THE STATES COMMON MEASURES PROJECT

Final Report

Revised: December 3, 2009

Prepared by: Steven DeGabriele, Massachusetts Department of Environmental Protection,
Susan Peck, Massachusetts Department of Environmental Protection and
Tara Acker, Northeast Waste Management Officials Association.

United States Environmental Protection Agency State Innovation Grant,
May 31, 2006 – May 31, 2009

An electronic copy of this report and other supporting materials can be found at
http://www.newmoa.org/hazardouswaste/measures/index.cfm

TABLE OF CONTENTS

The Common Measures Project Appendices listed below are available at:

http://www.newmoa.org/hazardouswaste/measures/index.cfm

ACKNOWLEDGEMENTS

The Project Management Team wishes to acknowledge the following project state participants and the NEWMOA support team:

PROJECT STATE PARTICIPANTS	
California Environmental Protection Agency Larry Matz, Project State Lead	New Hampshire Department of Environmental Services Rudy Cartier Tara Goodrich Sara Johnson Paul Lockwood, Project State Lead
Colorado Department of Public Health and Environment Shawna Morgan Mira Neumiller Kathy Stewart, Project State Lead Beth Williams Joyce Williams	New York State Department of Environmental Conservation Tom Killeen, Project State Lead Ralph Itzo Victoria Schmitt
Connecticut Department of Environmental Protection Mark Jepsen Nicole Lugli Maryann Nusom-Haverstock Susan Zampaglione, Project State Lead	Rhode Island Department of Environmental Management Joseph Antonio Tom Armstrong Ron Gagnon, Project State Lead
Maine Julie Churchill, Project State Co-Lead Sara Lippert, Project State Co-Lead Jim Rodrigue Suzanne Watson	Vermont Department of Environmental Conservation Tom Benoit Gary Gulka, Project State Co-Lead Paul Van Hollebeke Dave Linari Sue Rivard Steve Simoes, Project State Co-Lead
Massachusetts Department of Environmental Protection Steven DeGabriele, Project State Lead Leah Lyons Susan Peck William Sirull Inspectors Keith Anderson Kenneth Atkinson Matthew Barber David Brown Martha Caldwell Cecellia Catalan Scott Fasulo Daniel Gavin Karen Golden-Smith Samuel Isaac Eric Johnson John Keating Katherine Mcarthur John Mills Veronica O'donnell S. Pilegi A. Pizzini John Siderwicz John Stapler Giles Steele-Perkins Juliana Vanderwielen	Washington State Department of Ecology Alison Chamberlain, Project State Lead Laurel Cain Scott Lamb Michelle Underwood
NEWMOA PROJECT SUPPORT TEAM	
William Cass Terri Goldberg Tara Acker Andy Bray Nate Bisbee Lois Makina	

EXECUTIVE SUMMARY

State environmental agencies have increasingly been facing the dual pressures to oversee ever growing numbers of pollution sources with fewer resources, and to demonstrate that agency compliance assurance efforts are yielding measurable results. The States Common Measures Project began in 2006 as a multi-state effort to address both these pressures by:

- Evaluating the performance of targeted business sectors using common measures and statistical approaches.
- Beginning to use the results to identify particularly effective strategies states employ to promote good environmental performance on the part of the regulated community.

Measurement projects such as the States Common Measures Project are not undertaken just for the sake of measuring something. These measurements are needed to determine if the facility performance in a state is "good enough" to meet the state's policy objectives for the regulatory program and to identify any oversight practices that appear to be associated with higher performance levels. To the extent that the findings are reliable, the states are able to use the findings to make better decisions about efficient and effective programs.

Under the project, funded through the 2005-2006 EPA State Innovations Grant Program, the ten participating states - California, Colorado, Connecticut, Maine, Massachusetts, New Hampshire, New York, Rhode Island, Vermont and Washington applied the measurement methodology developed by Massachusetts for the Environmental Results Program (ERP). ERP is an innovative approach to improving and measuring the environmental performance of selected business sectors or groups. ERP uses a unique combination of linked compliance assistance, compliance certification and statistical performance measurement that leverages traditional compliance assurance activities to achieve improved performance for the selected group.

The project was implemented in four phases over three years:

- Phase 1) Group Orientation and Capacity Building.
- Phase 2) Making Decisions on Groups and Indicators.
- Phase 3) Data Collection, Field Observer Training and Statistical Methods.
- Phase 4) Data Analysis and Reporting Results.

<u>Phase 1) Group Orientation and Capacity Building</u>

This phase of the project served to ground the participants in the basic concepts of ERP measurement. In order to ensure the reliability and comparability of results for all participating state, training was developed and provided on the following topics:

- Data quality assurance.
- The use of indicators and statistics to measure performance.
- The characteristics of "good" indicators.

- Data collection methodologies.
- The characteristics of the group being measured.
- The indicators used to measure performance.

This background was provided prior to selecting the groups and indicators in the belief that it would lead to agreement on groups, indicators and data collection methodologies that were conducive to reliable measurement.

Phase 2) Making Decisions on Groups and Indicators

In Phase 2, project states applied the "lessons learned" about measurement to select two groups to measure and to define the set of indicators that all states would use to evaluate the performance of each group. While the grant commitment was to complete measurement on at least one group, the states decided, through a combination of "summer work" assignments for each state, in person meetings, and telephone conference calls, to measure two groups. Colorado, Connecticut, Maine, Massachusetts, New Hampshire, New York, Rhode Island, and Vermont all agreed to measure the performance of Small Quantity Generators of Hazardous Waste (SQG). In addition, the project states decided that to the extent that there was sufficient time during the grant period and a state had the capacity to take on an additional measurement project, project states could choose to work on the auto body sector as well. The project states successfully agreed on a set of common indicators for both groups. The indicators included both "regulatory requirements" and desirable best practices that went "beyond compliance."

Phase 3) Data Collection, Field Observer Training and Statistical Methods

Data collection was the focus of Phase 3. In this phase, participating states developed and implemented common approaches for identifying the universe of facilities in each group, and for selecting a random sample of facilities to inspect. The project states also agreed on the ways the results would be analyzed and presented. An automated data analysis and presentation tool developed for the Massachusetts and Colorado ERP programs, the "ERP Performance Analyzer," was augmented for use in this project.

A "sample-size calculator" tool developed for the Massachusetts ERP program was used to analyze the effect of sample size on the precision and reliability of the results. The project states decided that 56 was the optimal number of inspections for each state to conduct in light of available resources. In order to ensure data comparability across states, a common inspection checklist was developed for both sectors. All of the individuals that conducted field observations were trained in the use of the checklist and data quality assurance procedures. All eight states that participated in measuring SQG performance completed their inspections by the fall of 2008. However, due to budget, time constraints and universe composition, the number of inspections successfully completed per state varied from a low of 22 to a high of 57. Upon completion of the SQG field observations, the data collected on the inspection checklists was entered into the ERP Performance Analyzer. In addition to the SQG inspections, New York and Washington State completed a portion of their planned auto body inspections. However

this data was not collected and analyzed as a part of the States Common Measures Project.

Phase 4) Data Analysis and Reporting Results

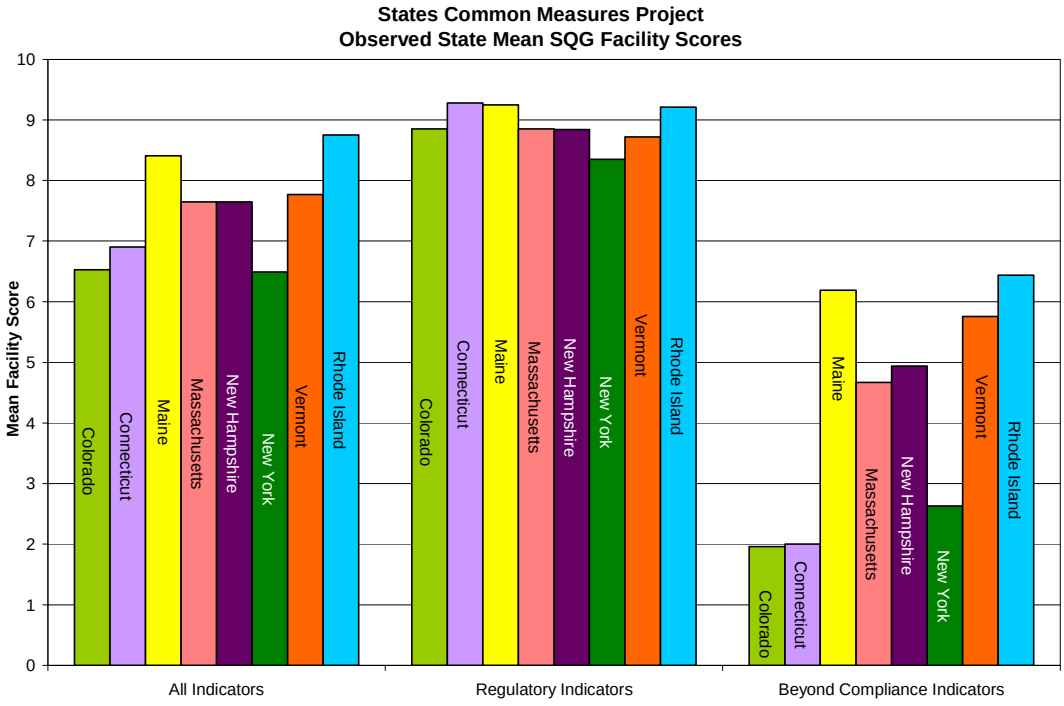
The final phase of the project involved analysis and presentation of the results. The ERP Performance Analyzer was used to calculate and display two different measures of SQG performance for each state:

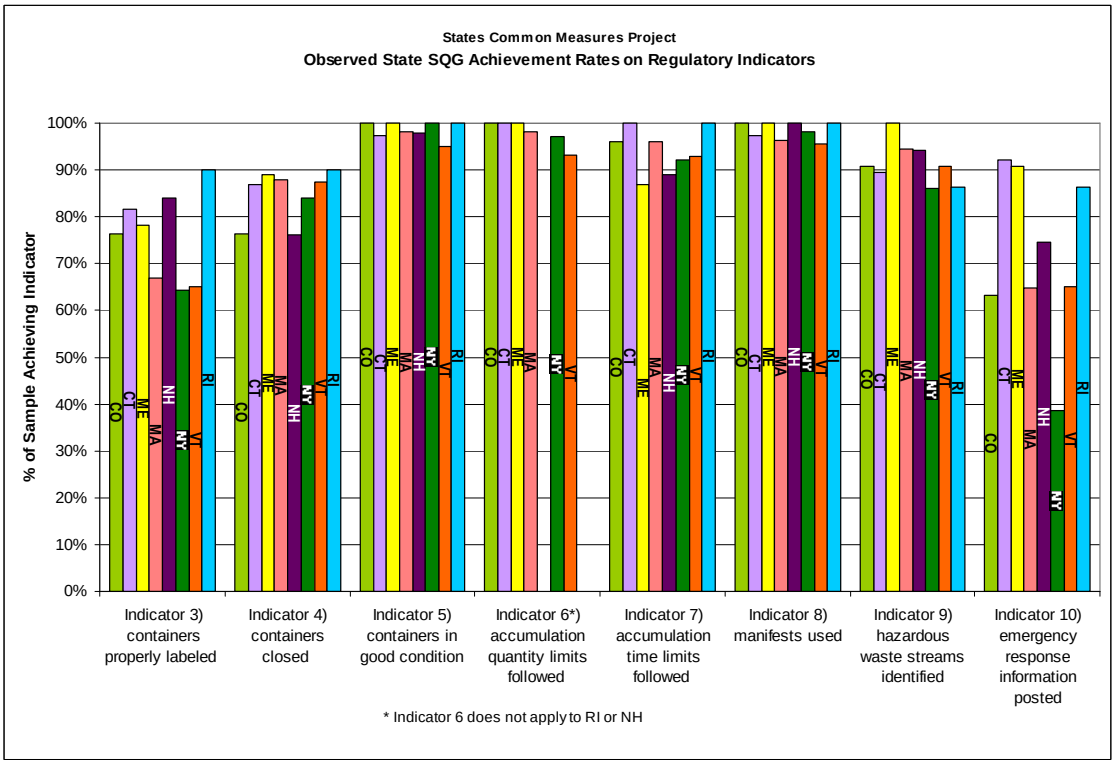
- The SQG mean facility score for "all indicators," "regulatory indicators," and "beyond compliance indicators." The facility score is the proportion of applicable indicators that the facility successfully achieved. It is measured on a scale of 0 – 10. A score of 10 indicates that the facility successfully achieved 100% of the indicators. A score of 0 indicates that the facility did not achieve any of the indicators. The mean facility score is the average score for all facilities in the state.
- The achievement rate on each indicator – the percentage of inspected facilities in the state that achieved the indicator.

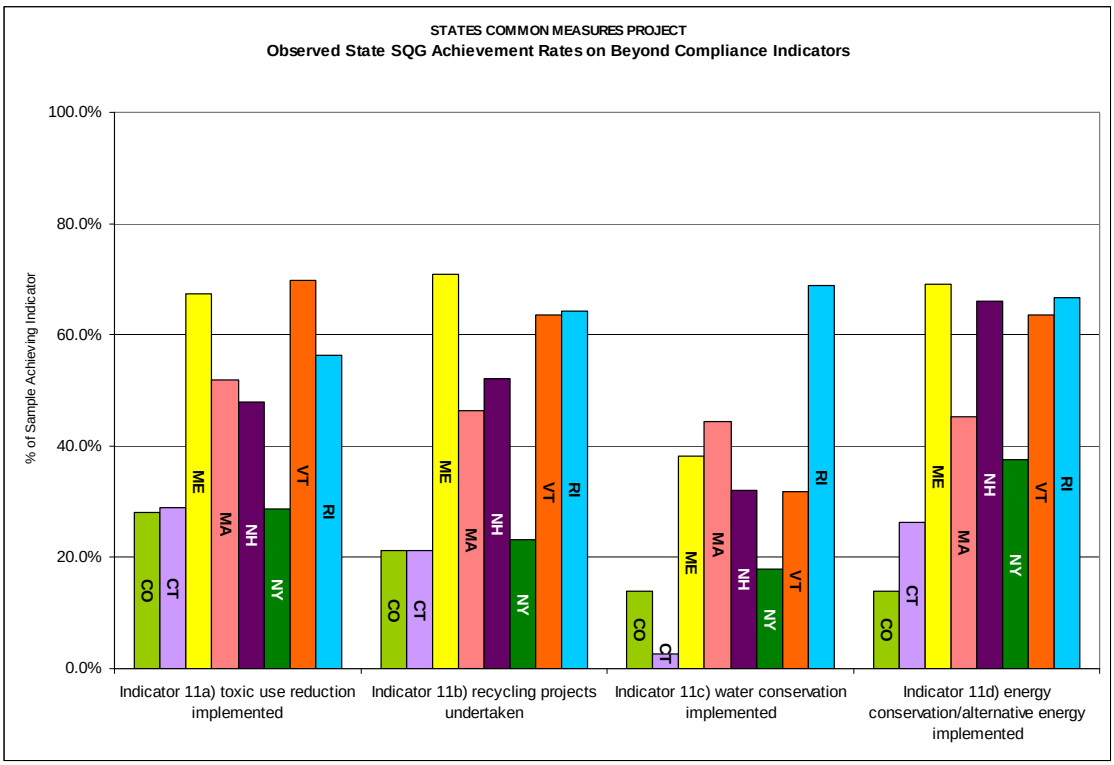
The ERP Performance analyzer also calculated the confidence intervals for the observed mean SQG facility score and the achievement rates in each state, and identified statistically significant differences between them at 3 confidence levels: 85%, 90% and 95%.

The charts below show observed mean facility scores for each state, and the observed achievement rates on each regulatory and beyond compliance indicator. Overall performance for regulatory indicators was relatively high across the states and no state felt the overall results were cause for serious concern. However, project states felt certain indicators warranted consideration of options for improvement. There were statistically significant differences in state SQG performance on the SQG regulatory indicator mean facility scores and on five of the eight individual regulatory indicators.

The beyond compliance performance was lower overall and showed wider variation among the participating states. There were statistically significant differences identified in state SQG beyond compliance indicator mean facility scores and on all four individual beyond compliance indicators.


States Common Measures Project
Observed State Mean SQG Facility Scores
Mean Facility Score
10
9
8
7
6
5
4
3
2
1
0
Colorado
Connecticut
Maine
Massachusetts
New Hampshire
New York
Vermont
Rhode Island
All Indicators
Regulatory Indicators
Beyond Compliance Indicators


States Common Measures Project
Observed State SQG Achievement Rates on Regulatory Indicators
% of Sample Achieving Indicator
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Indicator 3) containers properly labeled
Indicator 4) containers closed
Indicator 5) containers in good condition
Indicator 6*) accumulation quantity limits followed
Indicator 7) accumulation time limits followed
Indicator 8) manifests used
Indicator 9) hazardous waste streams identified
Indicator 10) emergency response information posted
* Indicator 6 does not apply to RI or NH



Participating states provided descriptions of the amount and nature of compliance and beyond compliance assistance provided, compliance inspection triggers and frequency, and enforcement tools and reporting requirements in place during the three years prior to the project. This program design information was compared to the performance results to identify if there were any oversight practices among the states that could be associated with higher performance rates.

This analysis indicated that onsite compliance assistance and active beyond compliance assistance programs appear to be associated with higher performance levels. The other program design attributes did not appear to be related to SQG performance levels.

Conclusion

Over a three year period, the ten project states were able to use the same set of common measures to evaluate the environmental performance of a common group of facilities. The project also created a replicable template that can be used by other agencies to build the capacity to measure group performance and to use the information to identify the most efficient and effective strategies for promoting better environmental performance.

This effort has already paved the way for other ERP-type measurement projects. Current activities being considered or under development include:

- The same project states selecting additional sectors to analyze

- The development of a second States Common Measures Project proposal which would build on the work of this project to do more in depth analysis of the relationship between program design and high SQG performance. It would also create more robust energy efficiency, pollution prevention, solid waste recycling, and water conservation beyond compliance performance indicators.
- A six-state initiative in EPA Region V to develop and implement a region-wide ERP for auto body shops that will include the use of the indicators developed under this project. The six-state initiative is funded through an EPA State Innovation Grant awarded in the spring of 2009.
- An EPA Region 1 and EPA Office of Enforcement and Compliance Assurance initiative that uses the indicators developed by this project to measure auto body performance in Massachusetts.

Achieving the full benefits of the States Common Measures Project requires the widespread adoption of ERP-type measurement across environmental agencies. The challenge going forward is to take meaningful steps to capitalize on the potential created by this project. The project states recommend that EPA provide the key leadership and financial support needed to:

1. Promote and expand the use of ERP-type measurement in both "core" and other work in states and EPA to:

    - Look within and beyond individual states to identify and adopt the most effective and efficient environmental performance improvement strategies.
    - Allow states the flexibility to deploy resources based on measured performance.
    - Promote the use of ERP-type measurement to routinely make environmental program priority and resource allocation decisions.

2. Support the development of an ERP Training Institute to codify this work into a formal ERP measurement curriculum.

0. INTRODUCTION

Due to two significant demands: 1) the need to effectively and efficiently improve the environmental performance of large groups of facilities with limited agency resources and 2) calls to demonstrate that agency compliance assurance efforts are yielding measurable results, a number of states began to actively employ a wide variety of traditional and innovative approaches to environmental compliance, enforcement, and assistance. These initiatives involved experimenting with various combinations of regulatory and non-regulatory tools to drive environmental performance improvements within identified regulated sectors and groups. Although state agencies have collected information about various aspects of their activities and the general performance of certain sectors, seldom has there been an accurate basis for drawing group performance conclusions and limited ability to compare the differences in group performance levels between two or more states working on a common sector or group. As a result, in 2006, the Massachusetts Department of Environmental Protection (MassDEP), building on the successful measurement approach developed for its Environmental Results Program[1] (ERP), designed and implemented the States Common Measures Project to apply the ERP measurement approach across more than one state. This project was funded through the 2005-2006 EPA State Innovations Grant Program.

0.1 THE OVERALL GOALS OF THE PROJECT

- Improve the ability of state environmental agencies to evaluate the performance of targeted business sectors including developing and implementing performance measures and using statistical approaches to analyze and report the results.

- Improve the ability of state environmental agencies to identify and adopt effective and efficient environmental performance improvement strategies based on those results.

[1] ERP, first developed by the MassDEP, is an innovative approach to improving and measuring the environmental performance of selected business sectors or groups. ERP uses a unique combination of linked compliance assistance, compliance certification and statistical performance measurement that leverages traditional compliance assurance activities to achieve improved performance for the selected group. For more information, go to: www.erpstates.org

The Anticipated Outcomes of the Project at Inception

Short-Term

- Agreement on groups to be measured and performance indicators by which they will be measured.
- Increased knowledge of measurement options, the role of data quality in measurement, data collection issues and results presentation.

Intermediate-Term

- Presentation of project states' performance data for the selected group(s).
- General comparison of performance levels of group(s) and differing state performance improvement strategies.

Long-Term

- Increased acceptance by states and EPA of ERP-type measurement.
- Common group performance data to support adoption of more effective and efficient group performance improvement strategies.

0.2 FUNDING AND SUPPORT

The MassDEP, in partnership with the Northeast Waste Management Officials Association[2] (NEWMOA), applied as the lead state agency for this multi-state project under the 2005-2006 EPA State Innovation Grant (SIG) program.

The SIG awarded by EPA was $255,000 over a 3 year period. In addition, MassDEP contributed $15,000 for upgrades to the ERP Performance Analyzer, an automated data analysis and presentation tool. The project was completed on budget and on time. *See Section 3.4 for more information on the ERP Performance Analyzer.*

0.3 OVERVIEW OF PROJECT PHASES

The project was implemented in 4 phases over a 3 year period:

- Phase 1 included group orientation on project goals, exploration of choices for groups to measure, types of measures and data quality training.
- Phase 2 involved making decisions on industry groups, developing common definitions and indicators, and reviewing data choices and implications.

[2] NEWMOA is a non-profit, non-partisan interstate governmental association. Their membership is composed of state environmental agency directors of the pollution prevention, hazardous and solid waste, and waste site cleanup programs in CT, ME, MA, NH, NJ, NY, RI, and VT. NEWMOA's mission is to develop and sustain an effective partnership of states to explore, develop, promote, and implement environmentally sound solutions for the reduction and management of materials and waste, and for the remediation of contaminated sites, in order to achieve a clean and healthy environment. For more information visit: www.newmoa.org

- Phase 3 involved applying the statistical and data quality assurance procedures to sample selection and data collection.
- Phase 4 involved data analysis and reporting of the results.

0.4 PROJECT ORGANIZATION

0.41 Project Management Team's Responsibilities

- Overall planning, organizing, directing, staffing and controlling of project.
- Managing outreach and training to participating states.
- Developing consensus concerning:
  - Business sectors to be measured.
  - Environmental performance indicators to be used.
  - Statistical and other methodologies to be used.
  - Data quality objectives.
- Developing and sharing ERP information tools and resources with states.
- Managing the development of a data management strategy in consultation with states.
- Managing the collection, analysis, and reporting of data.
- Reporting results to EPA.
- Implementing the project Quality Assurance Project Plan (QAPP).
- Submitting quarterly reports to EPA on project status.

Project Management Team		
Individual	**Role in Project**	**Organizational Affiliation**
Steven DeGabriele	Project Manager/MA State Lead	MassDEP
Susan Peck	Senior Project Analyst/Project Quality Assurance Officer	MassDEP
William Cass	Support Services Manager	NEWMOA
Tara Acker	Senior Management Consultant	NEWMOA

0.42 Quality Assurance Officer's Responsibilities

- Maintaining the QAPP.
- Distributing the QAPP and maintaining the distribution list.
- Conducting readiness reviews.
- Developing data management and analysis procedures.
- Overseeing quality assurance and quality control of data.

0.43 NEWMOA's Responsibilities

- Serving as a clearinghouse for project information.
- Supporting the group and indicator selection process.

- Organizing measurement and statistical methodology training.
- Providing direct support to individual states.
- Supporting the collection and analysis of performance data from participating states.
- Drafting final project report.

NEWMOA Directors (as of May 2006)	
Individual	**Organizational Affiliation**
Michael Harder	Connecticut Department of Environmental Protection
Robert Kaliszewski	Connecticut Department of Environmental Protection
Stephen K. Davis	Maine Department of Environmental Protection
Ron E. Dyer	Maine Department of Environmental Protection
Steven DeGabriele	Massachusetts Department of Environmental Protection
Sarah Weinstein	Massachusetts Department of Environmental Protection
Anthony Giunta	New Hampshire Department of Environmental Protection
Frank Coolick	New Jersey Department of Environmental Protection
Michael DiGiore	New Jersey Department of Environmental Protection
Stephen B. Hammond	New York Department of Environmental Conservation
David R. O'Toole	New York Department of Environmental Conservation
Jeff Sama	New York Department of Environmental Conservation
Ron Gagnon	Rhode Island Department of Environmental Management
Terrence Gray	Rhode Island Department of Environmental Management
P. Howard Flanders	Vermont Department of Environmental Conservation
Gary Gulka	Vermont Department of Environmental Conservation

0.44 EPA's Responsibilities

- Assisting with coordination of the participating states.
- Assisting with steering the project and ensuring that it remains on track.
- Assisting with statistical analysis including, if deemed necessary, third party review.
- Providing contractor support on the development of the Quality Assurance Project Plan (QAPP) through CrowEnvironmental and support with comparing the Common Measures Project auto body indicators to the new area source rule through Industrial Economics, Inc.

Environmental Protection Agency Participants		
Individual	**Role in Project**	**Organizational Affiliation**
Marge Miranda	EPA Grant Manager	USEPA Region I
Beth Termini	EPA Project Liaison	USEPA Office of Policy, Economics and Innovation (OPEI) and Region I

0.45 Involvement Level of Project States

States participating in the States Common Measures Project could designate their level of involvement as follows:

- Participating states agreed to fully participate in all 4 phases of the project for at least one business sector.
- Learning states agreed to participate on conference calls, selected meetings and related project activities to learn more about measurement approaches to increase overall state capability to develop and implement performance measurement-based programs.

Project States
Connecticut Department of Environmental Protection
Colorado Department of Health and Environment
Massachusetts Department of Environmental Protection
Maine Department of Environmental Protection
New Hampshire Department of Environmental Services
New York State Department of Environmental Conservation
Rhode Island Department of Environmental Management
Vermont Department of Environmental Conservation
Washington State Department of Ecology
California Environmental Protection Agency
Notes: • *New York and Colorado began the project as "learning states" and changed their status to "participating states."* • *California completed phases 1 and 2 of the project as a "learning state."*

0.46 Contractor Support

The project used contractors to help with preparing the QAPP, statistical training, enhancing the analytical database and assistance with meetings.

Contractors
CrowEnvironmental: ➢ Provided a one day training for project states on the use of ERP-type statistical methods, data collection instruments, sampling procedures and data quality indicators.
TetraTech, EM Inc: ➢ Enhanced an existing custom MS-Access-based application, originally designed by TetraTech, known as the ERP Performance Analyzer. The ERP Performance Analyzer is used to perform statistical and graphical analyses of data from ERP compliance inspections and/or self certifications. The output from the system is used to describe the environmental performance of different business sectors at a point in time, changes in performance over time and differences in performance across regulatory jurisdictions.

0.5 INDUSTRY GROUPS SELECTED FOR THE PROJECT

Participating states selected small quantity generators of hazardous waste (SQGs) as well as auto body shops as the groups for measurement. States believed that having a sector with one environmental medium (hazardous waste) that was relatively easy to understand and had common requirements across the states due to federal program standards, as well as one sector that was multi-media (air, hazardous waste and wastewater) with fewer common requirements across states, would expand the analysis and allow for greater learning. The project required participating states to complete all four phases for at least one sector, but depending on resources, participating states could elect to work on more than one sector. *See Sections 1.3 and 2.0 for more information about the process for selecting the project groups.*

The States Common Measures Project completed work on all four phases for the small quantity generator sector and completed work through phase 3 for the auto body sector.

0.6 QUALITY ASSURANCE MEASURES

A Quality Assurance Project Plan (QAPP) was developed for the Common Measures Project that included a detailed overview of all aspects of project management, data generation and acquisition, assessment/oversight and data review and evaluation. The QAPP is a mandatory component of any EPA-funded State Innovation Grant project that involves the collection and analysis of environmental data. It ensures that the data that is collected is accurate, precise, complete and reliable for its intended purpose. *See Appendix A for QAPP.*

In order for this project to be successful, data from each participating state had to be of the same, known quality. Quality issues were anticipated for each phase of the project and were identified and discussed with the states, and the QAPP was distributed to all project states. Participating states were asked to implement the required data quality assurance and analysis procedures and agree upon a set of quality criteria before beginning any data collection work. States were also required to sign a certification statement of data quality to verify that all data submitted for the project met the data quality standards described in the QAPP and in the project's training materials. *See Section 3.3 for more information on steps taken to ensure field data quality.*

0.7 REPORTING

As part of project reporting requirements, quarterly updates were provided to EPA by the Project Manager detailing all project activities. *See Appendix B for the EPA Quarterly Reports.*

The project also reported to the NEWMOA Directors and the States ERP Consortium[3] to keep both groups regularly informed throughout the project.

[3] The States ERP Consortium is a voluntary organization of 19 states and EPA formed in 2006 to promote the use of the Environmental Results Program (ERP). The primary goals of the Consortium are to: 1) share

Project state leads were responsible for keeping the key agency officials within their state informed of the project status.

0.8 ORGANIZATION OF FINAL REPORT

The final report has the following sections:

- SECTION 1: Group Orientation and Capacity Building (Phase 1).
- SECTION 2: Decisions on Groups, Indicators and Data Choices (Phase 2).
- SECTION 3: Data Collection, Field Observer Training and Statistical Methods (Phase 3).
- SECTION 4: Data Analysis and Reporting Results (Phase 4).
- SECTION 5: Auto Body Sector.
- SECTION 6: Project Conclusion and Recommendations
- APPENDICES

information and tools, 2) expand support for ERP within states and EPA, 3) report and communicate ERP results to key audiences, 4) identify new applications for ERP and 5) identify strategies to accelerate and achieve economies of scale in automating ERPs. For more information go to: www.erpstates.org

SECTION 1: GROUP ORIENTATION AND CAPACITY BUILDING (PHASE 1)

This Section of the Report Covers:

On June 23, 2006, the Project Management Team organized a project kick-off and training meeting involving approximately thirty stakeholders that included project states, staff from EPA New England, EPA OPEI and NEWMOA. The purpose of the meeting was to include a review of project goals and expected outcomes; discuss data quality choices, characteristics and limitations; understand indicators and a preliminary discussion of candidate groups for measurement. The meeting was opened by Ira Leighton, Deputy Regional Administrator, EPA New England, who expressed the importance of the project and EPA's commitment to its success.

As part of the orientation, states were asked to consider the following:

- What Would Success Look Like at the End of the Project and Into the Future?

- What Fears and Potential Barriers Did They See to Achieving the "Future State" Vision?

- What Were The Potential Misconceptions?

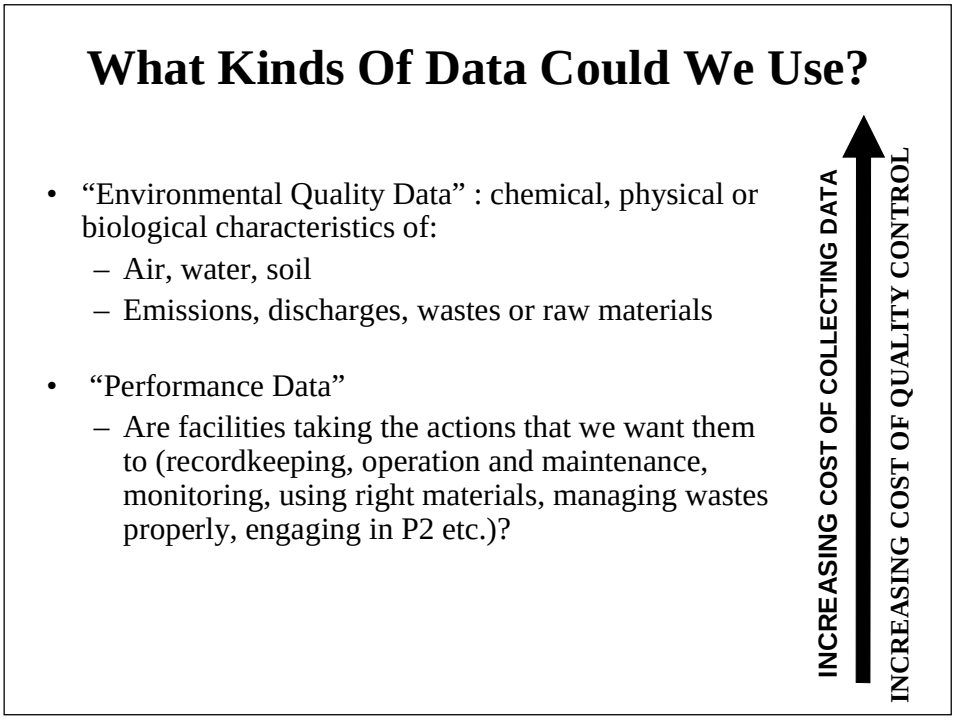
    - The Project used ERP-type measurement, and was NOT developing a full ERP program for the selected industry group(s).
    - The Project used ERP-type measurement, and was NOT using other measurement approaches or creating a new measurement approach.

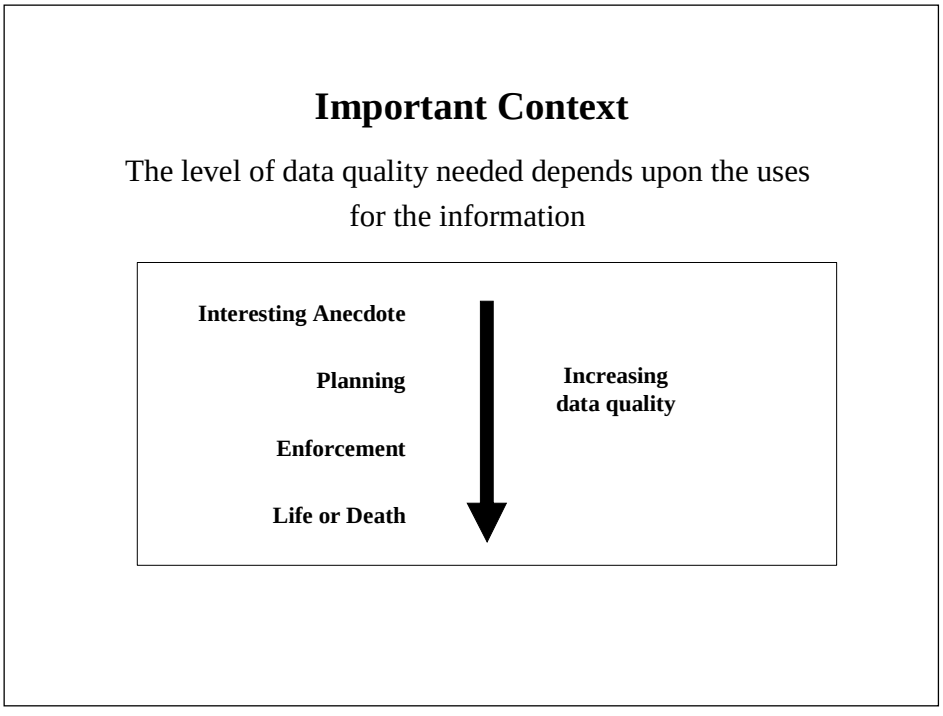
After identifying expected outcomes, potential barriers and misconceptions, states were ready to begin the evaluation of data choices. The following sections provide a summary of the information reviewed.

1.1 UNDERSTANDING DATA QUALITY CHOICES, CHARACTERISTICS AND LIMITATIONS

1.1.1 What IS Data, Anyway?

- Data is what is used to measure the indicators of environmental performance of selected groups.





Where Can The Data Come From?

PRIMARY DATA

- Data we collect in the field – inspections, sampling, surveys
- Data we we collected previously -- file or data base review

SECONDARY DATA

- Data submitted to us by someone else (e.g report review)
- Data collected and analyzed by others

New vs. Existing Data

- Data collected specifically for the project
- Data collected previously for other purposes

What Are The Key Factors Influencing Data Quality?

Data Quality Indicators	The Quality of Data Collection & Analysis
• Precision • Sensitivity • Representativeness • Comparability • Completeness • Bias	• Verification • Validation • Integrity

1.1.2 Data Quality Indicators and Quality of Data Collection and Analysis[4]

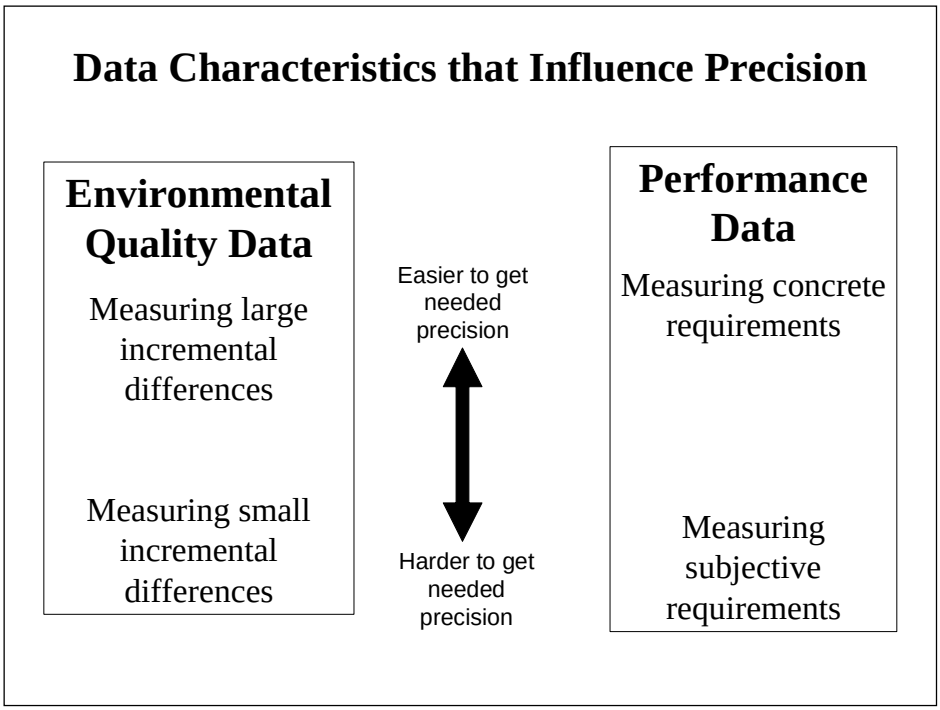
Data Quality indicators used in this project are presented below.

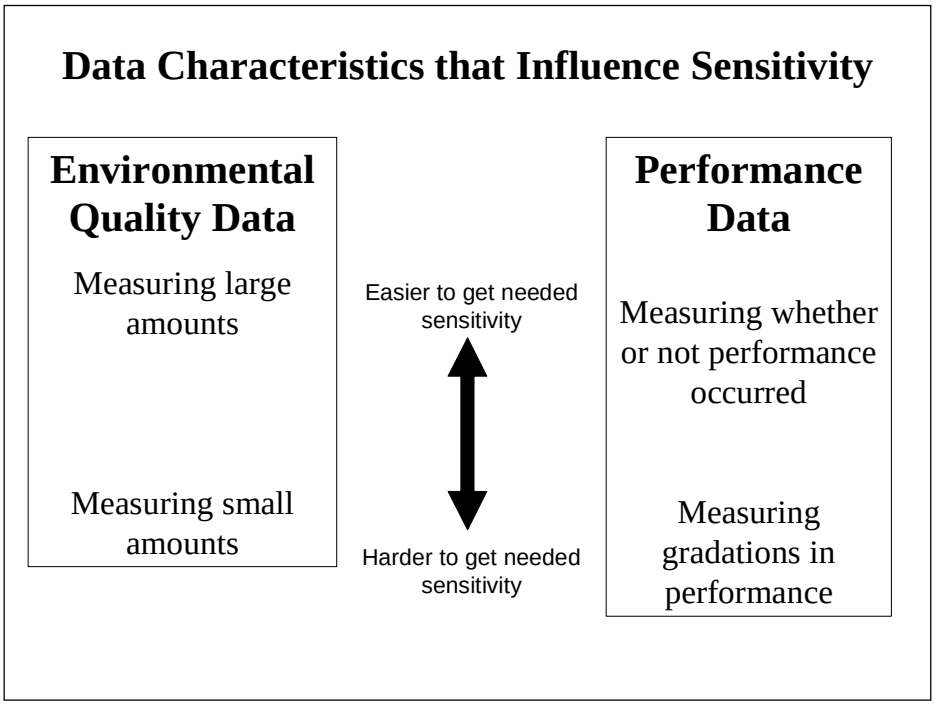
- Precision is the measure of agreement among repeated measurements of the same property under identical or substantially similar conditions.
- Sensitivity is the capability of a method or instrument to discriminate between measurement responses representing different levels of the variable of interest.
- Representativeness is the measure of the degree to which data suitably represent a characteristic of a population, parameter variations at a sampling point, a process condition, or an environmental condition.
- Comparability is a qualitative expression of the measure of confidence that two or more data sets may contribute to a common analysis.
- Completeness is a measure of the amount of valid data obtained from a measurement system, expressed as a percentage of the number of valid measurements that should have been collected.
- Bias is systematic or persistent distortion of a measurement process that causes error in one direction.
- Data validation is an analyte and sample matrix-specific process to determine the analytical quality of a specific data set.

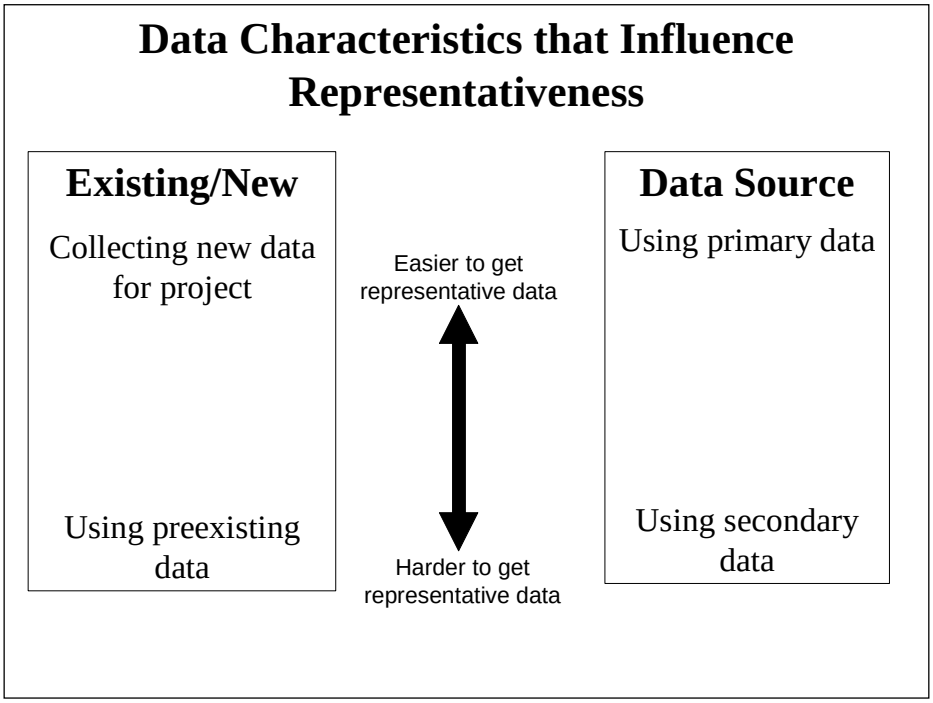
[4] Source: EPA Introduction to Data Quality Indicators
http://epa.gov/quality/trcourse.html

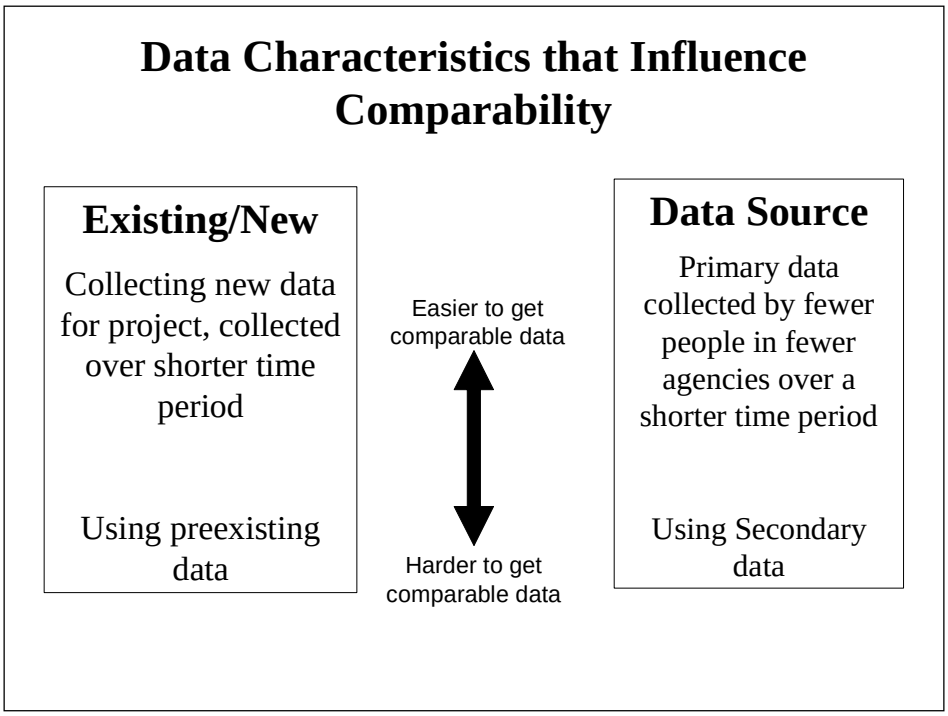
- Data verification refers to the procedures needed to ensure that a set of data is a faithful reflection of all the processes and procedures used to generate the data.
- Lack of integrity affects all aspects of data interpretation, especially data used for decision making.

The charts below were used in the training for project states to describe the data characteristics that can affect data quality:






Data Characteristics that Influence Representativeness
Existing/New
Collecting new data for project
Using preexisting data
Easier to get representative data
Harder to get representative data
Data Source
Using primary data
Using secondary data


Data Characteristics that Influence Comparability
Existing/New
Collecting new data for project, collected over shorter time period
Using preexisting data
Easier to get comparable data
Harder to get comparable data
Data Source
Primary data collected by fewer people in fewer agencies over a shorter time period
Using Secondary data

1.2 UNDERSTANDING INDICATORS

The project states next evaluated the issues associated with using indicators to measure performance.

1.2.1 What is an environmental performance indicator?

- An environmental indicator is a chosen characteristic that describes environmental performance.
- A select set of indicators will be used to measure the environmental performance of a group(s).

1.2.2 How are indicators used to measure group environmental performance?

- When combined with statistical measurement methods, the use of indicators can provide an understanding of the performance of a group at a point in time. Changes in performance can also be measured over time.

1.2.3 What kinds of indicators may we select?

- Activity measures, e.g., was the necessary equipment installed.
- Outcome measures, e.g., how much pollution was created.
- Regulatory measures. e.g., are requirements being met.
- Beyond compliance measures, e.g., measures being followed even if they are not required.

1.2.4 What are some of the key choices and considerations?

- Indicators should reflect the "most important" environmental performance practices (requirements) for the group. Although "higher value" indicators that relate to potential or actual emissions, discharges or releases may be preferable, they may not be feasible in some cases.

- Who should pick the indicators: Agency staff and mangers? The agency and representatives of the regulated group? Should the general public be asked?

- Should the indicators be limited to regulatory requirements or should beyond compliance practices be included?

- How many indicators should there be for the group?

- Is data available for the indicators and what is the quality of the data?

- Will new data be required?

- A test of the value of our selected indicators is answering the question "If I inspected a facility and I determined it was in compliance with the indicators, I would feel confident that its overall environmental performance is high."

1.2.5 Are there other issues to consider in using indicators?

- Those practices (requirements) not covered by the indicators that are not measured.

- Does using indicators mean that those practices (requirements) not covered are unimportant and maybe should be eliminated?

- What if the group has good performance as measured by the indicators, but the overall compliance rate is lower?

- Indicators provide information that allows the agency to focus its limited resources on the problem areas identified by the indicators and to not worry or to limit resources expended on areas of good performance.

- Group performance measurement using indicators allows the agency to report on the both the overall environmental performance level of the group and performance related to the individual indicators.

This session raised many important questions about indicators. The purpose of this exercise was to have states grapple with the types of indicators that could be selected. It was also agreed that final selection of indicators would not limit an individual state from measuring other activities beyond those indicators selected for the project.

1.3 PRELIMINARY DISCUSSION OF CANDIDATE GROUPS FOR MEASUREMENT

The last component of the kick-off meeting involved a preliminary discussion of candidate groups to measure. States were asked to identify a list of groups that they were currently working on. This list included at least two types of groups: those based on a particular environmental regulatory program such as hazardous waste generators and those based on an industry type such as autobody shops. Next, the Project Management Team asked states to identify other groups that they were interested in evaluating that were not reflected on the initial list. *See Section 2 for the group preference list.*

In order to select candidate groups for measurement, states needed to understand the characteristics and the implications of selecting a group.

The chart below shows how the type and size of a group affect data quality:

Complexity of Group

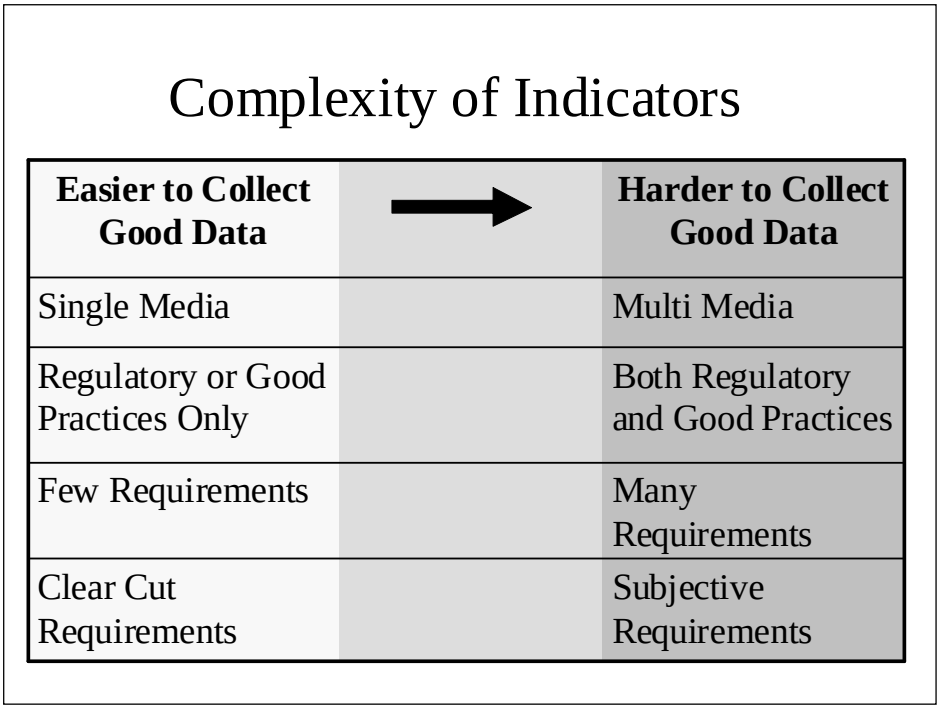
Easier to Collect Good Data	→	Harder to Collect Good Data
Similar Facilities		Diverse Facilities
Small Universe		Large Universe

The chart below shows how the level of experience with a group can affect data quality:

Experience With Group

Easier to Collect Good Data	→	Harder to Collect Good Data
Existing Group		New Group
Existing Data System		No Data System
Existing Common Definitions		No Common Definitions
Known Universe		Unknown Universe
Existing Checklists		No Checklists
Historical Data		No Historical Data

The chart below shows the complexities to consider when selecting indicators:



Other Factors to Consider in Selecting a Group

- The similarity of regulatory requirements within a group across state.
- The number of facilities across the states, e.g. were there enough facilities in each of the states to make the analysis worthwhile.
- Whether states include federal as well as state regulatory requirements.
- The environmental importance of the group, e.g., does group candidate for measurement have the potential for a significant impact on the environment if it is performing poorly.
- Whether the measurement can be linked to an environmental outcome, e.g. pounds of pollution reduced.
- Whether the investment of time working on a small universe yields sufficient environmental benefit.
- Whether the group is so "problematic" that it may be too difficult to use in a common measures project, i.e. resource intensive because of follow-up enforcement actions that would need to be undertaken.
- Difficulty in identifying a universe, e.g., not regulated by the state as a group and has no existing information, or because the state has not focused resources on identifying all the facilities subject to the program so many are "outside the regulatory system."
- Whether the group to be measured is subject to a mandatory or voluntary program.

The purpose of this session was to review these considerations. There were no final decisions made during this meeting because the purpose was to provide background and context for selecting a group or groups to measure. States were given a "summer work

assignment" which asked them to identify preferred groups to measure and to develop a preliminary list of possible indicators and the data quality issues associated with those indicators to determine whether their group of choice was viable for the measurement project. *See Section 2 for summer work assignments.*

1.4 FURTHER TRAINING ON DATA QUALITY AND STATISTICAL APPROACHES

The Project Management Team contracted with a firm to provide further data quality training on the use of ERP-type statistical methods, data collection instruments, sampling procedures and data quality indicators. This meeting took place on September 28, 2006 and the course addressed the following issues:

- Overview of the six Data Quality Indicators that should be considered in any measurement project: precision, bias, completeness, representativeness, comparability and sensitivity.

- Use of statistics in measuring innovative policy approaches, particularly focusing upon the techniques used in the Environmental Results Program (this two-part session included an introduction to statistics, benchmarking and comparisons).

- Identification of the sampling frame, also known as establishing the universe of facilities.

- Principles of good data collection, including the characteristics of good data collection instruments.

- Quality considerations specific to the use of secondary data.

- An interactive session where states worked in smaller groups on applying the concepts that were learned.

See Appendix C for Introductory Training and Quality Measurement information.

1.5 OBSERVATIONS AND LESSONS LEARNED FROM PHASE 1

√ The kick-off meeting was helpful to states in defining goals and objectives for the project. This first meeting was a critical step in establishing common understandings and a vision for the 3 year project.

√ Some of the state participants initially thought the project was going to involve a full ERP, i.e. certification, outreach and measurement program. Phase 1 helped clarify that the States Common Measures Project would focus on performance measurement across states, not compliance certification or other ERP techniques.

- √ Some participants noted that the choice of group(s) to measure for the project would be more helpful if determined in advance as it would help the state to decide their level of participation and what resources they could invest in the project.

- √ The "participating state" and "learning state" categories were useful in encouraging more states to become involved in the project and ultimately to participate in measurement. For example, New York and Colorado began as learning states and then became participating states.

- √ Asking the NEWMOA State Program Directors to sign letters of commitment was helpful in generating support for the project within the state agencies.

SECTION 2: DECISIONS ON GROUPS, INDICATORS AND DATA CHOICES (PHASE 2)

This Section of the Report Covers:

As a first step in selecting the group (or groups) to measure and choosing indicators to describe the performance of the group (or groups), the Project Management Team designed and distributed two summer work assignments. The first assignment asked states for a preliminary identification of preferred groups to measure. The second assignment asked states for a preliminary list of possible indicators for each preferred group and to consider the data quality issues associated with those indicators.

2.1 SELECTING THE GROUP OR GROUPS TO MEASURE

2.1.1 Summer Work Assignment 1: Group Preference Checklist

This assignment asked states to identify their group preferences. States reviewed a list of over 30 groups compiled from a brainstorming exercise and were asked to check off which groups they **would like** to select for the Common Measures Project, which groups they **would not like** to select and which groups they were **very interested in** selecting. The purpose of the exercise was to narrow down the list of groups to measure based on state interest. Below is the Group Preference Checklist:

STATES COMMON MEASURES PROJECT
GROUP PREFERENCE CHECKLIST:

STATE: **GROUP**	**Check if your state would do this group**	**Check if your state would not do this group**	**Also Check if this is a group your state is very interested in doing**
Small quantity hazardous waste generators			
Large quantity hazardous waste generators			
TSDFs			
Air Operating Permits			
Marinas			
Soils Recyclers			
Underground Storage Tanks			
Colleges and Universities			
Hospitals			
Photo Processors			
Commercial Offsite HW Recyclers			
Printers			
Stage 2 Programs			
NPDES Majors			
NPDES Minors			
Active Landfills			
Exterior Lead Paint Contractors			
K-12 Schools			
Auto Salvage / Junk Yards			
Dry Cleaners			
Auto body			
Auto repair			
Dental clinics			
Used oil handlers and recyclers			
Electronics recyclers			
Furniture strippers			
Radiator repair			
Metal fabricators			
Spray booth operators			
Platers			
Boat builders			
Small automotive touch up operations			
Portable minor air sources eg. wood chippers, stump grinders)			
Municipal operations: POTWS, DPWs, Water Utilities			

The Project Management Team compiled a summary of the responses and the top five groups that states were **"very interested in doing":** Small Quantity Generators of Hazardous Waste (SQGs), Auto Body Shops, Underground Storage Tanks (UST), Dental Offices and Stage II Programs. After further discussion, "Stage II Programs" was not selected as a measurement group and was not included in the second work assignment. *See Appendix D for individual state responses to group preference checklist.*

2.2 SELECTING THE INDICATORS

2.2.1 Summer Work Assignment 2: Group and Indicator Packet

The second work assignment asked states to complete a Group and Indicator Packet for each of the groups that they were interested in measuring. The purpose of this assignment was to:

- Further refine thinking about the top four groups in order to select final group(s) to measure.
- Begin to develop preliminary indicators for those selected groups.
- Identify the data issues associated with those indicators.

The Group and Indicator Packet consisted of two parts: a) The Group Evaluation Chart and b) The Indicator Evaluation Charts. States were asked to complete these charts based on the data quality training from the project kick-off meeting as well as the follow-up training each state received in September 2006.

2.2.2 Part A: The Group Evaluation Chart

The Group Evaluation Chart asked states to identify a group and consider a series of issues that would influence selection of the group for measurement:

STATES COMMON MEASURES PROJECT

Group and Indicator Packet:

(Fill out (electronically) one packet for EACH group your state would be willing to participate in measuring. A packet consists of one Group Evaluation Chart and one set of Indicator Evaluation Charts. Send completed packets as an attachment to Bill Cass wcass@NEWMOA.org by Tuesday August 22, 2006)

GROUP EVALUATION CHART:

(Enter your state, name of group, and level of interest in the group on the first line. Then mark the appropriate box for each consideration)

STATE		GROUP			__ Would do	__ Most Interested
Easier to Collect Good Data		***Easier***	***Somewhat difficult***	***More difficult***	***More Difficult to Collect Good Data***	
Consideration		**Complexity of Metrics**				
Single Media					Multi Media	
Regulatory or Good Practices Only					Both Regulatory and Good Practices	
Few requirements					Many Requirements	
Clear Cut Requirements					Subjective Requirements	
Complexity of Sector						
Similar Facilities					Diverse Facilities	
Small Universe					Large Universe	
Prior Experience With Sector						
Existing Sector					New Sector	
Existing Data System					No Data System	
Existing Common Definitions					No Common Definitions	
Known Universe					Unknown Universe	
Existing Checklists					No Checklists	
Historical Data					No Historical Data	
Similar requirements across states					Dissimilar requirements across states	
Other Considerations						
More Desirable/Easier Sector		***Most desirable***	***Somewhat desirable***	***Less desirable***	***Less Desirable/Harder Sector***	
Even distribution of facilities across states					Uneven distribution of facilities across states	
Important environmental concern					Lesser environmental concern	
Regulated by the states and EPA					State only sector	
Ability to link to an environmental outcome					Inability to link to an environmental outcome	
Facilities are sophisticated about environmental regulation					Facilities are unsophisticated about environmental regulation	
No language barriers					Language barriers	
Other (list:)						

Data from the chart was aggregated and the Project Management Team held a conference call on November 6, 2006 to discuss the results and to determine the states' readiness for selecting final groups for measurement. Prior to the call, states were asked to seek their organization's input on candidate groups. *See Appendix E for aggregated data on group evaluation chart.*

A follow-up call was held on November 21, 2006 and states selected **small quantity generators of hazardous waste** and the **auto body sector** as the measurement groups.

2.2.3 Part B: The Indicator Evaluation Chart

The second part of the work assignment asked states to develop a list of draft indicators for the small quantity generator of hazardous waste group and for the auto body group. States also evaluated each draft indicator to determine how well it would meet the data quality standards of precision, sensitivity, representativeness, completeness, bias and validation. States were also asked to identify the likely data source (e.g. new inspection, prior inspection, report) for each indicator.

Note: States interested in analyzing only one group, prepared draft indicators for just that group.

Below is the Indicator Evaluation template:

STATES COMMON MEASURES PROJECT
Group and Indicator Packet
INDICATOR EVALUATION CHART

Complete one indicator chart for each indicator you are proposing for the group
List your state and group name on each page
For each indicator:

-Provide a number and description (e.g., 1. Stores hazardous waste in properly labeled drums), indicate data you would use to measure it and where that data comes from (e.g. prior inspection data from compliance and enforcement files, new inspection or site visit, report, survey etc.).
-Mark the appropriate box for each "data quality indicator" and briefly describe any important data quality issues related to that data quality indicator (e.g. representativeness might be poor because using data from prior inspections and these were not chosen randomly).

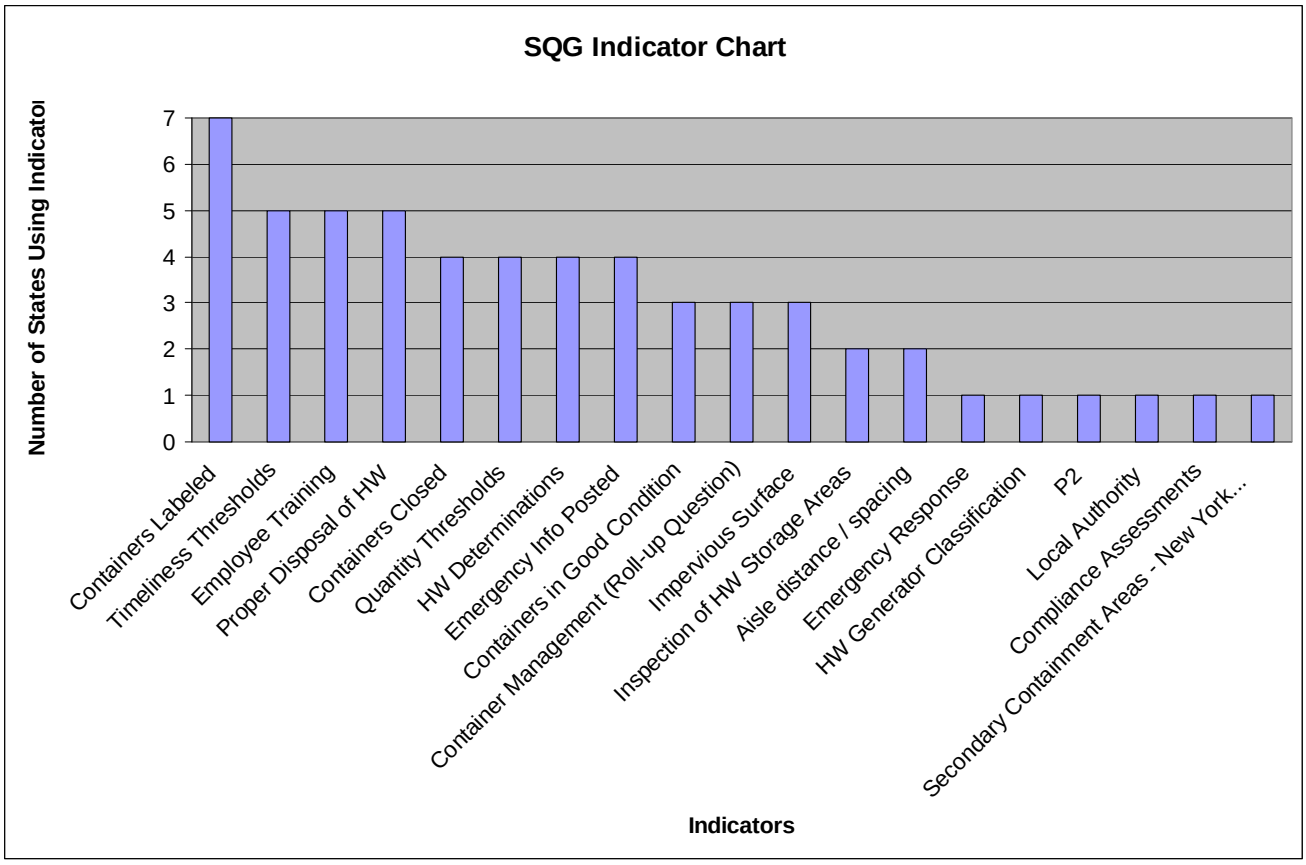
STATE		*GROUP*	
#	Indicator Description		
	Data Source		

Data Quality Indicator	**Good**	**Okay**	**Problem**	**Issue:**
Precision				
Sensitivity				
Representativeness				
Completeness				
Bias				
Validation				
Other issue:				

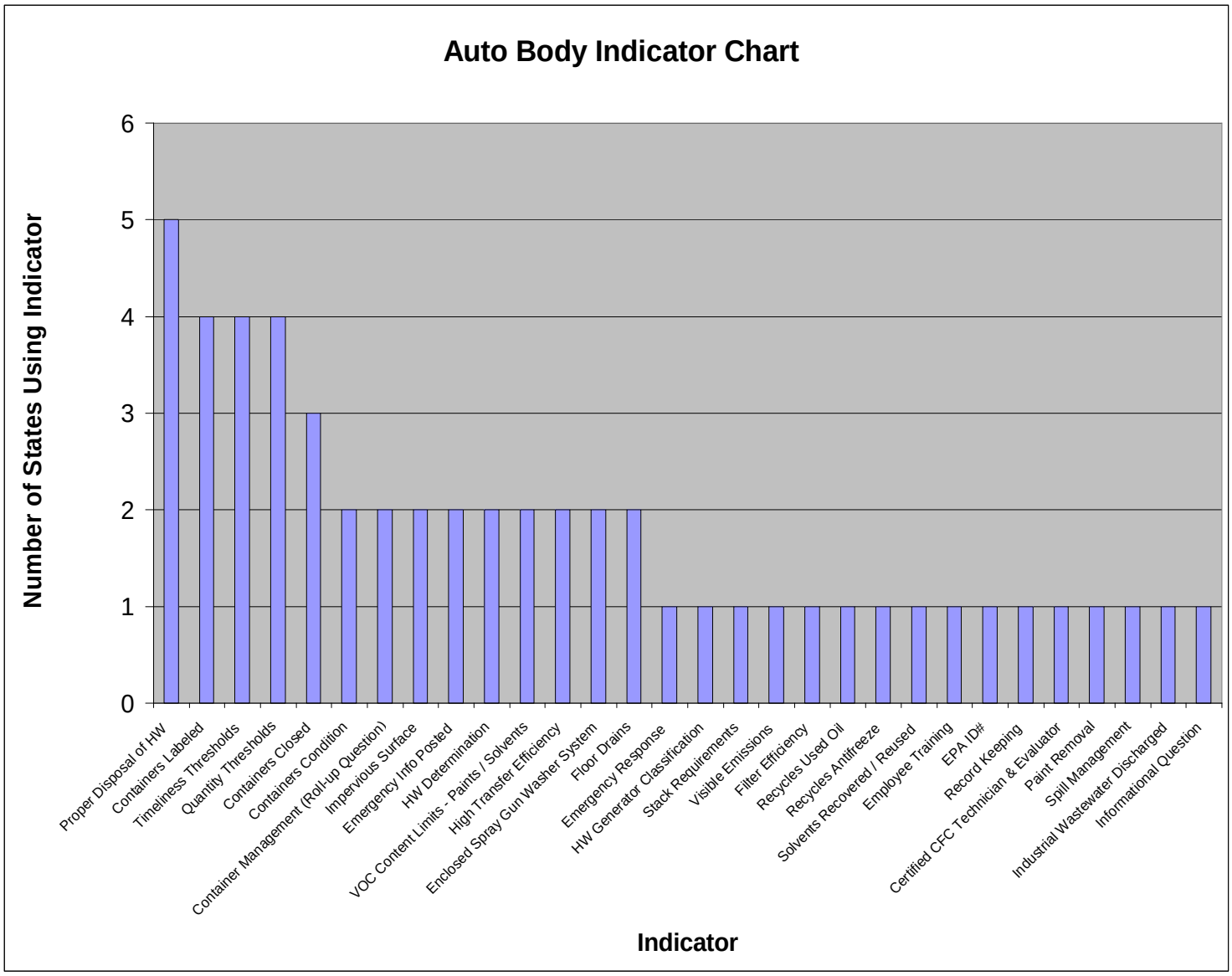
See Appendix E for aggregated data on indicator evaluation chart.

Once the states completed the Indicator Evaluation Chart for the small quantity generator of hazardous waste group and the auto body group, the data were aggregated and organized by indicator category.

Below is the SQG Indicator Chart from information provided by seven states [CO, CT, MA, ME, NH, NY, VT]:



Below is the Auto Body Indicator Chart from information provided by six states [CO, MA, ME, NH, NY, RI]:



2.2.4 Additional Data Gathering from States

The Project Management Team contacted each state individually to gather additional information to help prepare states for making group definition and data quality decisions. The following data was collected:

- Final commitment of group(s), i.e., which states were analyzing the auto body group and which were analyzing the SQG group.
- Individual state definitions for the groups.

- Identification of waste types within those groups, e.g. for SQGs, use only RCRA requirements or include used oil.
- An estimate of the state's universe size for each group.
- The data source that would be used, i.e., state would collect new data or use existing data or a combination.
- Timeframe for data collection.
- Other data quality issues (including those identified in the work assignment).
- Any major program changes in the last few years for those groups, e.g. regulatory changes.
- Whether there was agency support for measurement of the group.

See Appendix F for data quality considerations when selecting groups.

2.3 FINALIZING THE GROUP AND INDICATOR DEFINITIONS

On March 13, 2007, the States Common Measures Project held a workshop on finalizing group and indicator definitions for the small quantity generator sector and the auto body sector.

Note: Because of time constraints, states selected draft indicators for the SQG sector only. Work on the auto body sector was deferred to a later date. See Section 5 for Auto Body Sector.

2.3.1 Common Definition for Small Quantity Generators of Hazardous Waste

The project states evaluated similarities and differences in individual state definitions of SQGs. This evaluation included comparing generation rates, accumulation quantity limits and accumulation time limits across states. States decided to define SQGs for purposes of the project using the federal RCRA generation rate [*any generator that generates between 100 kg/mo and 1000 kg/mo]* and by the federal RCRA waste types [*excludes PCBs and waste oil generators]*.

Once the SQG universe definition was established, states began to develop a set of performance indicators. In cases where a project state had a different accumulation quantity limit and/or accumulation time limit, it was agreed that state would use its own accumulation standard.

Below is a chart that highlights differences in the project state's SQG accumulation quantity limit and accumulation time limit:

State	Accumulation Quantity Limit	Accumulation Time Limit	Project SQG Generation Rate
CO	6000 kg	180 or 270 (if >200 mi from TSDF	100 - 1000 kg / month
CT	1000 kg	180 days	100 - 1000 kg / month
MA	6000 kg	180 days	100 - 1000 kg / month
ME	3000 kg	90 days	100 - 1000 kg / month
NH	NA	90 days	100 - 1000 kg / month
NY	6000 kg	180 days	100 - 1000 kg / month
RI	NA	90 days	100 - 1000 kg / month
VT	6000 kg	180 days	100 - 1000 kg / month

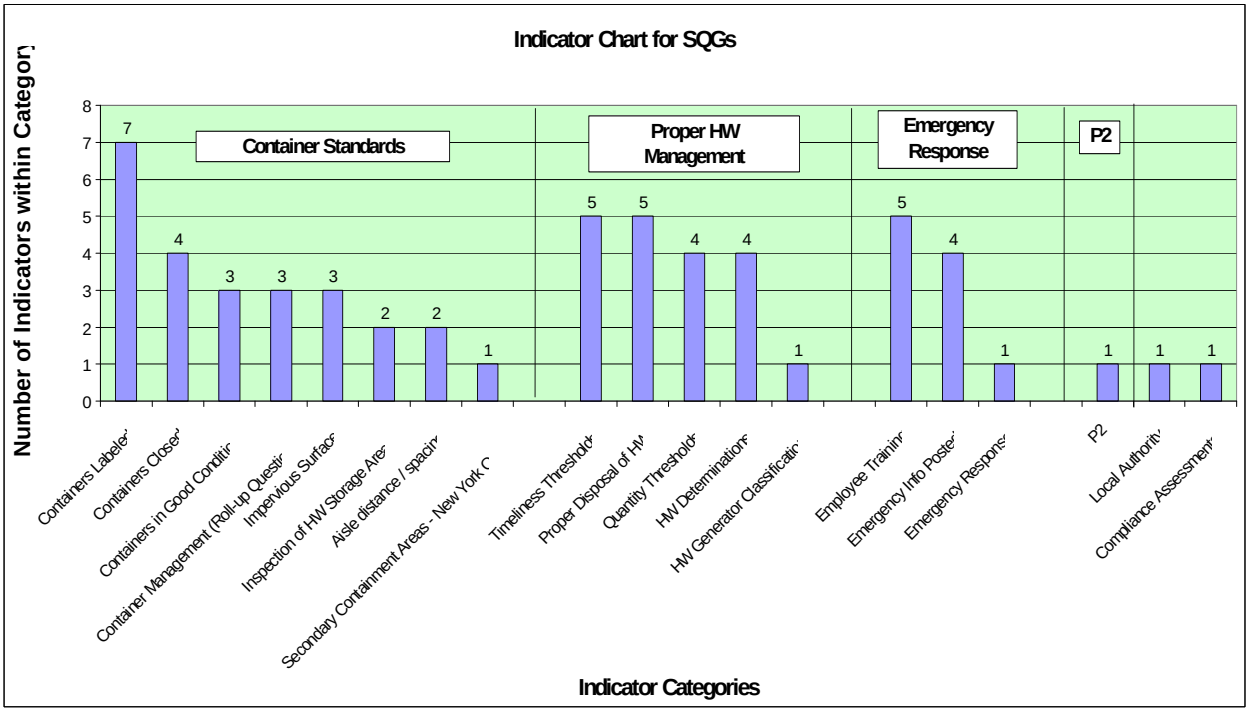
As noted in the chart below, four states [CO, CT, RI, VT] had SQG definitions that included only RCRA waste types. The other four states [MA, ME, NH, NY] also included PCBs and/or used oil in their definition of SQG. These four states agreed not to count PCBs and used oil toward SQG determinations to ensure a common sector definition for the project.

	CA	CO	CT	MA	ME	NH	NJ	NY	RI	VT
Waste Type		RCRA ONLY	RCRA Only	RCRA, PCBs greater than 50 ppm and waste oil (small number of PCB & waste oil, can be culled out)	RCRA & PCBs (PCBs can be culled out)	RCRA and PCB contaminated oil (used oil is recycled. PCBs less than a fraction of a percent and can be culled out)		RCRA and PCBs greater than 50 ppm (PCB SQGs are approximately 5-7% of universe and can be culled out)	RCRA Only (new rules for used oil - soon to be exempt)	RCRA, petroleum contaminated wastes, but not waste oil

2.3.2 Common SQG Performance Indicators

Results from the summer work assignment, (i.e., the indicator evaluation chart) demonstrated a preliminary interest in the following indicator categories for the SQG sector: container standards, proper management of hazardous waste, emergency procedures, and beyond compliance practices.

SESSION 1:
SELECTING SQG INDICATORS



The Project Management Team used this data to develop an exercise where states were asked to compare indicator language submitted by each state to identify common characteristics, potential gaps and quality issues that needed to be addressed. Once the language was reviewed, states were asked to either "accept" or "reject" the indicator category. *See Exhibit 2.1 which shows each state's SQG indicator language related to the container management category.*

Exhibit 2.1: States Common Measures Project: States' SQG Indicators

CATEGORY	Indicator	CA	CO	CT	MA	ME	NH	NJ	NY	RI	VT	Total Number of States Using Indicator	Select Indicator Category (Y / N / M)
CONTAINER MANAGEMENT	**Containers Labeled**		All hazardous waste **containers** are properly **labeled**	Are all **containers** of hazardous waste **labeled** with the words "Hazardous Waste" and a description of their contents.	All hazardous waste **containers** are properly **labeled**	All hazardous waste **containers** are properly **labeled**	All **containers** of hazardous waste are **labeled** with the words, "Hazardous Wastes."		All hazardous waste **containers** are properly **labeled** (for both accumulation and storage areas).		All hazardous waste **containers** are properly **labeled**	**7**	
	Containers Closed		All hazardous waste **containers** are kept **closed** except when being filled.			All hazardous waste **containers** are **closed**			All hazardous waste **containers** are **closed** (both accumulation and storage areas).		All hazardous waste is stored in **closed containers**	**4**	
	Containers in Good Condition		All hazardous waste **containers** are in **good condition**		All hazardous waste **containers** are in **good condition**				All hazardous waste **containers** are in **good condition** (for both accumulation and storage areas).			**3**	
	Container Management (Roll-up Question)			Is the hazardous waste being stored in **containers** that are **closed** and **free of significant damage**			Hazardous wastes are stored in **containers** that are **closed** and **free of damage or deterioration**.				All hazardous waste is **stored under cover and protected from freezing**	**3**	
	Impervious Surface		The facility operates in a way that **minimizes** the potential for **releases** of **hazardous wastes.** .i.e. HW containers are stored on a **crack free surface** that will contain leaks or spills and there is no evidence of significant releases outside the building.		All hazardous waste **containers** are stored on a **crack free surface** that will contain leaks or spills						All hazardous waste is **stored** on an **impervious surface**	**3**	
	Inspection of HW Storage Areas		All hazardous waste **containers** are **inspected weekly** checking for leaks or deterioration.	Are **inspections** of hazardous waste **storage areas** being conducted and documented?								**2**	
	Aisle distance / spacing			Is there at least three feet of **aisle space** between containers			There is a **distance of at least 2 feet of access (aisles)** space on at least one side of each HW container.					**2**	
	Secondary Containment Areas - New York Only								For Facilities located over sole source aquifers, **containers storage is within secondary containment areas** if more than 185 gallons are accumulated for storage. – For New York Only			**1**	

See Appendix G for a complete list of states' SQG indicators by indicator category.

For each indicator category (e.g., container management, proper hazardous waste management, emergency response, pollution prevention (P2), draft language was selected or developed and then tested against the six factors that influence data quality: precision, sensitivity, representativeness, completeness, bias, and validation. Each state also identified whether the state planned to use existing or new data for the indicator. Below is the container management section of the worksheet that project states used to develop common indicator language.

	SQG Indicator Worksheet								
CATEGORY	Indicator	Precision	Sensitivity	Representativeness	Completeness	Bias	Validation	FINAL DRAFT INDICATOR	Data Source
CONTAINER MANAGEMENT	Containers Labeled	Precision Yes ___	Sensitivity Yes ___	Representativeness Yes ___	Completeness Yes ___	Bias No ___	Validation Yes ___		New _____ Old _____
	Containers Closed	Precision Yes ___	Sensitivity Yes ___	Representativeness Yes ___	Completeness Yes ___	Bias No ___	Validation Yes ___		New _____ Old _____
	Containers in Good Condition	Precision Yes ___	Sensitivity Yes ___	Representativeness Yes ___	Completeness Yes ___	Bias No ___	Validation Yes ___		New _____ Old _____
	Container Management (Roll-up Question)	Precision Yes ___	Sensitivity Yes ___	Representativeness Yes ___	Completeness Yes ___	Bias No ___	Validation Yes ___		New _____ Old _____
	Impervious Surface	Precision Yes ___	Sensitivity Yes ___	Representativeness Yes ___	Completeness Yes ___	Bias No ___	Validation Yes ___		New _____ Old _____

At the end of the workshop, participating states agreed on draft SQG indicator language.

2.3.3 Indicator Sign-Off Process

Shortly after the workshop, a SQG Final Draft Indicator Acceptance Worksheet was sent to all project states to:

- Review final indicator language,
- Identify the data source for those indicators,

- Share the information with appropriate staff,
- Sign-off on the final list of indicators and compliance verification strategies.

Below is the container management section of the SQG Final Draft Indicator Acceptance Worksheet:

SQG FINAL DRAFT INDICATOR ACCEPTANCE WORKSHEET						
Category	**Sub Indicator Category**	**Final Draft Indicator**	**Default Federal Requirements for SQGs**	**Compliance Verification**	**Accept Indicator Wording & Accept Compliance Verification**	**Data Source**
CONTAINER MANAGEMENT	Containers Labeled	Are all hazardous waste containers properly labeled with the words "hazardous waste" and clearly marked with the date for which accumulation began?	§ 262.34 (2) The date upon which each period of accumulation begins is clearly marked and visible for inspection on each container; (3) While being accumulated on-site, each container and tank is labeled or marked clearly with the words, "Hazardous Waste"	Field observer will determine: • If all containers have labels • If labels are marked with both items • If labels are clear and legible	Accept Indicator Yes ( ) No ( ) Accept Compliance Verification Yes ( ) No ( )	New ( ) Existing ( ) Combo ( )
	Containers Closed	Are all hazardous waste containers closed unless waste was being added or removed?	§ 264.173 (a) A container holding hazardous waste must always be closed during storage, except when it is necessary to add or remove waste.	• Field observer will confirm that all containers are closed at the time of inspection unless waste was being added or removed	Accept Indicator Yes ( ) No ( ) Accept Compliance Verification Yes ( ) No ( )	New ( ) Existing ( ) Combo ( )
	Containers in Good Condition	Are all hazardous waste containers in good condition, (i.e., free of severe rusting or apparent structural defects, and not leaking)?	§ 264.171 If a container holding hazardous waste is not in good condition (e.g., severe rusting, apparent structural defects) or if it begins to leak, the owner or operator must transfer the hazardous waste from this container to a container that is in good condition or manage the waste in some other way that complies with the requirements of this part.	• Field observer will perform visual inspection of conditions of all containers looking for leaks and/or severe corrosion, bulging, rusting or dents	Accept Indicator Yes ( ) No ( ) Accept Compliance Verification Yes ( ) No ( )	New ( ) Existing ( ) Combo ( )

By completing and returning this worksheet, states "signed-off" on the final indicators and the compliance verification method for each indicator.

The final list of SQG indicators and compliance verification methods can be found at Exhibit 3.1 in Section 3.3.

2.4 FINALIZING THE DATA COLLECTION METHODS

Initially, some states planned to use new field observation data and others planned to use existing inspection data. In order to ensure data quality standards were met during data collection, the Project Management Team required each to state sign a certification of data quality that the following conditions were met:

- Data collected under this project was representative of the population as a whole.
- The data had the precision expressed in the indicator definition and compliance verification procedures.
- The data had the sensitivity expressed in the indicator definition and compliance verification procedures.
- The data was free of bias due to field observer interpretation.
- The data was complete.

See Section 3.3 for the Certification of Data Quality template.

The Project Management Team established standards to ensure a high level of data quality for new and existing data.

2.4.1 Standards for Collecting "New" and Existing Data

New data had to meet the following standards:

- Facilities inspected had to be selected randomly from the complete population of facilities in the study universe, using the agreed upon methodology.
- The field observers had to fully understand what was meant by each indicator, how "pass" or "fail" on each indicator was determined, and how to collect the data.
- The field observer had to complete the "checklists" accurately.
- Each checklist had to be filled out in its entirety.

2.4.2 Standards for "Existing" Data

Existing data had to meet the following standards:

- Facilities included were to be selected randomly, using the agreed upon methodology from the complete population of facilities in the study universe.
    - By extension this meant that if a state was using only existing data rather than a combination of existing or new data, every facility in the universe would have to have been inspected within a relatively short time period - at a minimum they would have to have been inspected since the last significant changes, if any, in the state's regulatory requirements and/or oversight procedures.
    - Similarly if a state was using a combination of existing and new data, the existing inspections would have to have taken place since the last significant change, if any, in the state's regulatory requirements or oversight procedures.

- The state had to have information on each indicator for each facility and had to fill out a complete checklist for each facility.

- The state had to be certain that the definition used by the field observer to "pass" or "fail" an indicator was the same standard used to pass or fail on indicators in the project; or that there was sufficient information on file to fill out the checklist accurately.

- The state had to be certain that there was no bias in the interpretation of "pass" or "fail."

Because of the relatively small number of statistically valid inspections needed for this project, all participating states decided that they were able to collect new data for the small quantity generators of hazardous waste sector. The data collection period was from Federal Fiscal Year 2006 to Federal Fiscal Year 2008.

2.5 FINALIZING THE ANALYTICAL AND REPORTING PROCEDURES

Once the project states selected final indicators and identified their data source for those indicators, they were presented with options for analyzing and reporting the results of the SQG sector. Below are basic concepts and measurement approaches that states agreed upon:

2.5.1 Data Elements to be Used

- Applicability Data: The questions field observers answered to determine if the facility is in or out of the universe.

- Performance Indicator Data: The individual facility behaviors the project states decided to measure -- the checklist questions the field observers answered to ascertain

performance. *See Exhibit 3.1 in Section 3.3 for the list of SQG indicators. See Exhibit 5.1 in Section 5.4 for the list of Auto Body indicators. They include both regulatory requirements and voluntary beyond compliance practices.*

2.5.2 What the Project Will Measure

- State SQG Mean Facility Score on 1) All Indicators, 2) Regulatory Indicators and 3) Beyond Compliance Indicators: A facility score is the proportion of the performance indicators to which a particular facility is subject, which the facility was observed to achieve (i.e., was either in compliance with the regulatory requirement or was implementing the beyond compliance practice). This is expressed as a number from 1 to 10: A score of "1" means that the facility was achieving 10% of the applicable indicators, a score of "5" signifies that the facility achieved 50% of the applicable indicators, and a score of 10 means that the facility achieved 100% of the applicable indicators. The state mean facility score is the average score achieved by all facilities in the sample. Scores can be calculated for all or a subset of indicators.

- State SQG Achievement Rate By Indicator: The percentage of the facilities in the state that were achieving each performance indicator. This was calculated for each indicator for each state.

- State SQG Facility Score Distribution: The percentage of facilities in the state that achieved each facility score (e.g., the percentage of facilities that achieved facility score of 1, the percentage of facilities that scored 2, the percentage of facilities that scored 3, etc.).

- Interstate Comparisons: Statistically significant differences in state SQG performance on achievement rates and mean facility scores at varying confidence levels.

2.6 OBSERVATIONS AND LESSONS LEARNED FROM PHASE 2

- √ The degree to which there were differences in state definitions of the selected groups was surprising. Summer work helped focus discussion, identified, and resolved definitional differences.
- √ It was helpful to not pick indicators before statistical training. This allowed for context on how data will be used. This also helped in the design of the indicators.
- √ A state participant suggested it would be helpful to have a more in depth training module available to help explain important measurement concepts to project participants as well as other decision makers within their state.
- √ It is beneficial to have a common approach to universe identification when doing multi-state comparisons.
- √ States that initially planned to use existing data on groups decided it was more efficient and effective to get new data for the project than try to meet the data standards needed to use existing data.

SECTION 3: DATA COLLECTION, FIELD OBSERVER TRAINING AND STATISTICAL METHODS (PHASE 3)

This Section of the Report Covers:

The third phase of the project involved applying statistical and data quality assurance procedures to sample selection and data collection. There were three general data collection objectives for Small Quantity Generators of Hazardous Waste (SQGs):

- Collect data on enough SQGs to be able to say that their performance reflects the performance of all of the SQGs in the state with enough precision and reliability for the results to be useful to state decision makers.
- Pick facilities randomly so as to not introduce bias into the results.
- Collect the data in the same way with the same interpretations so the results are comparable from state to state.

3.1 DETERMINING A REASONABLE SAMPLE SIZE FOR DRAWING STATISTICAL CONCLUSIONS ABOUT THE SQG GROUP

The States Common Measures Project had two primary analytical goals: 1) benchmarking SQG performance in each individual state and 2) comparing SQG performance across states to identify performance differences that do not appear to be due to chance. Since states could not inspect every SQG, statistics had to be used to draw conclusions from inspections of a sample of each state's facilities. The sample had to be large enough to provide enough certainty that the observed results reflected actual conditions, precise enough to provide the states with useful information, but small enough to be done with available resources.

3.1.1 Sample Size Needed to Benchmark an Individual State's Group Performance

The actual number of inspections needed to draw conclusions about a universe from a smaller "sample" depends on the following four factors:

- **Universe Size:** The required sample size INCREASES with the increase in the total number of SQGs in the state.
- **Confidence Level:** The required sample size INCREASES with the increase in the desired level of certainty that the population selected IS reflective of the population as a whole -- the "confidence level." Fewer inspections would be required if states felt they would be able to make decisions based on results that had a one in ten chance (a 90% confidence level) that the population selected was not representative than if they could only tolerate a one in twenty chance (a 95% confidence level) that the population selected is not representative of the whole.
- **Confidence Interval:** The required sample size INCREASES with the increase in the required precision of the results. When drawing conclusions about a population from a smaller sample, the actual performance must be expressed as a range around the "observed" value for the sample. This range is called the "confidence interval." For example, if the inspectors "observed" that 70% of SQGs were in compliance with labeling requirements, and the confidence interval was 10%, the true compliance rate for the entire population of SQGs would be somewhere between 60% and 80%.

  The number of inspections needed declines with declines in the minimum level of precision that is required. Fewer inspections would be needed if, for example the states felt that they could base decisions on a confidence interval of 20% than if they felt they needed a confidence interval of 5%. A higher percent is less precise.

  It is important to note that confidence intervals and confidence levels are also related. For a given sample size, the higher the confidence level, the larger the confidence interval. For example, one can be 99% certain that one has properly estimated a person's age if one guesses that they are somewhere between 1 and 100 years old. One might be only 90% certain (have a one out of ten chance of being wrong) if one guessed that their age was between 20 and 50.
- **Observed performance**: The required sample size DECREASES the closer the actual performance is to either end of the scale. This happens because a score cannot be greater than 100% or less than 0% -- there is less total room for variation in the result at either end of the scale. Therefore:
  - A 50% compliance rate requires the largest sample size.
  - 70% or 30% compliance rates require a smaller sample size.
  - 1% or 99% compliance rates require the smallest sample size.

The "sample-size calculator"[5] developed for the Massachusetts ERP Program was used to calculate the sample sizes that would be required to benchmark each state's performance

[5] The "sample-size calculator" is an excel based tool that may be obtained by contacting Susan.Peck@state.ma.us

at various confidence levels, confidence intervals and assumptions about the observed compliance rates. The chart below shows the results of the analysis.

Sample size (# of facilities) needed to benchmark an individual state's performance								
		Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7
***Confidence Level**: % certainty that the result reflects the true population and is not due to chance)*		***90%***	***90%***	***90%***	***90%***	***95%***	***95%***	***95%***
***Confidence Interval/Margin of Error:** the actual percentage of facilities in compliance falls somewhere within + or - the listed percentage points of the observed percent compliance*		***+ / - 10%***	***+ / - 5%***	***+ / - 5%***	***+ / - 5%***	***+ / - 10%***	***+ / - 5%***	***+ / - 5%***
estimated compliance rate of the universe		***50%***	***50%***	***80%***	***90%***	***50%***	***50%***	***80%***
State	***SQG Universe Size***							
NH	*190*	34	88	68	45	50	112	91
VT	*500*	38	124	87	53	60	176	129
CO	*800*	39	136	93	55	62	202	142
RI & CO	*1000*	39	141	95	56	63	213	148
CT	*1640*	40	149	99	57	65	232	157
RI	*2000*	40	152	100	57	65	238	159
MA	*2704*	41	160	101	58	67	246	163
NY	*10500*	41	162	104	59	67	264	170

The highlighted columns show the level of certainty (confidence level) and precision (confidence interval) that are possible for what project states felt was a reasonable number of inspections per state (between 34 – 67):

- The lowest range of inspections per state (column 1) – 34 for New Hampshire, the state with the smallest universe to 41 inspections for New York, the state with the largest) assumed the observed compliance rate would be 50%, used a 90% confidence level and a confidence interval for the measured value of plus or minus 10%.
- 45 inspections in New Hampshire and 59 inspections in New York would result in a confidence level of 90% and a confidence interval of + /- 5%, if the observed compliance rate 90% (as shown in column 4). While 50 and 67 inspections would result in a 95% confidence level and a confidence interval of + / – 10% if the observed compliance rate is 50% (as shown in column 5).
- Notice that all states would need sample sizes of over 90 and as many as 264, to obtain a confidence interval of + / – 5% AND a 95% confidence level (as shown in column 7).

- As shown in column 2, 3, 6 and 7, achieving a confidence interval of + / – 5% would require all states to do a minimum of about 100 inspections unless the observed compliance rate was 90% (as shown in column 4). Even if the confidence level dropped to 85 % (not shown) states would have to complete between 69 (NH) – 106 (NY) inspections assuming an observed compliance rate of 50%.

Decision about sample size for benchmarking performance

The project states initially agreed to a minimum sample size of between 34 – 41 inspections for each state, depending on the state's universe size. Inspecting this number would allow the project to meet the minimum project goal of benchmarking the performance of each state with a minimum level of precision (+ or – 10%) and with a reasonable level of confidence (90%) and assuming 50% observed compliance rates on each indicator. If performance levels were greater than 50%, the precision of the estimate would increase. As demonstrated in the chart above, states with smaller universes, such as NH would need to complete a smaller number of inspections than states with a larger universe such as NY.

Note: As discussed below, additional inspections are required to identify statistical differences in SQG performance between any two states.

3.1.2 Sample Sizes Needed to Compare Performance Levels Between States

In addition to benchmarking an individual state's performance, the States Common Measures Project also compared SQG performance results between states.

The issues that affect sample size are different when comparing performance levels between states. They are as follows:

- The **confidence level** (as described above – the likelihood that the observed difference actually exists).
- The **observed performance rates** of the two states (as described above).
- The **power –** this is a new concept, it is the likelihood that the results do not *miss* a statistically significant difference that is in fact there.
- The **magnitude of the statistically significant differences** that can be detected.

Note: Unlike with benchmarking an individual state's performance, the number of inspections needed for comparing performance across states does not depend on sample size. See Phase 4 for SQG performance results.

The Massachusetts ERP sample-size calculator was used to calculate sample sizes needed for various assumptions about confidence level, compliance rates, power and the magnitude of the differences that the project states wanted to detect. The results are shown in the chart below.

Sample size needed to detect whether a given-sized difference in performance level between two states is statistically significant								
		Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8
Confidence Level: *% certainty that a difference of the size listed below is not due to chance*		***90%***	***90%***	***90%***	***90%***	***95%***	***95%***	***95%***	***95%***
Power: % *certainty that a smaller difference than the given difference IS due to chance (in other words that you are not missing a true difference)*		***80%***	***80%***	***90%***	***90%***	***80%***	***80%***	***90%***	***90%***
estimated compliance rate of the universe	***State A***	***50%***	***90%***	***50%***	***90%***	***90%***	***70%***	***50%***	***70%***
	State B	***50%***	***70%***	***50%***	***70%***	***70%***	***70%***	***50%***	***70%***
# of facilities needed to determine that a finding that State A's performance is **20 percentage points higher** than State B's performance is not due to chance		56	34	82	49	46	65	107	95
# of facilities needed to determine that a finding that State A's performance is **15 percentage points higher** than State B's performance is not due to chance		100	60	146	88	82	115	190	137
# of facilities needed to determine that a finding that State A's performance is **10 percentage points higher** than State B's performance is not due to chance		225	135	328	197	185	260	428	309

The highlighted columns show the level of certainty (confidence level) and precision (confidence interval) that are possible given the number of inspections the project states said they could realistically afford to complete (between 34 – 67). The table shows that:

- If the observed compliance rates are in the 50% range, 56 inspections would be needed to detect a 20% point difference, at a 90% confidence level.
- Unless states inspected more than 80 facilities each, the highest difference that states could realistically hope to detect is 15%, and that would only be achieved if both states had relatively high observed compliance rates (as shown in column 2 with 90% and 70% estimated compliance rates).

Decision on sample size for comparing performance across states

More inspections were required to compare performance levels across states than to benchmark performance within a state. The project states decided that by inspecting at least 56 SQGs per state, the results would provide sufficient precision at a 90% confidence level. Using a higher confidence level would allow the project to detect smaller differences in performance but would require more inspections and resources. If, after conducting inspections, observed performance levels were greater than 50%, it would be possible to detect whether smaller differences in state performance levels were statistically significant.

3.2 UNIVERSE IDENTIFICATION AND RANDOM SAMPLING

If the facilities are randomly selected, statistical methods can be used to generalize the field observation findings to the sector as a whole and to determine if the observed differences among states are likely to be due to chance, or reflect actual differences in performance. Random samples are the foundation of all statistical analyses. Without them, bias can be introduced into the results. For example, if a state chose to target facilities that they suspected had problems (something field observers frequently do in order to maximize the likelihood of identifying and resolving environmental problems), the results from that state would likely show worse performance than the results from a state that selected facilities randomly. Selecting facilities randomly ensures a comparable analysis of results across states.

An SQG Universe Identification and Random Sampling Record was developed and distributed to states to document each state's methodology for universe identification and random sample selection. Rather than the Project Management Team establishing a single method for universe identification, each state used and documented its own identification approach. The Quality Assurance Officer determined that the approaches used by the project states were consistent and appropriate.

The Project Management Team also required that all states use the same methodology for generating a random sample of facilities to inspect. In cases where a facility turned out to be "not applicable," e.g., the facility closed or generated more or less than the 100-1000 kg of RCRA hazardous waste per month, all states agreed to replace this facility by inspecting the next facility on their randomized list.

The Quality Assurance Officer for the project provided step-by-step instructions to the states for generating a random list of facilities from their universe. *See Appendix H for How to Generate a Random Sample.*

SQG UNIVERSE IDENTIFICATION AND RANDOM SAMPLING SELECTION RECORD		
STATE		
	METHODOLOGY	**DATE PERFORMED**
UNIVERSE IDENTIFICATION		
RANDOM SAMPLING SELECTION	Pick One: ***State Common Measures Project: Methodology for Generating A Random Sample, June 8, 2007*** Alternative Method (describe):	
PROJECT LEAD		

3.3 STEPS TO ENSURE FIELD DATA QUALITY

The Project Management Team had to ensure that the project states were measuring the same activities in the same way so that the results were truly comparable from state to state. This was a challenge as there were multiple states with different regulations and different approaches to performing inspections. There also were different environmental concerns and different levels of detail regarding compliance with the different state standards. For example, a state that interpreted "good condition" for a hazardous waste container as absolutely no chipped paint, ding, or rust of any kind, would observe lower performance than a state that interpreted "good condition" as containers that were not actively leaking at the time of the inspection. Similarly, record review procedures could affect observed compliance rates. For example, states that reviewed the past 3 years of records would likely find more violations than states that reviewed the past 6 months of records. The Project Management Team had to minimize these differences as much as possible before any data collection could occur.

In order to ensure consistency with each state's approach to data collection the Project Management Team developed a Common Measures Checklist using the final SQG indicators that were selected by the project states. States had the option to use either the Common Measure Checklist or use their own checklist so long as it included the Common Measures Checklist questions exactly as they were phrased. Alternative checklists had to be approved by the Quality Assurance Officer before use.

In order to further ensure consistency with each state's approach to data collection, the Project Management Team presented mandatory data collection phone training on May 23, 2007. The training was provided by state field and enforcement experts with over twenty-five years of hazardous waste experience, the Project Manager and the Quality Assurance Officer. All project state leads, and as many field data collectors as possible were required to attend the training. The state lead was required to certify, when submitting SQG data, that all field observers who collected the data either participated in the training or were trained by the state lead. Attendance records were also kept to document that all field observers received the data collection training. *See Appendix I for Training Attendance Log Template.*

The training included:

- Careful review and discussion of each indicator, specifically reviewing the wording and intent.
- Refining language as necessary.
- Agreeing upon procedures for interpretation.
- Reviewing decision rules for determining whether or not the facility was achieving the indicator.

Minor adjustments were made to the checklist as a result of the training. *See Exhibit 3.1 below for the States Common Measures Project SQG Performance Checklist.*

Exhibit 3.1: States Common Measures Project SQG Performance Checklist

Date of Visit: _____Agency/Regional Office: _____

Field
Observer:_____Tel._____

Facility
Name:_____

Facility
Address:_____

Name of Contact Person:
_____

Telephone number of Contact Person:
_____

Generator ID Number: _____ SIC:_____

APPLICABILITY

1	**Does this facility generate more than 1000 KG per month of RCRA hazardous waste?**	**Y** **N**
	IF YES, STOP Submit form as is	
	Inspector will inventory all hazardous waste streams and question facility on maximum monthly generation rate for each and compile the total. If the facility generated more than 1000 Kg in any one month, the answer is YES. If needed the inspector will review one year of manifests to determine the quantities of hazardous waste shipped off-site and at what frequencies.	
	NOTE: When calculating thresholds do not consider "universal" wastes such as fluorescent lights or computers. Note also that the threshold is for amount generated in a month, which may or may not be the same as the amount of waste shipped in a month.	
1a	**Does the facility generate more than the Federal small quantity generator threshold for an acutely hazardous waste?**	Y **N**
	IF YES STOP Submit the form as is	
2	**Does this facility generate less than 100 KG per month of RCRA hazardous waste?**	**Y** **N**
	See answer to 1 above.	
	Note that in order to answer YES, the facility must be below the 100 KG for all 12 months. The answer is NO if the facility exceeded the 100 KG threshold in any one month.	
	IF YES, STOP **Submit form as is**	
	IF NO, CONTINUE, COMPLETE THE PERFORMANCE CHECKLIST	

Exhibit 3.1 Continued

ANSWER ALL QUESTIONS UNLESS SPECIFICALLY DIRECTED TO SKIP A QUESTION

If there are any doubts about verification of conformance with an indicator, inspector should discuss it with state project lead.

NOTE: ALL QUESTIONS APPLY TO RCRA HAZARDOUS WASTE ONLY. THEY DO NOTAPPLY TO "UNIVERSAL WASTES" SUCH AS COMPUTERS OR FLOURESCENT BULBS OR TO STATE- ONLY WASTES

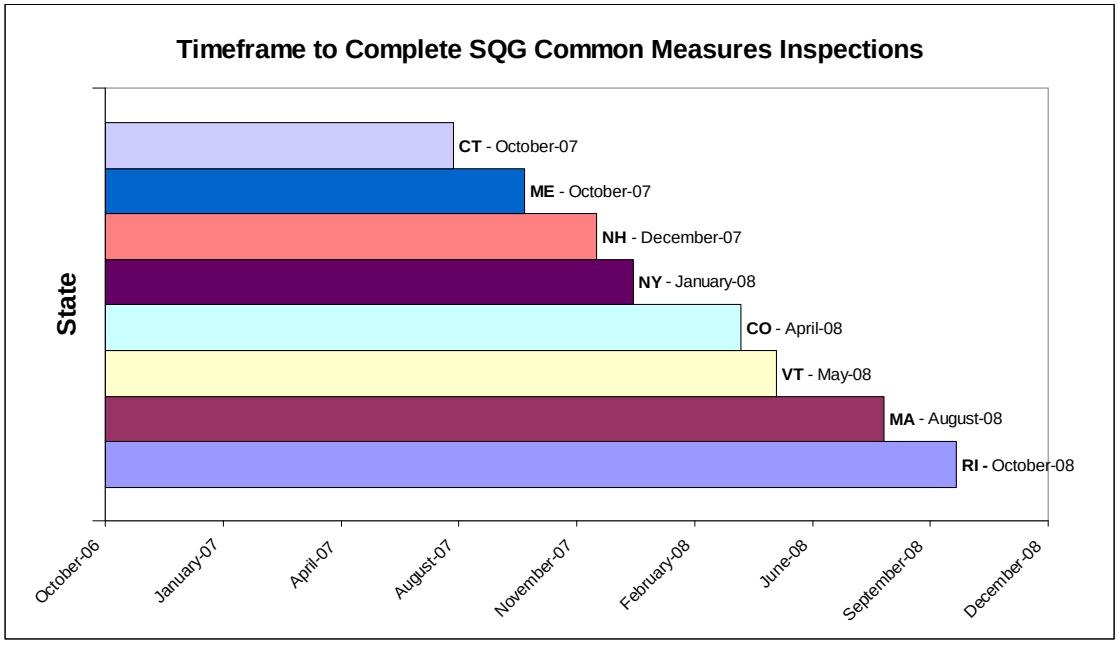
	Container Management Indicators *Note: Container management questions apply to central accumulation areas only. They do not apply to satellite areas or laboratories*		
3	**Are all hazardous waste containers properly labeled with the words "hazardous waste" and clearly marked with the date on which accumulation began?** Inspector will determine if facility has made a "complete and genuine effort" to ensure that all containers have labels, are marked with both items and are clear and legible *Note: inspector will use best professional judgment to determine what is "clear" and "legible" and whether the facility has made a "complete and genuine effort*"	**Y**	**N**
4	**Are all hazardous waste containers closed unless waste is being added or removed?** Inspector will confirm that all containers are closed at the time of inspection unless waste is being added or removed. **Note: "closed" means if the containers were tipped, nothing would spill. "Funnels" are acceptable as long as they are closed.**	**Y**	**N**
5	**Are all hazardous waste containers in good condition, (i.e., free of severe rusting or apparent structural defects, and not leaking** Inspector will perform visual inspection of conditions of all containers looking for leaks and/or severe corrosion, bulging, rusting or dents. *Note: inspector will use best professional judgment to determine what is "severe." There should be no imminent threat.*	**Y**	**N**
Proper Hazardous Waste Management Indicators			
6	**At the time of the inspection has the facility accumulated more than _____* kg of RCRA hazardous waste onsite?** **NOTE: Fill in the state's accumulation limit applicable to facilities that generate between 100 and 1000 kg of RCRA Hazardous Waste per month* Inspector will inventory all containers and tanks accumulating hazardous waste, noting their volumes and contents. A review of manifests may provide information on the weights of different hazardous waste streams so the total weight accumulated can be calculated.	**Y**	**N**

7	**At the time of the inspection, does the facility have any RCRA hazardous waste onsite that has been accumulated onsite for more than _____*days?** **No labels** **NOTE: Fill in the state's accumulation time limit applicable to facilities that generate between 100 and 1000 kg of RCRA Hazardous Waste per month.* Inspector will verify conformance based on dates on containers which detail when accumulation begins (if no containers are labeled, then inspector should circle **no labels**). *Note: inspector must use common sense to verify compliance. For example, if 49 out of 50 drums at a facility are dated to indicate compliance with the timeliness threshold and all contain the same wastes, chances are the facility is in compliance with the timeliness threshold for storage.*	**Y**	**N**
8	**Does the facility use a hazardous waste manifest to ship its hazardous waste when a manifest is required?** Inspector will look at one (1) year of manifest records as well as ask the facility if they have kept three (3) years of records. The inspector should look for gaps in the shipments or shipping records *Note: the word "look" in this case means that the inspector will confirm that all shipments have been manifested and not that each and every manifest was filled out correctly*.	**Y**	**N**
9	**Has the facility identified all of its hazardous waste streams?** Conformance will be determined based on: review of production processes, type of wastes generated at these processes and whether or not they have been characterized as hazardous waste. *Note: this is something inspectors do routinely. Interns may need to bring this information back to the state project lead.*	**Y**	**N**
Emergency Response Indicator			
10	**Has the facility posted the current name and telephone number of the emergency coordinator, the location of fire extinguishers and spill control material, and if present, fire alarm, and the telephone number of the fire department, unless the facility has a direct alarm?** Visual inspection of all elements listed and inquiry into whether the information is up to date. *Note: emergency information only has to be posted by 1 phone to be in conformance with this indicator for the purposes of this project*.	**Y**	**N**

Exhibit 3.1 Continued

Pollution Prevention Indicators			
11	**Has the facility taken one or more actions to reduce toxics, conserve water or energy over the past three years?** Inspector will ask the facility manager **IF NO, STOP. Performance Checklist is Complete.**	**Y**	**N**
11a	**IF YES:** **Has the facility implemented toxic use reduction over the past 3 years?** **IF NO, SKIP** to Question 11b Toxics Use Reduction includes any of the following types of changes to the production process: o substitution/replacement of a toxic raw material with a non-toxic or less toxic substance o substitution/reformulation of an existing end-product for one that is non-toxic or less toxic upon use, release or disposal o redesign, modification or modernization of production equipment (including integral or closed loop recycling or filtration) to reduce the amount of raw toxic material needed in the production process o improved operation and maintenance of the production process or equipment (e.g. housekeeping practices, system adjustments, product and process inspections), so that less raw toxic material is required in the production process	**Y**	**N**
11a1	***IF YES, BRIEFLY DESCRIBE the toxics use reduction projects***		
11b	**Has the facility undertaken recycling projects over the past three years?** **IF NO, SKIP** to Question 11c	**Y**	**N**
11b1	**IF YES, BRIEFLY describe the recycling projects:**		
11c	**Has the facility implemented water conservation projects over the past three years?** **IF NO SKIP TO Question 11d**	**Y**	**N**
11c1	***IF YES, briefly describe the water conservation projects:***		
11d	**Has the facility implemented energy conservation/alternative energy projects over the past three years?** **IF NO, STOP CHECKLIST IS COMPLETE**	**Y**	**N**
11d1	***IF YES, BRIEFLY describe the energy conservation/alternative energy projects***		

States collected data over a period of several months. The chart below shows the timeframe in which states completed inspections:



Each state project lead was responsible for the completeness of their state's data. Copies of completed checklists were sent to the Quality Assurance Officer along with the Certification of Data Quality signed by the state project lead. The certification statement is provided below:

COMMON MEASURES PROJECT

Data Quality Certification for Small Quantity Generators of Hazardous Waste

I, _____, Project Lead from the state of _____,
Print name State Name

certify that the enclosed field observation checklists meet the data quality standards described in the Common Measures Data Quality Training Workshop on September 28, 2006 and the Common Measures Inspector Training Conference Call on May 23, 2007.

_____ _____
State Lead's Signature Date

3.4 RECIPT AND ANALYSIS OF COLLECTED FIELD DATA

All inspection data was collected by the states and submitted to MassDEP for analysis.

3.4.1 The ERP Performance Analyzer

To facilitate data analysis, States Common Measures Project State Innovation Grant funds and funds from the Massachusetts Department of Environmental Protection were used to contract with TetraTech, EM Inc., an environmental consulting group, to enhance an existing custom MS-Access-based application, originally designed by TetraTech, known as the ERP Performance Analyzer.[6] The ERP Performance Analyzer is used to perform statistical analyses and create graphical presentations of data from ERP compliance inspections and/or self certifications. The application works with Excel, Power Point, and JMP Statistical Discovery Software. The output from the system is used to describe the environmental performance of different business sectors at a point in time, changes in performance over time and differences in performance across regulatory jurisdictions. The enhanced ERP Performance Analyzer was used to conduct the analyses and present the results in Section 4 of this report.

In order to accommodate data analysis needs for the Common Measures Project, the following enhancements to the ERP Performance Analyzer were completed:

- Increased Automation of Chart Generation
  The charts are created seamlessly by the ERP Performance Analyzer to eliminate the need for transferring output from the ERP Performance Analyzer to separate Excel spreadsheets.
- Expand Export Functionality to Include Compliance Scores
  The database is now designed to allow the user to specify which indicators to include in statistical analyses, charts and tables. This task also involves modification of the graphical user interfaces (GUI).
- Update and Expand Functionality of Statistical Software
  The automated statistical software was updated to accommodate new or improved statistical methodologies. These updates have made the Common Measures Project database more consistent with the Excel calculators developed by EPA and made available through their ERP Resources Library website, as well as current publications in the statistical literature that address optimal methods for estimating and comparing proportions.

One of the many benefits of the ERP Performance Analyzer is that it is highly adaptable and can be used by states for future ERP-type measurement work. Currently, the states of Colorado and Washington are using this software for their ERP work.

[6] A description of the ERP Performance Analyzer can also be found in the States ERP Consortium Guide to Reporting ERP Results (Appendix F) will be posted at www.erpstates.org.

3.4.2 Quality Assurance Procedures for Data Analysis and Reporting

The Quality Assurance officer established and implemented procedures for:

- Data entry into the ERP Performance Analyzer.
- Analysis and presentation of the data.
- Quality assurance procedures for data entry, analysis, and presentation.

See Appendix J for Quality Assurance Procedures.

3.5 OBSERVATIONS AND LESSONS LEARNED FROM PHASE 3

√ Some states found many facilities that were initially identified as SQGs were actually conditionally exempt, out of business or in some cases Large Quantity Generators. These states felt that the project was useful because it helped them clean-up classification issues and find and address problem facilities.

√ Despite a variety of individual state SQG definitions and procedures, project states were able to agree on definitions and uniform field verification methods for the project.

√ States were able to agree to data collection standards that would work for both a seasoned inspector and an intern.

√ It was possible to give effective training to staff with widely diverse backgrounds and responsibilities.

√ It was possible to give effective training for a large group over the phone.

√ Some states observed that it would have been more efficient and effective to define their enforcement response strategy before going out into the field.

√ Connecticut's training and intern program structure and approach was viewed as a useful model for other states because it showed how a state could do more field observations than could otherwise be possible with agency staff alone.

√ States found that involving hazardous waste field staff in the design of the indicators improved buy-in and the quality of data collected.

SECTION 4: DATA ANALYSIS AND REPORTING RESULTS (PHASE 4)

This Section of the Report Covers:

The final phase of the project involved comparing performance levels across participating states using the common set of regulatory and beyond compliance indicators. The goals of this analysis were to: 1) establish performance levels for each state, 2) identify statistically significant differences in performance scores (differences not due to chance) and 3) explore existing compliance and beyond compliance strategies being used by states and how they may influence performance levels.

4.1 HOW SQG PERFORMANCE WAS MEASURED

4.1.1 SQG Performance Indicators

Below are the final SQG indicators that states used to gather baseline data from field observations. The project used 8 regulatory indicators and 5 beyond compliance indicators.

Note: The first two questions on the checklist were screening/definitional questions designed to determine if a facility was an SQG.

STATES COMMON MEASURES PROJECT SQG INDICATORS

REGULATORY INDICATORS

3. Are all hazardous waste containers properly labeled with the words "hazardous waste" and clearly marked with the date on which accumulation began?

4. Are all hazardous waste containers closed unless waste is being added or removed?

5. Are all hazardous waste containers in good condition, (i.e., free of severe rusting or apparent structural defects, and not leaking)

6. At the time of the inspection has the facility accumulated more than _____* kg of RCRA hazardous waste onsite?

7. At the time of the inspection, does the facility have any RCRA hazardous waste onsite that has been accumulated onsite for more than _____*days?

8. Does the facility use a hazardous waste manifest to ship its hazardous waste when a manifest is required?

9. Has the facility identified all of its hazardous waste streams?

10. Has the facility posted the current name and telephone number of the emergency coordinator, the location of fire extinguishers and spill control material, and if present, fire alarm, and the telephone number of the fire department, unless the facility has a direct alarm?

BEYOND COMPLIANCE INDICATORS

11. Has the facility taken one or more actions to reduce toxics, conserve water or energy over the past three years?

11a. Has the facility implemented toxic use reduction over the past 3 years?

11b. Has the facility undertaken recycling projects over the past three years?

11c. Has the facility implemented water conservation projects over the past three years?

11d. Has the facility implemented energy conservation/alternative energy projects over the past three years?

4.1.2 How SQG Indicators were Analyzed

Data from each state were used to estimate the mean facility achievement score on "all indicators," "regulatory indicators" and "beyond compliance indicators" and the SQG achievement rate on each individual indicator.

How Each State's Mean SQG Facility Score was Calculated

- The proportion of the performance indicators to which a particular facility is subject, which the facility achieved. This is expressed as a number from 1 to 10: A score of "1" means that the facility was achieving 10% of the applicable indicators, a score of "5" signifies that the facility achieved 50% of the applicable indicators, and a score of 10 means that the facility achieved 100% of the applicable indicators.

  Note: Scores can be calculated on all or a subset of indicators. For example, separate scores can be calculated for "regulatory indicators" and "beyond compliance indicators." The scores can be calculated for each facility in each state.

How Each State's SQG Achievement Rate on each Indicator was Calculated

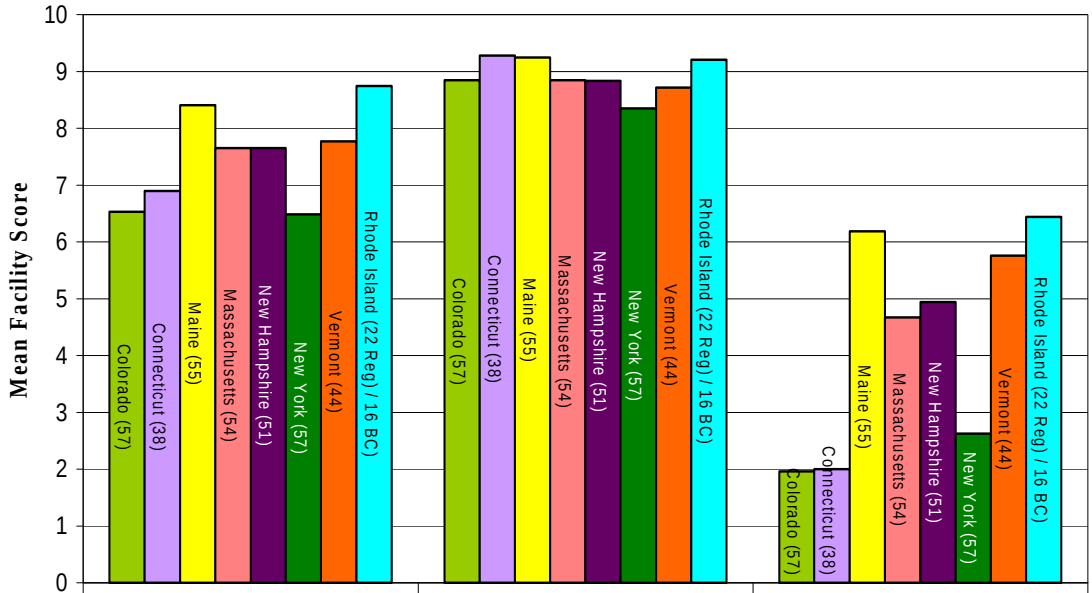
- The percentage of the facilities that were "achieving" (behaving in the desired way) each performance indicator (i.e., the facility was complying with the regulatory requirement or implementing the beyond compliance practice). This percentage was calculated for each applicable indicator for each state.

4.2 STATE-BY-STATE OBSERVED SQG PERFORMANCE RESULTS

4.2.1 Raw Observed SQG Performance Scores

Observed mean SQG facility scores ranged from 6.53 to 8.75 for "all indicators", 8.85 to 9.28 for "regulatory indicators" and 1.96 to 6.44 for "beyond compliance indicators."

States Common Measures Project
Observed State Mean SQG Facility Scores

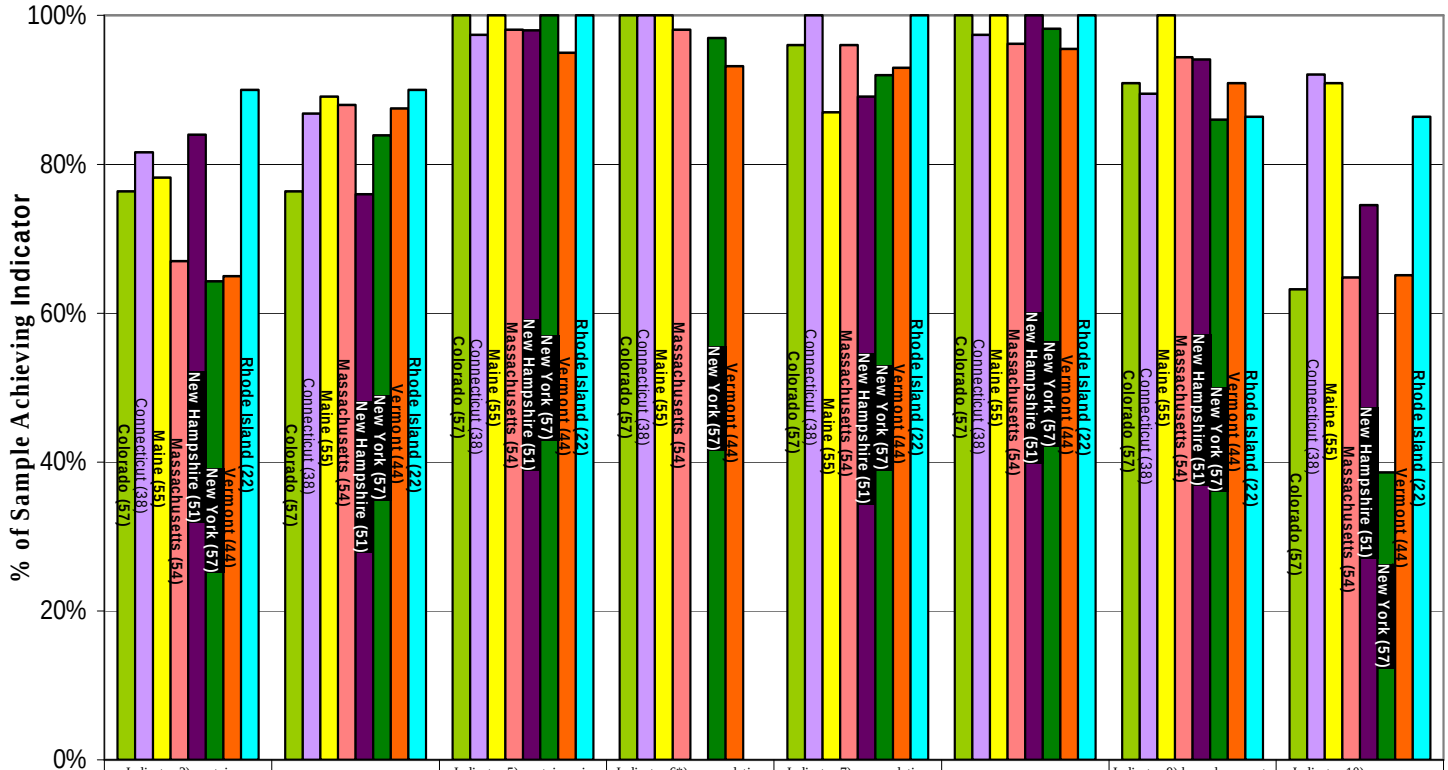


	All Indicators	Regulatory Indicators	Beyond Compliance Indicators
Colorado (57)	6.53	8.85	1.96
Connecticut (38)	6.90	9.28	2.00
Maine (55)	8.41	9.25	6.19
Massachusetts (54)	7.65	8.85	4.67
New Hampshire (51)	7.65	8.84	4.94
New York (57)	6.49	8.35	2.63
Vermont (44)	7.77	8.72	5.76
Rhode Island (22 Reg) / 16 BC)	8.75	9.21	6.44

#s in parentheses indicate # of facilities included in the sample

Observed SQG achievement rates for the individual regulatory indicators were relatively consistent across states. The greatest variation occurred with **Indicator 3**: Containers Properly Labeled, **Indicator 4:** Containers Closed and **Indicator 10:** Emergency Response Information Posted:

States Common Measures Project
Observed State SQG Achievement Rates on Regulatory Indicators

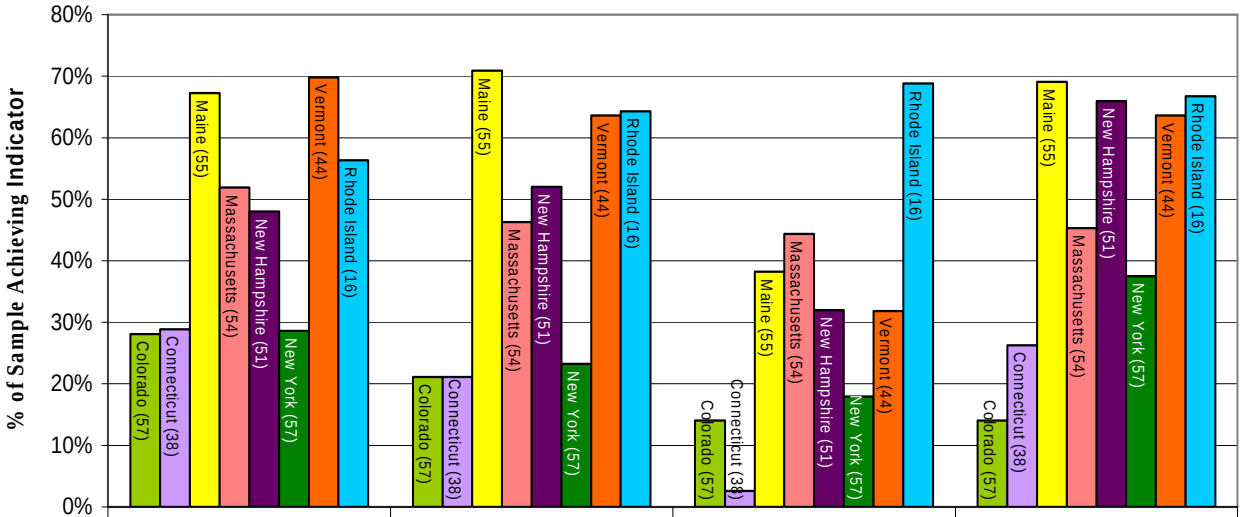


	Indicator 3) containers properly labeled	Indicator 4) containers closed	Indicator 5) containers in good condition	Indicator 6*) accumulation quantity limits followed	Indicator 7) accumulation time limits followed	Indicator 8) manifests used	Indicator 9) hazardous waste streams identified	Indicator 10) emergency response information posted
Colorado (57)	76.4%	76.4%	100.0%	100.0%	96.0%	100.0%	90.9%	63.2%
Connecticut (38)	81.6%	86.8%	97.4%	100.0%	100.0%	97.4%	89.5%	92.1%
Maine (55)	78.2%	89.1%	100.0%	100.0%	87.0%	100.0%	100.0%	90.9%
Massachusetts (54)	67.0%	88.0%	98.1%	98.1%	96.0%	96.2%	94.4%	64.8%
New Hampshire (51)	84.0%	76.0%	98.0%		89.1%	100.0%	94.1%	74.5%
New York (57)	64.3%	83.9%	100.0%	97.0%	92.0%	98.2%	86.0%	38.6%
Vermont (44)	65.0%	87.5%	95.0%	93.2%	93.0%	95.5%	90.9%	65.1%
Rhode Island (22)	90.0%	90.0%	100.0%		100.0%	100.0%	86.4%	86.4%

* Indicator 6 does not apply to RI or NH
#s in parentheses indicate # of facilities included in the state's sample

Observed achievement rates for beyond compliance indicators had much more variation.

States Common Measures Project
Observed State SQG Achievement Rates on Beyond Compliance Indicators



#s in parentheses indicate # of facilities included in the state's

	Indicator 11a) toxic use reduction implemented	Indicator 11b) recycling projects undertaken	Indicator 11c) water conservation implemented	Indicator 11d) energy conservation/alternative energy implemented
Colorado (57)	28.1%	21.1%	14.0%	14.0%
Connecticut (38)	28.9%	21.1%	2.6%	26.3%
Maine (55)	67.3%	70.9%	38.2%	69.1%
Massachusetts (54)	51.9%	46.3%	44.4%	45.3%
New Hampshire (51)	48.0%	52.0%	32.0%	66.0%
New York (57)	28.6%	23.2%	17.9%	37.5%
Vermont (44)	69.8%	63.6%	31.8%	63.6%
Rhode Island (16)	56.3%	64.3%	68.8%	66.7%

4.3 STATISTICAL ANALYSIS OF SQG PERFORMANCE RESULTS

4.3.1 Interpreting Observed SQG Results

The data in the above charts represent observed SQG performance levels in each state. Had all of the SQGs in each state been evaluated, these results could provide definitive information about each state's SQG performance and any differences among them. However, as is often the case, only a portion of the SQGs could be inspected. Since the facilities were randomly selected, statistical methods were used to generalize the field observation findings to the sector as a whole and to determine if the observed differences among states are likely to be due to chance, or reflect actual differences in performance.

When statistics are used to estimate the performance of an entire population from a sample, the conclusions are presented as a range rather than as a point (e.g. 75% + or – a confidence interval of 10% which means that the estimated achievement rate is between 65% and 85%).

4.3.2 Statistically Significant Differences in SQG Mean Facility Scores

Exhibit 4.1 below presents statistically significant differences in mean facility achievement scores, at a 90% confidence level with the associated confidence interval, across all states and categorized by group, (i.e., all indicators, regulatory indicators and beyond compliance indicators):

HOW TO READ THESE CHARTS: The chart below presents the number of facilities that were evaluated, the observed performance score and associated confidence interval range in each state. States are listed from highest to lowest SQG performance and statistically significant differences between states are shown by arrows. An arrow connecting a state to another state below on the list indicates the existence of a statistically significant difference. The lack of an arrow between any two states means the observed differences may be due to chance. A state may use these charts to compare its SQG performance with any other state's SQG performance.

Exhibit 4.1: Statistically Significant Differences in State SQG Mean Facility Scores

STATE	# of Facilities	Observed Mean Facility Score	Confidence Interval (90% confidence level)
All Indicators			
RI	*22*	*8.75*	*8.21 - 9.21*
ME	55	8.41	8.12 - 8.69
VT	44	7.77	7.23 - 8.27
NH	*51*	*7.65*	*7.17 - 8.10*
MA	54	7.65	7.18 - 8.09
CT	38	6.90	6.50 - 7.29
CO	57	6.53	6.25 - 6.81
NY	57	6.49	6.04 - 6.94
Regulatory Indicators Only			
CT	*38*	*9.28*	*8.99 - 9.53*
RI	22	9.21	8.94-9.44
ME	55	9.25	9.10 - 9.39
MA	54	8.85	8.57 - 9.10
CO	57	8.85	8.61 - 9.06
NH	*51*	*8.84*	*8.51 - 9.14*
VT	44	8.72	8.27 - 9.11
NY	57	8.35	7.99 - 8.67
Beyond Compliance Indicators Only			
RI	16	6.44	4.86 - 7.88
ME	55	6.19	5.34 - 7.00
VT	44	5.76	4.86 - 6.64
NH	*50*	*4.94*	4.08 - 5.81
MA	54	4.67	3.78 - 5.58
NY	56	2.63	1.93 - 3.40
CT	38	2.00	1.42 - 2.65
CO	57	1.96	1.45 - 2.53

Note: If the number of facilities is different from the number of inspections a state completed, it is due to incomplete data available for that indicator.

Using the "All Indicators" chart above as an example:

- Rhode Island and Maine's SQG mean facility scores were statistically significantly higher than Vermont, New Hampshire, Massachusetts, Connecticut, Colorado and New York.
- Vermont and Massachusetts' SQG mean facility scores were statistically significantly higher than Connecticut, Colorado and New York.
- New Hampshire's SQG mean facility scores were statistically significantly higher than Colorado and New York.

4.3.3 Statistically Significance Differences in State SQG Achievement Rates for Individual Indicators

Exhibit 4.2 below presents statistically significant differences in achievement rates across all participating states by individual regulatory indicators. There were no statistically significant differences in achievement rates for indicators 4, 5 and 8.

Note: Arrows extending from one state to another indicate a statistically significant difference between the two states at a 90% confidence level. The lack of an arrow between any two states means the observed differences may be due to chance.

Exhibit 4.2: Statistically Significant Differences in State SQG Achievement Rates for Regulatory Indicators

STATE	# Facilities	Observed Achievement Rate	Confidence Interval (90% confidence level)
Indicator 3 (Containers properly labelled)			
RI	20	90.0%	73.8% - 96.6%
NH	50	84.0%	73.8% - 90.7%
CT	31	81.6%	69.3% - 89.7%
ME	55	78.2%	67.8% - 85.9%
CO	55	76.4%	65.8% - 84.4%
MA	51	66.7%	55.2% - 76.4%
VT	40	65.0%	52.0% - 76.1%
NY	55	63.6%	52.6% - 73.4%
Indicator 4 (Containers closed)			
RI	20	90.0%	73.8% - 96.6%
ME	55	89.1%	80.3% - 94.3%
MA	51	88.2%	78.8% - 93.8%
VT	40	87.5%	76.5% - 93.8%
CT	38	86.8%	75.3% - 93.4%
NY	55	83.6%	73.9% - 90.2%
CO	55	76.4%	65.8% - 84.4%
NH	50	76.0%	64.9% - 84.4%
Indicator 5 (Containers in good condition)			
CO	55	100.0%	95.3% - 100.0%
ME	55	100.0%	95.3% - 100.0%
NY	55	100.0%	95.3% - 100.0%
RI	20	100.0%	88.1% - 100.0%
MA	51	98.0%	91.7% - 99.6%
NH	50	98.0%	91.5% - 99.6%
CT	38	97.4%	89.0% - 99.4%
VT	40	95.0%	86.0% - 98.3%
Indicator 6* (Accumulation quantity limit followed)			
CO	57	100.0%	95.5% - 100.0%
ME	55	100.0%	95.3% - 100.0%
CT	37	100.0%	93.2% - 100.0%
MA	54	98.1%	92.1% - 99.6%
NY	57	96.5%	89.9% - 98.8%
VT	44	93.2%	84.1% - 97.2%

* Note: Indicator 6 does not apply to NH or RI

STATE	# Facilities	Observed Achievement Rate	Confidence Interval (90% confidence level)
Indicator 7 (Accumulation time limit followed)			
CT	35	100.0%	92.8% - 100.0%
RI	21	100.0%	88.6% - 100.0%
CO	57	96.5%	89.9% - 98.8%
MA	53	96.2%	89.2% - 98.7%
VT	43	93.0%	83.8% - 97.2%
NY	52	92.3%	83.9% - 96.5%
NH	46	89.1%	79.3% - 94.6%
ME	53	86.8%	77.3% - 92.7%
Indicator 8 (Manifests used)			
CO	57	100.0%	95.5% - 100.0%
ME	55	100.0%	95.3% - 100.0%
NH	51	100.0%	95.0% - 100.0%
RI	22	100.0%	89.0% - 100.0%
NY	56	98.2%	92.4% - 99.6%
CT	38	97.4%	89.0% - 99.4%
MA	53	96.2%	89.2% - 98.7%
VT	44	95.5%	87.2% - 98.5%
Indicator 9 (Hazardous waste streams identified)			
ME	55	100.0%	95.3% - 100.0%
MA	54	94.4%	86.9% - 97.8%
NH	51	94.1%	86.2% - 97.6%
VT	44	90.9%	81.2% - 95.9%
CO	57	91.2%	83.1% - 95.7%
CT	38	89.5%	78.5% - 95.2%
RI	22	86.4%	70.3% - 94.4%
NY	57	86.0%	76.8% - 91.9%
Indicator 10 (Emergency response information posted)			
CT	38	92.1%	81.8% - 96.8%
ME	55	90.9%	82.5% - 95.5%
RI	22	86.4%	70.3% - 94.4%
NH	51	74.5%	63.4% - 83.1%
VT	43	65.1%	52.6% - 75.9%
MA	54	64.8%	53.7% - 74.6%
CO	57	63.2%	52.3% - 72.8%
	57	38.6%	28.7% - 49.5%

Exhibit 4.3 below presents statistically significant differences in achievement rates across all participating states by individual beyond compliance indicators.

Note: Arrows extending from one state to another indicate a statistically significant difference between the two states at a 90% confidence level. The lack of an arrow between any two states means the observed differences may be due to chance.

Exhibit 4.3 Statistically Significance Differences in State SQG Achievement Rates for Beyond Compliance Indicators

STATE	# Facilities	Observed Achievement Rate	Confidence Interval (90% confidence level)
Indicator 11 (Any reduction or conservation measures conducted over past 3 yrs)			
RI	16	87.5%	68.4% - 95.8%
ME	55	81.8%	71.8% - 88.8%
VT	44	84.1%	73.1% - 91.1%
NH	51	78.4%	67.7% - 86.3%
MA	54	72.2%	61.3% - 81.0%
NY	56	48.2%	37.6% - 59.0%
CT	38	47.4%	34.7% - 60.4%
CO	57	40.4%	30.3% - 51.2%
Indicator 11a (Toxics use reduction implemented)			
VT	43	69.8%	57.4% - 79.8%
ME	55	67.3%	56.3% - 76.7%
RI	11	56.3%	36.5% - 74.2%
MA	54	51.9%	40.8% - 62.7%
NH	50	48.0%	36.8% - 59.4%
CT	38	28.9%	18.6% - 42.1%
NY	56	28.6%	19.8% - 39.3%
CO	53	28.1%	19.4% - 38.7%
Indicator 11b (Recycling projects undertaken)			
ME	55	70.9%	60.0% - 79.8%
RI	14	64.3%	42.6% - 81.4%
VT	44	63.6%	51.2% - 74.5%
NH	50	52.0%	40.6% - 63.2%
MA	54	46.3%	35.6% - 57.4%
NY	56	23.2%	15.3% - 33.6%
CO	57	21.1%	13.6% - 31.1%
CT	38	21.1%	12.3% - 33.7%

STATE	# Facilities	Observed Achievement Rate	Confidence Interval (90% confidence level)
Indicator 11c (Water conservation implemented)			
RI	16	68.8%	48.2% - 83.9%
MA	54	44.4%	33.9% - 55.6%
ME	55	38.2%	28.2% - 49.3%
NH	50	32.0%	22.3% - 43.5%
VT	44	31.8%	21.6% - 44.1%
NY	56	17.9%	11.0% - 27.7%
CO	57	14.0%	8.1% - 23.3%
CT	38	2.6%	0.6% - 11.0%
Indicator 11d (Energy conservation/Alternative energy implemented)			
ME	55	69.1%	58.1% - 78.2%
RI	15	66.7%	45.5% - 82.7%
NH	50	66.0%	54.4% - 75.9%
VT	44	63.6%	51.2% - 74.5%
MA	53	45.3%	34.5% - 56.5%
NY	56	37.5%	27.7% - 48.5%
CT	38	26.3%	16.4% - 39.4%
CO	57	14.0%	8.1% - 23.2%

4.3.4 The Effect of Confidence Levels, Number of Inspections and Confidence Intervals on the Usefulness of the Data for Decision Making

Measurement projects such as the States Common Measures Project are not undertaken just for the sake of measuring something. These measurements are needed to determine if the facility performance in a state is "good enough" to meet the state's policy objectives for the regulatory program and to identify any oversight practices that appear to be associated with higher performance levels. To the extent that the findings are reliable, the states are able to use the findings to make better decisions about efficient and effective programs.

When making choices on the basis of data, decision makers generally consider two factors. One is the confidence level in the results -- the likelihood that the observed data accurately reflects the conditions in the world. The other is the precision of the results -- the confidence interval, or the range above and below the observed value within which the group's performance actually falls. In general, the larger the number of inspections, the greater the precision of the results and the confidence that the results accurately reflect conditions in the world.

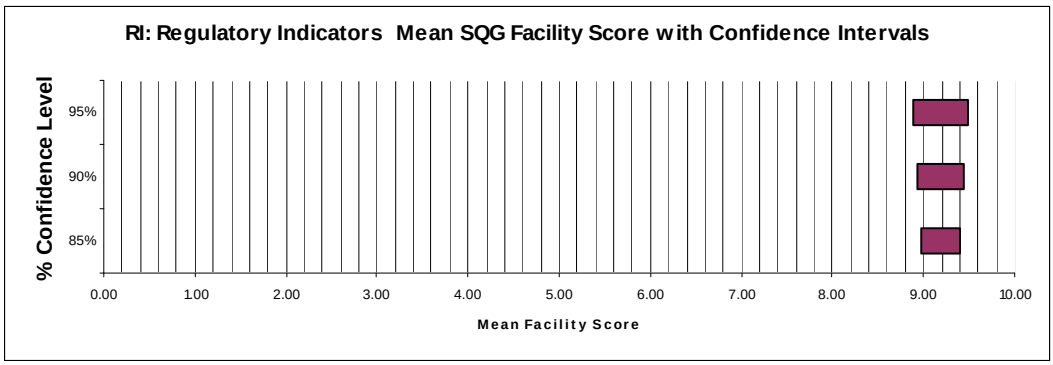
In many fields, a 90% confidence level (which indicates that there is a 10% chance that the observed results do not reflect the actual performance of the group) is considered the lowest "acceptable" level for drawing conclusions about the behavior of a group, and whether that behavior is "statistically different" than that of another group. However, achieving a precise measurement at a high level of confidence may require more inspections than a state can "afford" and perhaps may need. A state may be comfortable being only 85% certain that their results are within the specified confidence interval, if the consequences of being wrong are not serious. A state may be able to base a decision on a very wide confidence interval if "good enough" or "not good enough" performance is within that wide range, or the state only needs to be able to identify very large differences in performance levels. Furthermore, policy decisions sometimes need to be based on whatever amount of information can be obtained with the level of resources available to measure.

To explore the issue of what is "good enough" data, the States Common Measures Project analyzed the data at three different confidence levels. The results (shown in the charts below) illustrate how confidence intervals are affected by confidence levels, the numbers of inspections and the observed performance. This provides some insight into the level of confidence and precision and therefore the number of inspections that a state decision maker may "need" to make choices about program design.

Note: The confidence interval is calculated from a formula based on three factors: the confidence level, the observed performance rate and the sample size. The relationship among these three factors is the same regardless of the indicator being measured or the universe size. Therefore, the charts below use different indicators from different states to illustrate this relationship.

The Relationship between Confidence Levels and Confidence Intervals

As illustrated in the chart below, as the confidence level increases, so does the confidence interval. An analyst can be more certain of his or her result the wider the range within which it can fall. For example, one is likely to be 99% certain that one has properly estimated a person's age if one guesses that the person is somewhere between 1 and 100 years old. However, one would be less certain if he or she guesses that the person's age is between 40 and 60 years.

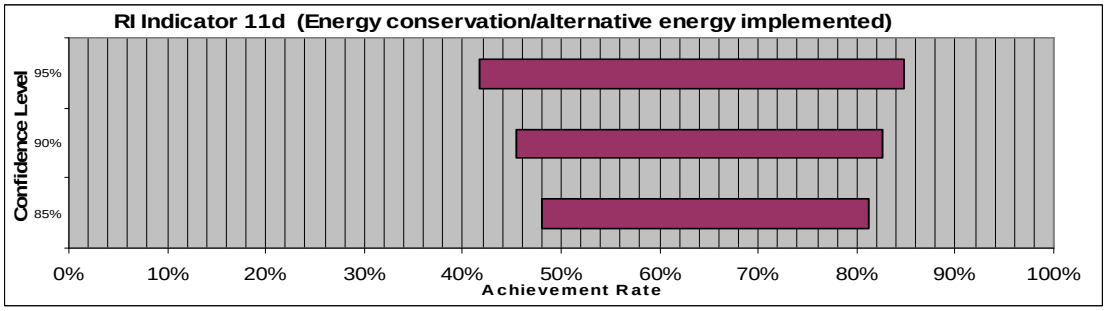


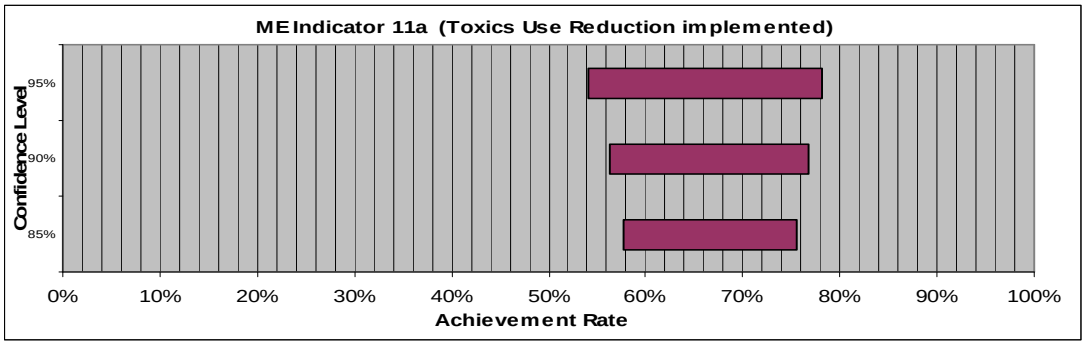
The above chart presents data collected from 22 facilities. The observed mean facility score was 9.2. The decision maker can be:

- 95% certain that the true mean facility score is between 8.88 and 9.48
- 90% certain it falls between 8.94 and 9.44
- 85% certain it falls between 8.98 and 9.41

The Relationship between Confidence Interval Breadth and Confidence Level at Different Numbers of Inspections:

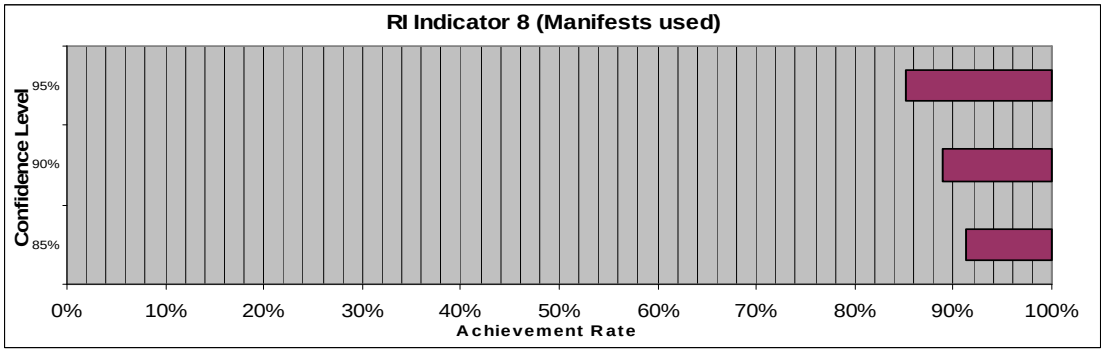
The confidence intervals are shown below for two states with an observed SQG achievement rate of 67% on an indicator. RI inspected 15 facilities, while Maine inspected 55 facilities. As can be seen, the confidence interval in RI was much wider than that of Maine.



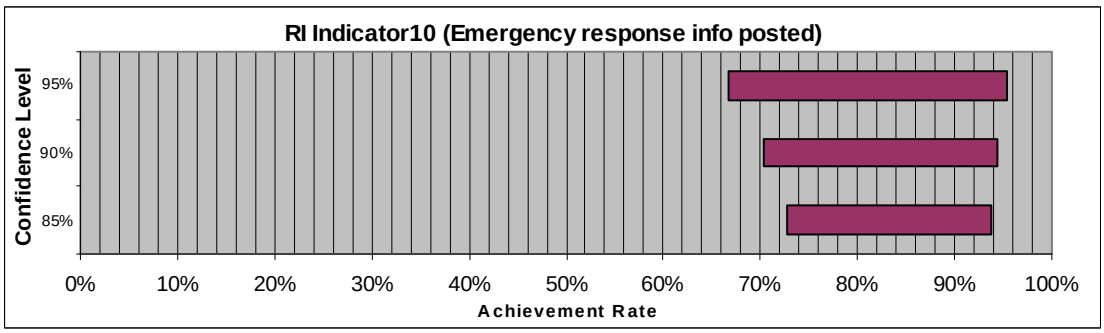


The Relationship between Confidence Interval Breadth and Confidence Level when the Observed Results are Nearer to 50% than to 0% or 100%

The charts below show the confidence intervals for the observed achievement rates on two different indicators in RI with the same number of facilities. The confidence intervals are much wider with an observed SQG achievement rate of 86% than with an observed SQG achievement rate of 100%.

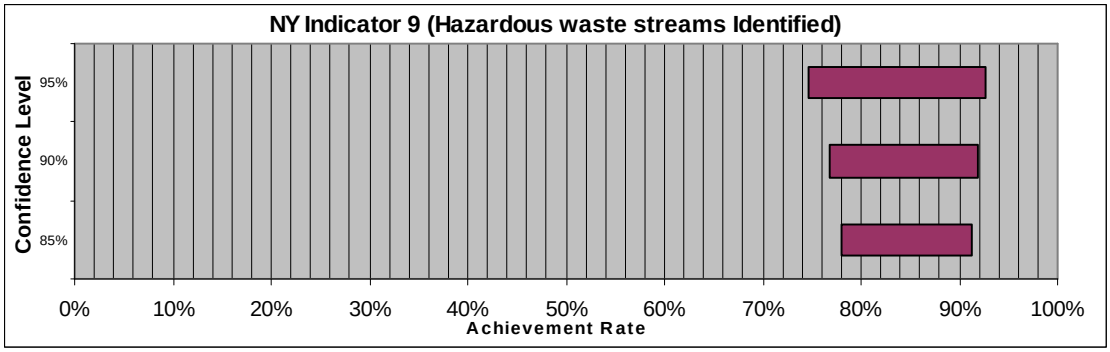


Note: The above chart presents data collected from 22 facilities. The observed achievement rate is 100%.



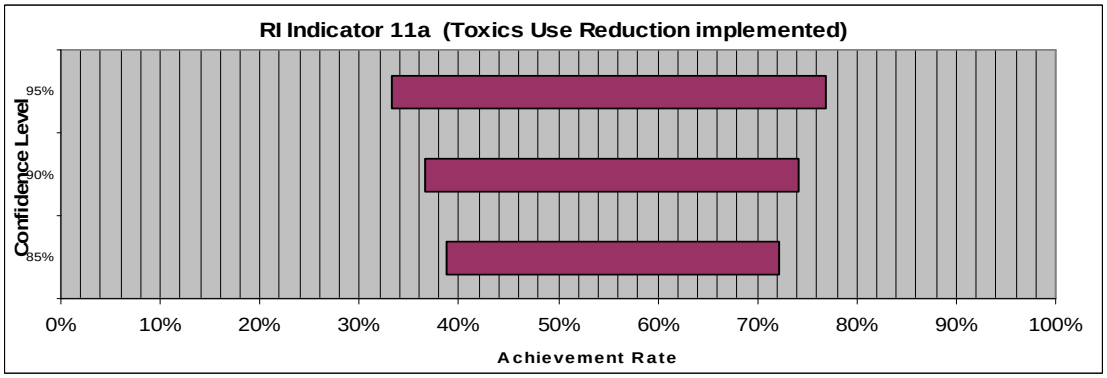
Note: The above chart presents data collected from 22 facilities. The observed achievement rate is 86%.

The chart below shows that conducting more inspections at the same observed SQG achievement rate (86%) results in a narrower confidence interval.



Note: The above chart presents data collected from 57 facilities. The observed achievement rate is 86%.

Finally as illustrated by the chart that follows, the wide confidence intervals at performance rates of around 50% may not pose a problem for a state decision maker even at low numbers of inspections. The upper bound of the observed SQG achievement rate for RI's indicator 11a is around 74% at a 90% confidence level. Even with as few as 16 inspections, the data may be precise enough for the decision maker to determine that compliance is not "good enough."



Note: The above chart presents data collected from 16 facilities. The observed achievement rate is 56.3%.

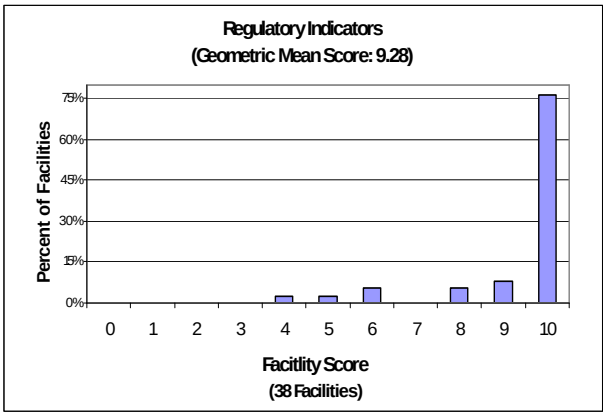
Conclusion:

In sum, while more inspections can help increase the precision and certainty of results, in many instances a state decision maker can still glean useful information at relatively low numbers of inspections, by looking at the confidence intervals at different confidence levels. *See Appendix K for the confidence intervals for each state's observed SQG achievement rates and mean facility scores at the 85%, 90%, and 95% levels. See Appendix L for the statistically significant differences in each state's SQG achievement rates and mean facility scores at 85%, 90%, and 95% confidence levels.*

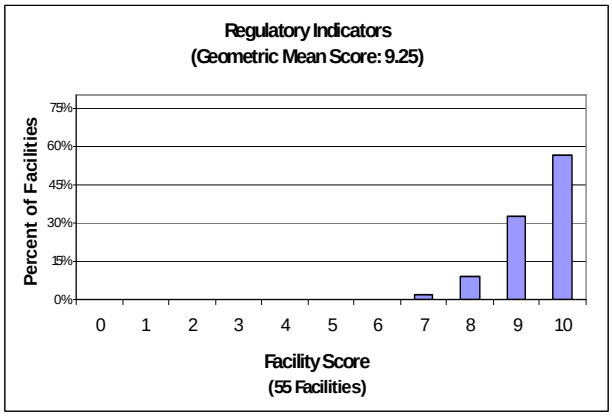
4.3.5 Another Look at the Data: Distribution of a State's SQG Facility Scores

At 9.28 and 9.25 respectively, Connecticut's and Maine's observed mean SQG facility scores on regulatory indicators was virtually the same, and indicated that facilities are in compliance with more than 90% of the applicable requirements. From this information alone one might conclude that SQG performance was equivalent in the two states. The histograms below, which show the percent of inspected facilities in each state that achieved a given facility score, provide a different perspective. 76% of the Connecticut SQGs had a score of 10 (meaning they were in compliance with all applicable regulatory indicators) while 56% of the Maine SQGs had a perfect score. On the other hand, all of the Maine SQGs had scores of 7 or above (meaning they all were complying at least 70% of the applicable indicators), whereas 11% of Connecticut SQGs had scores below 7. In light of such data, a state might ask itself whether it needs to take steps to address a small pocket of "poor" performance, or a state might ask itself whether it should implement strategies to increase the percentage of facilities achieving "perfect" scores. *See Appendix M for histograms for each state's facility scores for "all," "regulatory" and "beyond compliance" indicator groups.*

Connecticut



Maine



4.4 EXPLORATION OF STATE ACTIVITIES INFLUENCE ON SQG PERFORMANCE RESULTS

4.4.1 Introduction

Overall the mean SQG facility scores for regulatory indicators were relatively high across states and ranged from 7.99 to 9.53 (taking into account confidence intervals). No state felt the overall results were cause for serious concern, but nevertheless warranted consideration of options for improvement in particular areas (this can be seen by looking at achievement rates for individual indicators):

- SQG achievement rates on 5 individual regulatory indicators [#5: containers in good condition, #6: accumulation quantity limits followed, #7: accumulation time limits followed, #8: manifest used and #9: hazardous waste identified] out of 8, for all states, were deemed acceptable and ranged from 70.3% to100% (taking into account confidence intervals).

- SQG achievement rates on 3 individual regulatory indicators [#3: containers properly labeled, #4: containers closed and #10: emergency response procedures followed] out of 8, for all states, were notably lower and had much more variability. The SQG achievement rate ranged from 29% to 97% (taking into account confidence intervals).

  Note: The lower performance on indicators 3, 4 and 10 was consistent with historical observations of the experienced hazardous waste experts who participated in the project.

Overall mean SQG facility scores for beyond compliance indicators were lower and had much more variation ranging from 1.42 to 7.88 (taking into account confidence intervals).

Project states explored whether there was anything about measured SQG performance levels that could be attributed to what a state was doing before the Common Measures Project. A meeting was held on December 2, 2008 to capture compliance and beyond compliance activities performed by states to help project states explore possible root causes for measured performance differences. The project states reported on the following categories of activities occurring between June 2004 and June 2007 (prior to the Common Measures Project data collection period):

- Regulatory compliance assistance provided to the SQG sector.
- Beyond compliance assistance provided to the SQG sector.
- Percentage of SQG universe typically inspected per year.
- Most common inspection triggers.
- Who conducted compliance inspections between June 2004 and June 2007.
- Typical SQG enforcement actions.
- SQG reporting requirements.
- Other influences that may have affected observed SQG performance.

4.4.2 Observations from Comparing State Activities to Measured SQG Performance Results

After reviewing the performance data, the project states explored the following questions to see if there was a relationship between the data reported and the performance differences among states:

1) Did the Nature and Amount of Regulatory Compliance Assistance Provided Between June 2004 and June 2007 Influence Performance?

Regulatory Compliance Assistance Provided to SQG Sector (June 2004 – June 2007)	
CO	Minimal prior to Aug 2007-includes website, guidance documents and quarterly workshops available to all hazardous waste generators. Announced on website only. Generators Assistance Program offers amnesty. Since Aug 2007, required a self cert checklist to be submitted.
CT	On-site assistance during inspections and through handouts, SQG guidance manual and other applicable guidance.
MA	Generic fact sheets on SQG compliance and sector specific workbooks & fact sheets on MassDEP website. Mass Office of Technical Assistance (OTA) provides technical assistance upon request. All SQG enforcement actions are referred to OTA.
ME	Verbal assistance via inspections and telephone; Written compliance assistance via manuals; Written and verbal assistance via seminar or meeting formats; Assistance via Electronic media including web site development and availability of guidance manuals, DEP resources, inspection process, compliance issues, etc.
NH	SQG certification classes, Haz Waste topics training classes, Assistance site visits, assistance of web, RCRA Hot Line.
NY	Technical Assistance at Trade Association Meetings, SQG hotline, SQG compliance guide on web, workshops, P2 assistance guide.
RI	Respond to direct requests for assistance, proactive ERP auto body/auto salvage assistance (includes compliance assistance materials).
VT	Targeted assistance for RCRA-regulated community rather than to SQGs exclusively. Fact sheets and industry sector guides on web. On-line RCRA tutorial, compliance assistance workshops, on-site compliance assistance audits, over the phone assistance.

a) Observations Based on Mean SQG Facility Scores for Regulatory Indicators

- States that reported providing onsite regulatory compliance assistance: CT, MA, ME, NH, VT, RI.
- States that reported not providing onsite regulatory compliance assistance: CO, NY.

    - The mean SQG facility scores in all states that reported providing onsite regulatory compliance assistance except Vermont were statistically significantly higher than New York.
    - The mean SQG facility scores in three of these states [RI, CT, ME] were statistically significantly higher than Colorado.

    See Section 4.3.2, Exhibit 4.1: Statistically Significant Differences in State SQG Mean Facility Scores.

    - **It appears that there is a relationship between providing on-site regulatory compliance assistance and higher measured SQG performance.**

b) Observations Based on Individual Regulatory Indicators

    - NH reported that at every SQG certification and all training classes, pre-made labels were distributed as part of their certification program.
    - NH had the 2nd highest SQG achievement rate [84%] with indicator 3: containers properly labeled, and was statistically significantly higher than 3 other states [MA, VT, NY].
    - RI had the highest SQG achievement rate [90%] with indicator 3: containers labeled and was statistically significantly higher than 4 other states [NH, MA, VT, NY]. RI did not hand out labels (also note that RI did fewer inspections and had wider confidence intervals).

    See Section 4.3.3, Exhibit 4.2: Statistically Significant Differences in State SQG Achievement Rates for Regulatory Indicators.

    - **States felt that performance differences related to indicator 3 may warrant further evaluation.**

2) Did the Nature and Amount of Beyond Compliance Assistance Provided Between June 2004 and June 2007 Influence Performance?

Beyond Compliance Assistance Provided to SQG Sector (June 2004 – June 2007)	
CO	Very little prior to Common Measures Project – hope to improve through ERP and self-certification roll-out.
CT	None-other than on-site suggestions during the inspection.
MA	Information on Mass DEP's website, referral to OTA as part of enforcement and OTA direct assistance. Topics Covered: P2, Water Conservation, EMS.
ME	Phone calls, emails and site visits to encourage facilities to reduce their environmental and carbon footprint. The Office of Innovation & Assistance works with the RCRA group on referrals as well as suggestions for sector assistance focus. Topics Covered: P2, Energy Conservation, Water Conservation, Air Emission Reductions (both Toxics and Green House Gases), Chemical Use Reductions, and Environmentally Preferable Purchasing.
NH	NH Pollution Prevention Program and Small Business Technical Assistance Program provide site visits, conferences, outreach & education to promote "beyond compliance" activities. Topics Covered: P2, Energy Conservation, Water Conservation, EMS & EPA programs such as "Lean & Energy."
NY	None
RI	Through ERP and individual requests. Topics Covered: P2, Water Conservation.
VT	Assistance with toxics use/hazardous waste reduction planning. Topics Covered: P2

a) Observations Based on Mean SQG Performance Scores of Beyond Compliance Indicators

- States that reported having active beyond compliance programs: MA, ME, NH, RI, VT.
- States that reported not having active beyond compliance programs: CO, CT, NY.

    - The mean SQG facility scores in **all** states with active beyond compliance programs were statistically significantly higher than in **all** states without active programs.

    See Section 4.3.2, Exhibit 4.1: Statistically Significant Differences in State SQG Mean Facility Scores.

    - **It appears that there is a relationship between active beyond compliance programs and higher measured SQG performance.**

b) Observations Based on Individual Beyond Compliance Indicators

Toxic Use Reduction:

- States that reported providing toxic use reduction assistance: MA, ME, NH, RI, VT.
- States that reported not providing toxic use reduction assistance: CO, CT, NY.

    - The state SQG achievement rates for indicator 11a: toxic use reduction implemented, were statistically significantly higher in **all** states that reported having toxic use reduction assistance than in **all** states that reported having no toxic use reduction assistance.

    See Section 4.3.3, Exhibit 4.3: Statistically Significant Differences in State SQG Achievement Rates for Beyond Compliance Indicators.

    - **It appears there is a relationship between states reporting that they provided toxic use reduction assistance and higher measured SQG performance.**

Water Conservation:

- States that reported providing water conservation assistance: MA, ME, NH, RI.
- States that reported not providing water conservation assistance: CO, CT, NY, VT.

    - The state SQG achievement rates for indicator 11c: water conservation procedures implemented, were statistically significantly higher in **all** states that reported having water conservation assistance than in **all** states that reported not having water conservation assistance.

    See Section 4.3.3, Exhibit 4.3: Statistically Significant Differences in State SQG Achievement Rates for Beyond Compliance Indicators.

    - **It appears that there is a relationship between states reporting that they provided water conservation assistance and higher measured SQG performance.**

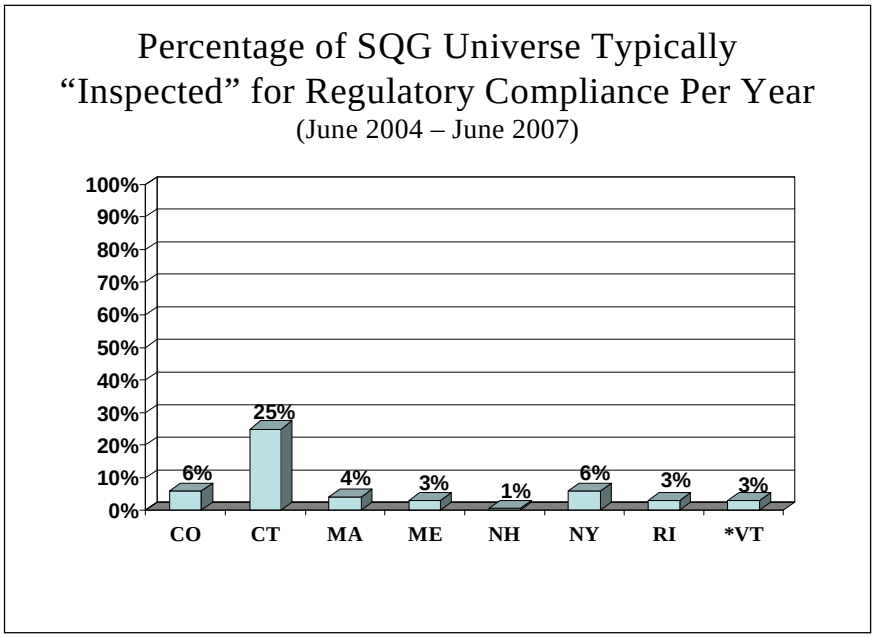
Energy Conservation:

- States that reported providing energy conservation assistance: ME, NH.
- States that reported not providing energy conservation assistance: CO, CT, MA, NY, RI, VT.
    - ME, RI, NH, and VT had the highest SQG achievement rates for indicator 11d and there were no statistical differences among them.

- ME's and NH's SQG achievement rates were statistically significantly higher than four other states' SQG achievement rates (MA, NY, CT, and CO).
- RI's and VT's SQG achievement rates were statistically significantly higher than three other states' SQG achievement rates (NY, CT, CO).

o **Based on the mixed results, the relationship between energy conservation assistance and measured SQG performance may warrant further evaluation.**

3) Did the Frequency of Inspections between June 2004 and June 2007 Influence Performance?



Note: Historically VT's goal has been to inspect 5 – 10% of this universe annually. Their hazardous waste generator database only identifies the current status of a facility, which is not necessarily the status of the facility at the time of inspection, e.g. an SQG facility turns out to be a CESQG or has gone out of business. For this reason, the actual percentage of SQGs inspected annually during the 2004 to 2007 period was closer to 3.5%.

- CT reported conducting substantially higher numbers of inspections [25% of universe] than all other states [3-6% of universe].

- CT had the 2nd highest mean SQG facility score on regulatory indicators. However, this score was not statistically significantly different than that of the other two states with high mean facility scores [RI which ranked first and ME which ranked third].

o **The data do not appear to support a relationship between the frequency of inspections and the measured SQG performance.**

4) Did the Most Common Inspection Triggers between June 2004 and 2007 Influence Performance?

Most Common SQG Inspection Triggers
(June 2004 – June 2007)

- "Routine" – with regular frequency, e.g., once every 2 years [None of the States]
- Complaint [CO, MA, NH, VT]
- Inspector Discretion [CO, MA]
- Other:
  - CT Special Initiative to Inspect all SQGs
  - NH SQGs Who Failed to Become Certified
  - MA targets all SQGs that are a major air source and/or major water source
  - Vermont has a policy to inspect once every 10 years

- **The data do not appear to support a relationship between most common inspection triggers and the measured SQG performance.**

5) Did Who Conducted SQG Compliance Inspection between June 2004 and June 2007 Influence Performance?

SQG "Compliance Inspections"
(June 2004 – June 2007)

Who Conducted SQG Inspections?

- All states used compliance inspectors with the exception of CT.
- CT inspections were mostly conducted by interns. Compliance inspectors accompanied interns during the first several weeks for training, periodically throughout the initiative, and would also return to facilities to conduct full inspections when significant violations were found.

What Unit?

- Most states used their hazardous waste group to complete compliance inspections.
- Other:
  - MA does not have a dedicated hazardous waste unit. Conducts primarily multi-media inspections and single media inspection for certain categories of sources
  - RI used Compliance & Inspections Unit which completes mostly RCRA inspections

Note: Staff who conducted the Common Measures field observations may or may not be the same staff noted above.

- **The data does not appear to support a relationship between who conducted the inspections and the measured SQG performance.**

6) Did the Type of SQG Enforcement Actions between June 2004 and June 2007 Influence Performance?

Typical SQG "Enforcement" Actions
(June 2004 – June 2007)

CO	Compliance Advisories (informal enforcement) and Penalties
CT	Warning Letters, Notices of Violation, Consent Orders
MA	Notices of Non-compliance, Orders and Penalties
ME	Informal enforcement including Letter of Warning (LOWs), and Notice of Violation (NOVs), and Formal enforcement including Consent Agreements (CAs) and formal legal actions civil and criminal actions.
NH	Warning Letters, Notice of Non-Compliance, Consent Orders
NY	Notices of Violation (informal), Consent Orders
RI	Letters of Non-Compliance, Formal Enforcement
VT	Notice of Alleged Violation Letters, Formal Enforcement

- **The data do not appear to support a relationship between the type of enforcement actions and the measured SQG performance.**

7) Did the Nature and Amount of SQG Reporting Requirements between June 2004 and June 2007 Influence Performance?

SQG Reporting Requirements
(June 2004 – June 2007)

*CO	No reporting requirements before August 2007. From August 2007 to present, self certification required
CT	No reporting requirements
MA	One time Notification only
*ME	Annual reporting
*NH	All SQGs must provide quarterly activity reports detailing their wastes generated, permit status, etc.
NY	No reporting requirements
*RI	Biennial reports
VT	Vermont requires SQGs (and LQGs) to pay an annual "generator fee." To facilitate payment of this fee, the Agency sends each generator a letter identifying a presumed generator status based on either manifest records from the previous year or the generator's notified status. Each generator is required to respond to the Agency letter and verify their actual generator status.

* **Routine Summary Reports**

- **The data do not appear to support a relationship between the nature and amount of reporting requirements and the measured SQG performance.**

8) Other Influences that May Have Affected Observed SQG Performance

- The chart below describes additional factors that states reported that may have affected performance.

Other Factors That May Have Affected SQG Performance (June 2004 – June 2007)

CO	Colorado is required to announce inspections at least 24 hours in advance. In addition to the advanced warning, they also performed a self certification and had the checklist to go by before we came out. This allowed the facility to prepare for and know what we were looking for on our inspections
CT	Since the inspections were conducted over a 4 year time period, the word got around that we were conducting SQG inspections- this may have had an impact on the level of compliance for the later inspections, but it is difficult to measure if there was any real change in behavior
MA	ERP Dry Cleaners, Photo Processors, Printers and other targeted groups. Enforcement against "SQGs" that are identified through report reviews to be LQGs.
ME	No information provided
NH	As we have found during ERP projects, few SQGs expect compliance inspections and are sometimes unconcerned about what they consider to be "frivolous" regulations, such as maintaining proper aisle space for drums. NH uses manifest list to generate SQG universe.
NY	Identified a number of SQGs over the past few years when reviewing manifests which had been sent out as CESQGs incorrectly. Follow-up inspections led to the discovery of violations as the facility was not familiar with the regulations. Once the inspection/enforcement cycle ends we hope the performance of these SQGs is improved.
RI	ERP Auto Body, ERP Auto Salvage, Clean Marina Program
VT	SQGs of routinely generated hazardous waste and users of more than 1000 lbs of toxic substances are required to prepare toxics use and hazardous waste reduction plans every three years and annual progress reports

4.5 EXPLORATION OF POSSIBLE AREAS OF BIAS IN SQG PERFORMANCE RESULTS

1) Did SQGs in States with Lower Accumulation Quantity Limits or Shorter Accumulation Time Limits Have Poorer Performance?

State	Accumulation Quantity Limit	Accumulation Time Limit	Project SQG Generation Rate
CO	6000 kg	180 or 270 (if >200 mi from TSDF	100 - 1000 kg / month
CT	1000 kg	180 days	100 - 1000 kg / month
MA	6000 kg	180 days	100 - 1000 kg / month
ME	3000 kg	90 days	100 - 1000 kg / month
NH	NA	90 days	100 - 1000 kg / month
NY	6000 kg	180 days	100 - 1000 kg / month
RI	NA	90 days	100 - 1000 kg / month
VT	6000 kg	180 days	100 - 1000 kg / month

Accumulation Time Limit:

- Three states [ME, NH, RI] had stricter accumulation time limits
    - ME and NH had the two lowest SQG achievement rates for indicator 7: accumulation time limits followed.
    - CT's SQG achievement rate was statistically significantly higher than NH's and ME's SQG achievement rates.
    - CO's SQG achievement rate was statistically significantly higher than ME's SQG achievement rate.
    - RI has an observed SQG achievement rate of 100% on indicator 7: accumulation time limits followed.

- **Based on the mixed results, this area may warrant further evaluation.**

Accumulation Quantity Limit:

- Two states [NH, RI] had no accumulation quantity limits.
- Two states [CT, ME] had lower quantity limits than four other states [CO, MA, NY, VT].

    - ME, CT and CO all had an observed achievement rate of 100%.
    - Both CO and ME were statistically significantly higher than VT.

    - **There does not appear to be a relationship between the accumulation quantity limits and measured SQG performance**

2) Did Who Conducted the Field Observations for the Common Measures Project Create Any Bias in the Results?

- States that reported using hazardous waste staff to conduct Common Measures Project inspections: CO, MA, NH, NY, VT.
- States that reported not using hazardous waste staff to conduct Common Measures Project inspections: CT, ME, RI.

    - RI scored the highest mean SQG facility score on both regulatory and beyond compliance indicators.
    - CT scored the 2nd highest mean SQG facility score on regulatory indicators and the 2nd lowest mean SQG facility score on beyond compliance indicators.
    - ME scored the third highest mean SQG facility score on regulatory indicators and the second highest mean SQG facility score on beyond compliance indicators.

    - **The observed SQG performance differences on regulatory indicators between the three states that did not use hazardous waste staff, and the five states that did use hazardous waste staff, may be due in part to differences in background of the field observer.**

4.6 NEXT STEPS

The project states had great interest in furthering this analysis by:

- Developing a model for root cause analysis methods/techniques and training state participants on the use of those techniques.
- Sharing compliance assurance strategies that appear to be most effective in improving performance results in the SQG sector.
- Developing beyond compliance indicators in energy conservation, water conservation, pollution prevention and/or recycling suitable for application in a variety of environmental programs.

This led project states to prepare and submit an FY 2009 State Innovation Grant Program application for a Common Measures Project 2. Unfortunately this project was not funded. States are currently exploring alternative mechanisms for furthering this evaluation work.

5. THE AUTO BODY SECTOR

This Section of the Report Covers:

Work on the auto body sector was completed through Phase 3 – applying the statistical and data quality assurance procedures to sample selection and data collection. The Project Management Team did not collect and/or analyze field observation data for the auto body group. As noted previously, the States Common Measures Project grant commitment was to complete measurement on at least one group and the states decided to measure the performance of Small Quantity Generators of Hazardous Waste (SQG). In addition, the project states decided that to the extent that there was sufficient time during the grant period and a state had the capacity to take on an additional measurement project, project states could choose to work on the auto body sector as well. New York and Washington State completed a portion of their planned auto body inspections.

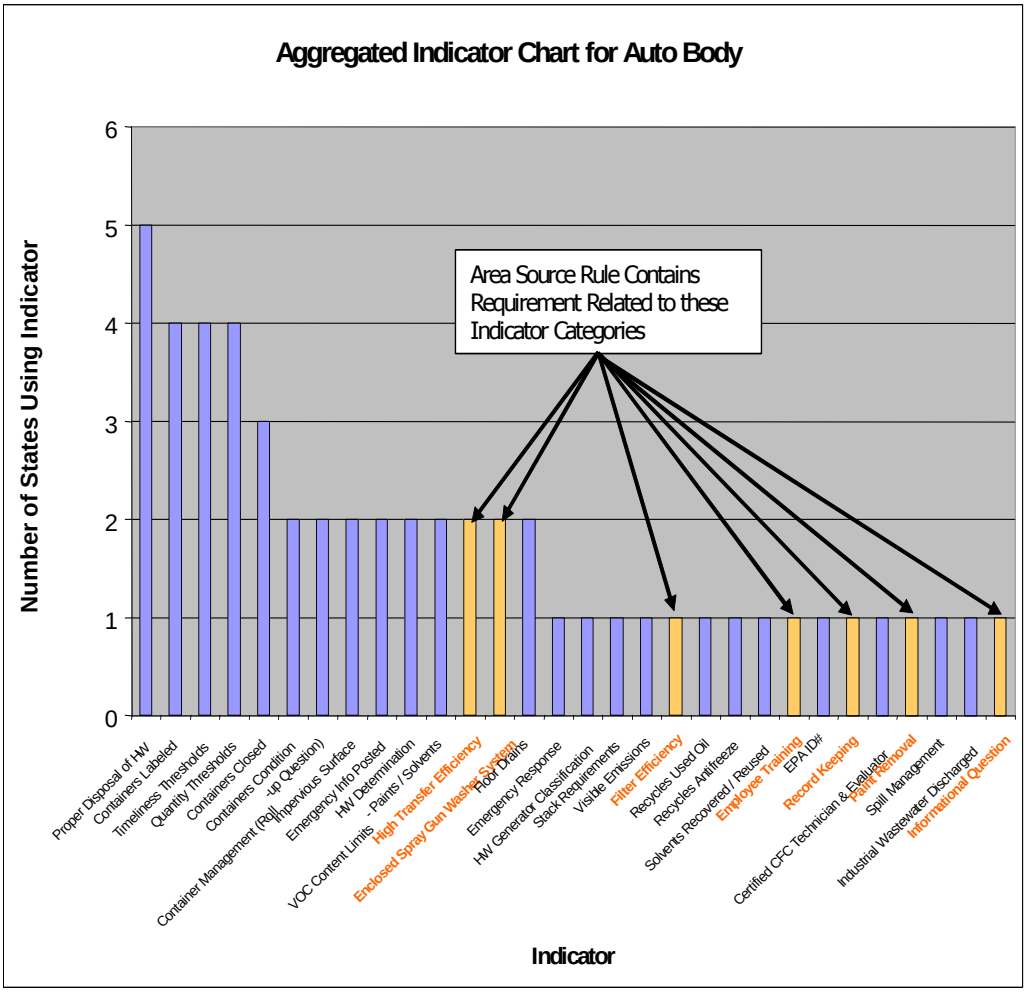
5.1 COMMON DEFINITION FOR THE AUTO BODY SECTOR

On October 23, 2007, a workshop was held for project states to review and finalize the definition of the auto body group and to develop performance indicators to measure the group. As a result of the meeting, the auto body sector was defined as any commercial or academic motor vehicle operation involving collision repair, vehicle painting, paint stripping or sanding, body work, antique restoration, or student training on any on these areas, where the work is performed inside a building or structure.

5.2 COMMON PERFORMANCE INDICATORS FOR THE AUTO BODY SECTOR

During the October 23, 2007 meeting states also reviewed a draft list of auto body indicators developed from the summer work assignment. See Section 2 of the report for summer work assignments. Using this list, states discussed the desired degree of alignment to the proposed EPA paint striping and miscellaneous coatings area source rule.

Overlay of Area Source Rule Requirements and State ERP Indicators for Auto Body Shops



The states agreed to use many of the new area source air indicators for the Common Measures Project. States also agreed to use the hazardous waste indicators that were

previously selected for the Common Measures SQG sector, as well as to develop industrial waste water and pollution prevention indicators.

A conference call was held on November 20, 2007 to further discuss the auto body indicators. Project states and EPA were asked to review draft language prepared by the Project Management Team and approve, reject or modify the draft indicators. Below is an example of how the data was collected:

DRAFT AIR INDICATORS (from AREA SOURCE RULE):	
1. Are all spray-applied coatings applied using an HVLP spray gun or an equivalent high transfer efficiency technology?	
NH	OK
NY	OK
RI	OK
VT	OK
WA	OK
EPA	OK
2. Have the painters been properly trained in high transfer efficiency within the last five years?	
NH	Suggests waiting to develop training question until Area Source Rule released to know what is and is not going to be required.
NY	Training not yet a regulatory requirement in NY. Suggests question be a BMP with space to provide specific type of training.
RI	Training is not a regulatory requirement in RI.
VT	Suggests adding the word high transfer efficiency "**Spraying"** within the last five years.
WA	I think we should define the word "properly" by describing in detail exactly what should be included in the training (as EPA elaborates on below) We don't think it should be required until the date that the NESHAP or our state specifies that it's required. We want to include it in the checklist now, but with a date in the future that they will need to comply. This way we don't need to rewrite our checklist later and the shops know what is coming.
EPA	Training consists of hands-on and classroom training that includes initial and refresher training in the following: Surface preparation; spray gun setup and operation for different types of coatings to improve transfer efficiency and minimize coating usage and overspray; routine spray booth and filter maintenance; safety precautions; and environmental compliance.

As a result of this discussion, draft auto body indicators were revised over several months and final auto body indicators were issued in February 2008. *See Exhibit 5.1 in Section 5.4 for the list of Auto Body indicators.*

5.3 DETERMINING A REASONABLE SAMPLE SIZE FOR DRAWING STATISTICAL CONCLUSIONS ABOUT THE AUTO BODY GROUP

5.3.1 Sample Size Needed to Benchmark an Individual State's Sector Performance

The Massachusetts ERP "sample-size calculator" was used to calculate the sample sizes that would be required to benchmark each state's performance at various confidence levels, confidence intervals and assumptions about the observed compliance rates. The chart below shows the results of the analysis.

Sample Sizes for Benchmarking Performance		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Confidence Level: *% certainty that the observed result reflects actual conditions/ is not due to chance*		***85***	***85***	***90***	***90***	***90%***	***95%***	***95%***	***95%***	***95%***	***95%***
Confidence Interval/Margin of Error: *the actual percentage of facilities in compliance falls somewhere within + or - the listed percentage points of the observed percent compliance*		***+/- 5%***	***+/- 10%***	***+ / - 10%***	***+ / - 10%***	***+ / - 5%***	***+ / - 15%***	***+ / - 10%***	***+/- 7%***	***+ / - 5%***	***+ / - 5%***
Observed "Good Performance" Rate of the Sample		***80%***	***50%***	***50%***	***80%***	***80%***	***50%***	***80%***	***90%***	***50%***	***80%***
State	***Auto Body Universe Size***										
NH	*500 –750 (used 500)*	135	47	**60**	40	129	39	55	62	217	165
VT	*235*	85	42	**53**	37	100	36	49	54	146	120
WA	*1096 – 1514 (used 1096)*	118	49	**64**	42	150	41	58	66	284	201
RI	*336*	95	45	**56**	38	114	38	52	58	179	142
NY	*900 (Regions 4 and 9)*	116	49	**63**	41	145	41	58	65	269	193
NY	*6000 (entire state)*	130	51	**67**	43	168	42	61	70	361	236

The highlighted columns show the level of certainty (confidence level) and precision (confidence interval) that can be achieved for a "realistic" number of inspections per state (e.g., between 36 – 70 inspections):

- Column 6 presents the lowest range of inspections per state at the most conservative observed "good performance" level (50%): a minimum of 36 for Vermont, the state with the smallest universe and a maximum of 42 inspections for New York, the state with the largest. This number of inspections would allow a state to be 95% confident that the actual result has a margin of error of +/- 15% of the observed result, however this 30% confidence interval may be a little too broad for some states. Alternatively, as shown in column 4, with just one more inspection per state, and IF the observed performance level turned out

to be 80%, a state would be able to say that it was 90% confident that the actual result was + /- 10% of the observed result.

- Increasing to 53 inspections in Vermont and to 63 inspections in Regions 4 and 9 in New York as shown in column 3, would allow them to say they were 90% certain that the actual performance of the universe was + or – 10% of the observed result if the observed performance was 50%. If the performance level was above 50%, the confidence interval would be narrower and/or the confidence level would be higher. See column 8 which shows that with one or two more inspections per state, if the observed a good performance rate was 90% states would be able to say they were 95% certain that the actual performance of the sector was + /- 7% of the observed result.
- Notice that regardless of the confidence level or observed performance rate, states would need sample sizes of at least 85 (for Vermont), 95 for New Hampshire, and over 100 for all other states to obtain a confidence interval of + / – 5%, regardless of the confidence level and observed performance.

Decision about sample size for benchmarking performance

In order to benchmark performance in an individual state, a minimum sample size of between 53 – 67 inspections for each state would be needed, depending on the state's universe size. Inspecting this number would benchmark the performance of each state with a minimum level of precision (+ or – 10%) and with a reasonable level of confidence (90%) and assuming 50% observed compliance rates on each indicator.

In addition to benchmarking an individual state's performance, sample size estimates were calculated for project states that wished to compare auto body performance results between states.

5.3.2 Sample Size Needed to Compare Performance Levels between States

The chart below shows the sample sizes needed for various assumptions about confidence level, compliance rates, power and the magnitude of the differences states want to detect. The cells with between 34 – 65 inspections are highlighted. Note that unless states inspect more than 80 facilities each, the highest difference that could realistically be detected is 15%, and that would only be if both states had relatively high compliance rates.

Sample size needed to detect whether a given-sized difference in performance level between two states is statistically significant

		Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8
Confidence Level: *% certainty that a difference of the size listed below is not due to chance*		***90%***	***90%***	***90%***	***90%***	***95%***	***95%***	***95%***	***95%***
Power: % *certainty that a smaller difference than the given difference IS due to chance (in other words that you are not missing a true difference)*		***80%***	***80%***	***90%***	***90%***	***80%***	***80%***	***90%***	***90%***
estimated compliance rate of the universe	***State A***	***50%***	***90%***	***50%***	***90%***	***90%***	***70%***	***50%***	***70%***
	State B	***50%***	***70%***	***50%***	***70%***	***70%***	***70%***	***50%***	***70%***
# of facilities needed to determine that a finding that State A's performance is **20 percentage points higher** than State B's performance is not due to chance		56	34	82	49	46	65	107	95
# of facilities needed to determine that a finding that State A's performance is **15 percentage points higher** than State B's performance is not due to chance		100	60	146	88	82	115	190	137
# of facilities needed to determine that a finding that State A's performance is **10 percentage points higher** than State B's performance is not due to chance		225	135	328	197	185	260	428	309

Decision on sample size for comparing performance across states

The Project Quality Assurance Officer recommended that **each state inspect at least 56 facilities** for the auto body sector. By inspecting at least 56 auto body facilities per state, the results would provide sufficient precision at a 90% confidence level, as well as allow the project to say that, for example, a 20% difference between two states was statistically significant and as such is not due to chance and may be due to differences in state programs.

5.4 UNIVERSE IDENTIFICATION AND RANDOM SAMPLE SELECTION RECORD

On November 6, 2007, the Project Management Team conducted a conference call to discuss the auto body universe identification methodologies proposed by participating states. The result of the call and subsequent discussions was the establishment of a performance standard for universe identification that included each state's approach to identifying their universe. The approach for universe identification and random sample

selection used for the auto body sector was similar to the approach used in SQG sector. *See Section 3.2 of this report.*

STATE		
	METHODOLOGY	**DATE PERFORMED**
UNIVERSE IDENTIFICATION	*Check all that apply:* _____Use of phone books _____Web-based searches including yellowpages.com _____Manifest System / Review of hazardous waste shipment data _____List from OSHA _____List from Auto Body Associations _____List from Dept. of Business Regulation's Licensee Program _____List from Department of Motor Vehicles _____ Info USA (please list SIC/NIAC Codes): _____ _____ Dunn and Brad Street (please list SIC/NIAC Codes): _____ _____ Other Electronic Business Databases (please list): _____Other Method (please describe):	
RANDOM SAMPLE SELECTION	Pick One: _____ ***State Common Measures Project: Methodology for Generating A Random Sample Power Point, June 8, 2007*** _____ ***Alternative Method (describe):***	
PROJECT LEAD		

A Common Measures Auto Body Performance Checklist was developed using the final auto body indicators that were selected by states. On January 31, 2008, the Project Management Team facilitated a data collection training workshop to review each indicator and to agree upon procedures and decision rules for determining whether or not the facility was in conformance with the indicator. *See Exhibit 5.1 below for the States Common Measures Project Auto Body Performance Checklist:*

Exhibit 5.1: States Common Measures Project Auto Body Performance Checklist

ANSWER ALL QUESTIONS UNLESS SPECIFICALLY DIRECTED TO SKIP A QUESTION.

Answer all yes/no questions even if the indicator is not a requirement in your state**. In such cases, please consider the indicator a Best Management Practice and evaluate whether the facility is engaged in the practice.**

If there are any doubts about verification of conformance with an indicator, inspector should discuss it with state project lead.

Date of Visit: _____Agency/Regional Office: _____

Field Observer:_____Tel._____

Facility Name:_____

Facility Address:_____

Name of Contact Person:
_____

Telephone number of Contact Person:
_____

Generator ID Number: _____ SIC:_____

AIR INDICATORS

		Y	N
1.	**Are all spray-applied coatings applied using an HVLP spray gun or an equivalent high transfer efficiency technology?**		

Visual inspection of spray gun for HVLP mark. If no HVLP mark on spray gun, inspector will ask facility for any documentation that demonstrates spray gun achieves the transfer efficiency of an HVLP spray gun.

If no documentation is available, inspector will collect spray gun manufacturer's name:

_____
and model number_____

Note: If flow is 15-26 cubic feet per minute and PSI at orifice is less than 10 lb per square inch, then likely to be an HVLP spray gun

Other equivalent high transfer efficiency technology examples include electrostatic application, airless spray gun, air assisted airless guns.

2.	**Does the facility have a high transfer efficiency painting training in place?** **IF YES,** Check all that Apply: ____Surface preparation; ____Spray gun setup and operation for different types of coatings to improve transfer efficiency and minimize coating usage and overspray; ____Routine spray booth and filter maintenance; ____Safety precautions; ____Environmental compliance; ____Other: (please describe) _____ *NOTE: Examples of training can include, but is not limited to, hands-on and/or classroom training. Training can also consist of initial and/or refresher courses.* **IF NO, SKIP** to Question 3		
2a.	**If yes, is this training documented?** Inspector will ask to see documentation. *NOTE: Since many states have not implemented this as a requirement, the compliance verification will be to determine if the facility is following the practice of "documenting." If the facility produces documentation, then the requirement is met. We will not evaluate the "quality" of documentation.*	**Y**	**N**
3.	**Does the facility use ventilated sander (dustless vacuum) equipment that captures paint dust and body filler, or an overhead capture system?** Inspector will perform visual inspection. *NOTE: Equipment should be easy to identify. Vacuum sanders have vacuum hose hooked up to sander head.*	**Y**	**N**
4.	**When sanding, does the facility keep the shop doors closed to avoid releasing dust outdoors?** Inspector will perform visual inspection of doors when shop is sanding. If no sanding activity is done during site visit, inspector will ask facility if they keep the shop doors open when sanding.	**Y**	**N**

5.	**Are all spray-applied coatings applied in an enclosed ventilated spray booth or preparation station?** Inspector will perform visual inspection to confirm that spray booths and prep stations, where coatings are applied for **full vehicles**, are fully enclosed with a full roof, **four complete walls** or side curtains and an exhaust fan. Spray booths or prep stations where coatings are applied on **vehicle components only**, i.e. not full vehicle, must be fully enclosed with a full roof and at least **three complete walls** or side curtains and an exhaust fan. **IF NO, SKIP** to Question 6	**Y**	**N**	
5a.	**If yes, is the spray booth and/or station fitted with a particle filter(s) on the exhaust?** Inspector will perform visual inspection of pipe and filter as well as ask the facility if station is fitted with a particle filter(s) on the exhaust. *NOTE: May see wall fan and small box with filter*.	**Y**	**N**	
6.	**If the facility uses a spray booth or prepstation, is it fitted with a type of filter technology or system that has been demonstrated to achieve at least 98-percent capture of paint overspray (this could include polyester fiber or fiberglass filters)?** Inspector will ask for any documentation, e.g. on filter package or from distributor, of the filter efficiency. *NOTE: Visual inspection alone may be too difficult to determine compliance.*	**Y**	**N**	**NA**
7.	**Is all paint spray gun cleaning done with a fully enclosed spray gun washer or in a manner that avoids creating an atomized mist or spray of gun cleaning solvent?** Inspector will ask facility to demonstrate how guns are cleaned. Acceptable methods of spray gun cleaning include: Hand cleaning of parts of the disassembled gun in a container of solvent; flushing solvent through the gun without atomizing the solvent and paint residue, by using a fully enclosed spray gun washer, or by a combination of these non-atomizing methods. *NOTE: Spraying into the air is an unacceptable cleaning method. Also, any waste solvents that are collected must be kept in a closed container to avoid release/evaporation to the air.*	**Y**	**N**	

Record Keeping Indicators

NOTE: ANSWER ALL YES/NO RECORD KEEPING QUESTIONS EVEN IF NOT REQUIRED IN YOUR STATE. IF ANY OF THE RECORDS ARE NOT REQUIRED, CONSIDER THEM A BEST MANAGEMENT PRACTICE AND EVALUATE WHETHER THE FACILITY IS ENGAGED IN THE PRACTICE. THE EXCEPTION IS IF THE INDICATOR HAS AN "NA" (NOT APPLICABLE) OPTION TO CIRCLE.

8	**Does the facility have MSDS or formulation data supplied by manufacturer for all the solvents and coatings that they use?** Inspector will ask to see documentation. If no physical documentation is available, inspector will confirm if facility uses a color matching computer technology on-site that has formulation data in it, and if so, ask to see data.	**Y**	**N**	
9.	**Does the facility have documentation of the amount of coatings used that contain chromium, lead, cadmium, nickel, and manganese (especially hexavelent chromium, most common in corrosion control undercoats and red, orange, and yellow paint colors) and the metals content of these coatings?** Inspector will ask to see documentation. *NOTE: Some facilities do not use coatings that contain chromium, lead, cadmium, nickel, and manganese. In this case, please circle NA.*	**Y**	**N**	**NA**
10.	**Does the facility use paint strippers containing Methylene Chloride (MeCl)?** Inspector will ask facility. **IF NO, SKIP** to Question 11	**Y**	**N**	
10a.	**If yes, does the facility keep records to document annual usage?** Inspector will ask to see documentation.	**Y**	**N**	
10b.	**If the facility uses MeCI for paint stripping, is there a written MeCl minimization plan?** Inspector will ask to see minimization plan.	**Y**	**N**	

HAZARDOUS WASTE INDICATORS

NOTE: ANSWER ALL YES/NO RCRA QUESTIONS EVEN IF NOT REQUIRED IN YOUR STATE. IF ANY OF THE RCRA QUESTIONS ARE NOT REQUIRED, CONSIDER THEM A BEST MANAGEMENT PRACTICE AND EVALUATE WHETHER THE FACILITY IS ENGAGED IN THE PRACTICE. THE EXCEPTION IS IF THE INDICATOR HAS AN "NA" (NOT APPLICABLE) OPTION TO CIRCLE.

ALL QUESTIONS APPLY TO RCRA HAZARDOUS WASTE ONLY. THEY DO NOTAPPLY TO "UNIVERSAL WASTES" SUCH AS COMPUTERS OR FLOURESCENT BULBS OR TO STATE- ONLY WASTES

11.	**What is the maximum amount of RCRA hazardous waste the facility generates in a month _____** ***NOTE: Includes satellite accumulation***	
12.	**What is the facility's hazardous waste generator status under state's classification system, (e.g., CESQG, SQG, LQG):**	
	<u>Container Management Indicators</u> *NOTE: Container management questions apply to central accumulation areas only. They do not apply to satellite areas or laboratories*	
13.	**Are all hazardous waste containers properly labeled with the words "hazardous waste" and clearly marked with the date on which accumulation began?** Inspector will determine if all containers have labels; if labels are marked with both items; if labels are clear and legible *NOTE: inspector will use best professional judgment to determine what is "clear" and "legible" and whether the facility made a "genuine and complete effort" to meet this requirement.* *Applies to central accumulation areas only, not laboratories or satellite areas.*	**Y N**
14.	**Are all hazardous waste containers closed unless waste is being added or removed?** Inspector will confirm that all containers are closed at the time of inspection unless waste was being added or removed. *NOTE: "closed" means if the containers were tipped, nothing would spill "Funnels" are acceptable if they are closed.* *Applies to central accumulation areas only, not laboratories or satellite areas*	**Y N**
15.	**Are all hazardous waste containers in good condition, (i.e., free of severe rusting or apparent structural defects, and not leaking)?** Inspector will perform visual inspection of conditions of all containers looking for leaks and/or severe corrosion, bulging, rusting or dents. *NOTE: inspector will use best professional judgment to determine what is "severe." There should be no imminent threat.* *Applies to central accumulation areas only, not laboratories or satellite areas*	**Y N**

	Proper Hazardous Waste Management Indicators			
16.	**At the time of inspection, has the facility exceeded the state's time limits for the amount of RCRA hazardous waste that can be stored on-site by this category generator (excludes satellite accumulation)?** As applicable, inspector will verify conformance based on dates on containers which detail when accumulation begins (if no containers are labeled, then inspector will circle: **no labels**. *NOTE: Some states have no time limits depending on the generator category. If this is the case circle NA.*	**Y**	**N**	**NA**
17.	**At the time of the inspection, has the facility exceeded the state's accumulation limits for hazardous waste for this category of generator (excludes satellite accumulation)?** As applicable, inspector will inventory all hazardous waste accumulated on-site in containers and tanks to determine total weight of waste being accumulated at one time. *NOTE: Some states have no accumulation limits. If this is the case circle NA.*	**Y**	**N**	**NA**
18.	**Is a manifest required for this type of facility to ship hazardous waste?** **IF NO, SKIP** to Question 18b	**Y**	**N**	
18a.	**If yes, does the facility use a hazardous waste manifest to ship its hazardous waste when a manifest is required?** Inspector will <u>look</u> at one (1) year of manifest records as well as <u>ask</u> the facility if they have kept three (3) years of records. If there are gaps in the shipments or shipping records, the inspector will ask facility to explain gaps. Inspector will use professional judgment to determine if explanation of gaps is valid. *NOTE: The word "look" in this case means that the inspector will confirm that all shipments have been manifested and not that each and every manifest was filled out correctly.*	**Y**	**N**	
18b.	**If a hazardous waste manifest is not required, does the facility document its hazardous waste shipments, e.g. non-hazardous manifest, bill of lading, other documentation?** Inspector will ask to see documentation. *NOTE: Circle NA for facilities that use a hazardous waste manifest.*	**Y**	**N**	**NA**

19.	**Has the facility identified all of its hazardous waste streams?** Conformance will be determined based on: review of production processes, type of wastes generated at these processes and whether or not the facility has made hazardous waste determinations on all waste streams. *NOTE: This is something inspectors do routinely. Interns may need to bring back process or other information to discuss with the state project lead* in order to answer this question.	**Y**	**N**
	Emergency Response Indicators		
20.	**Is the facility required to have emergency procedures?** **IF NO, SKIP** to Question 20b	**Y**	**N**
20a.	**If yes, is the facility in compliance with the applicable requirements for emergency procedures for this category of generator?** Compliance is based on individual state requirements for this category of generator.	**Y**	**N**
20b.	**If emergency procedures are not required, does the facility have emergency procedures in place?** Visual inspection and inquiry into whether the information is up to date. NOTE: Emergency procedures can include: The facility posting the current name and telephone number of the emergency coordinator; Posting the location of fire extinguishers and spill control material, and if present, fire alarm; Posting the telephone number of the fire department, unless the facility has a direct alarm.	**Y**	**N**
21.	**Does the facility have an employee training program that teaches employees proper hazardous waste management procedures?** Inspector will ask facility to describe program as well as ask to see any documentation. The compliance verification will be to determine if the facility is following the practice of having an employee training program. If the facility demonstrates its existence, then the requirement is met. We will not evaluate the "quality" of the training program.	**Y**	**N**
22.	**Is there any indication of spills in or near the shop?** Inspector will check for stains on the ground and in and around manholes, leaking tanks and containers and/or pooled liquids.	**Y**	**N**

	INDUSTRIAL WASTEWATER INDICATORS (IWW) *NOTE: The following questions are related to industrial wasterwater discharges only, i.e. process wastewater, and not sanitary wastewater.*			
23.	**Does the facility discharge industrial wastewater to surface water?** **IF NO, SKIP** to Question 24	**Y**	**N**	
23a.	**If yes, is facility in compliance with the applicable requirements for the discharge(s)?** Based on individual state requirements for the discharge(s).	**Y**	**N**	**NA**
24.	**Does the facility discharge industrial wastewater to a sewer system?** Sewer system includes sanitary, stormwater or combined sewers **IF NO, SKIP** to Question 25	**Y**	**N**	
24a.	**If yes, is facility in compliance with the applicable requirements for the discharge(s)?** Based on individual state requirements for the discharge(s).	**Y**	**N**	**NA**
25.	**Does the facility discharge industrial wastewater to groundwater (e.g., discharge to an on-site septic system, drywell, etc.)?** **IF NO, SKIP** to Question 26	**Y**	**N**	
25a.	**If yes, is facility in compliance with the applicable requirements for the discharge(s)**	**Y**	**N**	**NA**
26.	**Does the facility have any unsealed floor drains?** Inspector will perform visual inspection of floor drains to see if sealed. If no floor drains, answer no. **IF NO, SKIP** to Question 27	**Y**	**N**	
26a.	**If yes, is the facility in compliance with the state standard for discharges to unsealed floor drains?** Based on individual state requirements for the discharge(s)	**Y**	**N**	**NA**
	POLLUTION PREVENTION INDICATORS			

27.	**Has the facility taken one or more actions to conserve water the past three years?** Inspector will ask the facility manager to describe any initiatives in the past three years. **IF NO, SKIP** to Question 28	**Y**	**N**
27a.	**IF YES**, Briefly Describe The Water Conservation Projects:		
28.	**Has facility taken one or more actions to conserve energy over the past three years?** Inspector will ask the facility manager to describe any initiatives in the past three years. **IF NO, SKIP** to Question 29	**Y**	**N**
28a.	**IF YES**, Briefly Describe The Energy Conservation/Alternative Energy Projects:		
29.	**Has the facility taken one or more actions to reduce toxics the past three years? Check all that apply:** ____ Use water-based or low-solvent coatings (primers, basecoats and painting)? ____ Attempt to avoid use of coatings that contain toxic metals (chromium, lead, cadmium, nickel, and manganese) by asking suppliers for alternative formulations? _____ Avoid use of methylene-chloride based paint strippers? _____ Recycle any solvents? _____ Use recycled solvent for gun cleaning? _____ Have an inventory system in place to prevent products from going out of date? ____non-solvent based putty/fillers ____Other **IF NO, STOP. Performance Checklist is Complete.**	**Y**	**N**
29a.	**IF YES**, Briefly Describe The Toxic Use Reduction Projects:		

5.5 Auto Body Sector Conclusion

The States Common Measures Project successfully developed and implemented common measurement protocol including sector definition, indicators, and a checklist for a complex sector, that lacked a uniform federal program, was multi media, and involved many small sources. Other states and federal projects have adopted the indicators developed under this project to measure performance of the auto body sector.

This success demonstrates that the common measurement approach has very wide application across the many and varied Federal, state, and local environmental programs.

6. THE PROJECT CONCLUSION AND RECOMMENDATIONS

Over a three year period, the ten project states were able to use the same set of common measures to evaluate the environmental performance of a common group of facilities. The project also created a replicable template that can be used by other agencies to build the capacity to measure group performance and to use the information to identify the most efficient and effective strategies for promoting better environmental performance.

This effort has already paved the way for other ERP-type measurement projects. Current activities being considered or under development include:

- The same project states selecting additional sectors to analyze
- The development of a second States Common Measures Project proposal which would build on the work of this project to do more in depth analysis of the relationship between program design and high SQG performance. It would also create more robust energy efficiency, pollution prevention, solid waste recycling, and water conservation beyond compliance performance indicators.
- A six-state initiative in EPA Region V to develop and implement a region-wide ERP for auto body shops that will include the use of the indicators developed under this project. The six-state initiative is funded through an EPA State Innovation Grant awarded in the spring of 2009.
- An EPA Region 1 and EPA Office of Enforcement and Compliance Assurance initiative that uses the indicators developed by this project to measure auto body performance in Massachusetts.

Achieving the full benefits of the States Common Measures Project requires the widespread adoption of ERP-type measurement across environmental agencies. The challenge going forward is to take meaningful steps to capitalize on the potential created by this project. The project states recommend that EPA provide the key leadership and financial support needed to:

1. Promote and expand the use of ERP-type measurement in both "core" and other work in states and EPA to:

    - Look within and beyond individual states to identify and adopt the most effective and efficient environmental performance improvement strategies.
    - Allow states the flexibility to deploy resources based on measured performance.
    - Promote the use of ERP-type measurement to routinely make environmental program priority and resource allocation decisions.

2. Support the development of an ERP Training Institute to codify this work into a formal ERP measurement curriculum.

APPENDICES

The Common Measures Project Appendices listed below are available at:

http://www.newmoa.org/hazardouswaste/measures/index.cfm

A. Quality Assurance Project Plan (QAPP)
B. EPA Quarterly Reports
C. Introductory Training and Quality Measurement Slides
D. Individual State Responses to Group Preference Checklists
E. Aggregated Data on Group and Indicator Evaluation Charts
F. Data Quality Considerations When Selecting Groups
G. Complete List of States' SQG Indicators by Indicator Category
H. How to Generate a Random Sample
I. Training Attendance Log
J. Quality Assurance Procedures
K. Confidence Intervals for each State's SQG Achievement Rate and Mean Facility Score at the 85%, 90%, and 95% Levels
L. Statistically Significant Differences in each State's SQG Achievement Rate and Mean Facility Scores at 85%, 90%, and 95% Confidence Levels
M. Histograms for each State's Facility Scores for "All," "Regulatory" and "Beyond Compliance" Indicator Groups