on Mental Health at the White House and, subsequently, as a congressional fellow in the U.S. House of Representatives. He has edited or co-authored 23 books and more than 350 scientific publications in health services research, science policy, research career development, and the diagnosis, classification, and treatment of mental disorders. Dr. Pincus has had a particular research interest in the practice of evidence-based medicine, quality improvement and the relationships among general medicine, mental health, and substance abuse, developing and empirically testing models of those relationships. He has led major health policy and services research and research training projects totaling more than $150 million in external funding. Dr. Pincus has chaired committees for NIH and served on several Institute of Medicine/National Academy of Sciences committees, including Crossing the Quality Chasm in Behavioral Health and Cancer Care for the Whole Patient, as well as the World Health Organization's International Classification of Disease Oversight Committee, World Psychiatric Association Section on Economics, Behavioral Measurement Advisory Panel of the National Committee for Quality Assurance, and numerous other national and international committees. He is chair of the Medicaid Task Force for the Measurement Applications Partnership, authorized under the Affordable Care Act; and he is co-chair of both the National Quality Forum Behavioral Health Steering Committee and the World Health Organization International Classification of Diseases 11th Revision (ICD-11) Committee on Quality and Patient Safety.

Lisa Saldana, Ph.D., is a Senior Research Scientist at Oregon Social Learning Center. She is currently the Principal Investigator (PI) on the Stages of Implementation Completion for Evidence-Based Practice, an NIMH-funded R01 that examines the successful implementation of interventions in community settings. She also is working on NIH-funded research grants focusing on the economic evaluation of evidence-based practices and is a Co-Investigator on an NIDA-funded Translational Drug Abuse Prevention Center at the Oregon Social Learning Center (OSLC). Dr. Saldana is a Co-Investigator on a large real-world implementation of two linked evidence-based programs in a large multisite child welfare system. In collaboration with colleagues, she developed the Stages of Implementation Completion (SIC) and Cost of Implementing New Strategies (COINS) implementation tools. Previously, she was a Co-Investigator on a large-scale trial awarded to Patricia Chamberlain evaluating “what it takes” to implement an evidence-based practice (MTFC) for youth in foster care in communities with barriers to implementation. This trial produced the SIC and COINS tools. Dr. Saldana also has a strong intervention development background. She was the PI of an NIDA-funded K award in which she developed the Families Actively Improving Relationships (FAIR) program, an intensive intervention for child welfare involved families with comorbid substance abuse, and currently is serving as PI on an efficacy trial of this intervention. She has collaborated with a number of other evidence-based practice developer groups and has been involved in development activities ranging from evaluation to creation of intervention principles and strategies. More recently her collaboration has strongly focused on development of implementation strategies to assist in increasing the uptake of successful practice. She and colleague Patricia Chamberlain are about to launch a large rollout of the R^3 Supervisor Strategy as a means of infusing the use of evidence-based techniques throughout a statewide child welfare workforce.

Jeff Schiff, M.D., M.B.A., is the Medical Director for Minnesota Health Care Programs at the Minnesota Department of Human Services. This includes Minnesota’s Medicaid and Minnesota Care programs. He has served as medical director since June 2006. His work focuses on the development and implementation of evidence-based benefit policy, the advancement of improved care delivery models, and the improvement of clinical quality. Specific areas of interest include the development of policy to enhance the role of primary care and the provision of patient-centered medical homes; the use of collaborative intrastate processes to implement quality improvements across the health care system; and the use of claims and clinical data to report and improve health outcomes. The Minnesota Health Services Advisory Council, a physician-based policy advisory body for the development and implementation of evidence-based policy, is organized through his office. Dr. Schiff is the current chair of the Medicaid Medical Directors network. In the network, he has co-chaired a 22-state project to measure the status of early elective delivery in these states. He has chaired a Patient-Centered Outcomes Research Institute (PCORI) advisory group to identify perinatal topics to be considered for future funding. He has served as the co-chair of the Agency for Healthcare Research and Quality (AHRQ) National Advisory Council Subcommittee on Children's Healthcare Quality Measures for Medicaid and Children's Health Insurance Program (CHIP) programs. This subcommittee was charged with the identification of initial core set of children's health care quality measures for voluntary use by Medicaid and CHIP Programs across the country. He has served as co-chair of the CMS expert panel on birth outcomes subcommittee that recommended Medicaid measures to improve prenatal and birth outcomes. Dr. Schiff has served as president of the Minnesota Chapter of the American Academy of Pediatrics. His past administrative experience includes Medical Direction for Minnesota's Emergency Medical Services for Children program. He practices clinically in pediatric emergency medicine.

Sarah Hudson Scholle, Dr.P.H., M.P.H., is Vice President of Research and Analysis for the National Committee on Quality Assurance (NCQA). Her research on measurement of practice systems and patient-centered care has informed the development and evaluation of NCQA's Patient-Centered Medical Home program. She led a recent effort that led to the development of the Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey for patient-centered medical homes and currently heads a study to assess opportunities for incorporating shared decision making in accountable care organizations. Dr. Scholle leads one of seven Centers of Excellence in the AHRQ/CMS Pediatric Quality Measurement Program. She has extensive experience with measurement using multiple data sources including administrative claims and eligibility files, surveys, and most recently electronic health record data. Her research has focused on both content issues as well as technical issues such as reliability and validity of measures, as evidenced by her publication record. Dr. Scholle received her master's degree in public health from Yale University and her doctorate in public health from the Johns Hopkins Bloomberg School of Public Health.

Anne Sheridan is the Executive Director of the Maryland Governor's Office for Children (GOC). The GOC works to promote health and wellness among youth, and seeks to address the priorities of Maryland's children. Previously she served as the manager of the Maryland No Kid Hungry Campaign at Share Our Strength and co-chaired the Maryland Partnership to End Childhood Hunger where she helped to implement a statewide plan to meet partnership goals. Sheridan has years of community relations and field operations experience, working as Corporate Director for Community Relations at Harrah's Entertainment, Inc., as well as with the Kerry-Edwards Campaign in 2004, and as principal of the Dewey Square Group. Sheridan is a graduate of Georgetown University's School of Foreign Service and received her B.A. in Economics and Government from Simmons College in Boston, Massachusetts.