CONNECTED VEHICLE PILOT
Deployment Program

Preparing an Effective Performance Measurement Plan

John Halkias, Systems Management Team Leader, FHWA

ITS Joint Program Office

Agenda

- Purpose of this Technical Assistance Webinar Series
  - To assist early deployers of connected vehicle technologies to conduct Concept Development activities

- Webinar Content
  - Performance Measurement Concepts, Challenges and Potential Solutions
  - Stakeholder Q&A
  - How to Stay Connected

- Webinar Protocol
  - Please mute your phone during the entire webinar
  - You are welcome to ask questions via chatbox at the Q&A Section
  - The webinar will be recorded except the Q&A Section
  - The webinar recording and the presentation material will be posted on the CV Pilots website within a week

CV Pilot Deployment Program Goals

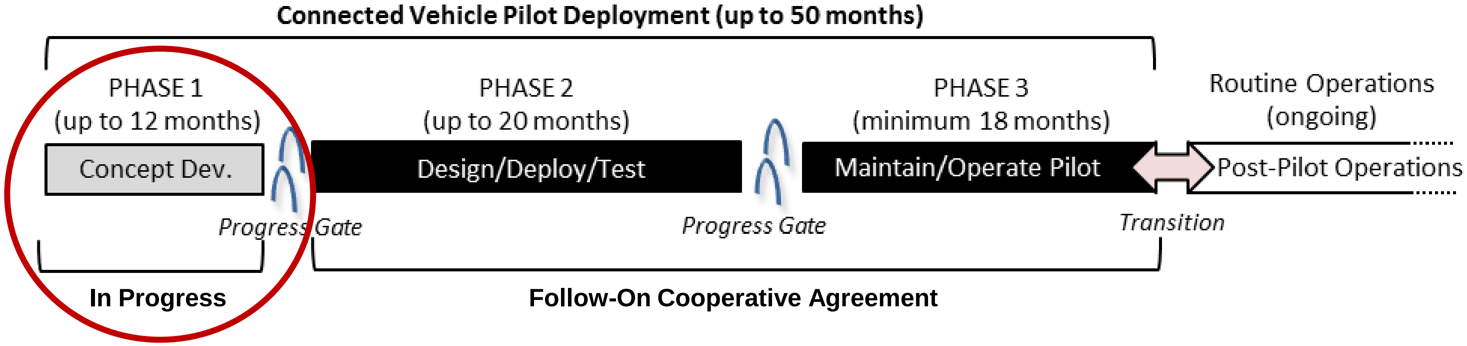
Sites Selected – 2015 Awards

- Reduce the number and severity of adverse weather-related incidents in the I-80 Corridor in order to improve safety and reduce incident-related delays.
- Focused on the needs of commercial vehicle operators in the State of Wyoming.

- Improve safety and mobility of travelers in New York City through connected vehicle technologies.
- Vehicle to vehicle (V2V) technology installed in up to 10,000 vehicles in Midtown Manhattan, and vehicle to infrastructure (V2I) technology installed along high-accident rate arterials in Manhattan and Central Brooklyn.

- Alleviate congestion and improve safety during morning commuting hours.
- Deploy a variety of connected vehicle technologies on and in the vicinity of reversible express lanes and three major arterials in downtown Tampa to solve the transportation challenges.

Deployment Schedule



- Overall Deployment Schedule
  - Phase 1: Concept Development
    - Creates the foundational plan to enable further design and deployment
  - Phase 2: Design/Deploy/Test
    - Detailed design and deployment followed by testing to ensure deployment functions as intended (both technically and institutionally)
  - Phase 3: Maintain/Operate
    - Focus is on assessing the performance of the deployed system
  - Post Pilot Operations (CV tech integrated into operational practice)
- Public webinars to share the concept development activities from the three sites
    - Concept of Operations Webinar (February – March 2016)
    - Performance Measurement Webinar (May – June 2016)
    - Deployment Plan Webinar (August 2016)

Webinar Objectives

- Define and distinguish between performance measurement, system deployment impact evaluation, and independent evaluation
- Establish need for performance measurement and evaluation
- Identify common challenges and issues with performance measurement and evaluation

Performance Measurement

- Performance measurement is a means of assessing the progress made towards attaining established goals
  - Goals can be financial (e.g., cost levels), operational (e.g., reduction in travel time), etc.
  - Broad stakeholder consultation required in establishing appropriate and realistic goals

- Performance measurement isn’t only about collecting data but using the data to understand the system

- Performance measurement is a part of an overall transportation system management

Monitoring and Reporting the Impact of the Deployment

- System deployment impact evaluation conducted by the system deployer is the process of interpreting results to understand the impacts that investments and policies have had on performance
  - Assesses robustness, effectiveness, usability, and acceptance of the application as deployed

Independent Evaluation

- Performance evaluation is conducted by an ***independent party*** who has no vested interest or stake in the project
  - Evaluation will be projected over time and geographic scope, and for varying market adoption rates of application and driver compliance rates

Motivation

- Provides the basis for evaluating the impacts of the deployed system
- Enables an agency to improve its internal operations
- Allows decision-makers to provide accountability for public expenditure
- Enhances the decision-making process for both short-term and long-term transportation investments
- Helps identify the location and severity of problems (e.g., congestion)
- Provides means to inform the travelling public of the effectiveness of deployed system
- Helps to determine how the transportation system is performing with respect to goals, overall vision and adopted policies

Types of Performance Measures

- Quantitative performance measures
  - Provide numerical estimates as an evidence of how a transportation system is performing
  - Enables comparison with established targets to determine progress/regress
  - Usually verifiable and yield similar results for repeated trials (when everything else is kept constant)
  - Can be continuous (e.g., average travel time, average speed, etc.) or discrete (e.g., average vehicle throughput, average person throughput, etc.)

- Qualitative performance measures
  - Represent subjective perceptions and satisfaction levels of users or customers
  - Complement quantitative measures to help improve service
  - Examples: user satisfaction, driver compliance, driver frustration

Key Terminologies

- Control group
  - Group that does not receive any treatment
  - Serves as a comparison point to evaluate the magnitude and significance of treatment
- Treatment group
  - Group that receives the treatment or the intervention
  - Exposed to the application or strategy being tested
  - Similar to the control group with respect to all factors except the treatment
- Confounding Factors
  - Variable(s) that completely or partially accounts for the apparent association between an outcome and a treatment
  - Variable(s) other than the independent variables of interest
  - External to experiment; hence, not monitored
  - May result in erroneous conclusions on the impacts of treatment

Common Issues and Solutions

- Performance measures of convenience, not of value
  - Choosing the wrong measures can result in distorted view of the system
  - *Solution*: Use performance measures that can help to gain insights

- No agreement on what “good” looks like
  - Stakeholders not agreeing on what good performance targets to achieve
  - *Solution*: Stakeholders need to reach an agreement or consensus on the measures and targets.

- Stovepipe approach to data collection
  - Collecting different types of data (e.g., arterial/freeway data, weather data) with different time resolutions at different time periods and storing in separate databases without integration
  - *Solution*: A data management process should be developed and multi-source data should be integrated into a common format and database

Common Issues and Solutions (cont.)

- Data gaps
  - Failure in hardware can result in missing data
  - *Solution*: Data gaps can be addressed using simple heuristics and imputation techniques

- Limited data quality verification
  - Limited resources often lead to limited data quality checks and this may result in invalid performance measures
  - *Solution*: Both automated and manual data checks necessary

- Measurement Uncertainty Due to Equipment
  - Wear and tear and environmental conditions may impact the performance of data collection equipment
  - *Solution*: Regular calibration will ensure equipment accuracy. Diagnostic checks should be in place to detect equipment failures.

Common Issues and Solutions (cont.)

- Poor experimental design
  - A poor experimental design can lead to flawed analysis and inaccurate representation of the system performance
  - *Solution*: Experimental design must be carefully pre-planned to account for potential confounding factors

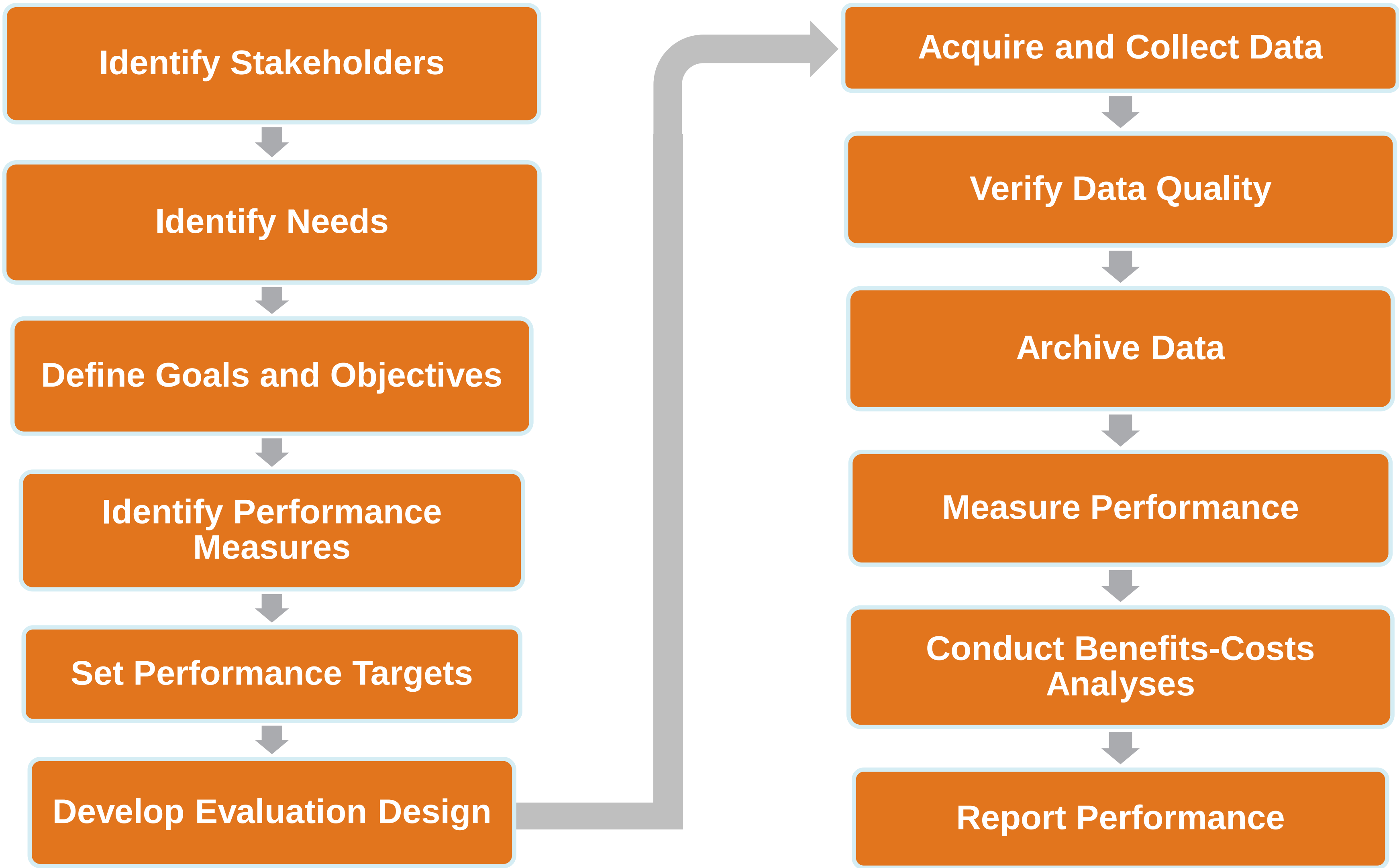
Potential Confounding Factors

- Change in weather or unusual weather events
  - Comparisons should be made between similar (adverse/non-adverse) weather conditions
- Construction/work zone activities during pre/post-deployment periods
  - Will alter traffic conditions and traveler behavior
- Unusually high/low crashes or incidents
- Change in travel demand
  - Significant changes will impact the performance measures
- Change in vehicle mix (e.g., size, in-vehicle technology, personal devices)
  - It may affect underlying traffic conditions and traveler behavior
- Change in truck percentage in the vehicle mix
  - Pilot is truck-centric; significant changes in truck percentage will impact performance measurement

Potential Confounding Factors (cont.)

- Concurrent deployment of synergistic or conflicting non-CV Pilot applications
  - Deployment of similar or conflicting applications should be delayed
- Change in on-time delivery criteria
- Changes in fuel prices and economy
  - This may influence traveler behavior and mode choice
- Change or shift in population
- Self-selection of participants
  - Individuals with special interests or motives may introduce bias
- Participants exploiting the limits of applications
  - Taking unnecessary risks

Process for Performance Measurement and Impact Evaluation



Source: Updated Freeway Management and Operations Handbook (DRAFT)

Performance Measurement Methods

- Field data and user survey data
  - Field data (connected vehicle messages and complementary weather, road sensors, signal control data, etc.) provide accurate estimates of system performance when integrated
  - Survey data helps to obtain user perspectives on transportation system
  - Refer Traffic Analysis Toolbox Volume VI for example techniques
- Analytical tools
  - Useful when measures cannot be observed directly
- Sketch planning tools
  - Useful for general order-of-magnitude estimates without in-depth engineering analysis
- Deterministic tools
  - Uses deterministic and static analytical procedures (e.g., HCM)
- Traffic simulation tools
  - Effective in evaluating dynamics of congestion in transportation systems
  - Capable of modeling variability in driver/vehicle characteristics
  - Capable of effectively controlling for confounding factors

Types of Impact Evaluation Designs

- Non-Experimental Design
- Randomized Experimental Design
- Quasi-Experimental Design

Non-Experimental Design

- Has no control group; hence, the weakest design type
- Involves the repeated measurement of one or more indicators (e.g., average vehicle throughput) over a specified time period
  - Time period may include both pre-deployment and post-deployment
- Impact of deployment is assessed by examining any changes in the post-deployment period given the trend in the pre-deployment period
- Does not account for confounding factors
  - May lead to false conclusions
- Examples: Time Series studies, Before/After studies, longitudinal studies

Randomized Experimental Design

- Study subjects are randomly assigned to the control group and the treatment group
- Randomization ensures that control and treatment groups are equivalent with respect to all factors except the deployment and are unbiased
- Control group serves as the “counterfactual” of what would have happened in the absence of the deployment, which is a key requirement in determining whether a deployment caused a particular outcome
- Classic design uses pre-test/post-test design
  - Ensures control and treatment groups are similar in both pre and post-deployment periods
  - Data for each group are collected for both pre-deployment and post-deployment periods
- Provides the most assurance that outcomes are the result of the deployment

Quasi-Experimental Design

- Approximation of randomized experimental design
  - Uses pre-test/post-test design ***but no random assignment***

- Control and treatment groups cannot be assumed to be similar
  - Agencies must assess the differences during the pre-test and account for the differences in the analysis
  - An assessment of the characteristics of members of the control and treatment groups is conducted during the pre-test period to determine the differences between the two groups

Potential Solutions for Controlling for Confounding Factors

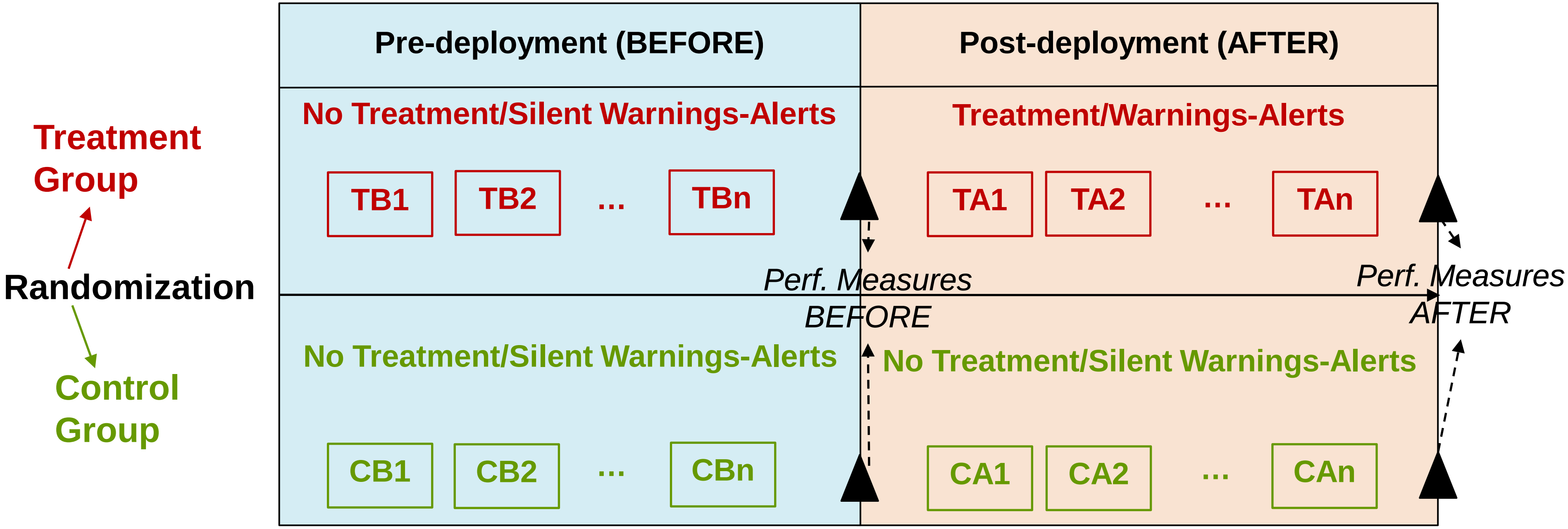
- Randomized experimental design
  - Pros:
    - Uses real-world data (real subjects)
    - Most effective in controlling for confounding factors
  - Cons:
    - Limited ability to provide system-wide measures

Potential Solutions for Controlling for Confounding Factors (cont.)

- Traffic simulation tools
  - Pros:
    - Can ensure consistency in weather, travel demand, travel patterns, vehicle mix, isolate impacts of construction activities, and applications
    - Provides cost-effective approach to calculating performance measures that cannot be easily measured in the field (e.g., queue length, throughput)
    - Provides system-wide measures
  - Cons:
    - Calibration of the models can be time consuming and expensive
    - Lack of driver behavior models in the presence of connected vehicle technology
    - Uses data generated by virtual environment

Potential Impact Evaluation Design

- Use randomized experimental design with treatment and control groups in both BEFORE and AFTER periods



LEGEND

TB1, TB2,...,TBn:	BEFORE data collection/measurement pts for treatment grp
CB1, CB2,...,CBn:	BEFORE data collection/measurement pts for control group
TA1, TA2,...,TAn:	AFTER data collection/measurement pts for treatment group
CA1, CA2,...,CAn:	AFTER data collection/measurement pts for control group

Stakeholder Q&A

- Please keep your phone muted
- Please use chatbox to ask questions
- Questions will be answered in the order in which they were received
- This Q&A section will neither be recorded nor posted to the website

STAY CONNECTED

Contact for CV Pilots Program:

Kate Hartman, Program Manager

Kate.hartman@dot.gov

Join us for the *Getting Ready for Deployment* Series

- Discover more about the 2015 CV Pilot Sites
- Learn the Essential Steps to CV Deployment
- Engage in Technical Discussion

Website: http://www.its.dot.gov/pilots

Twitter: @ITSJPODirector

Facebook: https://www.facebook.com/DOTRITA

February 2016 Webinars

Technical Assistance Webinars

- 2/10/2016, 2:30 – 4:00 pm EST
  SCMS Proof-of-Concept Interface Requirements for Connected Vehicle Deployments

Please visit the CV pilots website for the recording and the briefing material of the previous webinars.