

Validation of Direct Reading Methods (and how NOT to validate them)

Matthew L. Magnuson



1. Intro to EPA's National Homeland Security Research Center

2. Validation topics

- **Validation goals**
- **Performance Criteria**
- **Validation of entire process, not just instrumental component**
- **Importance of Third Party Testing**
- **EPA Technology Testing and Evaluation Program**

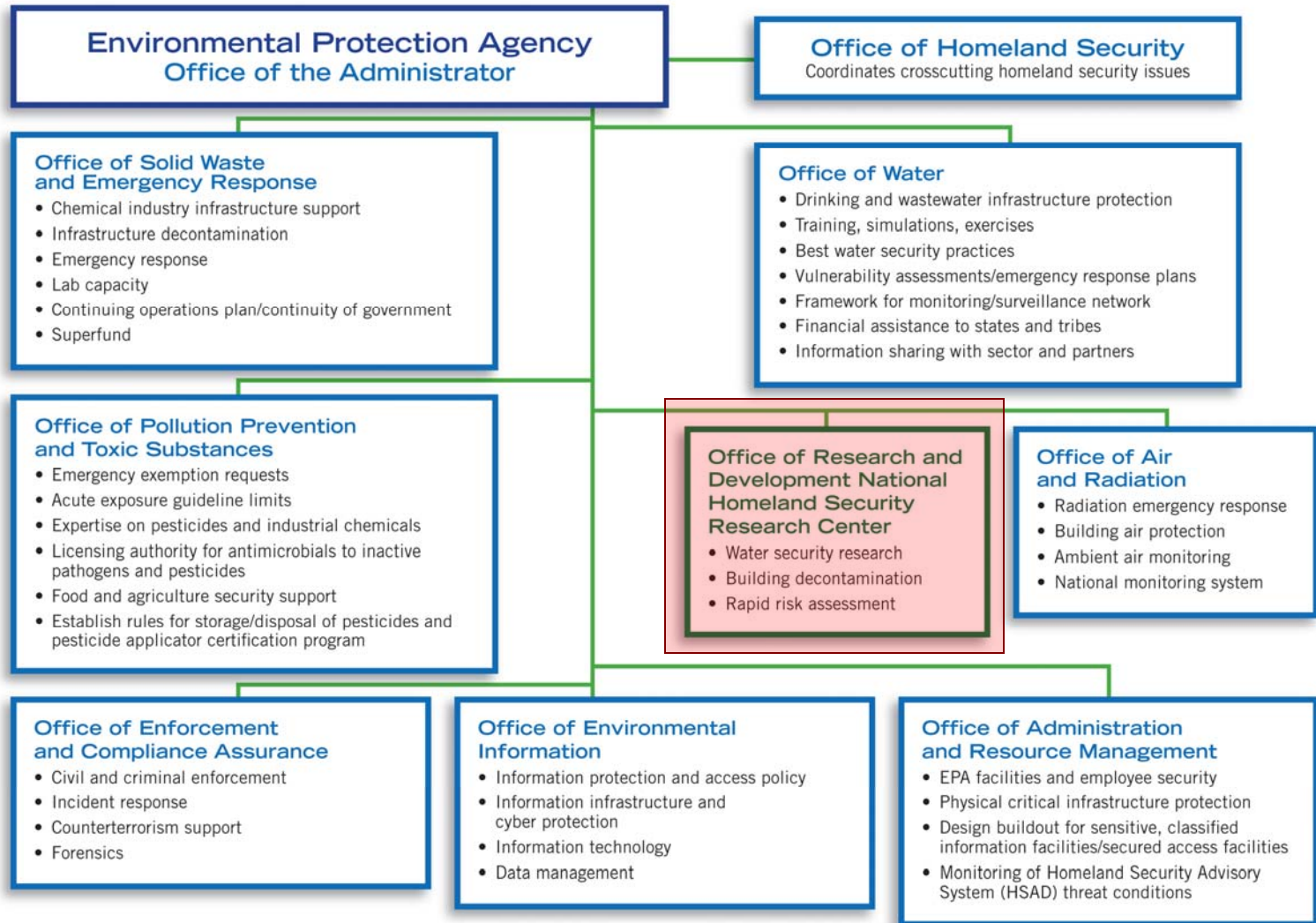
3. Validation related needs

- **Technical**
- **Policy**
- **Training**



1. Introduction to EPA's National Homeland Security Research Center

Homeland Security Roles at EPA



Authorities

The Public Health Security and Bioterrorism Preparedness and Response Act of 2002, together with Homeland Security Presidential Directives (HSPDs), charges EPA with protecting our nation's critical water infrastructure and decontaminating indoor and outdoor areas following an incident.

HSPD 7 – Defines EPA as Sector Specific Agency for drinking water and water treatment systems

HSPD 9 – Requires EPA to develop a comprehensive, and fully coordinated water quality surveillance/monitoring system and an interconnected laboratory network

HSPD 10 – Requires EPA to address risks from biological agents and develop strategies, guidelines, and plans to decontaminate persons, equipment, and facilities (classified)



EPA Roles in Homeland Security

- **Protecting water and water infrastructure**
- **Indoor and outdoor clean-up following attack or natural disaster**
- **Reducing vulnerability of the chemical & hazardous materials sector**
- **Research to protect water infrastructure & buildings**
- **Hazardous materials emergency response**

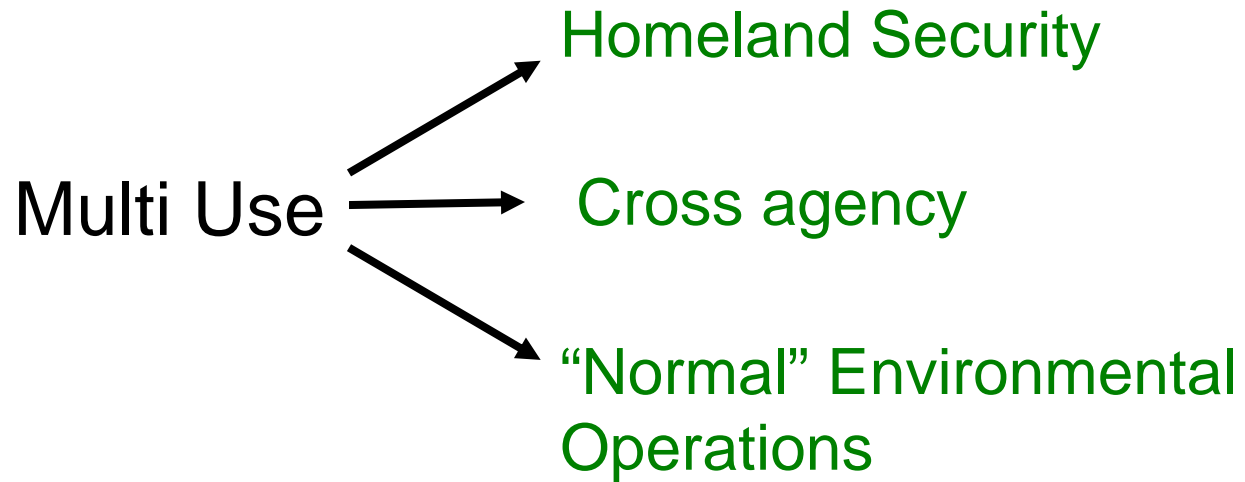




NHSRC Primary Areas of Focus (founded 2002)

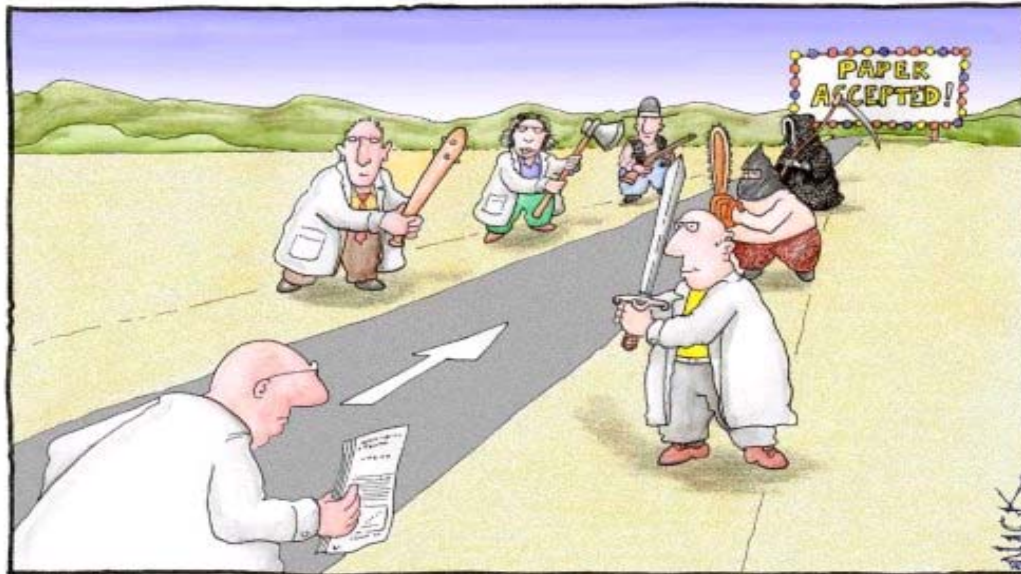
- **Water Infrastructure Protection Division** is charged with conducting research to prevent, detect, and respond to terrorist attacks on the nation's drinking water and wastewater systems.
- **Decontamination and Consequence Management Division** focuses on decontamination of buildings and outdoor environments, as well as the safe disposal of contaminated materials
- **Threat and Consequence Assessment Division** investigates human exposure to chemical, biological, and radiological contaminants to define dangerous levels of these contaminants and establish protective cleanup goals
- **Response Capability Enhancement Group** works directly with emergency responders and local governments to provide tools and information needed to make informed decisions in the event of an attack
- **Technology Testing and Evaluation Program** evaluates technologies that show potential for use in homeland security applications. These evaluations are used by water utility operators, building owners, emergency responders, and others to make informed decisions when purchasing security technology.

NHSRC Research Projects



NHSRC Products

- 125 reports and journal articles since 2003 (including classified)
- Results presented many other ways—stakeholder meetings, symposiums, workshops, etc.
- Products and research plans receive rigorous quality reviews



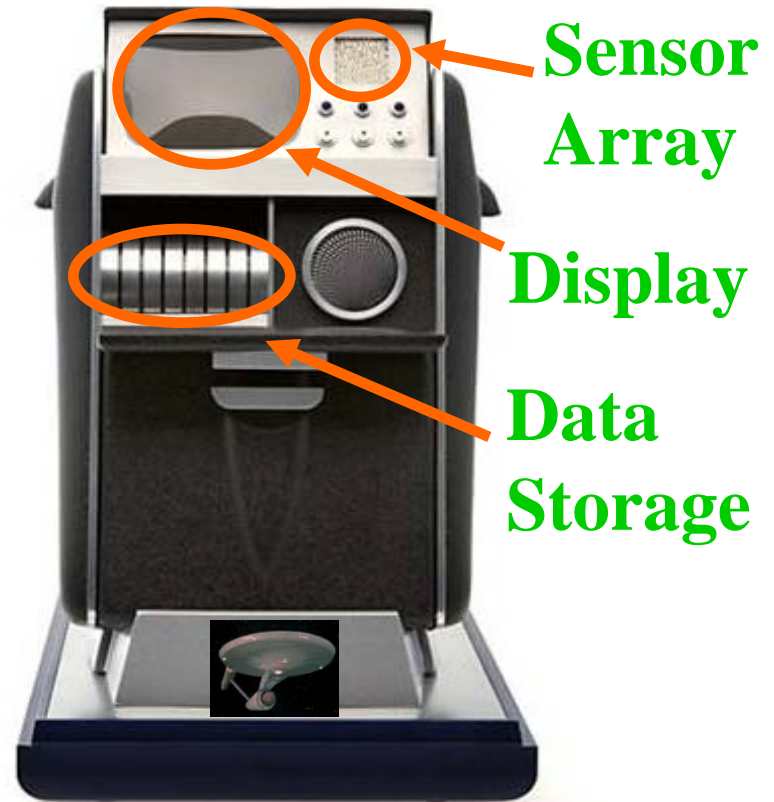
Most scientists regarded the new streamlined peer-review process as ‘quite an improvement.’

2. Validation Topics

Validation Topics

- **Desirable characteristics of DRM/DRI**

- Performs well
- Intuitive data
- Reliable
- Long lasting
- Easy to use and maintain
- Cheap (inexpensive)
- Easy to understand baseline in different situations
- Etc



Definitions of Validation

- **Validation:** ‘Confirmation by examination and provision of objective evidence that the particular requirements for a specified intended use are fulfilled.’ [ISO 17025 Rule 5.4.5.1]
- **Method Validation:** The process of establishing the performance characteristics and limitations of a method and the identification of the influences which may change these characteristics and to what extent. [Eurochem 1998]
- **Instrument validation:** The process of establishing that an instrument at any given moment is able to perform according to its design specification This process might be achieved for example by means of calibration or performance checks. [Eurochem 1998]

Validation Topics

Validation Goal: Data that is reliable and relevant.

- **Reliability:** Is the data produced by instrument technically sound?
- **Relevance:** Is the data useful and interpretable to achieve the goal?

Many DRMs attempt to do both!

Validation Topics

Reliability: Standards for Manufacturing and Performance

- **Manufacturing Standards and Certifications (the easier part)**
 - ISO and related manufacturing standards
 - Technical certifications, e.g. CE, UL, etc.
- **Performance Criteria and Standards (much harder)**
 - Few and far between
 - Only for targeted cases
 - Department of Homeland Security Standards database has 3825 entries. A handful are for detectors, mostly radiation, toxic gases, and personal alert safety systems.

Validation Topics

Reliability: Performance Characteristics (Metrics)

- The performance characteristics that should be validated will vary depending on the purpose of the procedure. Selecting and evaluating performance characteristics requires professional judgment.
- Four types of purposes:
 - Identification of unknown component(s)
 - Establishment of presence and/or absence of specified analyte(s) or classes
 - Quantitative analysis
 - Measurement of physical property

DRMs are designed to accomplish some or all these purposes. What performance characteristics are most applicable for each purpose?

Validation Topics

Reliability: Performance Characteristics (Metrics)

Six basic characteristics for DRM data:

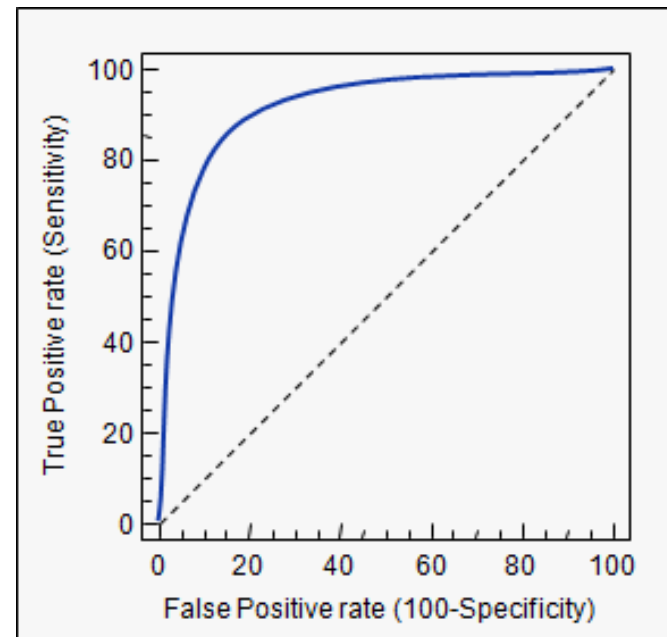
- Selectivity (e.g. in presence of interferents)
- Linearity (or other calibration model)
- Working range
- Limit of detection
- Limit of quantitation
- Accuracy (combination of precision and bias)

Some combination of characteristics may be applicable for each purpose. Which ones?

Validation Topics

Reliability: Performance Characteristics (Metrics)

- Characteristics are not independent, physically or statistically.
- It is necessary to understand the physical principles and apply an appropriate statistical test.
- Bayesian analysis and Receiver Operating Characteristic (ROC) curves are increasingly popular.



ROC curve (solid line)

Validation Topics

Reliability: Some sources are not always reliable

- Sales reps/advertising
- Internet
- Non-independent trade magazines
- Television
- Etc



Reliability: Third Party Testing

- For evaluating DRM performance
- Ensuring comparability between detector results
- Several types testing programs in existence
 - Performance
 - Technical Competence of Instrument
- Different organizations have different approaches

Reliability: EPA Technology Testing and Evaluation Program (TTEP)

Test, evaluate, and report on the performance of homeland security-related technologies

- building and outdoor area decontamination
- air, water, and wastewater treatment
- water security technologies
- detectors/monitors for air and water

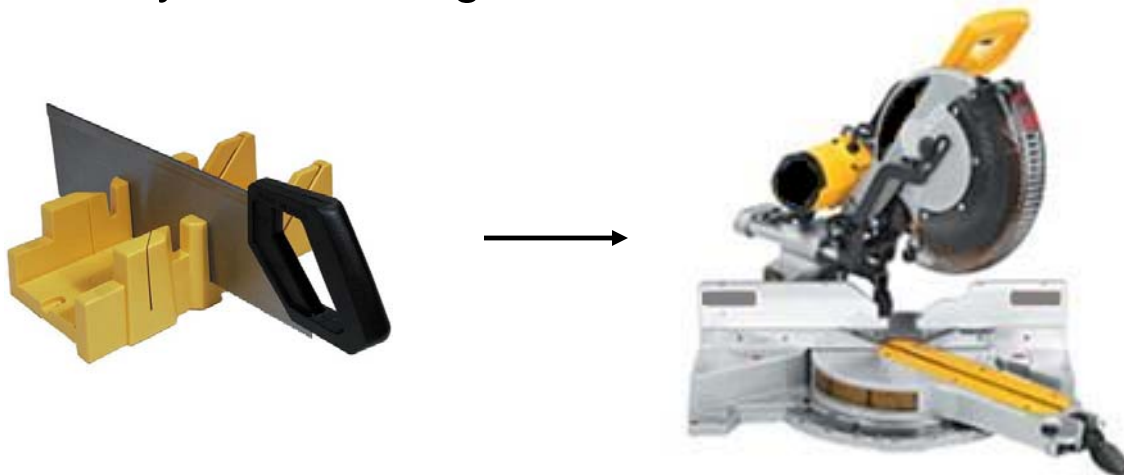
Desired Outcome:

Informed decision making by our customers to detect, contain, decontaminate, and manage hazardous chemical, biological, and radiological materials purposefully or accidentally introduced into structures, facilities, drinking water systems, or the environment.

Reliability: Technology Testing and Evaluation Program (TTEP)

Rationale

- Our customers rely on technology to protect human health and the environment
- In the event of an intentional or unintentional contamination event, need:
 - Tools for detecting, monitoring, treating, decontaminating, etc.
 - Information to assess the nature and extent of an event
 - Information to monitor the progress of a clean up and verify completion
- Continuously identify the evolving state-of-the art technology





Reliability: Technology Testing and Evaluation Program (TTEP)

Approach

- Primary focus is technology **testing** and **evaluation**
 - Test ***commercially available*** technologies using live agents under simulated field conditions
- Applied research and development testing
 - Determine if a technology not designed for homeland security-related uses could be applicable – stimulate development
- Information gathering
 - Information collection request (ICR)
 - Cataloging technologies that may be candidates for testing
 - Managing technology information to provide a ready source of information to emergency responders and consequence managers

Components

- Technology Testing
 - Voluntary and involuntary testing of technologies
 - De-emphasize vendor involvement in testing
 - Testing to a range of desirable performance characteristics, requirements, or spec's
 - Use existing test/QA plans when possible
 - Testing with live agents
- Products
 - User-oriented products for procurement/application decisions
 - Brief summary report at conclusion of effort (10 to 20 pages)
 - Technology briefs, side-by-side comparisons offered when possible
 - Reports note problems/deficiencies in performance



Reliability: Technology Testing and Evaluation Program (TTEP)

Components cont'd

- **Technical Peer Review**
 - Test plans and reports are reviewed by stakeholders and technical experts
- **Stakeholder Involvement**
 - Two stakeholder groups – Building Decontamination and Water Security
 - Environmental Response Technology Working Group (ERTG)
 - Taskforce on Research to Inform and Optimize CBR Response/Readiness (TRIO)
- **Information Sharing**
 - A single, quarterly newsletter containing information on all test and evaluation activities
 - Exhibiting at technical conferences and workshops

What Technologies Types Have Been/Are Being Tested?

- Building and Outdoor Area Decontamination
 - Decontamination for chemical, biological, and radiological (on going)
 - Detection and monitoring to support decontamination (on going)
 - Air cleaning/filtration (completed)
 - Wastewater treatment (completed)
- Water Security-Related Technologies
 - Detection and monitoring (on going)
 - Sampling (completed)
 - Drinking water treatment (completed)
 - Water system decontamination (planned)



Reliability: Technology Testing and Evaluation Program (TTEP)

Example: Air detectors

- Evaluate the performance of commercially available portable detectors for toxic industrial chemicals (TICs) and chemical warfare (CW) agents in air
- Target application is in guiding decontamination efforts, including:
 - identifying the nature of the contamination
 - identifying hot spots and uncontaminated areas
 - providing reliable response over a range of temperature and humidity conditions
 - minimizing false positives and false negatives caused by environmental interferences

Detectors Evaluated



Bruker RAID-M



Cerex UV Hound



**Envirochemics
ChemPro 100**



**Envirochemics M90-D1-C
CW Agent Detector**



**MSA Hazmat CAD Plus
formerly Microsensor
Systems, Inc**



SAIC S-CAD

Example: Testing description for air detectors

- Evaluation of performance characteristics under robust conditions
 - Challenges with TICs and CWAs in air at hazardous levels (e.g., immediately dangerous to life and health, IDLH)
 - **Controlled test conditions over range of temperatures (5 to 40 ° C) and relative humidity (<20 to 80 %RH)**
 - **Interferent vapors that might be encountered at the scene of a building contamination event**
- Assessment of operational factors (ease of use, alarms, reliability, maintenance, battery life, consumables, cost)

Example: Testing description (cont)

- Contaminants used in testing:
 - TICs: hydrogen cyanide (HCN), cyanogen chloride (ClCN), phosgene (COCl₂), arsine (AsH₃), and chlorine (Cl₂)
 - CWAs: Sarin (GB) and distilled mustard (HD)
- Interferences included latex paint fumes, air freshener vapors, ammonia floor cleaner vapors, exhaust hydrocarbons, and DEAE

Validation Topics

Reliability and Relevance: Operator integration

- ***Validation = Operator + Instrument***
- **Operator**
 - Is the operator technically familiar with instrument?
 - Does the instrument produce unambiguous data?
 - Is the operator able to interpret the data if instrument can not, or is known to be ambiguous?
 - Does the operator know if the instrument is producing ambiguous data?
 - Has the operator changed?
 - Is the new operator trained?
- **Instrument**
 - What is the instrument actually doing?
 - Does instrument perform as promised?
 - What do those results mean?



Validation Topics

Relevance: Validation of Entire Process, not just instrument component

- **Training is key component to validate operator**
 - Instrument operation
 - Data Interpretation
- Much training currently is focused on hazard/risk recognition, avoidance, or remediation



Relevance: Training Example-- Field Screening Workshop

Goals

- Identify the capabilities and limitations of field analytical instruments and tools that Hazmat teams may already possess, particularly for potential application to water contamination scenarios
- Understand how to interpret the results from field screening. Identify strategies for increasing the reliability of the results of field screening
- Understand the application of field screening with respect to reporting these results for all kinds of HAZMAT scenarios

Validation Topics

Relevance: Training Example-- Field Screening Workshop

The following organizations helped contribute to the contents of this workshop:

- United States (U.S.) Environmental Protection Agency (EPA): National Homeland Security Research Center/Water Infrastructure Protection Division and National Enforcement Investigations Center
- U.S. Department of Justice (DOJ): Federal Bureau of Investigation, Hazardous Materials Response Unit
- U.S. Department of Defense (DoD): Civil Support Teams, Technical Escort Unit
- City of Cincinnati
- City of Denver

Relevance: Training Example-- Field Screening Workshop

Workshop Format

- 6-8 hr workshop
- Target audience: HAZMAT responders
- Target instructors: To Be Determined (TBD)
- “Beta Version”- draft
- Two trial sessions: Denver and Cincinnati
- Four “Modules”



Module 1: Introduction

Module 2: Field Screening Considerations

Module 3: Results and Interpretation, Especially for Water

Module 4: Field Screening Exercises

3. Validation Related Needs

Validation Related Needs

Technical

- Technological revolution
- Technological evolution
- Getting most out of current technology -- software
- How does the instrument actually operate, particularly the software? – manufacturers won't often tell
- Standards for libraries and library software
- Standard for standards, e.g. what combination of performance characteristics are most applicable to each purpose?

Validation Related Needs

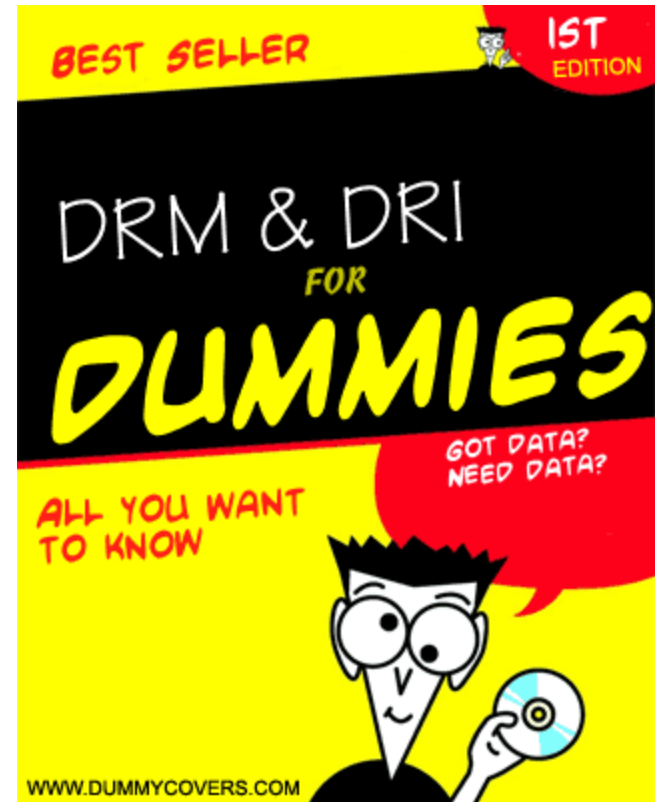
Policy

- Standards linked to funding:
 - Operator integration into DRM validation
 - Validation test plans
 - Comparable statistical interpretation of results
 - Instruments and instrument software
 - Libraries and library software
 - Standards for specific applications
- “Safe” disclosure route for manufacturers for software and hardware
- Increasing inventiveness – helping companies understand their potential market

Validation Related Needs

Training

- Targeted at audience for specific application
- Stressing entire process
 - Instrument fundamentals
 - Detection Strategy
 - Sample prep
 - Data Collection
 - Data Interpretation
 - Reporting of Results



Questions?

Matthew L. Magnuson

magnuson.matthew@epa.gov

513-569-7321

www.epa.gov/nhsrc

National Homeland Security Research Center

Cincinnati, OH; Research Triangle Park, NC; Washington, DC

Disclaimer: The U.S. Environmental Protection Agency funded, partially funded, managed, and/or collaborated in the research described in this presentation. It has been subject to an administrative review but does not necessarily reflect the views of the Agency. No official endorsement should be inferred. EPA does not endorse the purchase or sale of any commercial or non-commercial products or services.