



DEFENSE CENTERS
OF EXCELLENCE

For Psychological Health
& Traumatic Brain Injury

Strategies for Measuring Program Processes and Outcomes

CAPT Armen H. Thoumaian, Ph.D., USPHS

Aaron Sawyer, Ph.D.

Patrick High, Dr.P.H.

Debra Stark, M.B.A.

May 20, 2014



Webinar Details

- This webinar presentation has been pre-recorded
- A live question-and-answer session will be held at the conclusion of the presentation
- Questions may be submitted via the “Question” pod
- Audio for this presentation will be provided through Adobe Connect; there is no separate dial-in
- Closed captioning is not available for this event

Continuing Education Details

- Continuing education credit is not available for this event
- Sources for materials and additional training information:
 - Materials from this series are available at:
dcoe.mil/About_DCoE/Program_Evaluation.aspx
 - For information on other DCoE webinar and training series, visit:
dcoe.mil/Training/Monthly_Webinars.aspx
 - Materials for this webinar are available in the Files box

Presenters

CAPT Armen Thoumaian, Ph.D.
U.S. Public Health Service
Health Science Officer
Office of Policy, Programs and Integration,
DCoE

CAPT Armen Thoumaian is a scientist director in the Commissioned Corps of the U.S. Public Health Service with more than 30 years experience in health and mental health program design and evaluation.

In January 2012, CAPT Thoumaian joined the staff at the Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury (DCoE) to help design and implement program evaluation and improvement efforts in the Defense Department.

He holds a B.A. in Psychology and Sociology, an M.A. in General Experimental Psychology, and a Ph.D. in Social Welfare and Social Work, completing an National Institute of Mental Health fellowship in Community Mental Health.



USPHS Capt. Armen Thoumaian,
Ph.D.

Presenters

Aaron Sawyer, Ph.D.

Research Scientist, Contract Support for DCoE

Dr. Aaron Sawyer is a clinical psychologist with extensive expertise in intervention outcome research and program evaluation. He has delivered child, family and adult interventions for more than a decade, including specialization in trauma and experience working with military families. Dr. Sawyer holds an M.S. in Experimental Psychology and a Ph.D. in Clinical Psychology. He completed post-doctoral training at The Kennedy Krieger Institute/Johns Hopkins University and is a licensed psychologist.



Dr. Aaron Sawyer

Patrick High, Dr. P.H.

Epidemiologist, Contract Support for DCoE

Dr. Patrick High is an epidemiologist with over a decade of experience and has expertise in survey design, research methodology and program evaluation. His experience includes supporting the Office of the Undersecretary of Defense for Personnel and Readiness, Operations Research and Safety, and the Defense Suicide Prevention Office as an epidemiologist. Dr. High holds the degree of doctor of public health with specialization in Epidemiology and Biostatistics from the Uniformed Services University of the Health Sciences. He previously spent nine years in the Illinois Army National Guard.



Dr. Patrick High

Moderator/Presenter

**Debra Stark, M.B.A.,
Research Scientist, Contract Support for DCoE**

Ms. Debra Stark is a survey methodologist with 15-plus years of research experience. Her work includes program evaluation and monitoring, qualitative data analysis and survey instrument design. She has worked on public health services evaluation projects with various federal agencies, including the Department of Veterans Affairs and TRICARE Management Activity. Ms. Stark received an M.B.A. from Vanderbilt University.



Ms. Debra Stark

Overview and Objectives

This webinar will provide guidance on metrics and measurement strategies for program managers and others seeking to measure processes and outcomes during the course of program evaluations

At the conclusion of this webinar, participants will be able to:

- Explain how metrics and measurement strategies are used in program evaluation and improvement efforts
- Demonstrate knowledge of important considerations for selecting or developing measures
- Implement suggested guidance to collect and store evaluation data
- Identify common challenges that programs face in measuring processes and outcomes

Agenda

- Introduction to Metrics and Measurement
- Choosing Measurement Domains
- Best Practices in Measure Selection and Development
- Data Collection and Storage Considerations
- Common Challenges
- Conclusion
- Resources
- Feedback and Q&A Session

Introduction to Metrics and Measurement

Introduction



E.L. Thorndike

“Whatever exists at all, exists in some amount. To know it thoroughly involves knowing its quantity as well as its quality.”

– *E.L. Thorndike (1918)*

Image Source:

http://en.wikipedia.org/wiki/File:PSM_V80_D211_Edward_Lee_Thorndike.png

Key Terms

Measurement

- The process of collecting data

Metric

- Specific, quantifiable units of data used as indicators of activities and performance

Measure

- A tool used to collect data (e.g., test, questionnaire, survey, instrument)

Measurement in Military History

- The U.S. military has a long history of measurement, beginning in World War I with use of the Army Alpha and Beta test batteries
- Defense Department collects information through several efforts, such as:
 - Armed Forces Health Surveillance Center, Defense Manpower and Data Center, DoD Suicide Event Report and others
- Effective measurement is critical in program evaluation and improvement efforts and also in program selection and funding decisions



Image Source: <http://www.igs.net/~pballan/C4P1.htm>

Measurement is Everywhere Every Day

On the road

- Speed, tolls, compliance with traffic laws

On the web

- Website hits, connections or social media posts

On the phone

- Data usage, texts and phone calls

At home

- Cooking, bathroom scale, distance on treadmill

At work

- Productivity, performance, attendance, time cards, continuing education

At the doctor's office

- Blood pressure, pulse, temperature, weight, insurance utilization

By the government

- Census, taxes, travel

Basics of Classical Measurement Theory

One way to conceptualize measurement is as follows:

$$X = T + E$$

X = Data point

T = "True" value

E = Error

High quality measurement involves minimizing E so that X is as close to T as possible

Validity and Reliability



Validity – the degree to which a measure accurately represents the characteristic it is designed to measure



Reliability – the degree to which a measure's results are consistent over time and across situations

Quantitative Data Types

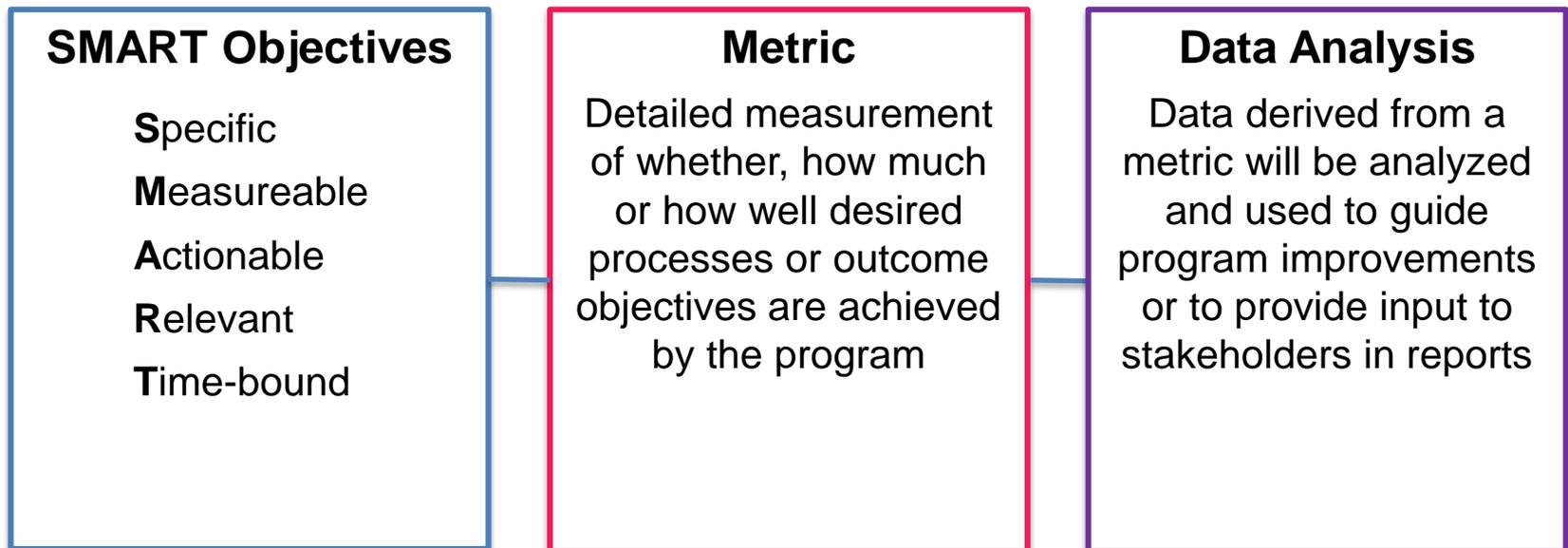
SCALE	DESCRIPTION	EXAMPLES
Nominal	Category or type	Gender, ethnicity, service branch
Ordinal	Rank order, units of variable size	Rankings, rating scales, scores on opinion and attitude surveys
Interval	Equal unit size with arbitrary zero point	Time on 12- or 24-hour scale, temperature, scores on standardized scale
Ratio	Equal unit size with meaningful zero point	Reaction time, pulse, symptom count, days present at work, number correct on test

Qualitative Data Types

SOURCE	DESCRIPTION	CHARACTERISTICS
Interviews	One-on-one conversation	Can be structured or semi-structured
Focus groups	Group conversation facilitated by moderator	Use structured protocol, groupings of similar individuals
Observations	Log or description of activity	Applied in consistent manner to minimize bias
After Action Reviews	Group review following activity	Focus on strengths and opportunities for improvement
Case Studies	In-depth longitudinal observations	Study of one individual, process or program

Establish Metrics for Processes and Outcomes

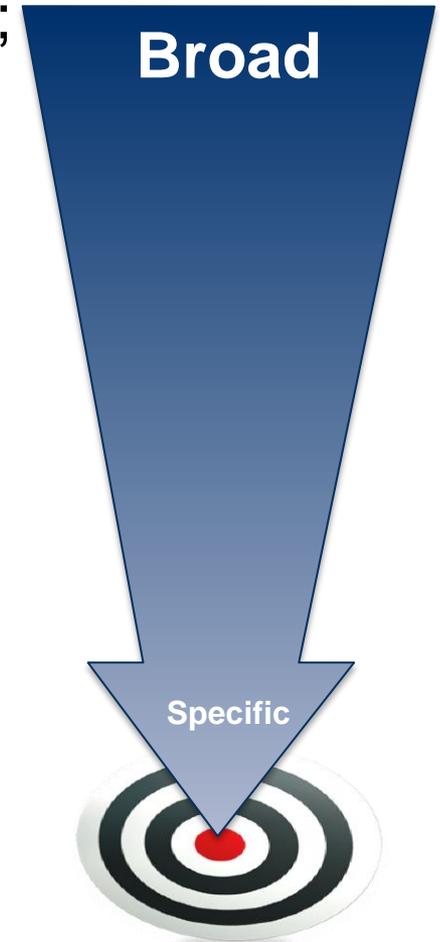
- Objectives drive development of metrics that reflect program processes or outcomes
- Data derived from metrics are analyzed to determine whether objectives are met



Choosing Measurement Domains

Mission, Goals and Objectives Drive Measurement Choices

- **Mission** – Purpose for program's existence; goals and objectives should support this mission
- **Goals** – Statements that outline what the program intends to accomplish
- **Objectives** – Descriptions of goals in terms of smaller units that can be measured; objectives should be SMART (Specific, Measurable, Achievable, Relevant, Time-Bound)



How Will Anyone Know That Expected Processes or Outcomes Have Occurred?

Identify the evidence needed to demonstrate key processes or outcomes are occurring and compare measured results to stated objectives

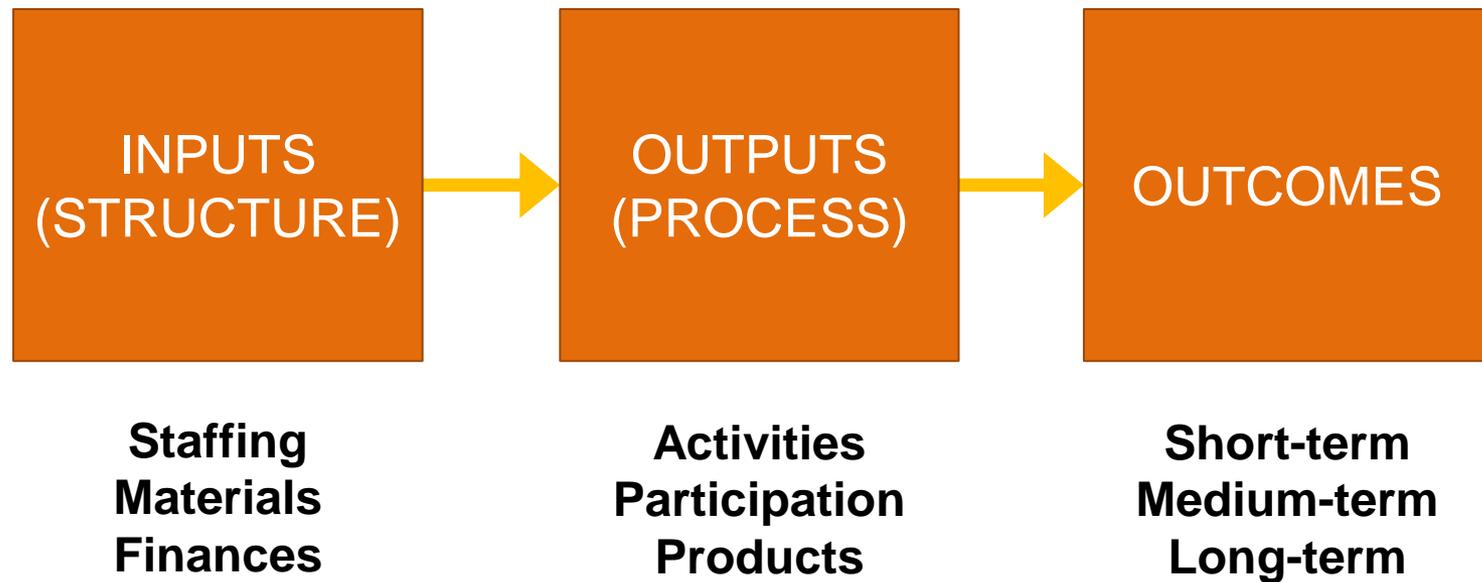
Example:

Objective: Program staff will provide 50 web-based training sessions to unit commanders, who will demonstrate increased knowledge of TBI symptoms

Measured Process: Program staff provided 35 web-based trainings to unit commanders during FY2014

Measured Outcome: Unit commanders showed an average 20 percent increase in knowledge of TBI symptoms from baseline to post-training assessment

Use Program Logic Model to Determine Measurement Domains



Tracking Inputs: What the Program Needs

Use forms, records, logs, audio or video recordings and other related materials available to track information relevant to staffing, materials and program finances

Examples include:

- Staffing – demographics, turnover, trainings, administrative meetings, staff training sessions
- Materials – number and age of computers, software, space availability, office supplies
- Finances – expenditures for staffing, materials, training, program activities
- Other inputs – meetings held with others

Tracking Outputs: Staff Activities

Use forms, records, logs, audio or video recordings, and other materials to track program delivery activities as well as promotion and outreach activities

Examples include:

- Delivering webinars, trainings or treatment
- Staffing a booth at a health fair or a help line
- Providing on-air interview (radio, tv, newspaper)
- Attending conferences
- Following up with participants
- Conducting assessments
- Making referrals for participants
- Handing out fliers
- Taking meeting minutes

Tracking Outputs: Participation and Products

Participation

- Attendance or program completion
- Program accessibility
- Downloads or web hits
- Demographics (age, gender, race, ethnicity)
- Service branch, rank and duty status
- Deployment status

Products

- Pamphlets
- Webcasts
- Mobile phone applications
- Units of service
- Scientific publication



Connect Outputs With Outcomes

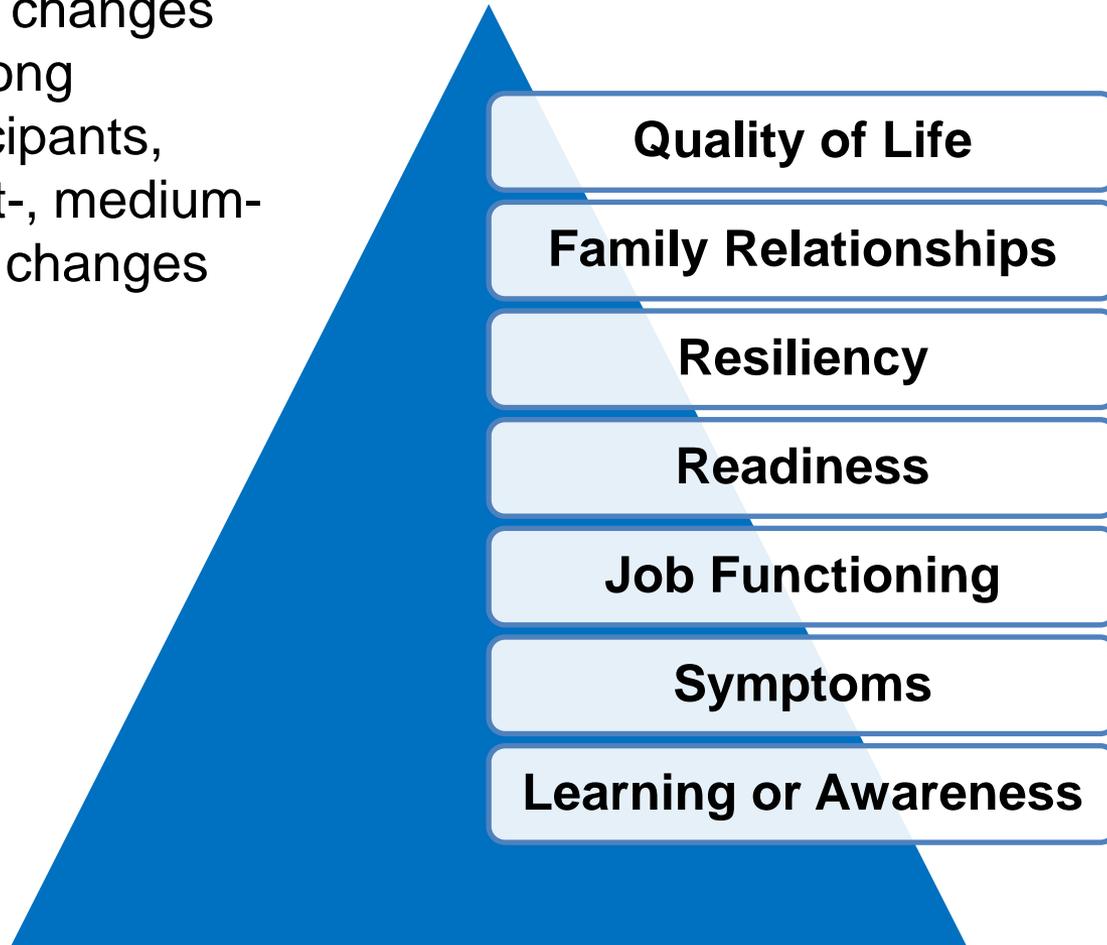
Tracking outputs provides “facts” about whether or not program processes produce the desired results

This will allow program staff to:

- Measure whether program processes result in desired outcomes
- Examine trends in activities, products, participation and outcomes over time
- Assess whether changes to the program result in improved outcomes

Tracking Outcomes: Measure Multiple Domains

Outcomes are changes that occur among program participants, including short-, medium- and long-term changes



Use Multiple Metrics Within Domains

Specific measures or metrics can be used to assess outcomes within a given domain

Resiliency

Coping skills

Problem-solving abilities

Self-esteem

Social support

Example Output and Outcome Measures: Domestic Violence Program

Tracking Outputs (Process):

- **Activities** – group sessions held, referrals made, outreach to service providers, community education
- **Participation** – calls to helpline, session attendance, participant demographics, participant referral source
- **Products** – pamphlets, posters, website
- **Program Satisfaction** – satisfaction ratings and comments, willingness to refer others

Example Output and Outcome Measures: Domestic Violence Program (continued)

Tracking Outcomes:

- **Awareness** – attitudes toward domestic violence, recognition of risk behaviors, knowledge of support and legal services
- **Quality of Life** – level of stress, rating on quality of life measure, optimism or hope for the future
- **Family Relationships** – frequency and severity of domestic violence incidents, parent-child conflict, partner communication quality

Best Practices in Measure Selection and Development

To Select or to Develop a Measure



- In general, it is best to select existing measures shown to be valid and reliable for the program's purpose and population
 - Resources for finding existing measures include published lists of validated measures, relevant scientific literature and expert consultation
 - To determine validity and reliability, consult user's manuals, measure websites and published research

- New or custom measures are most appropriate when assessing learning relevant to a specific program process (e.g., skills training)
 - Use best practices for item development

When Selecting Measurement Instruments, Consider...

- Validity and reliability for purpose and population
- Training and professional licensure required to administer
- Time, costs and usage license required to administer and score
- Availability of suitable alternatives
- Alignment with stakeholder interests and goals
- Overlap with other measures used in the continuum of services
- Practical or clinical utility in guiding program services

Best Practices in Item Development

- ✓ Seek a sixth to eighth grade reading level to accommodate varying language skills and cultural backgrounds
- ✓ Avoid overly complex language, including “double-barreled” questions, metaphors or culture-specific examples
- ✓ Use items that are “face valid.” Items should measure what they appear to measure
- ✓ Be consistent in the number and type of response choices offered (e.g., 5-point response scale, True-False, Yes-No)
- ✓ Seek input on content from multiple sources, such as previous measures, experts, stakeholders, staff and program participants

Best Practices in Item Development (continued)

- ❑ Avoid altering or shortening existing measures or use, which reduces reliability and validity
- ❑ Seek expert input when translating between languages
- ❑ Always pilot the measure before use



Item Development Example

Which of the following is the most accurate descriptive phrase representing an individual's probability of developing PTSD?

- A. A rolling stone gathers no moss
- B. Those who experience trauma are subject to increased hyperarousal triggered by memories of traumatic events over the course of time
- C. People who experience a traumatic event, who are predisposed to anxiety and who have little social support
- D. Certain people

Item Development Example (continued)

This language is more complicated than it needs to be

Which of the following is the most accurate descriptive phrase representing an individual's probability of developing PTSD?

Spell out acronyms

Avoid culture-specific metaphors

A. A rolling stone gathers no moss.

B. Those who experience trauma are subject to increased hyperarousal triggered by memories of traumatic events over the course of time

Too long relative to other responses and too complicated

C. People who experience a traumatic event, who are predisposed to anxiety and who have little social support.

Okay

D. Certain people

Too short relative to other responses and not descriptive enough

Data Collection and Storage Considerations

Begin with a Data Matrix

EXAMPLE 1: MENTAL HEALTH OUTCOMES FOR CLINICAL PROGRAM

Matrix Item	Metric 1	Metric 2	Metric 3
Data source	Participant	Provider	Commander
Data collection method	Self-report questionnaire	Functional rating	Functional rating
Who will collect data	Provider	Provider	Provider
Frequency	Pre-, Post-treatment, 3-month follow-up	Pre-, Post-treatment, 3-month follow-up	Pre- and Post-treatment
Who will input data?	Provider	Provider	Program manager
Who will analyze?	Program manager	Program manager	Program manager
How will data be used?	Treatment planning and outcome tracking	Outcome tracking and reports	Outcome tracking and reports

Begin with a Data Matrix (continued)

EXAMPLE 2: OUTPUTS AND OUTCOMES FOR NON-CLINICAL PROGRAM

Matrix Item	Participation Metric 1	Participation Metric 2	Learning Metric
Data source	Registration page	Website	Participants
Data collection method	Attendance on webinar day	Document downloads and web page visits	Post-test quiz results
Who will collect data	Provider	Provider	Provider
Frequency	On event day	Monthly	Post-event
Who will input data?	Event provider	Website manager	Event provider
Who will analyze?	Program manager	Program manager	Program manager
How will data be used?	Output tracking and reports	Output tracking and reports	Outcome tracking, reporting and continuing education

Develop a Data Collection and Storage Plan

A detailed plan must be developed to specify how data will be collected and stored, including:

- What data will be collected?
- Who will collect the data?
- When and how often will data be collected?
- How will data be cleaned and entered into storage system?
- How will quality assurance and accuracy be checked?



Inputs Required for Data Collection



- Time to complete collection and storage activities
- Training in administration of the metric and data entry
- Funding for measurement materials (e.g., paper, software, file cabinets)
- Support for quality maintenance

Protect Privacy and Confidentiality

Sensitive data should be collected and maintained in a way that ensures participants' privacy and confidentiality

- Personal identifiable information includes demographic information that can be used to identify a program participant
- Protected health information includes information about an individual's health status
- Many programs may be subject to the regulations under the 1996 Health Insurance Portability and Accountability Act

CONFIDENTIAL

Best Practices in Maintaining Privacy and Confidentiality

- ✓ **Secure data storage** – store data in locked rooms, locked file cabinets and/or password-protected computer systems
- ✓ **De-identification** – use participant identification numbers and separate any information that could be used to identify specific participants
- ✓ **Secure data transmission** – transfer data using encryption software or using a lockbox
- ✓ **Clear procedures** – create clear policies and standard operating procedures (SOPs) for data collection and storage
- ✓ **Staff training** – ensure all staff members who access data receive ongoing training and support
- ✓ **Anonymous data** – in rare circumstances, use forms or data collection procedures that do not include identifiers

Minimizing Risks

- Avoid unnecessary disclosures
- Prohibit staff members from taking data off-site
- Ensure data storage areas or computers remain locked
- Ensure policies and SOPs are clear and conduct regular audits



- Gain consent from participants, command, internal review boards, etc.
- Destroy data securely following mandatory storage period
- Track adverse events and lessons learned
- Do not collect unneeded data

Common Challenges

Special Considerations for Measurement in the Military

- Some issues, such as TBIs, make measurement processes more difficult; adaptations may be needed (e.g., breaks, alternative tools)
- Widely varying abilities and cultural differences must be considered when selecting measures
- Stigma and career concerns may prevent reporting of psychological health problems and TBIs; develop FAQs to address concerns and clarify how data will be used
- Comorbidity is common in clinical settings (e.g., between PTSD and TBI); measure both target outcomes and commonly co-occurring outcomes



Source: California National Guard

Threats to Validity and Reliability

When making measurement decisions and carrying out data collection and storing data, keep aware of factors that can increase error such as:

- Inconsistent procedures for measurement
- Attrition or incomplete data collection
- Use of measures in populations or circumstances for which they have not been validated
- Staff members may bias results by communicating expectations

Common Challenges in Measurement FAQ

- If my prevention program is successful, how do I measure something that did not happen as a result?
- I cannot find measures that have been shown to be reliable and valid for the purposes and population of my program
- My staff lack the resources, such as time, training, materials and funding to carry out high quality measurement

If My Prevention Program is Successful, How Do I Measure Something That Did Not Happen as a Result?

- Prevention should address both risk reduction and health promotion¹
- Non-clinical programs can and should measure processes and short-term outcomes related to longer-term goals
 - Measured processes and outcomes may include participation, referrals, learning, health behaviors, risk behaviors, etc.
- Programs or leadership may wish to use service-level databases to link processes and short-term outcomes to long-term outcomes but can only do so if initial data exist

I Cannot Find Measures That Have Been Shown to Be Reliable and Valid for the Purposes and Population of My Program

- First, consult with experts in your program area, such as researchers, consultants or other programs
- Second, for measures that have been validated for similar purposes and populations, pilot testing and focus groups may help to determine appropriateness for a program
- Third, it is often possible to adapt an existing measure using best practices for item development and then re-examine validity and reliability within the program
- Last, if no suitable measures exist, a custom measure may be developed

My Staff Lack the Resources, Such as Time, Training, Materials and Funding, to Carry Out High-quality Measurement

- Measurement is an important investment in a program's future; by measuring, program leaders can learn what is working and what needs improvement, and they can convey results to stakeholders who fund the program
- By carefully assessing time and budgeting, high-quality measurement is generally possible; over time, the results may be used to identify which processes are most critical and which can be eliminated or streamlined
- Many measurement materials are free or low-cost, and consultation and training may be readily available from colleagues or researchers

Conclusion

Key Takeaways

- ★ Programs can use data to establish clear connections between objectives, inputs, outputs and outcomes
- ★ Careful planning and attention to measurement quality ensures accuracy
- ★ Measurement is essential to document program processes and improve performance



Photo by Stewart Leiwakabessy

Resources

DCoE Program Evaluation Guide: www.dcoe.mil/Content/Navigation/Documents/DCoE_Program_Evaluation_Guide.pdf

Agency for Healthcare Research and Quality: <http://www.qualitymeasures.ahrq.gov>

National Quality Forum: www.qualityforum.org/Measures_Reports_Tools.aspx

Deployment Health Clinical Center: www.pdhealth.mil/clinicians/assessment_tools.asp

Defense and Veterans Brain Injury Center: [http://dvbic.dcoe.mil/diagnosis-assessment?audience\[0\]=3](http://dvbic.dcoe.mil/diagnosis-assessment?audience[0]=3)

National Center for Telehealth and Technology: www.t2.health.mil

Centers for Disease Control and Prevention: www.cdc.gov/healthyyouth/evaluation/resources.htm#4

Real Warriors Campaign Guidance on SF86: www.realwarriors.net/active/treatment/clearance.php

Institute of Medicine, 2014 Report :

Preventing Psychological Disorders in Service Members and Their Families: An Assessment of Programs, Chapter 5:

www.iom.edu/Reports/2014/Preventing-Psychological-Disorders-in-Service-Members-and-Their-Families.aspx

University of Wisconsin-Extension: www.uwex.edu/ces/pdande

Minnesota Department of Health: www.health.state.mn.us/divs/opi/qi/toolbox