

# Phase 2 Performance Measurement and Evaluation Support Plan (PMESP)

## Buffalo NY ITS4US Deployment Project

[www.its.dot.gov/index.htm](http://www.its.dot.gov/index.htm)

**Phase 2 Draft Update – September 25, 2023**

**Phase 2 As-Built Update -- September 25, 2025**

**Publication Number FHWA-JPO-21-878**



U.S. Department of Transportation



Produced by Niagara Frontier Transportation Authority (NFTA)  
U.S. Department of Transportation  
Intelligent Transportation Systems Joint Program Office  
Federal Highway Administration  
Office of the Assistant Secretary for Research and Technology  
Federal Transit Administration

## Notice

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

The U.S. Government is not endorsing any manufacturers, products, or services cited herein and any trade name that may appear in the work has been included only because it is essential to the contents of the work.

---

<b>1. Report No.</b> FHWA-JPO-21-878		<b>2. Government Accession No.</b>		<b>3. Recipient's Catalog No.</b>	
<b>4. Title and Subtitle</b> Phase 2 Revised Performance Measurement and Evaluation Support Plan (PMESP) – Buffalo NY ITS4US Deployment Project				<b>5. Report Date</b> September 25, 2025	
				<b>6. Performing Organization Code</b>	
<b>7. Author(s)</b> Mark Bradley (RSG), Johanna Zmud (Blue Door SR), Kyeongsu Kim (RSG), Aaiysha Khursheed (RSG), Deepak Gopalakrishna (ICF), Nayel Urena Serulle (ICF), Polly Okunieff (ICF), Adel Sadek (UB), Jordana Maisel (UB), Victor Paquet (UB)				<b>8. Performing Organization Report No.</b> Task 5 – Performance Measurement and Evaluation Support Plan	
<b>9. Performing Organization Name and Address</b> BNMC, 640 Ellicott Street, Buffalo, NY 14203 ICF International, 9300 Lee Highway, Fairfax, VA 22031 University at Buffalo, Amherst, NY 14228 Open Doors Organization, 8600 W. Catalpa Avenue, Chicago, IL 60656 RSG, 55 Railroad Row, Suite 101, White River Junction, VT 05001 ETCH, 4696 Smothers Road, Westerville, OH 43081 NFTA, 181 Ellicott Street, Buffalo, NY 14203				<b>10. Work Unit No. (TRAIS)</b>	
				<b>11. Contract or Grant No.</b> 693JJ321C000005	
<b>12. Sponsoring Agency Name and Address</b> United States Department of Transportation 1200 New Jersey Ave., SE Washington, DC 20590				<b>13. Type of Report and Period Covered</b> Performance Measurement and Evaluation Support Plan	
				<b>14. Sponsoring Agency Code</b>	
<b>15. Supplementary Notes</b> Elina Zlotchenko (USDOT ITS-JPO) is the Contracts Officer Representative (COR) and Sarah Tarpgaard (USDOT) is the Contracting Officer (CO).					
<b>16. Abstract</b> The Buffalo NY ITS4US Deployment Project seeks to improve mobility to, from and within the Buffalo Niagara Medical Campus by deploying new and advanced technologies with a focus on addressing existing mobility and accessibility challenges. Examples of the technologies to be deployed are electric and self-driving shuttles, a trip planning app that is customized for accessible travel, intersections that use tactile and mobile technologies to enable travelers with disabilities navigate intersections, and Smart Infrastructure to support outdoor and indoor wayfinding. The deployment geography includes the 120-acre Medical Campus and surrounding neighborhoods with a focus on three nearby neighborhoods (Fruit Belt, Masten Park, and Allentown) with underserved populations (low income, vision impaired, deaf, or hard of hearing, wheeled mobility device users and older adults).  This document describes the Performance Measurement and Evaluation Support (PMESP) Plan, originally drafted during Phase 1 and now updated during Phase 2 of the Complete Trip Deployment in Buffalo, NY. This PMESP lists the performance measures and targets based on deployment goals and use case scenarios, describes the confounding factors in measurement and proposed mitigation approaches, details the experimental design for each use case, and defines the proposed data collection plan for deployment to ensure the required data is collected for system measurement.					
<b>17. Keywords</b> Deployment, ITS, Intelligent Transportation Systems, Complete Trip, Self-Driving Shuttle, Open Trip Planner, Systems Integration, Performance Measurement, ITS4US.			<b>18. Distribution Statement</b> This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161		
<b>19. Security Classif. (of this report)</b> None		<b>20. Security Classif. (of this page)</b> None		<b>21. No. of Pages</b> 77	<b>22. Price</b> N/A

# Revision History

Name	Date	Version	Summary of Changes	Approver
Buffalo ITS4US Phase 2 Performance Measurement and Evaluation Support Plan	31 July 2023	0.1	Initial Draft	Kelly Dixon
Buffalo ITS4US Phase 2 Performance Measurement and Evaluation Support Plan	1 September 2023	0.2	Revised Draft	Kelly Dixon
Buffalo ITS4US Phase 2 Performance Measurement and Evaluation Support Plan	25 September 2023	0.3	Second Revised Draft	Kelly Dixon
Buffalo ITS4US Phase 2 Performance Measurement and Evaluation Support Plan	September 25, 2025	1.0	Final – As Built	



# Table of Contents

<b>1 Introduction.....</b>	<b>1</b>
1.1 Intended Audiences .....	1
1.2 Project Background .....	2
1.3 Scope.....	3
1.4 Performance Measurement and Evaluation Support Plan Purpose .....	4
<b>2 Goals and Objectives.....</b>	<b>7</b>
2.1 Deployment Goals and Objectives .....	7
2.2 Use Cases .....	8
2.2.1 Complete Trip Platform (CTP) .....	8
2.2.2 Human-Driven and Self-Driving Community Shuttle Service.....	9
2.2.3 Smart Infrastructure .....	10
<b>3 Performance Measurements and Targets .....</b>	<b>11</b>
3.1 Identification of Potential Performance Measures and Targets.....	12
3.1.1 PM #1: Improved ability of the CTP users to make satisfactory Complete Trips in the study area.....	14
3.1.2 PM#2 Usefulness of the CTP Registration and Trip Preferences Processes .....	16
3.1.3 PM#3 Usefulness of the CTP Trip Planning, Booking, Reporting and Review Processes .....	17
3.1.4 PM#4 Improved ability to find destinations efficiently using the CTP wayfinding functionality .....	18
3.1.5 PM#5 Improved ability to cross specific intersections safely using CTP smart signal functionality .....	20
3.1.6 PM#6 Provision of an efficient, reliable and safe new on-demand transit shuttle system	22
3.2 Correspondence between PMs and Use Cases .....	25
3.3 Relationship between Performance Measures and Technologies/Services/Components .....	27
3.4 Potential Constraints.....	28
<b>4 Confounding Factors and Mitigation Approaches.....</b>	<b>31</b>
4.1 Use of Non-Probability Sample Selection.....	31
4.2 Attrition Bias in Participation.....	31
4.3 Influence of Incentives on Travel Behavior .....	32
4.4 Extreme Weather Conditions .....	32
4.5 Progress on the Middle Main Road Improvement Project.....	33
4.6 The Availability of Automated (Driverless) Shuttle Technology that Meets Operating Regulations .....	33

4.7 Other External Factors ..... 34

**5 System Deployment Impact Analysis Design.....37**

5.1 Approach/Strategies for Focused Performance Analysis ..... 37

5.2 Experimental Design..... 42

    5.2.1 Pre-Post-Design with Comparison Group ..... 42

    5.2.2 Target Populations ..... 43

**6 Support to Independent Evaluation Effort .....45**

**7 Data Collection Plan .....47**

7.1 Data Needed ..... 48

7.2 Participant Recruitment Data Collection ..... 49

7.3 Baseline Data Collection ..... 50

7.4 Deployment Data Collection..... 51

7.5 Human Subjects Protection..... 52

7.6 Data Quality Check Approach..... 52

7.7 Data Sharing Framework ..... 53

**8 Performance Reporting .....55**

**9 Performance Measurement and Evaluation Support Schedule.....59**

**Appendix A. References .....63**

**Appendix B. Acronyms.....65**

## List of Tables

Table 1. Demographics of the targeted neighborhoods.....	3
Table 2. Metrics, Data Sources, and Targets for PM #1 - Improved ability of the CTP users to make satisfactory Complete Trips in the study area. ....	15
Table 3. Metrics, Data Sources, and Targets for PM#2: Usefulness of the CTP Registration and Trip Preferences Processes.....	16
Table 4. Metrics, Data Sources, and Targets for PM #3: Usefulness of the CTP Trip Planning, Booking, Reporting and Review Processes.....	18
Table 5. Metrics, Data Sources, and Targets for PM #4: Improved Ability to Find Destinations Efficiently using the CTP Wayfinding Functionality.....	19
Table 6. Metrics, Data Sources, and Targets for PM #5: Improved Ability to Cross Specific Intersections Safely using CTP Smart Signal Functionality.....	21
Table 7. Metrics, Data Sources, and Targets for PM #6: Provision of an efficient, reliable and safe new on-demand transit shuttle system.....	22
Table 8. Correspondence of Recommended PMs to Use Cases.....	26
Table 9. Correspondence of Performance Measures and System Components.....	27
Table 10. Outcomes, Use Cases, and Performance Measures.....	38
Table 11. Summary of Data Requirements for each Performance Measure.....	38
Table 12. Plans for Independent Evaluator Interactions ( <i>Source: Volpe IE</i> ).....	45
Table 13. Proposed Composition of Phase 3 Samples.....	50
Table 14. Phase 2 Performance Measurement and Evaluation Support Schedule (PMESS).....	59
Table 15. Envisioned schedule of events for Phase 3.....	61

## List of Figures

Figure 1. Buffalo Niagara Medical Campus relative to the neighborhoods of focus: Allentown, the Fruit Belt, and Masten Park. ....	3
Figure 2. Context Diagram of the Buffalo, NY ITS4US Deployment System.....	27
Figure 3. Data Sharing Framework.....	48
Figure 4. Example of Main Performance Measure Type Dropdown Menu.....	55
Figure 5. Example of Customization Dropdown Menu.....	55
Figure 6. Example of a Specific Report.....	56
Figure 7. Examples of PM visualization.....	57



# 1 Introduction

Buffalo, New York (NY) is one of five demonstration sites selected for U.S. Department of Transportation (USDOT) ITS4US Deployment Program, which seeks to integrate innovative technologies to improve mobility and accessibility for underserved populations. The Buffalo, NY project plans to deploy an integrated set of travel support services and systems within neighborhoods surrounding Buffalo Niagara Medical Campus (BNMC) to meet these goals.

The primary objective of the complete trip deployment is to address travelers' challenges to planning and executing a trip. The BNMC project team will demonstrate, quantify, and evaluate the impact of advanced technologies, strategies, and applications in achieving this objective. Performance measurement, done by the BNMC project team, and independent evaluation (IE) efforts conducted by USDOT Volpe team, provides validation of the outcomes achieved and lessons learned to the broader communities of interest.

This document presents the Performance Measurement and Evaluation Support Plan (PMESP) that identifies and defines performance measures that will be used to assess and analyze the success of the complete trip deployment in Buffalo, NY. It also describes the plan for collecting data to populate the performance measures as well as the plan for routinely and publicly reporting on the system performance.

## 1.1 Intended Audiences

This document has been written to serve the information needs of multiple and varied audiences. This includes:

- USDOT - Intelligent Transportation System (ITS) Joint Program Office (JPO), Office of the Secretary (OST), Federal Highway Administration (FHWA), and Federal Transit Administration (FTA).
- BNMC project team members: University of Buffalo (UB), ICF, Open Doors Organization, RSG (Resource Systems Group), and ETCH, as well as any independent evaluators or interested performance measurement and evaluation specialists.
- Niagara Frontier Transportation Authority: lead agency for deployment phases.
- Public sector partners: NY State Department of Transportation, the Niagara Regional Transportation Commission, and City of Buffalo.
- Other important stakeholders: neighborhood associations, medical, healthcare, community, and human service organizations, and the BNMC Transportation Operations Council.
- Industry professionals: industry professionals or practitioners interested in the project or perhaps considering deploying a similar system.

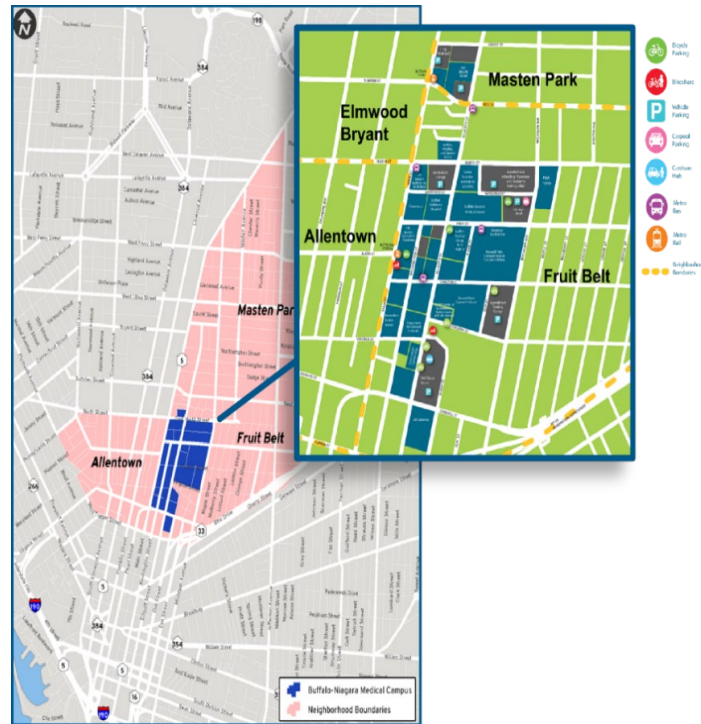
## 1.2 Project Background

The ITS4US Deployment Program is an effort led by the ITS JPO and supported by OST, FHWA, and FTA to identify ways to provide more efficient, affordable, and accessible transportation options for underserved communities that often face greater challenges in accessing essential services. The program aims to solve mobility challenges for all travelers with a specific focus on underserved communities, including people with disabilities, older adults, low-income individuals, rural residents, veterans, and limited English proficiency (LEP) travelers. This program will enable communities to build local partnerships, develop and deploy integrated and replicable mobility solutions to achieve complete trips for all travelers.

As one of the selected sites, Buffalo, NY is striving toward a sustainable future at all levels of society, incorporating actions in the community, government, and private entities in the area. Enabling community mobility and access to jobs, healthcare, and services to traditionally underserved populations is the primary motivation for all the regional partners involved in this deployment. As such, this deployment seeks to serve the specific transportation needs of BNMC visitors, employees, and residents of the focus neighborhoods that identify with the following groups:

- People with disabilities, including mobility/wheelchair users, visually impaired individuals, and hearing impaired individuals.
- Low-income individuals.
- Older adults.
- Limited- or non-English proficient persons.

The deployment location is in the downtown Buffalo, with a focus on travel to and from the Buffalo Niagara Medical Campus (BNMC). The deployment includes the 120-acre Medical Campus and surrounding neighborhoods, targeting three specific neighborhoods (i.e., Fruit Belt, Masten Park, and Allentown) as seen in Figure 1. The Buffalo Niagara Medical Campus includes nearly 9 million sq. ft. of clinical, research, office and development space. More than 16,000 people work or study at the BNMC and more than 1.5 million visit each year for health care and other services, generating significant transportation demand for the area, its visitors, and its employees.



Source: Buffalo, NY ITS4US

**Figure 1. Buffalo Niagara Medical Campus relative to the neighborhoods of focus: Allentown, the Fruit Belt, and Masten Park.**

The demographics of the surrounding neighborhoods are emblematic of the socioeconomic and racial divide that this deployment seeks to address, as depicted in Table 1 (Fruit Belt/Masten Park demographics versus Allentown and Buffalo MSA).

**Table 1. Demographics of the targeted neighborhoods.**

Geography (ACS 2018 tracts)	Percent 0-vehicle households	Percent population 65+	Percent poverty	Percent income <\$25k	Percent with a disability (18 to 65 yrs. old)	Total households	Total population
<b>Fruit Belt</b>	<b>47.0%</b>	<b>21.9%</b>	28.0%	<b>39.5%</b>	<b>20.0%</b>	<b>976</b>	<b>2,435</b>
<b>Masten Park</b>	<b>35.0%</b>	18.5%	34.7%	38.9%	15.2%	1496	3,208
<b>Allentown</b>	<b>18.4%</b>	6.2%	28.8%	17.4%	8.0%	1978	3,143
<b>Buffalo MSA</b>	<b>36.6%</b>	<b>12.0%</b>	<b>31.1%</b>	<b>30.7%</b>	<b>9.7%</b>	<b>110701</b>	<b>255,423</b>

### 1.3 Scope

The Buffalo, NY ITS4US project will be completed in three phases:

- Phase 1 – System Concept Development
- Phase 2 – System Design and Test

- Phase 3 – System Operation and Evaluation

This document was originally prepared as part of Phase 1 and has been revised during the initial stages of Phase 2 and finalized as an as-built at the conclusion of Phase 2. It defines a performance measurement and evaluation support plan for what is expected in Phase 3, when the proposed Buffalo, NY ITS4US system is operating, and evaluation data are being collected, analyzed, and reported.

The proposed system seeks to enhance the mobility accessibility for all travelers including those with disabilities, older adults, those with low-income, those who are limited- or non-English speaking. These individuals live across the region and particularly in the Fruit Belt and Masten Park neighborhoods. They include those who need access to the jobs and health care services at the Buffalo Niagara Medical Campus. To provide a seamless complete trip, the Buffalo, NY ITS4US system will provide:

- **Complete Trip Platform Application (CTP)** that is a trip planning app, which is customized for accessible travel.
- **Electric Shuttle Service (both human-operated (HDS) and self-driving shuttles (SDS))** that is integrated with the CTP and provides circulation in BNMC campus and adjacent neighborhoods.
- **Smart Infrastructure** improvements to digital features within and around BNMC, such as adding communication, connectivity and traveler information technologies to sidewalks and their adjacent loading/parking areas for transportation vehicles, bus shelters, intersections, and wayfinding technologies in indoor and outdoor venues.
- **Performance Dashboard** which monitors and reports on the above system components by ingesting log files from the other component and external data sources, then storing and analyzing the data to display current and historic data sets produced by the proposed system.

This PMESP is based on current best available information about the system and its goals as derived from a thorough review of project goals and the Phase 1 concept, reviews of related literature, experience with prior similar projects, information gathering efforts with sponsors, team members, and stakeholders, as well as preliminary system design and testing carried out in Phase 2. We have updated the plan based on the current system concept and development plans. We have built in contingencies using conservative assumptions about data availability and privacy restrictions where relevant.

## 1.4 Performance Measurement and Evaluation Support Plan Purpose

The purpose of the PMESP is to identify the information (that is, the performance measures) that is needed to evaluate whether the Buffalo, NY ITS4US system is meeting its objective of improving accessibility to the BNMC (for healthcare services and jobs) for travelers with disabilities and limited- or non-English speaking, older adults and low-income travelers (especially those without access to personal automobiles). Performance measures are also useful in monitoring progress towards meeting this objective. The PMESP presents the data collection plan that represents the project team's thoughts on the best possible methods and strategies for capturing the needed information. The successful execution of the PMESP depends on technologies being deployed and data being available for evaluation activities.

By following the PMESP to systematically collect data and conduct performance measurement and analysis activities, the Buffalo, NY ITS4US project team will be able to identify and quantify which deployed strategies, services, and/or components were of value in addressing the targeted complete trip challenges. Documenting the results with robust supporting data and analyses will enable others to understand and build upon the investments made in this project to facilitate incorporating complete trip strategies and goals more effectively in other future deployments.



# 2 Goals and Objectives

## 2.1 Deployment Goals and Objectives

According to the Phase 1 Concept of Operations (ConOps) (FHWA-JPO-21-860) document for the Buffalo, NY ITS4US Deployment Project, within and around the BNMC, there are accessibility-and pedestrian-related transportation and mobility challenges that hinder the BNMC and local/regional transportation agencies from providing safe and accessible transportation to residents and visitors. These challenges include traffic safety issues at street crossings, lack of accessible infrastructure in street rights of way, and lack of efficient and reliable circulation paths between facilities on the campus and surrounding neighborhoods. These problems affect pedestrian activity, especially from Fruit Belt and Masten Park neighborhoods, as well as for individuals with disabilities, older adults, and travelers with temporary mobility challenges (such as people carrying groceries or with double strollers). Street crossings, in particular, are challenging to people with disabilities and older adults. Common problems include the collection of water, ice, and snow at the bottom of curb ramps due to poor drainage; counter slopes at the bottom of curb ramps that make walking difficult and cause wheelchair users to bottom out; broken and irregular pavement in the crossing area; the use of brick and cobbles that makes wheelchair use very uncomfortable; deterioration of tactile warning signals; and poor contrast of crosswalks with the street surface. These issues and problems have a direct impact people's ability to reach BNMC's facilities and get to their appointments on time, or even make the trip/appointment at all—this particular issue was highlighted during several interviews with stakeholders, as noted in the ConOps document. In Buffalo, emerging mobility services (such as on-demand ride sharing or subscription-based vehicle services) are not viable options for many travelers with difficulties. For those who cannot drive a personal vehicle, transit continues to be the primary mode of choice.

The user needs assessment identified the following as desired changes in mobility and transportation opportunities.

- **Consistent, continuous trips to, from, and within the BNMC area:** Ensuring that there is an accessible and convenient path of travel to and from all origins/destinations.
- **Online and offline ways to receive real time information on services, and infrastructure usability and accessibility:** Enabling a comprehensive array of alternatives that includes online and offline visual, audio, and haptic notifications, or a combination thereof, would allow travelers to understand the current condition of their trip and how to best react to deviations from their original plans spontaneously.
- **Trip paths that are safe, accessible, and compatible with user-defined preferences and capabilities:** Proving reasonably leveled, slip resistant, smooth paths of travel.
- **Integrated, flexible, demand-responsive, end-to-end transit options for the community:** Ensuring integrated, complete, and comprehensive end-to-end services to travelers, including travelers with disabilities.

The proposed Buffalo ITS4US concept brings together inclusive design, integrated trip planning, and micro transit elements to provide complete trip support to travelers who use the system,

especially travelers with disabilities, and who are older, have lower income, or are limited in their English proficiency. Travelers using the complete trip system gain the following benefits from the system elements:

- A door-to-door travel planning app, or complete trip platform (CTP), that allows travelers to make efficient and effective transit trips to and from the deployment area.
- Turn-by-turn indoor and outdoor wayfinding guidance near the BNMC.
- For paratransit access line (PAL) eligible travelers, an alternative way to access shuttle services adequate for travelers with disabilities, as well as providing more flexibility and support for trips that are not supported by PAL.
- Increased safety and improved capability to cross specific intersections and use specific prioritized pathways for accessing BNMC campus entities.
- Access to a new community shuttle (CS) service with human-driven shuttle (HDS) and self-driving shuttle (SDS) that connects nearby neighborhoods to destinations and services within the deployment area.
- Enabling caretakers to manage and monitor trips for the travelers who are in their care.

## 2.2 Use Cases

Performance measures (PMs) assess whether the short-term outcomes of the deployment (e.g., more efficient trips) have occurred as well as progress towards achievement of the desired longer-term impacts. To prioritize and guide the development of PMs, use cases were identified that describe the *functions for which* the deployment technologies (i.e., CTP, SDS and HDS shuttle service, Smart Infrastructure) are expected to have the greatest benefits for underserved groups. There are five use cases associated with the CTP; four use cases associated with the SDS and HDS; and one associated with Smart Infrastructure.

Scenarios were also identified that describe *for whom* the deployment is expected to benefit the most. However, scenarios are not presented in this Section because they are not as useful as the use cases for guiding development of the performance measures.

### 2.2.1 Complete Trip Platform (CTP)

The following use cases are relevant to the CTP.

#### 2.2.1.1 Use Case 1: Register Profile and Preferences

This use case describes the processes and interactions with travelers to set up a CTP user account. The function enables the account holder to select their travel preferences for types of navigation triggers, wayfinding notifications and alert communications. The functions also enable users to identify their preferences for mode, accessibility needs, and link other accounts (such as a PAL account) with their CTP account.

### **2.2.1.2 Use Case 2: Generate Trip Plan and Book a Trip**

This use case consists of functions for a traveler to plan a trip by inserting their origin and destination. They may customize this trip by selecting general preferences (e.g., modes, maximum walking distance, shortest trip, fewest transfers), or if they log in to their account use an existing trip plan or set of preferences for travel and notification. The traveler can also adjust their trip preferences and save the updated trip plan. In addition, as an account holder authorized to use registered mobility services such as PAL or Shuttle, the traveler can generate a complete trip plan with a trip leg that includes reservations and confirmation with the mobility service (PAL Direct or Shuttle).

### **2.2.1.3 Use Case 3: Public Transport Services**

This use case describes the information provisions associated with accessing public transit mode options. These include NFTA bus, light rail, and PAL Direct, as well as Shuttle options that are included in these services. The services consist of hailing, boarding, traveling in and alighting these public transport vehicles.

Note: Accessibility issues are not included in this use case. They are addressed in Use Cases 7 and 8.

### **2.2.1.4 Use Case 4: Navigation**

This use case describes wayfinding and navigation on pathways to complete a trip. This use case consists of the use of the CTP when traveling including crossing intersections, traversing sidewalks, and wayfinding to and through indoor facilities.

### **2.2.1.5 Use Case 5: Reporting and History**

This use case describes information provided to the traveler on the CTP that is available for account holders about trips they completed. In addition, the traveler can submit trip obstacles and improvements made during their journey. This provides a crowd-source approach to collecting information on accessibility status, like elevator outages, paths in the trip plan, etc.

## **2.2.2 Human-Driven and Self-Driving Community Shuttle Service**

The following use cases are relevant to the human-driven and self-driving community shuttle service.

### **2.2.2.1 Use Case 6: Shuttle Service Reservation and Dispatch**

This use case describes several of the processes and functions of the Shuttle Operation Center (SOC), and especially those that will be applied and activated when receiving a traveler request for service by the shuttles system.

### **2.2.2.2 Use Case 7: Passenger Pick-up, Securement and Drop-off via the SDS**

This use case describes several of the processes and functions of the Shuttles Subsystem which will be applied and activated when a traveler boards the SDS, secures him- or herself (plus any

mobility aid mechanism) onboard the vehicle, travels on the SDS, and finally gets off the SDS at their final or intermediate destination.

#### **2.2.2.3 Use Case 8: Passenger Pick-up, Securement and Drop-off via the HDS.**

This use case describes several of the processes and functions of the Shuttles Subsystem, which will be applied and activated when a traveler boards an HDS, secures him- or herself (plus any mobility aid mechanism) travels on the shuttle, and finally gets off at their destination.

#### **2.2.2.4 Use Case 9: Manage Incidents**

This use case describes the processes and functions that will be activated by the shuttles subsystem to manage shuttle-related incidents. The incidents may involve the environment around the shuttle (e.g., inclement weather), the vehicle itself (a malfunction or accident), or travelers onboard the shuttle.

### **2.2.3 Smart Infrastructure**

The following use cases are relevant to the smart infrastructure.

#### **2.2.3.1 Use Case 10: Intersection Pedestrian Crossing (PedX) Request**

This use case describes the transmission of a PedX request message from the CTP to the traffic signal controller. This use case supports the intersection crossing system.

*Note: this use case is modeled after the NTCIP 1211 Signal Control and Prioritization architecture and use case scenarios. The terms Priority Request Generator (PRG), Priority Request Server (PRS) and Coordinator (CO) are use case actors in the NTCIP 1211 and will be used in this use case even as the PedX request is a simple request and does not include a request for priority.*

# 3 Performance Measurements and Targets

The primary expected outcome of the new complete trip technologies and services, described in the ConOps document and in the preceding sections, is that the travel experience of specific population groups, listed below, within the BNMC will be measurably enhanced in important ways in comparison to the baseline situation (travel prior to the deployment of the updated travel system), including:

- increased ease of planning in carrying out desired trips, particularly multi-modal trips
- increased reliability of travelers arriving at the trip destination without significant delays
- enhanced travelers' ability to make trips without exposing the traveler to new safety risks (and reducing existing safety risks, where feasible)

The focus of the complete trips technologies is on specific population groups:

- those with visual, mobility, and/or auditory disabilities
- older adults advancing beyond age 65
- people living in low-income households
- people with limited- or non-English-speaking proficiency

The complete trips technologies may also benefit other population groups that do not fall in any of the categories above, including people who are caregivers for persons who fit the categories above, who work at the BNMC and other people living in the neighborhoods surrounding the BNMC (Fruit Belt and Masten Park, which will offer the CS service).

The level to which the complete trips technologies enhance trip-making to BNMC may vary a good deal across different types of travelers. As one example, although improvements in wayfinding technology and messaging may be useful to all groups, they may provide more benefit to those with visual disabilities than to people without such disabilities. Therefore, in measuring the performance on the complete trips improvements, it will be important that data on trip-making is collected from a sufficient number of people in each of the specific population groups listed above so that the performance of the system can be assessed separately for different groups (as well as factors such as access to or familiarity with technology, familiarity with the environment, and assistance from a caregiver). From outreach efforts performed as part of this project, it has been made clear that some individuals fall in several of these categories simultaneously (Sener, et al., 2021). The System Deployment Impact Analysis Design (Chapter 5) discusses sample size issues in detail.

Most of the performance measures recommended in this Section do not focus on one specific type of traveler. In those instances, we assume that the measures will be relevant for most or all travelers, and that a key outcome of the performance analysis will be to investigate how the performance varies across different types of users.

## 3.1 Identification of Potential Performance Measures and Targets

Consistent with the ConOps document and the Data Management Plan (DMP), and supplemented by further consideration of data needs for the PMESP, the main sources and types of data that will be available are:

- **Data from the CTP app:** These data include any information on traveler characteristics and preference provided during the registration process, as well as data from each trip that is planned, booked, and/or carried out using the CTP app. This includes locations of trip ends, modes used on trips, travel times, and trip plans made using the app. It should be noted that for data privacy reasons, the locations of trip ends will not be exact addresses, but will indicate which area of the BNMC and which neighborhoods the trips are to and from.
- **Data from User Surveys:** As described in detail in later chapters, all end users will be asked to respond to various surveys, including (a) a baseline survey identifying traveler characteristics, recent travel history, and their ratings and satisfaction regarding relevant aspects of the pre-deployment travel situation, and (b) post-deployment surveys asking about usage of the CTP system and their ratings and satisfaction regarding the current travel situation and regarding specific system components.
- **Shuttle Reservations Management Data:** Data will be available on all reservations made for the Community Shuttle systems, including the time the reservation is made, the pick-up and drop-off locations, the quoted pick-up and drop-off time, the fare, and the CTP user ID (which will allow this data to be tied to the data from the CTP app and the surveys).
- **Shuttle Trip Operations Data:** Data will also be available on each trip made by the Community Shuttle systems. These data include the actual pick-up and drop-off times and stop locations for all passengers (identified by CTP user ID), which can be compared against the reservations data to assess on-time performance. Data will also record the start and end time and distance traveled on each vehicle-trip, and identify and classify any vehicle malfunctions, driver or steward interventions, crashes, near-crashes or other problems encountered during each trip.
- **Smart Infrastructure Operations Data:** These data include the operating functionality over time of the smart infrastructure (including outdoor and indoor wayfinding infrastructure and smart intersection controls). For the smart intersection controls, data on activation by specific CTP users will be available. For the wayfinding beacons, specific interaction data will not be collected, although data on the use of wayfinding will be available from the post-deployment survey data.

- **Additional data items in the Performance Management Dashboard:** Other very useful data for performance measurement include external factors that may affect travel and the ability for users to make complete trips using the system. These items will be monitored over time with data added to the performance measurement system at an appropriate frequency (e.g., daily or weekly):
  - Data on weather conditions, road construction and closures, facility closures, power outages, etc.
  - Data on operating conditions for key NFTA services, particularly those serving the BNMC.
  - Operating status of key equipment/facilities at the BNMC destinations (elevators, entrances, etc.), to the extent which it is available.

Baseline data for these external factors will be collected so that they can be compared against conditions over time during the post-deployment period. Even though we do not expect the system to be affected by these external factors, changes in those factors across the pre- and post-deployment periods can affect the users of the system and system performance.

The list above describes an extensive array of data sources filtering through the Performance Measurement Dashboard, upon which performance measures can be based. One particular type of survey-based data that may be new to readers is the use of a modified version of the Rapid Assessment of Product Usability and Universal Design (RAPUUD). It is comprised of a battery of questions to assess new technologies, with the questions designed to be particularly relevant to users with disabilities (Lenker, et al. 2011; Choi, et al. 2021). In short, they are designed to obtain subjective measurements of the technological feature or service along nine dimensions: (1) ease of use, (2) high independence, (3) low safety risk, (4) low risk of mistakes and “do overs”, (5) efficiency of time use, (6) low physical effort required, (7) low mental effort required, (8) low level of unwanted attention, and (9) low risk of embarrassment. The ratings are made on a 5-point Likert scale, with the midpoint representing “neutral”.

For many of the performance measures listed below, while we provide targets based on objective data, we also recommend obtaining subjective measures when and where it is best suited. Ideally, by the end of Phase 3, all components of the system would rate better (on average) for all of the subjective dimensions, but further confirmation is needed in Phase 3. The Agile development process conducted as part of the Phase 2 development of the app and mobility enabled the product owners (NFTA, BNMC) and other stakeholders to (1) provide a preliminary assessment of the performance of various aspects of the CTP and other system components as they are developed and introduced, and (2) to set realistic targets for what levels can be achieved in Phase 3.

Another key point to raise prior to recommending specific performance measures is that this project will target a minority of the population in a limited area of the region. So, we do not expect to see any measurable impacts on the entire system (NFTA, traffic counts, etc.). The focus will be on the core user experience and benefits of the CTP users, in most cases against a pre-deployment baseline, although some broader system-level (Tier 1) data can also be informative, such as the use and performance of the CS system (both self-driven and human-driven vehicles), as well as the number of paratransit (PAL) trips made to the BNMC campus.

Finally, while the experimental design and data collection plan are laid out in more detail in Sections 5 and 7, it is useful to provide a few key aspects of the proposed plan at this stage in order to clarify what is being proposed for the performance measures. Alpha and Beta testing of the deployed technologies will be done as part of the Agile development process with persons “outside” of the recruited study participants. Alpha testing will be done with a small, restricted group of internal project team staff, who will be asked to test the technologies in different use cases. Beta testing will be conducted with a wider set of participants, such as NFTA or BNMC staff and some individuals from the key stakeholder groups who will be asked to use the technologies for a specific duration (e.g., weeks or months). These individuals will be debriefed after use in one-on-one interviews or in focus groups and will also be asked to complete the baseline and post-deployment surveys as a means of piloting the administration of the surveys and retrieval of the data

In Phase 3, all study participants will be recruited in an attempt to reach 500 study participants. We anticipate that new CTP users will require a period of a few months to become familiar enough with the CTP and other system components to evaluate it on an informed basis. In Phase 3, we propose collecting travel data from the CTP continuously, supplemented by survey data at four points:

1. **Upon Deployment (baseline):** After recruitment, before using the CTP (starting when the system is deployed in Phase 3, providing a description of prior travel patterns).
2. **Initial Post-Deployment:** At the end of the initial period (3 to 4 months after a participant registers to use the system) .
3. **Mid-Term Post-Deployment:** Roughly half-way between deployment and the end of Phase 3 (6 to 8 months after a participant registers).
4. **Final Post-Deployment:** Near the end of Phase 3 (approximately one year after a participant register to use the system).

Any performance measurements that evaluate changes in travel behavior due to the use of the app will tend to compare data from measurement points 2, 3 and 4 to the point 1 baseline. Other than baseline conditions, there will be very little objective data at point 1, since the CTP will not have collected much data up to the point, although participants may provide useful information about past travel behavior.

### ***3.1.1 PM #1: Improved ability of the CTP users to make satisfactory Complete Trips in the study area***

The extent to which the CTP can improve the ability to make successful, satisfactory complete door-to-door trips within the study area depends on how well the system meets the primary user needs. The user needs assessment identified the following as desired changes in mobility and transportation opportunities:

- **Consistent, continuous trips to, from, and within the BNMC area:** Ensuring that there is an accessible and convenient path of travel to and from all origins/destinations.

- **Trip paths that are safe, accessible, and compatible with user-defined preferences and capabilities:** Proving reasonably leveled, slip resistant, smooth paths of travel.
- **Online and offline ways to receive real time information on services, and infrastructure usability and accessibility:** Enabling a comprehensive array of alternatives that includes online and offline visual, audio, and haptic notifications, or a combination thereof, would allow travelers to understand the current condition of their trip and how to best react to deviations from their original plans spontaneously.
- **Integrated, flexible, demand-responsive, end-to-end transit options for the community:** Ensuring integrated, complete, and comprehensive end-to-end services to all travelers, including travelers with disabilities.

We propose to directly measure the users' perceptions (ratings) of how well these goals and objectives have been met in terms of an improvement compared to the pre-deployment baseline situation. There will be one overall question corresponding to each of the four objectives listed above. The questions will be asked in both the baseline survey and the post-deployment surveys to measure changes compared to pre-deployment, as well as any changes over time during the deployment period. The questions will be asked using 5-point Likert scales, as those are easier for respondents to comprehend than questions with more complex scales. The targets are an average increase of one point or more relative to the baseline survey responses—see Table 2.

**Table 2. Metrics, Data Sources, and Targets for PM #1 - Improved ability of the CTP users to make satisfactory Complete Trips in the study area.**

Metric	Data Source	Target
System user ratings of the ease of making door-to-door trips to, from and within the BNMC.	Baseline and post-deployment surveys	An average increase of 1 point or more on a 5-point Likert scale
System user ratings of how safe door-to-door travel paths are for trips to, from and within the BNMC, including level, slip-resistant paths.	Baseline and post-deployment surveys	An average increase of 1 point or more on a 5-point Likert scale
System user ratings of the adequacy and usefulness of information for making trips to, from and within the BNMC.	Baseline and post-deployment surveys	An average increase of 1 point or more on a 5-point Likert scale
System user ratings of the ability to make trips using integrated transit services to, from and within the BNMC.	Baseline and post-deployment surveys	An average increase of 1 point or more on a 5-point Likert scale
System user ratings of the satisfaction with making a trip using the CTP app.	CTP user data	An average score of better than neutral on a 5-point Likert scale

There is also a metric based on a rating question asked in the CTP app after the completion of each trip. There will be no baseline data for this metric, so the target is a positive (better than neutral) average rating.

Although the performance metrics and targets are expressed as averages across all participants, it will be valuable to analyze them over time as a function of the characteristics of each user group (disability type, age group, income group, etc.). This would allow us to show how the perceived benefits of the complete trip system vary across different population groups.

The metrics above assess the overall usefulness of the CTP system. However, there are several different components and features of the CTP system, so it will be informative to also obtain user ratings and frequency of use specific to each key component. The more focused performance measures recommended below can provide more insight into why some people find the system more useful than others do.

### 3.1.2 ***PM#2 Usefulness of the CTP Registration and Trip Preferences Processes***

The registration process in the CTP app is important to the success and usage rate of the CTP system in two main ways. First, the respondents must find the registration process to be physically accessible (e.g. via smartphone, personal computer, kiosk or call center) and easy to complete (with assistance provided if needed). Second, the requirement for users to state their travel preferences must be seen as relevant and useful in terms of the CTP app providing trip options that they perceive as meeting their preferences. The two metrics recommended are based on answers to subjective questions asked in the post-deployment surveys—see Table 3.

The target is an overall positive (above neutral) score, although lessons can be learned by analyzing the variations in the user ratings across the user types.

The performance in terms of the trips meeting user preferences will be monitored over time during the post-deployment period, as it is possible that users may adjust their preferences once they have more experience using the system, and the system may also be able to adapt based on which trip options users have selected. It may also be interesting to ask in the pre-deployment survey how well respondents expect that such a system will be able to provide trip options that are customized to their needs and preferences, but at this stage we have not based the performance metrics on such a speculative question.

**Table 3. Metrics, Data Sources, and Targets for PM#2: Usefulness of the CTP Registration and Trip Preferences Processes.**

Metric	Data Source	Target
System user rating of the ease of registration process.	Post-deployment surveys	An average score of better than neutral on a 5-point Likert scale
System user ratings of the extent to which the options available during registration served their needs.	Post-deployment surveys	An average score of better than neutral on a 5-point Likert scale

Note that the Agile CTP app development process carried out during Phase 2 allowed us to test the methods for collecting the data for these metrics. In turn, that data (along with more focused feedback from beta-test users) will help the project team fine-tune the CTP registration process so that it can meet the performance targets. The user data will be analyzed against the targets in Phase 3 to note any significant discrepancies as Phase 3 progresses.

### **3.1.3 PM#3 Usefulness of the CTP Trip Planning, Booking, Reporting and Review Processes**

The travelers' primary interface with the CTP before making a specific trip will be through the trip planning and (for some) the community shuttle (CS) booking functions, so the perceived ease of use and usefulness of those functions will be important factors in the extent that the system is used in the long run. The route planning function in the CTP can apply to users of all modes, including those who use the new transit options, those who walk or bike to/from BNMC, and those who drive and park or who use existing NFTA transit services and then walk to their destination. In contrast, the CTP booking process is only relevant for those who use the CTP to reserve trips on the community shuttle (CS).

The first two metrics listed in Table 4 measure the perceived ease of using the route planning option and the satisfaction with the paths offered by the CTP. It should be noted that the CTP system—including internet and telephone options—can be useful to plan a route or book a transit trip even if the smartphone app is not used during the actual trip. This feature means that the system can be useful even for those who do not own smartphones.

The third metric is the percentage of CTP users who use the system to book transit trips at least once. A related measure for analysis will be the frequency with which users book transit trips, but this is not listed as a performance metric because it will depend to a large extent on how often users make trips for which the shuttle system is a useful option. The fourth metric is a subjective rating of the ease of booking CS trips via the CTP. In each case, the target is an overall positive (above neutral) score, although more detailed analysis will be done on the variations in the user ratings across the user types. Note that these are overall subjective ratings, while other PMs contain more objective measures of the efficiency of the paths, depending on the CTP functions used. Also, the ratings for booking transit trips will be averaged across the subset of CTP users who actually use the booking function.

The fifth metric uses the CTP user data to determine the percentage of users who have utilized the CTP functionality for reporting incidents or conditions encountered during the trip, such as elevator outages or treacherous sidewalk conditions. This “crowdsourcing” function could provide very useful real-time information to BNMC visitors, but the usefulness will depend a great deal on the extent to which users provide that information. The sixth metric is to measure the ease of using the reporting function (for those who have used it), and the final two metrics are for the fraction of users who have used the app function to review past trips planned in the CTP app and for the perceived usefulness of that functionality. These last four metrics correspond to Use Case 5.

Note that some of the objective metrics in this performance measure (PM) and in other PMs below are for the percentage of CTP users who make use of various CTP features. We have provided initial targets for those metrics that seem reasonable. Two key uses of such metrics will be (a) to be able to analyze which groups of users use the various types of CTP functionality, and

(b) to provide objective evidence to compare against user's self-reported CTP usage patterns reported in surveys. In Phase 3, the surveys can also be used to ask why some people choose **not** to use certain types of CTP functionality that are available to them.

**Table 4. Metrics, Data Sources, and Targets for PM #3: Usefulness of the CTP Trip Planning, Booking, Reporting and Review Processes**

Metric	Data Source	Target
System user ratings of the ease of planning a door-to-door trip.	Post-deployment surveys	An average score of better than neutral on a 5-point Likert scale
System user ratings of the satisfaction with the specific route/mode options provided by the CTP.	Post-deployment surveys	An average score of better than neutral on a 5-point Likert scale
The percentage of CTP users who use the system to book on-demand (CS) transit trips.	CTP User data and post-deployment surveys	At least 70% of CTP users who use transit in the baseline book a transit trip at least once
System user ratings of the ease of booking on-demand (CS) transit trips via the system.	Post-deployment surveys (among those who have booked transit trips)	An average score of better than neutral on a 5-point Likert scale
The percent of CTP users who use the CTP to report incidents or travel conditions during their trips.	CTP User data and post-deployment surveys	At least 50% of CTP users use the reporting functionality at least once
System user ratings of the ease of reporting incidents or conditions encountered during a trip in the CTP.	Post-deployment surveys (among those who have used the report function)	An average score of better than neutral on a 5-point Likert scale
The fraction of CTP users who use the function to review past trip history.	Post-deployment surveys	At least 50% of CTP users use the function at least once
System user ratings of the usefulness of reviewing past trip history in the CTP.	Post-deployment surveys (among those who have used the review function)	An average score of better than neutral on a 5-point Likert scale

### **3.1.4 PM#4 Improved ability to find destinations efficiently using the CTP wayfinding functionality**

The use of the CTP for both indoor and outdoor navigation is a unique and potentially very useful feature of the system. The usefulness of the wayfinding functionality may be an important

incentive for people to use the CTP system and to continue using it beyond the study duration. Table 5 includes four types of metrics for outdoor wayfinding and three analogous types for indoor wayfinding:

- The percentage of CTP users who elect to receive the wayfinding notifications during their trips.
- The self-reported frequency (fraction of trips made after planning in the CTP) in which the person actively uses the wayfinding information.
- The perceived usefulness of the wayfinding information in arriving at trip destination on time.
- The perception of various other aspects of using wayfinding (including safety, physical effort, mental effort, etc.) using the modified RAPUUD questions introduced earlier.

The latter questions will be asked only of people who report having used the wayfinding information and will be asked using a 5-point Likert scale with the target of a positive average score on each aspect.

The traveler needs and experience for outdoor versus indoor wayfinding may be different enough that separate survey questions and metrics for the two types of CTP functionality seem warranted. Also, different types of users may have different levels of need and use for wayfinding assistance. Those with visual disabilities and cognitive disabilities may find the information most helpful. As noted previously, the metrics can be analyzed to investigate differences across user groups, even though the performance measure applies to all users.

Note that the CTP route trace data will not indicate whether each user actually heeds and follows the directions given, as spatial information at that level of detail is not available from the CTP itself. Instead, various metrics based on users' self-reported data will be used to answer these questions, as well as supporting data from trip-specific ratings questions asked in the CTP upon completion of each trip.

**Table 5. Metrics, Data Sources, and Targets for PM #4: Improved Ability to Find Destinations Efficiently using the CTP Wayfinding Functionality**

Metric	Data Source	Target
The fraction of CTP users who elect to receive <b>outdoor</b> wayfinding notifications.	CTP Registration and User Data and/or post-deployment surveys	At least 70% of CTP users elect to receive outdoor wayfinding notifications
System user self-reported frequency of using <b>outdoor</b> wayfinding notifications.	Post-deployment surveys and CTP User Data	The most frequent response is that a person uses outdoor wayfinding for "most" or "all" of their trips to/from BNMC"

Metric	Data Source	Target
The fraction of CTP users who elect to receive <b>indoor</b> wayfinding notifications.	CTP Registration and User Data and/or post-deployment surveys	At least 70% of CTP users elect to receive indoor wayfinding notifications
System user self-reported frequency of using <b>indoor</b> wayfinding notifications.	Post-deployment surveys and CTP User Data	The most frequent response is that a person uses indoor wayfinding for "most" or "all" of their trips to/from BNMC"
System user ratings of how useful the <b>outdoor</b> wayfinding functionality is in reaching their trip destination on time.	CTP User Data and post-deployment surveys, for those who report using outdoor wayfinding	Average scores of better than neutral on a 5-point Likert scale
System user ratings of how useful the <b>indoor</b> wayfinding functionality is in reaching their trip destination on time.	CTP User Data and post-deployment surveys, for those who report using indoor wayfinding	Average scores of better than neutral on a 5-point Likert scale
User ratings of various dimensions of using the CTP <b>outdoor</b> wayfinding functionality using the RAPUUD method.	Post-deployment surveys, for those who report using outdoor wayfinding	Average scores of better than neutral on a 5-point Likert scale
User ratings of various dimensions of using the CTP <b>indoor</b> wayfinding functionality using the RAPUUD method.	Post-deployment surveys, for those who report using indoor wayfinding	Average scores of better than neutral on a 5-point Likert scale

### **3.1.5 PM#5 Improved ability to cross specific intersections safely using CTP smart signal functionality**

The ability to activate "smart signals" for pedestrians crossing remotely using the CTP app may be a very attractive feature of the system for some users. The smart signal functionality will only be available at specific intersections, and thus the metrics for this performance measure (listed in Table 6) must take location into account, considering only pedestrian trips traversing the relevant intersections. It is expected that this functionality may be especially useful for those using wheelchairs and those with visual disabilities as well as older travelers who require more time to cross. While analysis can be focused on specific user groups, a single set of metrics is proposed across all travelers who use the intersections.

The first and second metrics are the objective and self-reported percentages of users who make trips crossing the relevant intersections who also use the smart signal activation functionality for at least some of their trips. We will have objective data from the smart signal operations data and

CTP user data on the number of times that each CTP passes within an intersection with a smart signal and the number of times the user activates a signal. Post-deployment surveys can be used to help confirm any findings based on self-reported data from the post-deployment surveys.

CTP users who report having used the smart signal functionality will also be asked to report their overall perceived ease of use and responsiveness of the signal activation on a 5-point Likert scale, with targets of positive overall average scores.

Perhaps most importantly, CTP users who report using the relevant intersections will be asked about the perceived safety of crossing the intersections in both the baseline and post-deployment surveys, to detect any perceived change in crossing safety. This question will be asked for all who use the intersections, regardless of whether they report using the smart signal activation, because there may be external factors such as construction and weather that influence perceived safety, so it will be important to measure overall changes as well as being able to compare the perceived changes for those who use the smart signal activation to the perceived changes for those who do not use it. Although there may be some objective data on safety such as the number of pedestrian-related injuries at the intersections, the low numbers of such injuries make such crash data inconclusive when observed from a small number of locations for only one or two years. A longer data collection period and/or more sites would typically be needed to identify any statistically significant trend. The number of “near misses” would also be relevant, but collecting such data would require continuous video monitoring and interpretation of video, meaning that such data would tend to be expensive to collect and would still be rather subjective.

**Table 6. Metrics, Data Sources, and Targets for PM #5: Improved Ability to Cross Specific Intersections Safely using CTP Smart Signal Functionality**

Metric	Data Source	Target
The percent of CTP trips crossing at the relevant intersections that use the smart signal remote activation function.	CTP Usage Data	The smart-signal activation feature is used for at least 50% of CTP trips crossing the relevant intersections
Self-reported fraction of people who cross at the relevant intersections who use the CTP smart signal activation functionality.	Post-deployment surveys	At least 50% of CTP users who make trips at the relevant intersections report using the smart-signal functionality for at least some trips
Ratings of the ease of use and ratings of other aspects of activating the smart signals, using the RPUUD method.	Post-deployment surveys (asked only for those who report using the feature)	An average score of better than neutral on a 5-point Likert scale
Self-reported frequency of crossing at the intersections with smart signals.	Baseline and post-deployment surveys	A statistically significant increase in frequency of crossing

Metric	Data Source	Target
Perceived safety of crossing at the intersections with smart signals.	Baseline and post-deployment surveys	An average increase of 0.5 points or more on a 5-point Likert scale compared to baseline

### 3.1.6 *PM#6 Provision of an efficient, reliable and safe new on-demand transit shuttle system*

The final major component of the CTP system included in the performance measures is the introduction of an on-demand transit Community Shuttle (CS) system. The CS can only be used by requesting and booking a trip via the CTP—in the app, on-line or by telephone. The CS includes both a human-driven shuttle (HDS) service and a self-driving shuttle (SDS) service. While the HDS can provide door-to-door service, the SDS will be more restricted in where it can travel, providing service on one specific route and requiring users to board and alight at specific stop locations. So, while the SDS may provide somewhat shorter in-vehicle times than the HDS because of using more direct routes, it will also require some additional access and egress time to and from the specified stop locations.

The metrics listed in Table 7 are for the CS as a whole, including both the HDS and SDS routes.

**Table 7. Metrics, Data Sources, and Targets for PM #6: Provision of an efficient, reliable and safe new on-demand transit shuttle system**

Metric	Data Source	Target
Percent of CS trips that arrive at the <b>boarding</b> stop within the targeted time allowance of the scheduled arrival time (scheduled at the time of booking).	CTP User data and Shuttle Booking and Operations data	90% of trips arrive within the targeted service level allowance (SLA) determined at the start of Phase 3 after go-live.
Percent of CS trips that arrive at the <b>alighting</b> stop within the targeted time allowance of the scheduled arrival time (scheduled at the time of booking).	CTP User data and Shuttle Booking and Operations data	90% of trips arrive within the targeted service level allowance (SLA) determined at the start of Phase 3 after go-live.
System user ratings of how reliable the transit system is in reaching their BNMC trip destination on time.	Baseline and post-deployment surveys	An average increase of 0.5 points or more on a 5-point Likert scale compared to baseline
Self-reported frequency of using all transit services, including the CS as	Baseline and post-deployment surveys	A statistically significant increase in the number of

Metric	Data Source	Target
well as PAL services and NFTA bus and rail lines.		transit trips per person per month
Percent of transit trips made by PAL-eligible CTP users within the CS service area that are made via the CS versus regular PAL trips.	CTP User data and Shuttle Operations data, supplemented by PAL Direct bookings data	25% or more trips made in the BNMC area by PAL-eligible users use CS services
User ratings of CS service in terms of key service aspects – convenience, affordability and safety, as well as other aspects included in the modified RPUUD questions.	Post-deployment surveys (asked separately for the HDS and SDS for those who have used the type of service)	Average scores of better than neutral on a 5-point Likert scale
Cost efficiency of the HDS and SDS shuttle services in terms of operating cost per passenger trip.	Shuttle Operations data, supplemented by additional system cost data	To be determined at the start of Phase 3 based on an investigation into the performance of similar on-demand shuttle systems elsewhere and other NFTA services
Percent of CS bookings for which the earliest available pick-up time is within 30 minutes of the time the trip is requested.	User data and Shuttle Booking and Operations data	90% of CS users can book a pick-up time no more than 30 minutes after the time the trip is requested

The first metrics deal with reliability in terms of on-time performance at both the board and alighting stops, comparing the actual shuttle departure and arrival times with the times that were scheduled when the shuttle trip was booked. This can be measured at the vehicle-trip level using data from the Shuttle Operations data, and at the person-trip level using both the Shuttle Operations data and the CTP user data. According to the System Requirements Specification (SyRS) document completed in Phase 1, realistic service level target is that 90 percent of trips arrive within that margin as described in the System Design Document. These allowances will be in the range of 10 to 15 minutes and could be different for the SDS and HDS due to their different routing characteristics. In particular, the number of HDS vehicles will be monitored to match the supply (number of vehicles operating) with the supply.

The second type of metric is based on user ratings of how reliable the transit system is in reaching their BNMC trip destination on time. This will provide a subjective comparison to the preceding metrics based on objective on-time performance data. More importantly, it will also provide a comparison to the perceived reliability of traveling by transit in the baseline situation prior to the introduction of the CS. The target is a slight improvement in perceived ability to arrive on time at the destination.

The third metric is also based on comparison of baseline and post-deployment survey responses, asking CTP users how reliable the transit system is in reaching their BNMC trip destination on time. The hypothesis is that the CTP helps users better navigate their travel around the BNMC campus by providing real time information such as directions and wayfinding. The target is an average increase of 0.5 points or more on a 5-point Likert scale compared to baseline.

The fourth metric measures the percentage of transit trips made by PAL-eligible CTP users within the CS service area that are made using the CS via the CTP rather than using the existing PAL services and PAL Direct booking system. The advantage of using the CS is that the trip does not need to be booked a day in advance as is the case with PAL Direct. Other differences in using the CS are that it is not a dedicated system for PAL-eligible users only and that it requires users to learn a new booking procedure. The metric will be measured by the Shuttle Operations data, supplemented by data from the PAL Direct booking system (which will need to be analyzed by origin and destination to determine which trips could have been served by the CS). The target is set at 25% of eligible trips using the CS instead of regular PAL services. The substitution effect could be much higher eventually, but it is difficult to predict how flexible current PAL users are in their travel habits. The 25% target can be modified to correspond to actual behaviors identified early in the data collection.

The next set of metrics considers perceptions of key aspects of the CS service in the post-deployment surveys, on factors such as convenience, affordability and need to transfer, as well as the RAPUUD questions on aspects such as safety and mental and physical effort required. These will be asked during the post-deployment surveys, with a target of average scores above neutral on a 5-point Likert scale. The questions will be asked separately for the SDS and HDS systems, and asked only for those CTP users who have made trips using the specific type of shuttle. It will be interesting to see if the perceptions regarding the SDS change across the post-deployment surveys. It will also be interesting to include some questions regarding people's expected values along similar dimensions in the baseline surveys, as a comparison to the post-deployment responses and to help explain why some people may be most averse to using the SDS. Reasons for not wanting to use the SDS could also be asked directly in the baseline surveys.

Note that perceived safety is a key performance metric collected as one of the aspects of the user experience. There will also be objective information in the Shuttle Operations data about the timing and frequency of safety-related incidents and other types of service interruptions that occur in the HDS and SDS systems. Such information can be reported via the Performance Dashboard but is not designated as a separate performance metric at this stage.

As one of the final metrics, the cost-efficiency of the new on-demand transit services should be evaluated. In available information, the cost-efficiency for on-demand transit is most often quoted as the operating cost per passenger trip. Examples are Via systems in Arlington, Texas (TX) and West Sacramento, California (CA) that are reported to have operating costs per ride of \$8 to \$11, which they cite as being well below the "US national average of \$38 for traditional demand-responsive services" (Hazan, et al. 2019). RideCo (2021) report an operating cost of \$7 per trip for an on-demand shuttle service in San Antonio, TX, compared to \$11 per trip for the fixed-route transit routes that it replaced. These numbers are reported by private operators who picked promising locations to introduce services and are trying to expand their business, so they can be viewed as "best case" values. Those sources also report efficiency in terms of riders carried per vehicle per hour, with values in the range of 3.5 to 5.0. That measure depends more on demand

than “efficiency”, however, and may not be appropriate for the Phase 3 deployment period when all potential users may not yet be aware of the service and how to access it with the CTP.

Neither of the examples above use self-driven vehicles. Stopher, et al. (2021) very recently reported on the evaluation of a Waymo automated vehicle on-demand shuttle service in the Phoenix area. While the examples of surveys and analyses used in that study may be useful examples for the later phases of this study, it does not provide any measures of cost-effectiveness. We have not set a specific cost-effectiveness target at this point but will continue to investigate this during the soft launch Phase 3. At a minimum, it will be possible to compare the cost per passenger-trip for the HDS and SDS to the operating cost per trip for the various other NFTA and PAL services in the region.

The final metric in Table 7 is for the percentage of CS booking for which the earliest available pick-up time is within 30 minutes of the time of the request (even though the user may reserve a pick-up time that is later than the earliest available time). Because the CS reservation system will use a dynamic system that keeps track of all requests and because the system will have a limited capacity of 3 or 4 shuttle vehicles, the system will attempt to book multiple users for each vehicle trip. This means that the last rider that is scheduled for pick-up on a given vehicle-trip will tend to have a longer minimum wait from the time of request, all else equal. The time between request and first available pick-up may also be longer during peak demand periods. The stated target of 90% of bookings having a first available pick-up time within 30 minutes seems reasonable given the current planned capacity but will depend in practice on the level of demand across the times of day and days of the week. A more detailed analysis of the data will look at distributions of first available pick-up time across the days and weeks of the study.

In addition to these metrics, it may be important to also measure any impacts on the on-time performance and ridership of NFTA fixed route services that serve the BNMC and adjacent neighborhoods. No separate performance metrics or targets have been included, as any such effects are outside of the main scope of the project, and any relevant system data from NFTA may not be available within the schedule of the deployment. However, any relevant data available from NFTA will be presented and analyzed in the final Phase 3 project report.

## 3.2 Correspondence between PMs and Use Cases

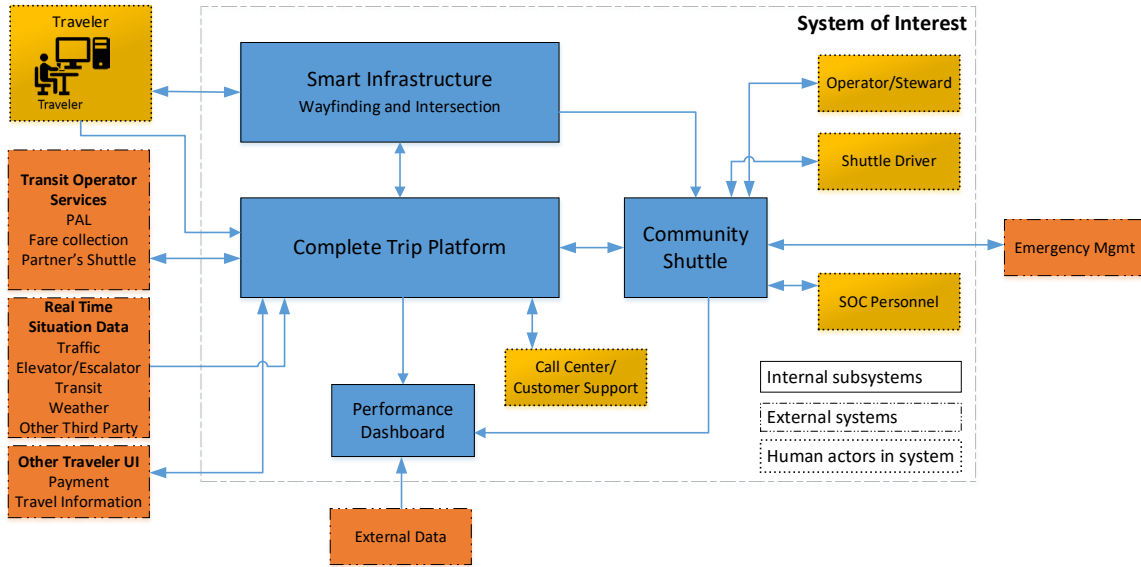
Table 8 provides a one-to-one correspondence between the six performance measures listed in Section 3.1 and the ten use cases listed in Section 2.2. Each of the use cases has at least two associated performance measures.

**Table 8. Correspondence of Recommended PMs to Use Cases.**

Performance Measure	UC1 CTP Reg- istration	UC2 CTP Planning and Booking	UC3 CTP Public Transport Trips	UC4 CTP Navigation	UC5 CTP Reporting and History	UC6 Shuttle Reserva- tions	UC7 Self- Driving Shuttle	UC8 Human- Driven Shuttle	UC9 Manage Incidents	UC10 Inter- section Crossing
1- Improved ability of the CTP users to make satisfactory Complete Trips in the study area	X	X	X	X	X	X	X	X	X	X
2- Usefulness of the CTP registration and trip preferences process	X	.	.	.	.	.	X	X	.	.
3- Usefulness of the CTP trip planning and booking process	.	X	.	.	X	X	X	X	.	.
4- Improved ability to find destinations efficiently using the CTP wayfinding functionality	.	X	.	X	.	.	.	.	.	.
5- Improved ability to cross specific intersections safely using CTP smart-signal functionality	.	.	.	X	.	.	.	.	.	X
6- Provision of an efficient, reliable and safe new on-demand transit shuttle system	.	X	X	.	.	X	X	X	X	.

### 3.3 Relationship between Performance Measures and Technologies/Services/Components

As depicted in Figure 2, the system of interest (SOI) consists of four main components: (1) the Complete Trip Platform (CTP), (2) the Community Shuttle (CS), (3) the Smart Infrastructure (SI), which includes indoor and outdoor wayfinding technology as well as smart intersection signal controls, and (4) the Performance Measurement Dashboard, sometimes also referred to as the Performance Dashboard or PMD.



**Figure 2. Context Diagram of the Buffalo, NY ITS4US Deployment System**

Source: Buffalo, NY ITS4US

Table 9 shows how the six Performance Measures correspond to the four main system components. All the measures correspond to the Performance Dashboard, as the measures will be calculated and presented via that component. All the measures also correspond to the CTP, as the CTP will be involved in the use of the Community Shuttle and the Smart Infrastructure. Three of the measures focus more closely on use of the Community Shuttle, while two focus on the use of the Smart Infrastructure.

**Table 9. Correspondence of Performance Measures and System Components**

Performance Measure	Complete Trip Platform (CTP)	Community Shuttle	SI: Indoor and Outdoor Wayfinding	SI: Smart Intersection Signal Control	Performance Dashboard
1- Improved ability of the CTP users to make satisfactory Complete Trips in the study area	X	X	X	X	X

Performance Measure	Complete Trip Platform (CTP)	Community Shuttle	SI: Indoor and Outdoor Way-finding	SI: Smart Intersection Signal Control	Performance Dashboard
2- Usefulness of the CTP registration and trip preferences process	X	.	.	.	X
3- Usefulness of the CTP trip planning and booking process	X	.	.	.	X
4- Improved ability to find destinations efficiently using the CTP wayfinding functionality	X	.	X	.	X
5- Improved ability to cross specific intersections safely using CTP smart-signal functionality	X	.	.	X	X
6- Provision of an efficient, reliable, and safe new on-demand transit shuttle system	X	X	.	.	X

### 3.4 Potential Constraints

One potential constraint may be the ability to get enough people to participate as CTP users, both overall and within specific population groups of interest. The target population is limited both in terms of geography and in terms of the types of people who would benefit the most from the system. Also, groups such as people with low-income, travelers with disabilities, and limited- or non-English-speaking persons tend to be difficult to recruit (“hard to reach”) for most types of surveys and research. Our proposed approach for addressing these constraints is described in more detail in the Section 5.2 Experimental Design, but in general terms consists of two main features:

- Use of targeted convenience sampling:** Rather than trying to get a representative sample of the full population, which would not be efficient for this project, the approach is to obtain sufficient participants within each key user group (type of disability, low income, residents of key neighborhoods, etc.) by contacting them in the most efficient way possible, such as through employers and intercept surveys at the BNMC, and through community-based organizations in the surrounding neighborhoods. As an example, we will want to be sure that the sample includes people who use the intersections with smart

signals, and that at least some of those people have potential difficulties crossing intersections, such as people with visual disabilities or those who use a mobility aid. Chapter 7 contains more detail on sample composition and sample sizes for the research design.

- **Use of incentives:** It is typical that monetary incentives are used in travel-related surveys to compensate people for the time required in participation and have been found to increase participation rates substantially. As this is a project for which we hope to retain participants for as long as possible over a period of several months, it may be most efficient to pay any incentives on a per-survey basis—once for each of the four intended surveys (one baseline and three post-deployment). Incentives can be paid via cash, check, or (most commonly) gift cards. Another possible type of incentive that has been suggested would be to give participants free travel on transit during the project. This may help to boost participation but would also be likely to confound the data on system performance, as discussed in Section 4. For this reason, free travel was ruled out as a type of incentive.

Although the target sample size of 500 study participants for Phase 3 remains desirable, the feasibility of reaching that target remains somewhat in question given the limited geographical size of the study area and the very targeted nature of the sampling towards specific types of travelers who visit the BNMC campus and/or access the NFTA light rail system. On-the-ground research and outreach to better gauge potential sample sizes was planned for Phase 1 but was not possible due to the COVID-19 pandemic. The project team is anticipating a final Phase 3 sample size in the range of 300 to 500, which would be adequate to measure system performance for key population segments.

A second potential constraint is respondents' data privacy concerns. Data privacy is primarily a function of the data privacy policies used in the CTP app and system. CTP users will have to opt into a data privacy agreement, and any data available from the CTP to measure system performance will need to adhere to that agreement. For example, it will not be possible to obtain exact locations of trip ends or trip paths, but it will be possible to obtain locations such as the part of the study area that each trip begins and ends. The current "zone" system includes 12 zones in the study area—8 separate areas within the BNMC, as well as each of the 4 adjacent neighborhoods (Fruit Belt, Masten Park, Allentown, and Bryant). Outside of the study area, spatial data on trip ends will be at the zip code level. Further steps are being taken to protect personally identifying information (PII) at each stage of data collection (including surveys) and processing, and PII will not be included in any of the performance metrics available via the Performance Dashboard.

Other constraints on gaining user participation and data could be a resurgence of the COVID-19 pandemic, or more transitory interruptions such as particularly long and/or intense winter storms, or renewed civil unrest, particularly around the time of national elections. These types of potential "interruptions" are discussed in the next chapter.



# 4 Confounding Factors and Mitigation Approaches

This section describes and discusses anticipated confounding factors that may affect the performance measures and associated analyses. The identified confounding factors and mitigation approaches will be reviewed and updated when the PMESP is updated based on the schedule proposed in Section 9.

## 4.1 Use of Non-Probability Sample Selection

The objective in recruiting participants for this project will not be to obtain a proportionally representative sample, but rather to obtain an adequate sample within each key user group as efficiently as possible. Use of targeted, non-probability sampling methods means that the distribution of the study sample across key dimensions (disability type, income level, race/ethnicity, residence area, etc.) may not match the distribution of the full populations of likely users of the system. To address this issue, the study participants will be weighted in analysis so that the weighted distributions match those of the target population as closely as possible. Such weighting is common in sample-based studies, but there are two aspects of this study that complicate the weighting process. First, the target population for the CTP system is not representative of the regional population as a whole, as it is focused on a particular geography centered around the BNMC and on particular sub-populations, such as those with disabilities. Second, the data that is usually used for weighting, such as the American Communities Survey (ACS), does not contain detailed data on disabilities, particularly on the type of disability. To address these complications, other data sources such as visitor data collected by BNMC, can be used to supplement the ACS data, so that the weighted data represents the full target population as closely as possible, given the limitations on sample size and demographic data. The preferred method for survey weighting is an entropy maximization method to minimize the variance in the resulting weights (Barratt, et al. 2020), which has been implemented in available software packages. In this study, given the limited sample size and limitations on data regarding the actual size of the target population groups, the weighting process will be applied and weighted results will be compared alongside unweighted results to determine the influence of weighting.

## 4.2 Attrition Bias in Participation

Another confounding factor is the desire to measure each participants' behavior over time, rather than taking just a single "snapshot" of each person at one point in time. Ideally in Phase 3, all participants will be recruited within the first six months, and then each participant would remain in the study until the end of Phase 3 and provide at least a full 12 months of data. In practice, there will certainly be some amount of "attrition", with some participants no longer participating by the time that Phase 3 ends. If attrition were a purely random unbiased process, it would not present an issue for analysis. However, there is a possibility that some bias will be involved, with those users who use the CTP least often and/or who find it to be the least useful being the most likely to stop participating in the study.

One way of mitigating this factor will be to prevent attrition as much as possible, by keeping participants engaged by means of frequent contact and by providing incentives for participation that are based on the length of time that people remain in the study. Some level of attrition is inevitable, however, so any attrition bias should be accounted for in analysis when computing performance measures. A simple method for doing this is to adjust participants' weights over time so that the weighted sample continues to represent the same total target population. A more sophisticated variation of that method (Brownstone and Chu, 1997) is to model the probability of dropping out of the study as function of respondent characteristics, including their use of the CTP system up until that point. Using the probabilities from applying such a model, the sample weights can be adjusted over time in a way that adjusts for attrition bias with respect to usage of the system, as well as other factors such as disability type, income, and demographics. If the attrition rate can be kept below a reasonable threshold such as 20%, then this level of bias adjustment would not be needed. If people drop out of the study early on in Phase 3, an approach is to replace the participant with a new participant in the same population sub-group.

### **4.3 Influence of Incentives on Travel Behavior**

As discussed in Section 5, we anticipate that it will be necessary to offer monetary incentives to obtain and maintain an adequate sample size for this study over the duration of the project. There is a possibility that offering incentives will encourage participants to use the system more than they would otherwise, with the perception that they are "being paid to use the CTP". It will be stressed to participants that as long as they respond to periodic surveys to show that they are still participating and they are providing truthful responses, receiving the incentive is not conditional on how frequently they use the CTP. An important use of the periodic surveys can be to ask people who are not using the CTP why they are not using it. With this understanding, the major reason for using the CTP will be the usefulness of the system, and not the incentive for participating, and the outreach and recruitment of participants will stress that fact. With that in mind, a monthly prize drawing has been selected as the form of incentive to be used. The prize drawing approach is less expensive to fund and administer than direct monetary incentives.

It has been suggested that a useful additional incentive for participation would be to offer free travel on transit for some period of time, whether that is only the CS or for NFTA/PAL services as well. This type of incentive is potentially more problematic because it provides a reason for participants to use transit more often, which would directly affect some of the metrics proposed in Section 3, for reasons that are not related to the benefits of the CTP system. If all participants were offered this incentive for the entire study period, there would be no way to measure the incremental effect of offering it. For these reasons, we have ruled out using free transit as an incentive for participation.

### **4.4 Extreme Weather Conditions**

The Buffalo region is susceptible to heavy snowstorms coming across Lake Erie, sometimes accompanied by high winds and very low temperatures. Such storms can occur between November and April and can affect travel conditions for up to several days. Snow or ice can make walking on streets and sidewalks treacherous, particularly for people with visual disabilities and/or using mobility aids, as well as older adult age groups. Snow and ice can also cause road closures and interruption of transit services and may cause building closures and interruption of services on the BNMC campus. Extreme weather can also take the form of summer heat waves with high humidity, which can be dangerous for older populations especially. When calculating performance

measures, it will be important to use supplemental data on weather conditions and related events such as road closures, transit service interruptions and building closures, especially in the BNMC and surrounding areas. Daily weather data can include maximum and minimum temperatures (including wind chill factor), average humidity level, and amount of rainfall and snowfall. An indicator of icy conditions would also be useful, if available. The most typical way that this data will be used for analysis will be to screen out days during which travel in the area is infeasible for many people, usually indicated by road closures, building closures and school closures. Another way that the data can be used is to classify days during which travel is still feasible (i.e., the campus is open and medical appointments are not cancelled) but travel conditions are treacherous for vulnerable populations. Although we have not included separate performance measures for such days, it will be informative to analyze the usage of the system on those days versus days with milder weather and “normal” travel conditions.

## 4.5 Progress on the Middle Main Road Improvement Project

The Smart Signals that can be remotely activated for intersection crossing are proposed for two key intersections: Ellicott & High and Best & Main. The Best & Main intersection is in the area for the proposed Middle Main Street road improvement project, which had been largely funded and through the design phase (<https://www.buffalony.gov/CivicAlerts.aspx?AID=1212>) as of February 2023. It will be a major construction project, including the addition of bicycle lanes and pedestrian-related improvements. The Ellicott & High intersection is not in the area of this project, so it is expected that the installation and testing of the smart signal technology can proceed as planned at that intersection during Phase 2 and be ready for Phase 3. There is some risk, however, that the Middle Main Street project construction will not be completed before Phase 3, so that the installation of the smart signals at Main & Best will need to be delayed and/or the ongoing construction will impede pedestrian crossings at the intersection.

We cannot recommend a full mitigation strategy for any delays due to ongoing construction of the Middle Main Street project. The best strategy is to install the smart signals at Best & Main, and to monitor and record the week-to-week progress of the project using variables that can be related in analysis to pedestrian usage of the intersection. Frequent video documentation of the intersection conditions will be a good source of such data. As a last resort, another intersection could be chosen instead of Best & Main, but there are no other intersections that are as critical in terms of pedestrian access to BNMC.

Due to delays in construction at the Best & Main intersection, the Smart Signal technology (PED-X) is being implemented only at the Ellicott & High intersection during this project. The user survey questions have been updated accordingly.

## 4.6 The Availability of Automated (Driverless) Shuttle Technology that Meets Operating Regulations

As described in more detail in other project documents, such as the SyRS document, there are several different technology options for the driverless shuttle, some of which do not meet current New York State operating requirements. After completing a review of technology options, ADASTEC's vehicle technology has been selected. The ADASTEC vehicle platform has a steering wheel and satisfies all Federal Motor Vehicle Safety Standards. Given this, the vehicle is

eligible for obtaining DMV's approval for an Autonomous Vehicle (AV) Demonstration on public roads, under New York State's (NYS's) current regulations for AV testing and demonstrations. Over the coming months, ADASTEC will be working on submitting an application to NYS DMV, requesting the issuance of a permit allowing them to operate autonomously on public roads, but with a human safety steward behind the wheel, as currently required by NYS law.

There are several stages planned for the testing of the Self-Driving Shuttles. The project team is currently working with ADASTEC to obtain documented safety and testing results from ADASTEC's prior deployments (e.g., in Michigan, Norway or Turkey). Second, before real-world deployment of the SDS on BNMC, ADASTEC will be conducting testing of the Autonomous Driving Software, in simulation, and using a simulated route of the planned deployment site on BNMC. Third, for the last three months of Phase 2, field testing of the SDS will be conducted on BNMC, without live passengers first and then with project participants / project owners onboard who would then provide feedback to the vendor regarding how comfortable the ride was, and whether they felt safe and secure during the ride. Satisfactory performance of the SDS during all these testing stages will be a requirement for operations to start in Phase 3. The SDS is in operation and was demonstrated as part of the Operational Readiness Demonstration (ORD) in September 2025.

## 4.7 Other External Factors

There are a wide variety of other factors that could confound the analysis of performance metrics. These include both episodic events and slower-moving trends, as listed below:

- Temporary opening or closure of major roads (particularly near BNMC)
- Temporary changes to NFTA transit service levels (particularly near BNMC)
- Temporary closing of buildings or facilities within BNMC
- Floods, hurricanes, tornadoes, earthquakes or major fires. Note that these are historically very rare in the Buffalo region
- Riots and/or curfews due to social unrest, elections, etc.
- Imposition of new social distancing regulations or other measures in response to a resurgence of COVID-19 or other health threats

Related trends:

- Trends in traffic levels and congestion on major roads (particularly near BNMC)
- Trends in visitation levels and employment levels at BNMC
- Trends in pedestrian and bicycle traffic at key intersections in/near BNMC
- Resident demographic trends in the neighborhoods surrounding BNMC
- Number of new cases and hospitalizations for COVID-19 or other emerging health threats

We do not currently anticipate any permanent substantive changes in the road system or NFTA or PAL transit systems in the BNMC area, so any occurrences in the upper list would likely be temporary, episodic events. Data on these types of events and trends are generally available from the appropriate agencies including NFTA/PAL, the City of Buffalo (traffic and pedestrian counts), University of Buffalo/BNMC, and the Census Bureau (ACS block group level estimates). Such data can be synchronized over time with the other data sources (CTP system data, weather data, etc.) and analyzed to look for correlations with performance measures that are monitored over time, similar to how weather data will be used in analysis.

Data on unpredictable events such as epidemics, social unrest or natural disasters can be added as needed. These may need to be treated differently depending on their severity as they could even lead to the postponement or interruption of the study. Other than these, the mitigation approaches for each confounding factor have been described above within this Section 4.



# 5 System Deployment Impact Analysis Design

This section describes the design for the ITS4US deployment impact analysis. It has been formulated as an outcome evaluation. An outcome evaluation measures the impact that an intervention or project has on program participants, service users, or a neighborhood or community (James Bell Associates, 2008). It is often the only way to demonstrate that a project has accomplished its intended objectives and has had an effect on participants.

## 5.1 Approach/Strategies for Focused Performance Analysis

The approach that will be used to focus the performance analysis is one that links expected outcomes based on use cases with performance measures, their associated indicators, which then drives the data required to populate those indicators, and sources for those data. Through the design of logic models, the 10 use cases are associated with expected outcomes. Outcomes are:

- More efficient Complete Trip planning and execution
- Enhanced PAL access and support for non-PAL trips
- Enhanced indoor/outdoor wayfinding to/from/on BNMC campus
- Trip planning and navigation functions improved through crowdsourced data
- Improved destination accessibility through shuttle service
- Safe and convenient shuttle use
- Positive perceptions of travel by SDS and HDS
- Safer intersection crossing.

These outcomes are both appropriate (i.e., the deployment could truly achieve the outcome) and measurable (i.e., there is a way to determine if the outcome has been achieved). Also, each outcome and use case are linked to the six performance measures identified in Section 3 and their associated metrics—see Table 10.

**Table 10. Outcomes, Use Cases, and Performance Measures**

Outcomes	Use Cases	Performance Measures
More efficient Complete Trip planning and execution	1- Register Profile and Preferences 2- Generate Trip Plan and Book a Trip	1. Ability to make Complete Trips in the study area 2. Usefulness of the CTP registration and trip preferences processes
Enhanced PAL access and support for non-PAL trips	3- Public Transport Services	3. Usefulness of the CTP trip planning and booking processes
Enhanced indoor/outdoor wayfinding to/from/on BNMC campus Trip planning and navigation functions improved through crowdsourced data	4- Navigation 5- Reporting and History	4. Improved ability to find destinations efficiently using the CTP wayfinding functionality
Improved destination accessibility through shuttle service Safe and convenient shuttle use Positive perceptions of travel by SDS and HDS	6- Shuttle Service Reservation and Dispatch 7- Reporting and History 8- Passenger Pick-up, Securement and Drop-off via the SDS/HDS 9- Manage Incidents	6. CS is an efficient, reliable, and safe mode of transportation
Safer intersection crossing	10- Intersection Pedestrian Crossing (PedX) Request	5. Improved ability to cross specific intersections safely

With the links between outcomes, use cases, and PMs determined, it is possible to specify data requirements and analytical approaches for the evaluation. Table 11 presents this information.

**Table 11. Summary of Data Requirements for each Performance Measure**

PM #1: Improved Ability to Make Complete Trips in the Study Area

MEASURE 1	Improved Ability to Make Complete Trips in the Study Area
Description	Ensuring that there is a safe, accessible path, compatible with user-defined preferences and capabilities to/from/within the BNMC and other desired origins and destinations. This requires real-time information on end-to-end travel options and infrastructure usability and accessibility
Hypothesis	The ability of system users to make Complete Trips will improve from the pre-deployment baseline situation.

<b>MEASURE 1</b>	<b>Improved Ability to Make Complete Trips in the Study Area</b>
Metrics	1- Accessibility ratings of door-to-door travel for trips to/from/within the BNMC 2- Safety ratings of door-to-door travel for trips to/from/within the BNMC 3- Rating of the adequacy and usefulness of information on making trips to/from/within the BNMC 4- Rating of the ability to make trips using integrated transit services to/from/within the BNMC 5- Rating of the accessibility of trips made using the CTP app
Data Needs	Average ratings on a Likert scale
Data Sources	(1-4) Surveys of system users, 5- User data from the CTP
Design	(1-4) Baseline and Post-Deployment Surveys, 5- CTP data extraction
Modeling/Tools	Users compared pre-and-post. Descriptive statistics of individual-based change by respondent type and user/non-user across data collection waves; chi-square tests
Risks	The respondent ID is missing so an individual cannot be tracked across data collection waves; there are too few users of the system. CTP functionality does not perform as planned

PM #2: Usefulness of the CTP Registration and Trip Preferences Process

<b>Measure 2</b>	<b>Usefulness of the CTP Registration and Trip Preferences Process</b>
Description	This addresses user needs around whether the CTP registration process is feasible to complete for the great majority of (potential) users, and whether the trip preference options that are presented are deemed as relevant and useful inputs for planning their travel
Hypothesis	People will want to register in the CTP because of its perceived usefulness
Metrics	1- Users' ratings of ease and convenience of registration process 2- Users' ratings of usefulness of the preferences input process
Data Needs	Average ratings on a Likert scale
Data Sources	Survey of target populations
Design	Post-Deployment Surveys. There is no baseline measure.
Modeling/Tools	CTP Users only. Descriptive statistics of ratings data by respondent type across data collection waves; chi-square tests
Risks	The respondent ID is missing so an individual cannot be tracked across data collection waves; Not enough recruited individuals register in the CTP; CTP functionality does not perform as planned.

PM #3: Usefulness of the CTP Trip Planning, Booking, Reporting and Review Process

<b>Measure 3</b>	<b>Usefulness of the CTP Trip Planning, Booking, Reporting and Review Process</b>
Description	This addresses user needs around the trip planning and booking processes
Hypothesis	People will plan and book trips in the CTP because of its utility and ease of use

Measure 3	Usefulness of the CTP Trip Planning, Booking, Reporting and Review Process
Metrics	1- Users' ratings of ease of planning a door-to-door trip 2- Users' ratings of satisfaction with the specific route/mode options provided by the CTP 3- Percent of CTP users who use the system to book on-demand (CS) transit trips 4- Users' ratings of the ease and convenience of booking on-demand (CS) transit trips 5- Percent of CTP users who use the CTP to report incidents and/or travel conditions 6- Users' ratings of the ease of reporting incidents and travel conditions in the CTP 7- Percent of CTP users who use the function to review past trip history 8- Users' ratings of the usefulness of reviewing their trip history in the CTP
Data Needs	Self-reported usage of app features and average ratings on a Likert scale
Data Sources	Surveys of target populations, supplemented with CTP user data for metrics (3) and (5)
Design	Post-Deployment Surveys. There is no baseline measure.
Modeling/Tools	CTP Users only. Descriptive statistics of frequency and ratings data by respondent type across data collection waves
Risks	The respondent ID is missing so an individual cannot be tracked across data collection waves; Not enough recruited individuals register in the CTP; Not enough registered individuals book a trip in the CTP; CTP functionality does not perform as planned

PM #4: Improved ability to find destinations efficiently using the CTP wayfinding functionality

Measure 4	Improved ability to find destinations efficiently using the CTP wayfinding functionality
Description	This addresses user needs around outdoor and indoor mobility and wayfinding for travelers in the target groups
Hypothesis	Wayfinding functionality both outdoor and indoor will be perceived as useful support for efficient trip making
Metrics	1- Fraction of CTP users who elect to receive outdoor wayfinding notifications 2- Self-reported frequency of using outdoor wayfinding functionality (1) Fraction of CTP users who elect to receive indoor wayfinding notifications 3- Self-reported frequency of using indoor wayfinding functionality 4- Users' ratings of usefulness of outdoor wayfinding functionality for reaching destination on time 5- Users' ratings of usefulness of indoor wayfinding functionality for reaching destination on time 6- Users' rating of outdoor wayfinding functionality on RAPUUD items 7- Users' rating of indoor wayfinding functionality on RAPUUD items
Data Needs	Self-reported usage of app features and average ratings on Likert scale
Data Sources	Surveys of target populations, supplemented with CTP user data for metrics (1) and (3)
Design	Post-Deployment Surveys. There is no baseline measure.
Modeling/Tools	CTP Users only. Descriptive statistics of frequency and ratings data by respondent type across data collection waves

<b>Measure 4</b>	<b>Improved ability to find destinations efficiently using the CTP wayfinding functionality</b>
Risks	The respondent ID is missing so an individual cannot be tracked across data collection waves; Not enough recruited individuals register in the CTP; CTP users do not use the wayfinding functionality; CTP wayfinding functionality does not perform as planned

PM #5: Improved ability to cross specific intersections safely using the CTP smart-signal functionality

<b>Measure 5</b>	<b>Improved ability to cross specific intersections safely using the CTP smart-signal functionality</b>
Description	This addresses user needs around outdoor mobility safety at intersections, particularly for travelers who have visual impairments, who are in wheelchairs, or older adults who need more time crossing
Hypothesis	At those intersections where it is deployed, the smart-signal functionality will facilitate safer crossing and so will be used
Metrics	1- The percentage of CTP trips crossing at the relevant intersections that use the remote activation function. 2- Self-reported fraction of people who cross at the relevant intersections that use the remote activation function. 3- Users' ratings of ease of use and other aspects of smart signal activation using the RPUUD method 4- Self-reported frequency of crossing at the intersections with smart signal functionality. 5- Users' ratings of perceived safety of smart signal activation functionality
Data Needs	Objective and self-reported frequency data and average ratings on Likert scales
Data Sources	CTP usage data for (1), surveys of target population for all others
Design	Post-Deployment Surveys for (2-5). Comparison against Baseline surveys for (4) and (5).
Modeling/Tools	CTP Users only. Descriptive statistics of frequency and ratings data by respondent type across data collection waves for (1-3). Descriptive statistics of individual-based change by respondent type across data collection waves for (4) and (5).
Risks	The respondent ID is missing so an individual cannot be tracked across data collection waves; Not enough recruited individuals register in the CTP; Not enough people cross/use the instrumented intersections; CTP smart intersection functionality does not perform as planned

PM #6: Provision of an efficient, reliable, and safe on-demand transit shuttle system

<b>Measure 6</b>	<b>Provision of an efficient, reliable, and safe on-demand transit shuttle system</b>
Description	This addresses the operations of both the HDS and SDS
Hypothesis	The CS will provide a level of service that will encourage usage

Measure 6	Provision of an efficient, reliable, and safe on-demand transit shuttle system
Metrics	1- Percent of CS trips that arrive at the boarding stop on time 2- Percent of CS trips that arrive at the alighting stop on time 3- Users' ratings of reliability of transit to reach destinations on time 4- Self-reported frequency of using the CS and other local transit 5- Percent of transit trips by PAL-eligible users made via the CS 6- Users' ratings of CS service using RPUUD items, including safety 7- Cost efficiency of HDS and SDS shuttle service 8- Percent of CS bookings for which the earliest pickup time is within an acceptable time from when the trip is requested.
Data Sources	(1), (2), (5), (7) Shuttle operations data. (3), (4), (6) Surveys of target population, (7) Additional system cost data
Design	Post-Deployment Surveys for (3), (4), (6). Comparison against Baseline surveys for (3) and (4).
Modeling/Tools	CTP Users only. Descriptive statistics of individual-based change by respondent type across data collection waves for (3) and (4). Descriptive statistics of frequency and ratings data by respondent type across data collection waves for all others.
Risks	The respondent ID is missing so an individual cannot be tracked across data collection waves; Not enough recruited individuals register in the CTP; Not enough people use the CS.

## 5.2 Experimental Design

This section describes the pre-post design that will guide data collection and analysis.

### 5.2.1 Pre-Post-Design with Comparison Group

The project team will conduct a pre-post design because it is known that “comparison” is necessary for documenting the changes that occur as the result of the deployment. The design is a pre-post-post-post because we are proposing to do three waves of post-deployment data collection, but for simplicity we will use the term “pre-post” in the following text. The pre-post design involves identifying an “event” that marks the beginning of an individual’s participation in the program. In this case, that event is when the individual registers in the CTP system during the Phase 3 deployment. Data about each participant’s pre-deployment travel patterns and perceptions are then collected just after CTP registration, which is the baseline assessment. Data is then collected during and after the deployment from the same participants. The follow-up data are then compared to the baseline data to identify whether each participant’s travel patterns and perceptions changed relative to the baseline pre-deployment data (as well as relative to earlier post-deployment surveys).

In Phases 3 there are four proposed surveys administered to each registered system user approximately three to four months apart, one happening pre-deployment and three waves of survey data collected post-deployment. See below for suggested flow.

- Pre-deployment

- Baseline: collected immediately following each CTP user's registration in the system as a basis for comparison; survey will ask users to respond to questions recalling their experiences prior to use of the CTP
- Post-Deployment
  - Initial: About 4 months after the person has registered to use the system
  - Mid-Term: About 8 months after the person has registered to use the system
  - Final: About 12 months after the person has registered to use the system

The ideal case would be for each participant to complete each survey. In practice, however, some degree of non-response is inevitable. For each survey, each participant will receive up to three e-mail reminders approximately one week apart, asking them to complete the survey if they have not already done so. Thus, the data collection period will last up to one month from the time a survey invitation is sent out. Not all participants will register at the same time however, and the surveys are triggered relative to the time of registering to use the system, so the exact schedule for sending out the survey invitations will vary across respondents.

In this design, we are interested in the amount of change that has occurred in terms of performance measures due to the deployment of the ITS4US system. Not all people recruited to participate in the evaluation study will end up being users of the ITS4US system. Those that do not use the system will be a comparison group (by default) on pre-post metrics against those that do use the system. Those that do not use the system will also be asked as to their reasons for not using it, although those data are not listed as metrics, as they are qualitative in nature.

The pre-post metrics of change apply to four of the metrics in PM#1, as well as two metrics in PM#5 and two metrics in PM#6. The other metrics are based only on post-deployment data, as they measure the usage and perception of specific features of the system—features which did not exist during the pre-deployment period.

## 5.2.2 Target Populations

The Buffalo ITS4US deployment seeks to serve the specific transportation needs of the following underserved groups:

- People with disabilities, including mobility/wheelchair users, visually impaired individuals, and hearing impaired individuals
- Low-income individuals
- Older adults (age 65+)
- Limited- or non-English proficient (LEP) persons
- Caregivers for people fitting the above groups
- People who work at the BNMC and who visit to receive medical services.

Individuals that fit one or more of these groups will be recruited to participate in the evaluation study. Individuals will be recruited to participate in the study using non-probability sampling strategies as described in Section 7. This is because the population groups of interest are limited in adequate numbers using probability sampling strategies such as address-based sampling (ABS) or random digit dialing (RDD). By design, the sampling strategies used will capture BNMC workers and visitors and residents of nearby neighborhoods.



# 6 Support to Independent Evaluation Effort

Over the course of Phases 2 and 3 of the program, the project team will provide access to all drafts and final reports to the independent evaluator (IE) for review and possible comment. This will be accomplished in the normal course of submission of the documents to ITS JPO. Performance against baseline measurements and targets will be publicly reported throughout Phase 3 in the form of summaries/dashboards of key measures. This information will be available to the IE to use in the Final Performance Results Assessment (which is a report Volpe will prepare during Phase 3 of the deployment). Any additional information requested by the IE such as data sources or availability of baseline data or items of clarification, will be provided, in cooperation with the ITS JPO.

In addition to providing access to site deliverables and performance measurements, the IE will conduct two sets of interviews, one during pre-deployment and one during post deployment. The project team will identify individuals for the interviews and provide contact information and possibly introductions, with the understanding that the IE will coordinate the logistics of the interviews. The project team will also identify key deployment partners (deployment agency staff, technology partners, universities, government or policymakers, or others). As of this writing, Table 12 represents the IE’s current draft plan for these interactions.

**Table 12. Plans for Independent Evaluator Interactions** (Source: Volpe IE)

Participant Role	Pre-Deployment Interviews	Post-Deployment Interviews	Questionnaire(s) - TBD
Time Frame	Phase 2 Sept-Oct '23	Phase 3 TBD	Phase 3 Date: TBD
Federal Program Managers (ITS JPO, FTA, other)	X	X	.
Deployment Managers (e.g., lead agency and other decision makers)	X	X	.
Deployment Partners (deployment agency staff, technology partners, universities, government or policy makers, others)	X	X	.
Project Stakeholders (transit agencies, healthcare organization, social service agencies, community organizations, etc.)	.	.	X
Government Entities (FTA Regional Offices, FHWA Resources Center, FHWA Division Offices, State DOT, County, City, etc.)	.	.	X

In subsequent drafts of the PMESP, this Section can serve as a repository of information regarding any agreements/decisions, documentation, and scheduling between the Buffalo site and the IE, with coordination by the COR and ITS JPO.

# 7 Data Collection Plan

The Data Collection Plan fits within the Data Curation Process and the Data Sharing Framework (see Section 7.7 Data Sharing Framework). The Data Sharing Framework (as well as the Metadata Update process as described in the DMP, Section 7) is driven by the data curation processes for each dataset, including generating the dataset formats and their appropriate metadata files. Each dataset will follow a similar curation process when collected or acquired by the system. Separate protocols for each dataset are being developed to ensure data governance rules and procedures are followed. Appropriate processing will be applied to ensure the data integrity, quality and security provisions. Data schema and storage systems are being developed and tools to load the data without loss of data fitness. The data integration processes implement the performance measurement analysis. Data is organized into views for each access and longevity studies (e.g., before / after analyses). The last step in the data curation process is readying the data sets for distribution. Description of each stage in the curation is shown in Figure 3 and described as follows:

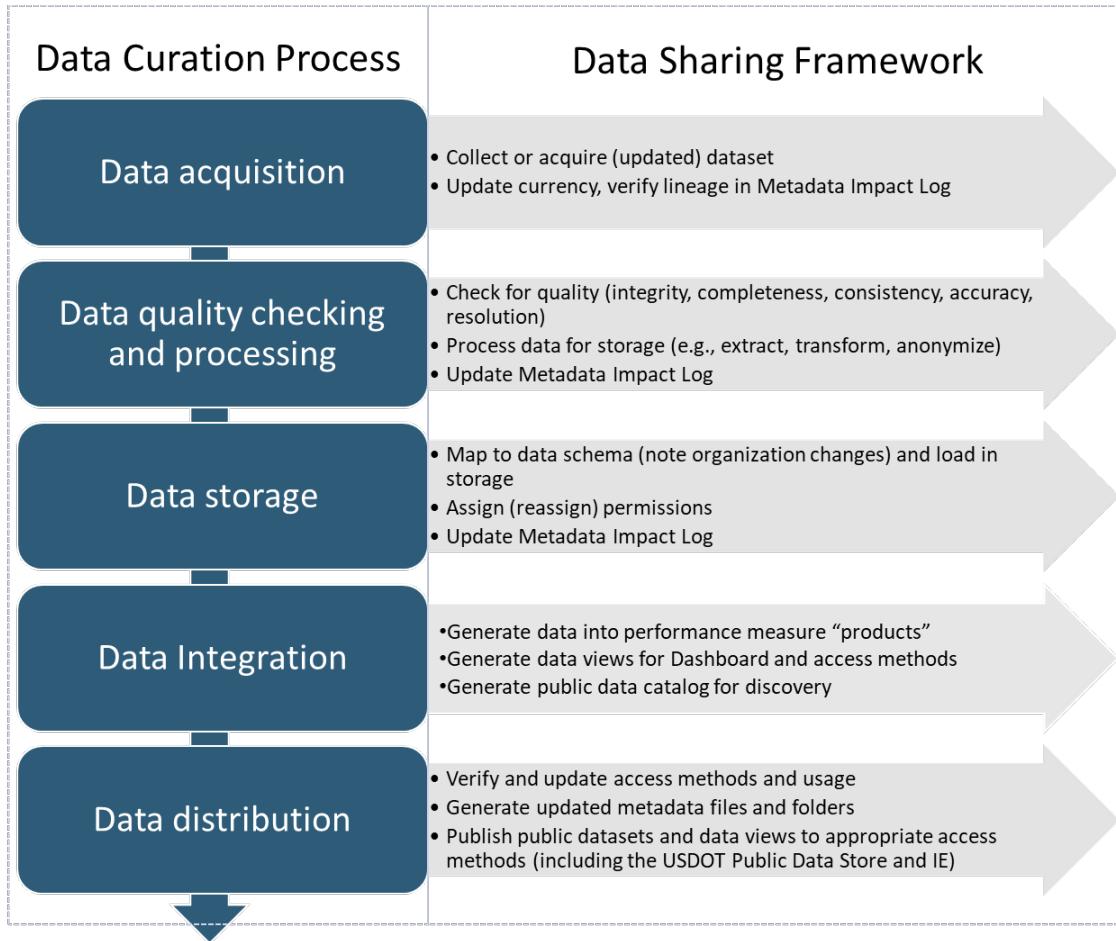
**Data acquisition:** data is collected, imported or acquired through surveys, operations or sensors.

**Data quality checking and processing:** data is checked for quality (integrity, completeness, consistency) as described in Section 7.6. Each dataset will undergo its own quality checking procedures. Data will then be extracted and transformed for loading into the appropriate data store. Data will be extracted and processed (e.g., anonymized or aggregated) for the appropriate permission level prior to being transformed for its assigned data store.

**Data storage:** data is stored in their appropriate data storage systems. Operational data will reside in their respective subsystem data stores, while datasets needed to derive performance measurements and their metrics will be forwarded to the Performance Measurement Dashboard (PMD) subsystem. Within the context of the PMD, data will be partitioned into the appropriate data store systems to secure user privacy and sensitive information. Any changes to the data store schema will flag an error, and those will need to be fixed prior to loading the data into the storage system.

**Data integration:** multiple dataset and partial datasets are needed to evaluate operational fidelity and system performance. Data analysis and business intelligence processes will be applied to the datasets to analyze and evaluate the information. In addition, data will be readied and transformed into data views (virtual collection of data that supports specialized analysis), feeds and other *data products* that are used for data distribution.

**Data Distribution:** as the last part of the data curation process, the datasets available for public and permission-based access are listed in separate data catalogs for discovery. Their metadata files are also associated with their dataset. The datasets, metadata and data catalogs are published for role-based access. The access method(s) and frequency of dataset publication used by the project team, USDOT Public Data Store and IE was developed during Phase 2.



**Figure 3. Data Sharing Framework**

Source: Buffalo, NY ITS4US

This curation process puts the data collection plan in context. The Data Needed (Section 7.1), Baseline Data Collection (Section 7.3) and Deployment Data Collection (Section 7.4) describe the performance measure, datasets and their timing. It identifies what data products are needed to enable metric management over time. These data products then provide information on what types of data quality checking and procedures (e.g., integrity, completeness and consistency) that need to be applied in the Data Quality Check Approach (Section 7.6). The final stage of the curation process identifies what data needs to be published and provided to the USDOT public data store and the IE staff so they can validate and evaluate the project against the goals and objectives.

## 7.1 Data Needed

Data is needed to populate the performance measures. The requirements for doing so are summarized in Table 11 in Section 5.1 for each performance measure.

The required data to evaluate the Buffalo NY ITS4US system deployment will be gathered at multiple times and for multiple purposes. All information, especially data identified as sensitive or protected information will be collected in accordance with the DMP.

## 7.2 Participant Recruitment Data Collection

Participants will be recruited to participate in the study using screening criteria to match the target population groups for the study:

- People with disabilities, including mobility/wheelchair users, visually impaired individuals, hearing impaired individuals, and cognitively impaired individuals
- Low-income individuals
- Older adults (age 65+)
- Limited- or non-English proficient (LEP) persons

It should be noted that caregivers for people with disabilities can also be recruited as participants.

Recruitment will occur early in Phase 3, prior to the point of the full deployment in Phase 3 (go-live). Three types of non-probability sampling approaches can be used to recruit participants.

- **Targeted outreach:** Recruitment will be done via community-based organizations. These include organizations based in the neighborhoods surrounding the BNMC, particularly the Fruit Belt neighborhood, as well as regional organizations such as Aspire of Western NY (WNY), WNY Independent Living Center, Inc., or the New York State offices of vocational rehabilitation. The project team also expects to sample NFTA paratransit (PAL) subscribers through the service's database. In addition, designated research staff may access the IDEA Center's database of participant contact information from past studies in order to recruit participants for the study. **By design, this sampling approach will capture people with disabilities and older adults who travel to/from the BNMC.**
- **Convenience sampling:** Interviewers can be stationed at various entrances to the BNMC campus and administer a short screening questionnaire in English or Spanish to determine whether each potential respondent fits a target population and if so, to describe the research opportunity (and any incentives for participating) and then to invite them to participate in the evaluation study. Note that this type of recruitment can also be done via staffed or unstaffed kiosks strategically located on the BNMC campus. **By design, this sampling approach will capture BNMC workers and visitors.**
- **Voluntary response sampling:** If needed, project staff can place English/Spanish door hangers on homes in the Fruit Belt, Masten Park, and Allentown neighborhoods that describe the research opportunity (and any incentives for participating) and invite people who have experienced traveling to the BNMC to volunteer themselves for the evaluation study. A toll-free phone number and a web link can be offered for people to contact the research team. This strategy may include some level of "snowball" sampling, in which recruited participants invite eligible friends or colleagues to participate. **By design, this sampling approach will capture people who live in nearby neighborhoods and travel to/from the BNMC.**

For all three sampling approaches, there will be no way to know if the resultant samples are representative of the target populations; however, there will be credible information about the

outcomes of the deployment for target population groups because of the use of an experimental design.

The plan is to recruit 300-500 people for the Phase 3 evaluation activities. The proposed composition of the sample is presented in Table 13. It is expected that many individuals will fall into more than one category, which is why the totals across the groups exceed the values of 300 to 500.

**Table 13. Proposed Composition of Phase 3 Samples.**

Sub-Sample	Disability	Older Adults	Low Income	Low English Proficiency	BNMC Staff	Total Sample
Target Sample Size	150-250	50-100	50-100	35-70	35-70	300 - 500

## 7.3 Baseline Data Collection

While not the only source of data for the identified PMs, surveys of users will provide much of the information for the performance evaluation.

Baseline survey data collection will occur when people are registered to use the CTP system. The registration process will trigger an e-mail to be sent to the user, with a link to complete the baseline survey. The baseline survey will collect the following types of information:

- Information on current trip-making to and from BNMC (frequency of travel, usual mode, usual purpose)
- Satisfaction with current travel options to BNMC (baseline performance metrics)
- Status of PAL registration and frequency of transit use
- Frequency of using NFTA transit services, and perception of the reliability of transit
- Frequency of crossing intersections where smart signals will be placed, and perception of safety at those locations
- Auto ownership and frequency of use
- Information about disabilities and mobility devices used
- Demographics (age, household size, income, employment, zip code)

Baseline data collection will also include, to the extent possible, secondary (non-survey) data that will be tracked throughout Phase 3. This will include BNMC visitor data, performance and ridership data on relevant routes collected by NFTA, PAL-direct booking data, and any traffic and pedestrian counts at relevant intersections collected by the City of Buffalo.

## 7.4 Deployment Data Collection

Deployment data collection will occur in two main forms: (1) system data provided via the CTP platform and supporting operations continuously over time, and (2) survey data collected from study participants at appropriate and opportune times. The continuous system data includes:

- **Performance Metrics from the CTP.** Data will include registrations, bookings, trip plan preferences, trips booked, trip modes used, trip times, trip origin and destination zones, and use of smart-intersection functionality. This data will be extracted at least weekly from the CTP using the applications programming interface (API) that provides access to the data. Data type: Operations.
- **Shuttle Reservations Management.** This function reports data on the processes by which travelers reserve the Community Shuttle to satisfy one leg of their trip, confirmation of the reservation, ETA of the shuttle, and location of the pick-up. Data type: Operations.
- **Shuttle Trip Reporting.** This function collects shuttle trip-related information (e.g., Shuttle utilization, hailing source, number of trips, miles traveled, shuttle operations, including incidents and delays). The function tracks problems encountered during the trip, any cases that required the human stewards to take over the control of the shuttle, and any crashes or near-crashes that were about to occur. Data type: Operations.
- **Data from the SOC.** Data include requests for service (to calculate on-time performance against actual service), time, pickup and drop-off locations, number of passengers, accessibility preferences, occupancy, and occupied securements. Data type: Operations.

Survey-based data collection will take place upon each CTP user's registration in the system, and then at three points in time after that roughly 4 months apart. Data items will include information to populate the subjective metrics for the performance measures. Data items and sources include:

- User ratings of the various CTP functionalities (e.g., registration system, trip planning, trip booking, shuttle services, smart intersections). Survey invitations will be sent out automatically upon triggering events or schedules via the CTP. Data type: Survey.
- In addition to typical user rated items, the project team will employ a modified version of the RPUUD method (described in Section 3) developed by project team members that is designed to investigate the specific needs and preferences of travelers with disabilities. Questions will be designed to obtain subjective measurements of technological features or services along up to nine dimensions (the exact ones used to depend on which features are being evaluated):
  - (1) ease of use,
  - (2) high independence,
  - (3) low safety risk,
  - (4) low risk of mistakes and "do overs",
  - (5) efficiency of time use,

- (6) low physical effort required,
- (7) low mental effort required,
- (8) low level of unwanted attention, and
- (9) low risk of embarrassment.

The ratings are made on a 5-point Likert scale, with the midpoint representing “neutral”. Data type: Survey.

Clearly, it will be most useful to ask participants only about aspects of the system that they have used, so the CTP usage information will be used in the survey logic so that each respondent only receives detailed survey questions about features they have used.

**Cost Data:** The cost to the user in terms of transit fare will be collected in the Shuttle Reservation Management data, and there are survey questions on the users’ perceived affordability of transit (a metric for PM #6).

System deployment and operating costs will also be recorded. There is a metric in PM #6 for the Community Shuttle operating cost per passenger trip, but otherwise there are no performance measures for system costs.

## 7.5 Human Subjects Protection

The project team has developed a human subject’s protection protocol, along with draft instruments (i.e., consent form, recruitment screener, baseline questionnaire, and deployment questionnaire) that has been submitted to the UB Institutional Review Board (IRB). The initial IRB application was submitted and approved during Phase 1. Revised versions of the questionnaires have been re-submitted for IRB approval in Phase 2, and approval has been received.

## 7.6 Data Quality Check Approach

Data quality checks will be conducted on survey and operations data when the system is active for all activities in accordance with the data curation process noted above. Data will be checked for quality (integrity, completeness, consistency), and each dataset will undergo its own quality checking procedures.

- Survey data quality checks will include a review of the raw data for significant anomalies (i.e., missing data or inappropriate answers). This will be followed by a review of descriptive statistics (i.e., frequencies, mean, median, mode), and then appropriate cross tabulations by significant variables (respondent type, population sub-group, transit type, shuttle type).
- Operations data quality checks will include matching IDs/references for consistency (especially when comparing/combining data from multiple systems), also time and location (e.g., time should be synchronized to GPS across systems to match); looking for missing data from log files, checking data consistency (e.g., when data changes like quarterly transit schedules), and checking for anomalies in datasets

- Data will then be loaded into the appropriate data store. Data will be anonymized or aggregated for the appropriate permission levels prior to being transformed for its assigned data store.

## 7.7 Data Sharing Framework

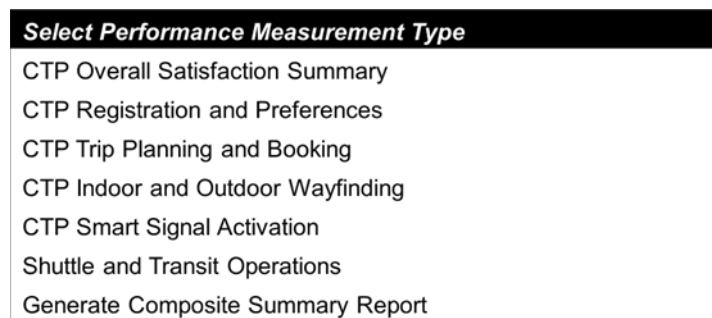
According to the data curation process, data are stored in their appropriate data storage systems. Operational data will reside in their respective subsystem data stores, while datasets needed to derive performance measurements and their metrics will be stored in the PMD. Survey data from the CTP surveys will be stored in the appropriate data store system.

As the last part of the data curation process, the datasets are made available for public and permission-based access and are listed in separate data catalogs for discovery. Metadata files will be associated with each dataset. The datasets, metadata and data catalogs are published for role-based access.



# 8 Performance Reporting

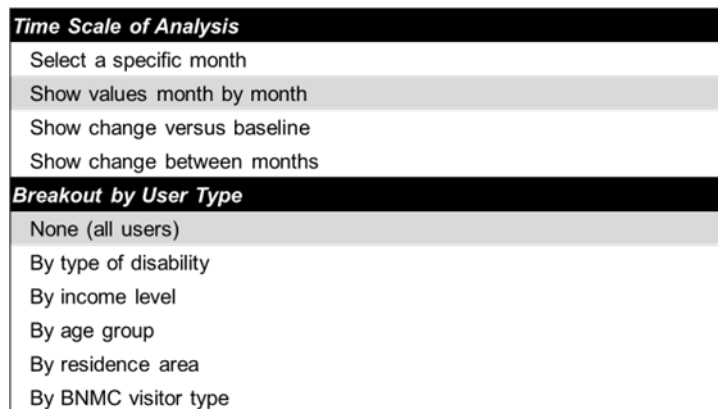
This Section proposes the summaries and dashboard to be created, including the options that will be available to users to create custom reports. Figure 4 shows a main selection screen, with one option for each of the eight groups of performance metrics that are described in Sections 3, 5, and 7. The interface for the Performance Dashboard is anticipated to be web-based, providing on-screen tabular results that can also be saved as HTML documents. There is also an option to generate a “Composite Summary Report” that will produce an HTML document with a summary across all the metrics of all performance measures.



Source: ITS4US Buffalo, NY, 2023

Figure 4. Example of Main Performance Measure Type Dropdown Menu.

It will also be possible to create customized reports. The two main types of customizations that are anticipated are the time-scale of the analysis and possible breakouts by different user types, as shown in example menus in Figure 5. Temporal options currently anticipated are to view a specific month, to view month by month over time, to show changes versus baseline conditions, and to show changes between two selected points in time. There are also options to select a specific user type to contrast performance measures across types.



Source: ITS4US Buffalo, NY, 2023

Figure 5. Example of Customization Dropdown Menu.

Figure 6 shows an example of a report that could be generated for the specific user selections shown in Figure 4 and Figure 5—the report for SDS Operations across all user types in a month-by-month format. The user can select the months to be displayed. Note that survey-based data such as the user ratings in the bottom section of the table, may not be available for every month or period requested, so some cells in the table may be empty. There may also be cases where measures are based on very small sample sizes—particularly when examining breakouts by user type. In those cases, the cells can be annotated with asterisks and/or footnotes to indicate that the performance measures are less reliable statistically.

Example report for Shuttle and Transit Operations		Values by Month		
User Type = All users	Target	May	June	July
% of CS trips that arrive at boarding stop within 5 min of schedule	>90%	87%	91%	92%
% of CS trips that arrive at alighting stop within 5 min of schedule	>90%	82%	87%	89%
Cost efficiency of CS service in terms of \$ per passenger-trip	<\$20	\$22	\$19	\$19
Increase in rating of using transit to arrive at the destination on time	>0.4	0.50%	0.40%	0.50%
Increase in total use of transit in the study area in trips/person-mo	>1.0	1.1	1	1.1
% of trips by Pal-eligible users in the study area made by CS	>25%	19%	23%	27%
<b>Average user rating of CS service:</b>				
Convenience	>3.0	3.4		3.5
Affordability	>3.0	3.1		3.1
<b>Assessment of CS Service Usability and Design:</b>				
High ease of use	>3.0	2.8		3
High independence	>3.0	3.8		3.6
Low safety risk	>3.0	2.7		2.9
Low risk of user mistakes	>3.0	3.7		3.4
High efficiency of time use	>3.0	4.1		3.9
Low physical effort required	>3.0	3.1		3.2
Low mental effort required	>3.0	3.9		3.7
Low level of unwanted attention	>3.0	4.1		4.2
Low risk of embarrassment	>3.0	3.9		3.7

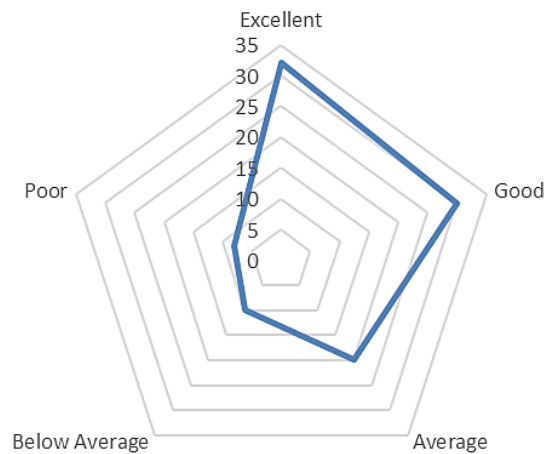
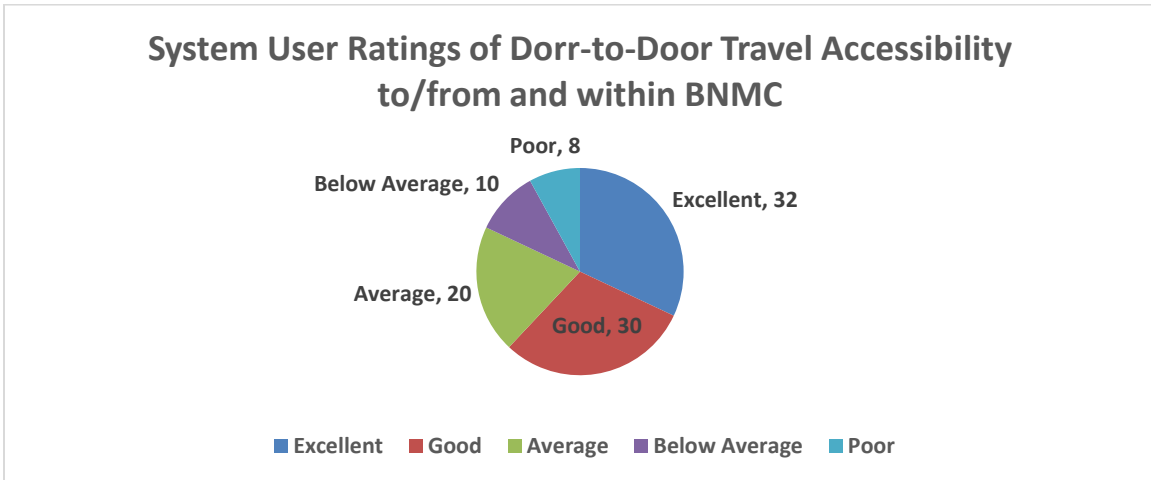
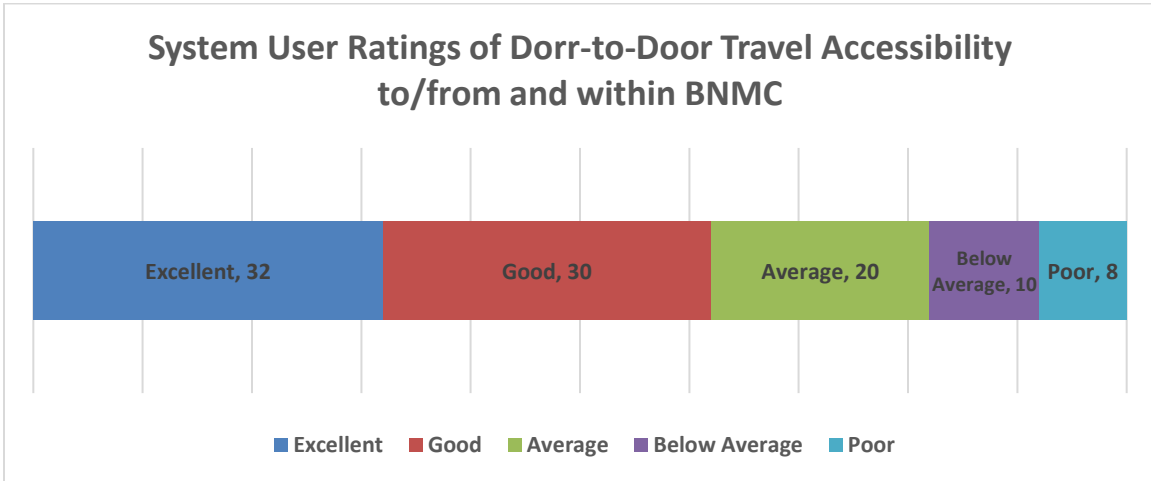
Source: ITS4US Buffalo, NY, 2023

**Figure 6. Example of a Specific Report.**

The exact format of the output table may depend on the combination of options chosen. For example, if the user selects a specific month or change between two months, then it can be possible to show the different user types selected (e.g., 3 income groups or 3 age groups) side by side within a single table. If, however, a user decides to look at measures for several time periods and several different user types simultaneously, then it would be necessary to create a separate table for each user type to avoid putting too much information into a single table.

Note that any of the performance metrics that are displayed in tabular form could also be displayed in graphical form, via bar charts or line charts. The project team will confer with users of the performance measurement output to determine which types of information may be most useful in graphical form, as well as to obtain feedback on other aspects of the performance dashboard formats and capabilities. Figure 7 shows some possible forms of graphical output. All three figures show the proportion of users that provide ratings in different categories (i.e., Excellent, Good, etc.). In the first two figures the amount of space in the shapes (bar and pie)

visually represent the share of users who gave a specific ranking. Figure 7 shows the same information where the sum of each rating is 100 (percent).



**Figure 7. Examples of PM visualization**  
 Source: ITS4US Buffalo, NY, 2023

The performance metric output will not contain any PII for study respondents, so can be distributed among project stakeholders as needed. The dashboard will also be used by the project team to evaluate system performance over time, and to provide a wide variety of performance information that can be included in the final analysis report to convey lessons learned.

# 9 Performance Measurement and Evaluation Support Schedule

As performance measurement relies on the collection and management of the various types of data, the PMESP schedule must be consistent with the DMP schedule. The current Phase 2 Performance Measurement and Evaluation Support Schedule (PMESS) is shown in Table 14. The initial items for creating the PMESS have been completed, and this document is the result of the PMESP Update. The dates are consistent with the Expected Data Milestones in the DMP.

**Table 14. Phase 2 Performance Measurement and Evaluation Support Schedule (PMESS).**

Activity	Start	End
<b>PMESS</b>	.	.
Initial PMESS	4/17/2023	5/19/2023
USDOT Review	5/15/2023	6/16/2023
Final PMESS	6/16/2023	6/30/2023
Monthly PMESS Progress/Risk Summary Report	Monthly	Monthly
<b>PMESP Update</b>	.	.
Initial revised PMESP	5/15/2023	7/31/2023
USDOT Review	8/1/2023	9/1/2023
Final revised PMESP	9/1/2023	10/2/2023
As-built PMESP	9/15/2025	9/29/2025
<b>HUAS Update</b>	.	.
Updated HUAS	11/3/2023	2/29/2024
Final Updated HUAS	4/1/2024	5/17/2024
<b>Performance Measurement Dashboard</b>	.	.
<b>Develop Routine for PMD Data Ingestion &amp; Processing</b>	5/22/2023	11/3/2023
w/ proxy Pre-Deployment (Baseline) Data	5/22/2023	11/3/2023
Survey	5/22/2023	7/5/2023
CTP	9/5/2023	11/3/2023
NFTA Cost Efficiency Data	8/23/2023	10/22/2023
Develop Analytical Methods	8/23/2023	11/3/2023
<b>Develop and Test Baseline Survey Instrument</b>	11/3/2023	1/22/2024
Develop Draft Final Online Survey	11/3/2023	12/3/2023
Test Survey Instrument (Web UI/Mobile UI)	12/4/2023	1/5/2024

Activity	Start	End
Develop Final Online Survey	1/5/2024	1/22/2024
Refine Survey Instrument	1/5/2024	2/5/2024
<b>Develop Routine for PMD Presentation Reports</b>	<b>4/22/2024</b>	<b>5/28/2024</b>
w/ proxy Pre-Deployment (Baseline) Data <i>Survey</i>	4/22/2024	5/22/2024
w/ proxy Post-Deployment Data <i>Survey</i>	4/22/2024	5/28/2024
<i>CTP</i>	4/29/2024	5/13/2024
<i>HDS</i>	5/6/2024	5/20/2024
<i>SDS</i>	5/13/2024	5/28/2024
<b>System Performance Reports (PMD Presentation)</b>	<b>11/3/2023</b>	<b>9/30/2025</b>
Baseline surveys Report	5/3/2024	9/30/2025
(proxy) Post-deployment surveys Report	5/13/2024	9/30/2025
(proxy) CTP Registration/User/Trip data Report	5/24/2024	9/30/2025
(proxy) CS (HDS/SDS) Booking and Operations Cost data Report	6/3/2024	Early Phase 3
<b>Baseline Survey Data Collection</b>	<b>Phase 3</b>	<b>Phase 3</b>
Initial Recruitment	Go-live+ 1 day	Go-live +90 days
Baseline Survey Data Collection	Go-Live+ 1 day	Go-Live +120 days

Table 15 shows an envisioned schedule of events in Phase 3. A full sample of study participants will be recruited during the initial months of Phase 3 as the final system components are brought into operation, and full baseline data will be collected. If needed, the PMESP can be further updated at that point to form the basis of the full system performance evaluation during the final 12 months of Phase 3. It is expected that the Final Report at the end of Phase 3 will document performance measures computed continuously across the year, to show how they vary over time with possible influences due to weather or other factors. If desired by USDOT, interim performance measure reports can be provided during Phase 3, albeit without the level of explanatory analysis and description that will be included in the Final Report.

**Table 15. Envisioned schedule of events for Phase 3.**

Event Title	Description	Date
Performance Measurement Review	Following the Data Review, a Performance Measurement Review will be conducted with USDOT and IE to ensure the measurement capabilities conform to the PMESP. (Note, the Agile development process commenced early in Phase 2 and will continue mid-way through Phase 3 when the complete system will be deployed and operational, so there is not yet data available for performance metrics by the end of Phase 2.)	By 4 months into Phase 3
Recruitment of Participants for Phase 3 and evaluation of baseline data	When the target sample size of participants for the full Phase 3 performance evaluation has been reached, a summary of the sample composition and weighting and the complete baseline data will be provided to USDOT for review.	By Go-live + 120 days
PMESP Update #2 (If needed)	If needed, a second update of the PMESP updated with any further changes coming from the Agile process and from the participant recruitment and baseline data collection.	By Go-live plus 180 days
Collection of post-deployment data	Continuous data collection from the CTP and other system components, along with periodic post-deployment user surveys.	The rest of Phase 3
Draft Final Analysis Report submitted	Draft Final Analysis Report submitted to USDOT. Content to be determined (TBD) based on USDOT guidance for Phase 3.	Near the end of Phase 3
Final Analysis Report submitted	Final Analysis Report submitted to USDOT.	End of Phase 3



# Appendix A. References

The following references were used during the development of this document.

- Barratt, Shane, Guillermo Angeris and Stephen Boyd (2020). Optimal Representative Sample Weights. Stanford University Research Paper. [https://stanford.edu/~guillea/papers/optimal\\_representative\\_sampling.pdf](https://stanford.edu/~guillea/papers/optimal_representative_sampling.pdf)
- Brownstone, David and Xuehao Chu (1997). Multiply Imputed Sampling Weights for Consistent Inference with Panel Attrition. In. Panels for Transportation Planning: Methods and Applications. Edited by Thomas Golob, Ryuichi Kitamura, Lyn Long. Springer Science and Business Media, 1997.
- Choi, Jimin, Jordana L. Maisel, Brittany Perez, Don Nguyen, and Victor Paquet (2021). “User Experiences with Two New Wheelchair Securement Systems in Large Accessible Transit Vehicles”. Transportation Research Record 2021, Vol. 2675(2) 150–161. National Academy of Sciences.
- Federal Transit Administration (2020). Mobility Performance Metrics (MPM) for Integrated Mobility and Beyond, FTA Report 0152, <https://www.transit.dot.gov/sites/fta.dot.gov/files/docs/research-innovation/147791/mobility-performance-metrics-integrated-mobility-and-beyond-fta-report-no-0152.pdf>
- Hazan, J., N. Lang and A. Wegsheider (2019). “On-Demand Transit can Unlock Urban Mobility”. Article by Via and Boston Consulting Group (BGC), November 7, 2019. <https://www.bcg.com/publications/2019/on-demand-transit-can-unlock-urban-mobility>
- James Bell Associates. (2008). Evaluation Brief: Conducting an Outcome Evaluation. Arlington, VA. December 2008
- Lenker, James A., Mahiyar Nasarwanji, Victor Paquet and David Feather (2011). “A tool for rapid assessment of product usability and universal design: Development and preliminary psychometric testing”. Work 39 (2011) 141–150 141. DOI 10.3233/WOR-2011-1160. IOS Press.
- RideCo (2021). “RideCo Case Study: On-demand transit reduces cost per passenger by 20%”. San Antonio, Texas, 2021. <https://www.rideco.com/case-study/san-antonio>
- Sener, I.N., C. Viggiano and A. Pulido (2021). Impacts of Transformational Technologies on Underserved Populations. Task 5B and Task 6: Technical Memorandum on Survey Data Analysis to Investigate Travel Behavior and Transportation System Impacts. TCRP Project B47. National Academy of Sciences.

- Stopher, P., T. Magassy, R Pendyala, et al. (2021). An Evaluation of the Valley Metro-Waymo Automated Vehicle RideChoice Mobility on Demand Demonstration. FTA Report No. 0198. Federal Transit Administration, August 2021.

# Appendix B. Acronyms

Acronym	Description
ABS	Address-Based Sampling
ACS	American Communities Survey
BNMC	Buffalo Niagara Medical Campus
CO	Contracting Officer
COR	Contracts Officer Representative
CS	Community Shuttle
CTP	Complete Trip Platform
DMP	Data Management Plan
ETA	Estimated Time of Arrival
FHWA	Federal Highway Administration
FTA	Federal Transit Administration
GOFS	General On-Demand Feed Specification
HDS	Human-Driven Shuttle
IE	Independent Evaluator (Volpe)
IRB	Institutional Review Board
ITS	Intelligent Transportation System
JPO	Joint Program Office
LEP	Limited English Proficiency
MPM	Mobility Performance Metrics
MSA	Metropolitan Statistical Area
MVP	Minimum Viable Product
NFTA	Niagara Frontier Transportation Authority
NY	New York
PAL	Paratransit Access Line
PedX	Pedestrian Crossing
PII	Personally Identifiable Information
PM	Performance Measure
PMD	Performance Measurement Dashboard
PMESP	Performance Measurement and Evaluation Support Plan
PRG	Priority Request Generator
PRS	Priority Request Server
RAPUUD	Rapid Assessment of Product Usability and Universal Design
RDD	Random Digit Dialing
SDS	Self-Driving Shuttle

Acronym	Description
SOC	Shuttle Operation Center
SOI	System of Interest
TBD	To Be Determined
TCRP	Transit Cooperative Research Program
UB	University at Buffalo
USDOT	U.S. Department of Transportation

U.S. Department of Transportation  
ITS Joint Program Office-HOIT  
1200 New Jersey Avenue, SE  
Washington, DC 20590

Toll-Free "Help Line" 866-367-7487  
[www.its.dot.gov](http://www.its.dot.gov)

FHWA-JPO-21-878



U.S. Department of Transportation