

WORKING P A P E R

New Tools for Assessing State and Local SNS Readiness

CHRISTOPHER NELSON
EDWARD W. CHAN
ELIZABETH M. SLOSS
LEAH B. CALDARONE
AMANDA POMEROY

WR-455-DHHS

February 2007

Prepared for the U.S. Department of Health and Human Services Office of the
Assistant Secretary for Preparedness and Response

This product is part of the RAND Health working paper series. RAND working papers are intended to share researchers' latest findings and to solicit informal peer review. They have been approved for circulation by RAND Health but have not been formally edited or peer reviewed. Unless otherwise indicated, working papers can be quoted and cited without permission of the author, provided the source is clearly referred to as a working paper. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors. **RAND®** is a registered trademark.

CONTENTS

CONTENTS.....	i
FIGURES.....	v
TABLES.....	vi
SUMMARY.....	vii
ACKNOWLEDGEMENTS.....	xiii
ABBREVIATIONS.....	xiv
1. INTRODUCTION.....	1
Current SNS Assessment System Lacks a Clear Operational Focus.....	1
RAND Was Asked To Develop Operational Components for the Assessment System.....	2
Assessment Development Process Was Guided by Explicit Design Criteria.....	2
Report Is Organized Around Answers to Key Questions.....	3
2. DETERMINING WHAT TO ASSESS.....	5
Weak Evidence Base Hampers Attempts to Identify Short List of Assessed Elements.....	5
Capabilities-Based Approach Aligns with Homeland Security Doctrine.....	5
Process Analysis Identified List of Candidate Capabilities for Assessment.....	6
Initial Assessment Development Focused on Failure-Prone Capabilities.....	7
3. HOW TO ASSESS CAPABILITIES: MAIN FEATURES OF THE PROPOSED SYSTEM.....	10
Logic of the Proposed System Derives from a Conceptual Model of PHEP.....	10
Small Scale Drills Test Operational Capabilities without Excessive Burdens.....	10
Standardized Metrics Promote Comparability.....	10
Modularity, Scalability, and Extensibility Increases Relevance to Quality Improvement.....	10
Embedding Drills in Routine and Small-Scale Events Might Increase Relevance.....	11
Format of Operations Manuals Designed to Promote Self-Assessment.....	12
Data Collected From Assessments Could Help Develop SNS Evidence Base.....	12
4. OVERVIEW OF ASSESSMENT MANUALS DEVELOPED THUS FAR.....	14
Staff Call Down.....	14
Site Activation.....	16
Facility Setup.....	17
Pick List Generation (RSS).....	18
Dispensing (POD).....	18
5. WHO SHOULD DO THE ASSESSING?.....	22
Expanded Number of SNS Assessments Will Make Site Visit-Based Assessment Difficult.....	22
Assessments Could Be Self-Administered or Administered by Other Health Departments.....	22
Post Hoc Reviews Would Create Smaller Burden But Offer Less Quality Assurance.....	23
6. ROLLING UP THE ASSESSMENTS INTO AN OVERALL EVALUATION.....	25
“Levels of Preparedness” Construct Is Anchored in the Conceptual Model.....	25
Each Level of Preparedness Would Require Different Assessments.....	27
Requirements for Assessment Frequency, Administration, and Review Might Vary By “Jurisdiction Tier”.....	27

Performance Thresholds and Scoring	29
7. AGGREGATING LOCAL ASSESSMENTS INTO STATE-LEVEL SCORES	32
Aggregation Options Depend on Degree of Formal Quantification in Roll Up	32
Option 1: Simple (unweighted) Mean.....	32
Option 2: Weighted Mean.....	32
Option 3: Using LHD Performance Thresholds	33
8. NEXT STEPS	34
Step 1: Field Testing	34
Step 2: Initial Roll Out and Ongoing Evaluation.....	35
Step 3: Set Performance Standards.....	35
Developing Assessments of Other Operational Capabilities	36
Linking Assessment with Quality Improvement	36
Assigning Consequences to Assessment Results.....	37
Improving the Evidence Base	37
Consider Efforts to Improve Coordination with Other PHEP Assessments.....	38
APPENDIX A: ANALYSIS USED TO SELECT SNS ELEMENTS AROUND WHICH TO BUILD ASSESSMENTS	39
APPENDIX B: OPERATIONS MANUALS FOR THE PROPOSED DRILLS	42
METRICS FOR STAFF CALL DOWN DRILL Operations Manual	43
METRICS FOR SITE ACTIVATION DRILL Operations Manual.....	54
METRICS FOR SETUP DRILL Operations Manual.....	63
PICK LIST GENERATION DRILL Operations Manual.....	69
TIMED METRICS FOR A POD EXERCISE Operations Manual	75
APPENDIX C: DRAFT MATHEMATICAL ROLL UP MODEL.....	86
Tab 1: POD Staffing Levels (POD staff).....	90
Tab 2: POD Timing (PODtimes)	91
Tab 3: RSS Timing (RSS times).....	92
REFERENCES	903

FIGURES

Figure 1: Conceptual Model of Public Health Emergency Preparedness and Response...	7
Figure 2: Candidate Capabilities for Assessment Development.....	8
Figure 3: Call Down Drill May Be Assembled Into Compound Drills	11
Figure 4: Performance Levels Related to Conceptual Model and Assessment Formats .	26
Figure A.1: Precedence Diagram with Crosscutting Capabilities Highlighted	40
Figure A.2: Precedence Diagram with Function-Specific Capabilities Highlighted	41
Figure D.1: Functions and Timing Post Locations	78

TABLES

Table S.1: Overview of Current Set of Operational Assessments	vii
Table 1: Content of Operations Manuals	13
Table 2: Overview of Current Set of Operational Assessments	15
Table 3: Options for Administration and Review of Assessments	23
Table 4: Operational Details on Assessment Frequency	28
Table G.1: Sample Checklist for Pick List Generation Drill Hot Wash	73
Table D.1: Sample Checklist for Dispensing Exercise Hot Wash	83

SUMMARY

State and local governments play the key role in receiving, distributing, and dispensing materiel from the Strategic National Stockpile (SNS). Assessing their capability to perform these functions is critical to the success of the SNS program.

The current system for assessing state and local SNS readiness – developed by the Centers for Disease Control and Prevention’s Division of the Strategic National Stockpile (CDC/DSNS) – draws primarily upon a written checklist tool (the “tool”) for evaluating SNS plans and related documents. The tool, administered by CDC/DSNS during one-day site visits, covers the functions and capabilities included in its main guidance document, *Receiving, Distributing, and Dispensing Strategic National Stockpile Assets* (Version 10.02). Items on the checklist are rolled up into an overall score that results in jurisdictions being rated on a 0-to-100 scale.

A growing consensus has emerged that the current checklist-based assessment approach needs to be supplemented to include assessments of jurisdictions’ ability to *implement* and *adapt* these plans in real-world situations. This report presents a set of operational drills designed to supplement the current checklist-based system.

Assessment Development Process Was Guided by Explicit Design Criteria

The assessment development process included extensive discussion with key stakeholders and end users, observation of SNS drills and exercises, and systematic process analysis. The assessments presented here seek to:

- Measure jurisdictions’ ability to implement and adapt preparedness plans
- Minimize assessment burdens
- Support both accountability and quality improvement.
- Cover a broad spectrum of jurisdictions.
- Be aligned with relevant guidance and doctrine

The report addresses the following issues:

- *What* SNS elements are most important to assess,
- *How* they should be assessed,
- *Who* should do the assessing
- How the assessments could be *rolled up* and *aggregated* into an overall evaluation of SNS readiness.

What SNS Elements Should Be Assessed?

Perhaps the best assessment of operational capability is a documented response to a real-life emergency or, short of that, a full-scale exercise that simulates the scale and stresses of real emergencies. However, large-scale emergencies are (fortunately) rare, and full-scale exercises impose heavy burdens on assessed jurisdictions and the CDC.

Focusing assessments around a small number of critical operational capabilities helps keep assessment burdens reasonable. We propose a capabilities-based approach to assessment as a complement to the capabilities-based planning philosophy articulated in Homeland Security Presidential Directive 8 (HSPD-8) and elsewhere. Instead of developing measures around specific response scenarios, the proposed assessment system focuses on “building block” capabilities that can be deployed, combined, and adapted in response to a broad spectrum of response scenarios.

In the absence of a rich empirical evidence base on the characteristics of effective SNS responses, we relied on engineering-style process analysis to identify the building blocks. The analysis (described in detail in Section 2 and Appendix A of the report) identified 14 candidate capabilities, some of which are specific to certain SNS functions (e.g., warehousing, distribution, dispensing) and others of which are applicable to a variety of SNS as well as other public health functions (e.g., staff call down, site activation).

We selected five building-block capabilities for initial assessment development:

- Staff call down
- Site activation
- Facility set up
- Pick-list generation
- Dispensing

Assessments for other capabilities could be readily added to the system.

How Should Key Capabilities Be Assessed?

We developed drill-based metrics around each of the building-block capabilities. The drills are summarized in Table S.1. Each drill has the following characteristics.

Small scale. The drills are small in scale, testing fairly narrow and specific processes. This allows the system to test important operational capabilities without imposing heavy burdens on state and local health departments. Small scale drills are also less likely to overwhelm state and local quality improvement efforts, thus increasing the likelihood that the assessments will be viewed as useful for internal as well as external accountability purposes.

Table S.1: Overview of Current Set of Operational Assessments

Assessment	Overview of Assessment Task	Metrics (required and optional)
Staff call down	Unannounced call down drill; respondents indicate ability to report, but no actual staff movement required (hypothetical assembly).	<u>Required:</u> <ul style="list-style-type: none"> ○ Calling completion time ○ Acknowledgement completion time ○ Contact rate ○ Assembly rate (hypothetical) <u>Optional:</u> <ul style="list-style-type: none"> ○ Actual assembly rate ○ More detailed information about assembly activities
Site activation	Unannounced calls placed to managers/owners of facilities used for SNS operations (e.g., warehouses, POD sites); respondents indicate availability of facility to be made ready for SNS use.	<u>Required:</u> <ul style="list-style-type: none"> ○ Calling completion time ○ Contact rate ○ Availability rate (hypothetical) <u>Optional:</u> <ul style="list-style-type: none"> ○ Actual activation time ○ More detailed information about key steps in activation process
Facility setup	Announced actual setup of a facility (e.g., RSS, POD).	<u>Required:</u> <ul style="list-style-type: none"> ○ Setup time ○ % items completed on jurisdiction-generated checklist <u>Optional:</u> <ul style="list-style-type: none"> ○ Time to complete critical subcomponents of setup process
Pick list generation (RSS)	Generate pick list in response to apportionment decision (pre-specified by jurisdiction), using CDC/DSNS inventory files.	<u>Required:</u> <ul style="list-style-type: none"> ○ Total elapsed time <u>Optional:</u> <ul style="list-style-type: none"> ○ Data importation time ○ Total time for managed inventory shipment
Dispensing (POD)	Run a set of 100 “express” (non-special-needs) patients through the POD.	<u>Required:</u> <ul style="list-style-type: none"> ○ Patient throughput (per hour) ○ Mean patient flow time ○ Dispensing accuracy <u>Optional:</u> <ul style="list-style-type: none"> ○ Throughput for multiple regimens ○ Throughput, flow time, and accuracy for special needs patients ○ Throughput using additional time posts ○ Station queue length ○ Diagnostic accuracy

Standardized metrics. Each drill includes standard metrics and provides enough standardization in the assessed tasks to ensure at least a modicum of comparability across jurisdictions and over time. Comparability is important for measures designed for accountability purposes and for attempts to track improvements over time.

Modularity, scalability, and extensibility. The drills can be run as standalone assessments or combined with other drills or assessment activities to form compound assessments. Thus, the assessment system can be adapted to meet state and local quality improvement needs without sacrificing policymakers' need for standardized – and therefore comparable – metrics that can provide an overall summary of performance. The drills are also extensible, allowing jurisdictions to add locally-relevant metrics to the standard set.

Embeddability. The drill-based assessments could also be embedded in more routine public health functions, thus reducing the need to assemble key staff solely for exercising. For example, jurisdictions might use annual refresher training for POD staff as an opportunity to assess the entire call down and assembly process.

Who Should Administer the Assessments?

The report presents options for who should administer the assessments, along with ideas for quality control.

Assessments Designed to Support Self-Assessment. Currently, CDC/DSNS administers SNS assessments to state health departments and CRI-related local health departments during site visits. However, increases in the number of CRI programs might increase the range of jurisdictions CDC/DSNS wishes to assess. Because current CDC/DSNS resource levels might not allow for site-visits to all CRI health departments, the operations manuals were designed to provide enough detailed guidance to support self-assessment.

Peer Jurisdictions Might Assess Each Other. As an alternative, assessments might be administered by other health departments, thus retaining some of the advantages of an external assessor while spreading the burdens beyond CDC/DSNS. Specifically, state health departments might be assessed by neighboring state health departments. LHDs could be assessed either by their state HD or by neighboring LHDs. Some attempt would have to be made to ensure that jurisdictions less adept at drilling and exercising are assessed by more adept jurisdictions.

Peer Jurisdiction Might Review Self-Administered Assessments. Given the present demands on health departments, however, asking them to conduct assessments on other jurisdictions may be unreasonable. A slightly less burdensome approach would be for jurisdictions to provide post hoc review of other departments' SNS self-assessments. Reviews might include interviews with key personnel and examination of documentation generated by the drills. This would also encourage the development of networks of health officials who could facilitate the diffusion of exemplary practices. These networks would also help foster coordination during real emergencies.

How Should the Assessments Be Rolled Up?

The report also discusses how the proposed assessments might be combined to generate an overall evaluation of state and local SNS readiness.

Conceptualizing “Levels of Preparedness.” The roll-up strategy is grounded in the following “levels of preparedness” construct:

- *Beginning.* The jurisdiction has acquired the *inputs* of preparedness (e.g., equipment, personnel) and has made satisfactory progress in implementing key preparedness *processes* (e.g., planning, training, exercising, organization/leadership), but has not demonstrated the ability to implement and adapt key capabilities in operational situations.
- *Developing.* The jurisdiction has acquired the necessary inputs and is engaging in adequate preparedness processes, *and* has demonstrated proficiency in performing separate tasks related to the building-block capabilities. However, the jurisdiction has *not* demonstrated an ability to execute core SNS functions *simultaneously*.
- *Secure.* The jurisdiction has demonstrated that it possesses adequate resources, is engaged in required preparedness processes, has demonstrated separate mastery of core capabilities, *and* has shown the ability to successfully *combine* crosscutting capabilities in the execution of the full spectrum core SNS functions.

Assessment Requirements Differ By Level and “Jurisdiction Tier.” Different combinations of assessments would be required for each preparedness level. Only the checklist tool would be required to attain the “Beginning” level. Separate administration of the capability-based assessments would be required to attain the “Developing” level. Finally, jurisdictions would be required to conduct full-scale exercises – what we call “compound drills” –to attain the “Secure” level.

We also recommend that assessment frequency vary by “jurisdiction tier.” Specifically, higher priority “Tier 1” jurisdictions could be assessed more frequently than other jurisdictions (see Table 4). We also suggest tier-related differences in who administers and reviews the assessments. For instance, while state and CRI-related LHDs might be required to complete their checklist assessments during CDC/DSNS site visits, Tier 2 and 3 jurisdictions might self-administer the checklist with peer review and random audits in order to ensure quality (see Section 5).

Combining Assessments Into An Overall Index. Finally, there must be a way to determine whether jurisdictions have performed adequately in all essential categories. The report provides two options:

- Set performance thresholds for each assessment and then determine the number of assessments that must be “passed” in order to achieve a given level of performance. CDC/DSNS would then determine what percentage of assessments must be passed to reach each level of preparedness.

- Use a mathematical algorithm (described in Appendix C) to combine metrics into a smaller set of time-based indexes that can be compared to the CRI 48-hour requirement.

How Could Assessments From Local Health Departments Be Aggregated Into Statewide Scores?

Finally, the report considers how (once each jurisdiction has completed its required assessments and the scores have been “rolled up” into an overall evaluation) jurisdiction-level evaluations could be combined into state-level ratings. The report considers both weighted and unweighted means as alternatives. However, both approaches assume that low scoring local health departments in a state are offset by high scoring departments, and vice-versa. Thus, the report also considers options that involve calculating the percentage of local health departments that exceeded a minimal performance threshold.

Next Steps

The proposed improvements to the SNS assessment system described above provide a roadmap. The report also describes a number of concrete steps that are required to move down that road, including:

- Field testing
- Initial rollout and ongoing evaluation
- Setting performance standards
- Developing assessments of additional capabilities
- Linking the assessments with quality improvement
- Assigning consequences to assessment results
- Improving the evidence base behind the assessments and standards
- Improving coordination between SNS and other assessment systems

The report provides a detailed description of key processes in each of the steps.

ACKNOWLEDGEMENTS

We gratefully acknowledge the large number of individuals who provided insight, guidance, and criticism at all stages of this project.

We begin by thanking the staff of the Centers for Disease Control and Prevention's Division of the Strategic National Stockpile (CDC/DSNS) for helping us to understand the operational realities faced by state and local health department and for critiquing various portions of the report and operations manuals. In particular, we thank Tim Quinn, Wrenn Slocum, Stephanie Dulin, Andy Demma, Patti Pettis, Mike Moore, Jason Zimmerman, Curt Mast, and Mike Staley.

We also thank the dozens of state and local health department officials who patiently tutored us on key aspects of the SNS process, allowed us to observe exercises, and offered comments on and criticism of draft documents. We owe a particular debt of gratitude to Steve Harrison and colleagues at the Virginia Department of Health for exceptionally detailed comments on earlier versions of the operations manuals. We also benefited early on from the insights of Mike Frasier and Katie Dunkle at the National Association of City and County Health Officers (NACCHO) and Steve Curren and Jim Blumenstock at the Association of State and Territorial Health Officers (ASTHO).

We also acknowledge the insights and assistance of RAND colleagues Melinda Moore, Karen Ricci, and Michelle Horner on various aspects of the project, and Nicole Lurie and Jeffrey Wasserman for providing insight, support, and encouragement at all stages of the project.

Finally, we thank William F. Raub and Lara Lamprecht at the U.S. Department of Health and Human Services for their support and guidance throughout this project.

We alone bear responsibility for the content of the report and the operations manuals.

ABBREVIATIONS

AAR	After Action Report
AHRQ	Agency for Healthcare Research and Quality
BERM	Bioterrorism and Epidemic Outbreak Response Model
CBP	Capabilities-based Planning
CDC	Centers for Disease Control and Prevention
CDC/DSNS	Centers for Disease Control and Prevention's Division of the Strategic National Stockpile
CRI	Cities Readiness Initiative
DHS	Department of Homeland Security
EOC	Emergency Operations Center
FEMA	Federal Emergency Management Agency
FSE	Full Scale Exercise
HD	Health Department
HHS/ASPR	Department of Health and Human Services' Office of the Assistant Secretary for Preparedness and Response
HSPD-8	Homeland Security Presidential Directive 8
IOM	Institute of Medicine
JIT	Just-in-Time
LHD	Local Health Department
MOU	Memorandum of Understanding
NIMS	National Incident Management System
NPG	National Preparedness Goal
OSHA	Occupational Safety and Health Administration
PHEP	Public Health Emergency Preparedness
PLG	Pick List Generation
POD	Point of Dispensing
RSS	Receipt, Staging, and Storage
SME	Subject Matter Expert
SNS	Strategic National Stockpile
TCL	Target Capabilities List
UASI	Urban Area Security Initiative
USPS	United States Postal Service
UTL	Universal Task List

1. INTRODUCTION

In the event of a terrorist attack or other large public health emergency, pharmaceuticals and other medical materials are available from the Strategic National Stockpile (SNS) to supplement and re-supply state and local public health departments' caches. State and local governments are responsible for requesting, receiving, staging, storing, distributing, and dispensing SNS materiel. Thus, assessing their capability to perform these functions is critical to the success of the SNS program. This report presents a set of new tools for assessing state and local SNS readiness.

Current SNS Assessment System Lacks a Clear Operational Focus

The current system for assessing state and local SNS readiness – developed by the Centers for Disease Control and Prevention's Division of the Strategic National Stockpile (CDC/DSNS) – draws primarily upon a written checklist tool (the “tool”) for evaluating SNS plans and related documents. The tool is administered during a one-day site visit by CDC/DSNS staff to state health departments and to local health departments in regions funded by the Cities Readiness Initiative (CRI) program. The tool covers the functions and capabilities included in CDC/DSNS's main guidance document, *Receiving, Distributing, and Dispensing Strategic National Stockpile Assets* (Version 10.02). Items on the checklist are rolled up into an overall score that results in jurisdictions being rated on a 0-to-100 scale.¹

However, a growing consensus has emerged – on the part of CDC/DSNS and the Department of Health and Human Services' Office of the Assistant Secretary for Preparedness and Response (HHS/ASPR) that while the checklist-based approach has done a reasonably good job of assessing the quality of state and local plans, other methods are required to assess jurisdictions' ability to *implement* and *adapt* these plans in real-world situations.

The current assessment system requires jurisdictions to conduct a full-scale SNS exercise in order to receive the highest (Green) rating; and CDC/DSNS provides technical assistance in the design and execution of these exercises. However, there are neither criteria nor scoring systems for evaluating performance on those exercises. Thus, jurisdictions can be rated “Green” even after poor performance on an exercise.

¹ Specifically, each checklist item is given a score of 0, 0.5, or 1 based on scoring guidance provided in the checklist for each item. Average scores are derived for each of the 12 (in the local checklist) or 13 (in the state checklist) functions (i.e., sections). In order to derive an overall score across all functions, each section is assigned a weight. An overall score is calculated as a weighted average by summing over all sections. In addition, jurisdictions receive written feedback on particular issues related to each of the functions assessed.

RAND Was Asked To Develop Operational Components for the Assessment System

In 2005, HHS/ASPR and CDC/DSNS asked RAND for help in developing a more operationally-focused assessment system. RAND's assistance began with an examination of issues and options for improvements in the assessment system (Nelson, Chan, Moore, Shea, & Sloss, 2005; Nelson & Sloss, 2005). This report builds upon that earlier work by describing a set of newly-developed operational assessments that can be integrated with the current checklist-based tool.² This proposed system can be expanded to include more assessments as they are developed.³

Assessment Development Process Was Guided by Explicit Design Criteria

The assessment development process was guided by the following design requirements, derived from discussions with CDC/DSNS, analysis of key SNS and CRI documents, and earlier HHS/ASPR-sponsored RAND research on measuring Public Health Emergency Preparedness (PHEP) (e.g., Nelson, Shugarman, Lurie, & Hartman, 2006). In particular, the assessments should:

- *Measure jurisdictions' ability to implement and adapt preparedness plans.* Most importantly, the proposed system must assess jurisdictions' ability to mobilize response infrastructure and resources and adapt response plans. Thus, the proposed system includes greater use of operational assessments (drills) than the current system.
- *Impose reasonable assessment burdens.* The best assessment of operational capability is a documented response to a real-life emergency or, short of that, a full-scale exercise that simulates the scale and stresses of real emergencies. However, large-scale emergencies are (fortunately) rare, and full-scale exercises impose heavy burdens on assessed jurisdictions and the CDC. Given this, the proposed system is based on a set of small-scale drills that test smaller, common-mode components of a full-scale response, with larger exercises conducted only occasionally.
- *Support both accountability and quality improvement.* The SNS assessment system must provide information to support both accountability decisions (which requires standard metrics that can be compared across jurisdictions and over time) and process improvement efforts (which requires more customized data collection). Thus, we sought to develop standard – or semi-standard – metrics for

² The work is part of a large set of studies of public health preparedness produced under contract with the Department of Health and Human Service's Office of the Assistant Secretary for Preparedness and Response (ASPR). Information on other projects can be found at <http://www.rand.org/health/centers/healthsecurity/>

³ The project's scope of work did not include setting standards and performance thresholds. Instead, the focus of the work was on developing yardsticks that could represent key dimensions of state and local readiness, with efforts to set performance thresholds coming later.

each of the drills, but also designed opportunities to adapt and combine the drill formats to fit a wide variety of state and local improvement needs.

- *Cover a broad spectrum of jurisdictions.* The current expansion of the CRI program will require assessment of a larger number of jurisdictions than in the past, at some of whom might have relatively little experience running drills and exercises. Thus, we sought to design assessments that could (if necessary) be reliably self-administered by a broad range of jurisdictions.
- *Be aligned with relevant guidance and doctrine.* Assessments will be more efficient and less frustrating to state and local jurisdictions if they are systematically aligned with relevant program guidance. Thus, we sought to align our efforts with CDC/DSNS's *Receiving, Distributing, and Dispensing Strategic National Stockpile Assets* (version 10.02) and the *Target Capabilities List*. At a more conceptual level, the proposed assessment system is grounded in the capabilities-based approach to emergency planning that underlies Homeland Security Presidential Directive 8 (HSPD-8)⁴ and the modular and scalable approach to emergency response articulated in the National Incident Management System (NIMS).

Report Is Organized Around Answers to Key Questions

The report is organized around answers to the following questions:

- *What SNS elements are most important to assess?* **Section 2** describes a conceptual framework (grounded in the principles of capabilities-based planning) used to generate a short list of SNS elements around which to develop assessments. Details on the analytical process used to operationalize that framework are provided in **Appendix A**.
- *How should those elements be assessed?* **Section 3** describes the main features of the proposed assessment system, while **Section 4** provides an assessment-by-assessment overview. Full operations manuals for each of the proposed assessments are provided in **Appendix B** along with spreadsheet-based data reporting tools that will facilitate calculation of required metrics and collection of key data.
- *Who should do the assessing?* **Section 5** outlines options for administering the assessments and maintaining quality control.
- *How should individual assessments be rolled up into an overall evaluation of each jurisdiction's SNS readiness?* **Section 6** proposes a schedule of assessments and describes how results of the operational assessments and checklist tool might be combined to form an overall rating of each jurisdiction's SNS readiness.

⁴ The principles of HSPD-8 are operationalized in the National Preparedness Goal (NPG), the Universal Task List (UTL), and the Target Capabilities List (TCL).

Appendix C contains a mathematical model that illustrates one of the roll up options described in the text of the report.

- *How might evaluations of local jurisdictions be aggregated into state-level summary scores?* **Section 7** reviews a small set of aggregation methods and identifies advantages, disadvantages, and data requirements of each.
- *What are the next steps in developing and rolling out the proposed assessment system?* **Section 8** briefly lays out recommendations for field testing, rolling out the assessments to jurisdictions, setting performance standards, linking assessments with state/local quality improvement efforts, attaching consequences to assessment scores, and improving the evidence base behind assessments and standards.

2. DETERMINING WHAT TO ASSESS

Weak Evidence Base Hampers Attempts to Identify Short List of Assessed Elements

As noted previously, RAND's charge was to develop assessments that test operational capabilities (i.e., the ability to actually implement key SNS capabilities and functions) but do not impose excessive burdens on assessed jurisdictions or CDC/DSNS. Keeping the operational assessments focused on a relatively small number of critical SNS elements is essential to simultaneously satisfying both of these criteria.

Yet, the rarity of these events, while fortunate in most respects, means that there is little empirical evidence on which to base identification of the most critical drivers of, or barriers to, successful SNS responses. Thus, unlike clinical medicine, where researchers often can observe thousands of similar individuals receiving different treatments for the same health conditions, preparedness researchers must identify key quality drivers based on a small number of seemingly disparate cases.

Capabilities-Based Approach Aligns with Homeland Security Doctrine

In the absence of a rich empirical evidence base we have relied on a conceptual model of preparedness and engineering-style process analysis to identify a short list of SNS capabilities that seem important enough to merit assessment.

The conceptual model is derived from capabilities-based planning (CBP), which is the foundation of HSPD-8 and a series of guidance documents flowing from this directive (e.g., the NPG, UTL, and TCL).⁵ The CBP approach is well-suited for situations like public health emergency preparedness, where there is considerable uncertainty about the requirements of unforeseen emergencies and disasters. Instead of preparing detailed plans and capabilities for each threat scenario, CBP suggests developing adaptable, robust "building block" capabilities that can be deployed, adapted, and combined to meet whatever response needs arise, and in whatever context.⁶

⁵ The capabilities-based approach to planning was developed by the U.S. national security community in response to the post-Cold War challenge of developing and maintaining readiness for an increasingly complex international security environment and increases in asymmetric threats (see, e.g., Davis, 2002a).

⁶ The problem of developing robust plans in the face of uncertainty is extremely prevalent in human life. For instance, in preparing for a lunch rush, the short-order cook must prepare materials and staff without knowing exactly which plates will be ordered, and in what quantity. Like the capabilities-based planner, the cook might prepare ingredients that appear in a wide variety of menu items, leaving until the last minute decisions about how to combine these ingredients in response to specific orders. The problem is also well-summarized in a quote attributed to former NASA Flight Director, Gene Krantz, in the film *Apollo 13*. Krantz is shown to remark, "I don't care what the thing was designed to do; I want to know what it *can* do." Here again, the art of preparedness lies in the ability to adapt previously developed and planned procedures to meet rapidly unfolding situations.

Public health emergency preparedness, then, can be viewed as having both a static and dynamic component.

- The *static* component lies in using plans, training, exercises, leadership and other processes to develop equipment and personnel into capacities that could be deployed in response to a broad spectrum of potential threats.
- The *dynamic* component lies in the “improvisational” capability to assess a situation and deploy, adapt, and combine building-block capabilities that can reduce morbidity and mortality.⁷

Figure 1 represents this capabilities-based understanding of preparedness (including, but not limited to, SNS preparedness) in terms of the familiar input-process-output-outcome framework common in the literature on program design and evaluation. Inputs and processes constitute the static portion of the model. *Inputs* (shown in the left-most box) are the “raw materials” of preparedness and include personnel and equipment. These inputs, in turn, are transformed via preparedness *processes/activities* (e.g., plans, training, exercising, organization/leadership) into *outputs*, which we conceptualize as the “building block” capabilities described above,⁸ and which comprise the dynamic component preparedness.

These building-block capabilities (shown as separate boxes) are deployed and assembled (represented in the diagram as a set of connected blocks) in response to emergencies. If the match between capabilities and ongoing situational requirements is sufficient, the *outcome* of the response (shown at the right-hand side of the diagram) will be a reduction in mortality and morbidity. As noted in the diagram this match is ensured by appropriate detection, reporting, and investigation, along with sound decision-making about which combination of capabilities is most likely to reduce mortality and morbidity.

Process Analysis Identified List of Candidate Capabilities for Assessment

As noted above, the selection of SNS capabilities to fill in the boxes in the diagram cannot be clearly guided by evidence, given the weak empirical base on the barriers and drivers of effective responses. Instead, we used critical path analysis (Kelly, 1961; McClain, Thomas, & Mazzola, 1992) to identify the “rate-limiting factors” in the SNS process – those things that, if not done in a timely fashion, could prevent delivery of medications to the public within 48 hours, as required by the Cities Readiness Initiative guidance (CDC, 2006). Examination of these factors helped identify capabilities most in need of assessment. Appendix A provides details on this analysis.

⁷ Writing of defense planning, Davis (2002b) observes that, “success . . . depends on the suitability of the building blocks and the organization’s prowess in quickly assembling and controlling their integrated application to missions.”

⁸ Taken together, the inputs and processes comprise what DHS (and before that FEMA) documents refer to these as the key “capability elements” (see DHS, 2005; FEMA, 1996).

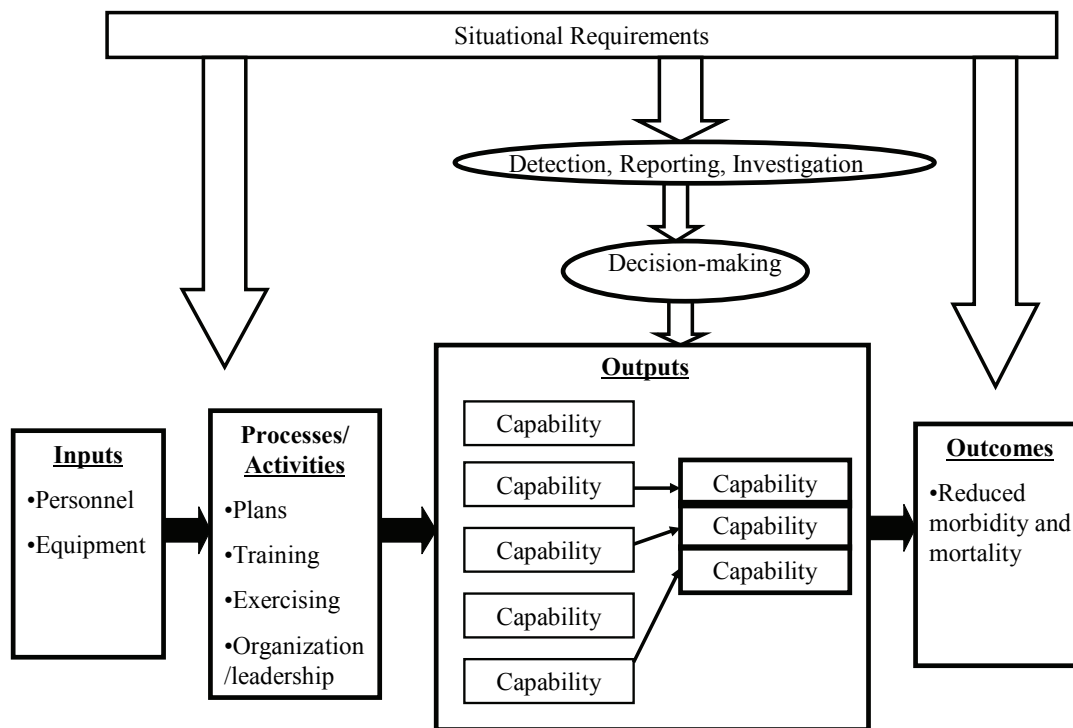


Figure 1: Conceptual Model of Public Health Emergency Preparedness and Response

Figure 2 shows a set of capabilities identified in the critical path analysis that might be candidates for assessment. The capabilities are represented as Lego®-like building blocks and are color-coded and grouped into three “bins.” The vanilla (as in “plain vanilla”) blocks represent the “crosscutting” capabilities described above (call down, assembly, and training), that can be deployed in the execution of multiple SNS functions. The gray blocks represent command and management capabilities (e.g., requesting SNS, apportioning resources). Similarly, building blocks for SNS-specific functions are grouped together (distribution; receiving, storing, and staging; and points of dispensing). The set of candidate capabilities focuses on those most directly involved in the movement of materiel and people and is not intended to be exhaustive (e.g., it does not include capabilities related to epidemiology, surveillance, lab testing, and so on).

Initial Assessment Development Focused on Failure-Prone Capabilities

Thus far, we have developed five capabilities-based drills from the list of capabilities shown above. Two focus on SNS-specific functions: dispensing and pick list generation (part of Receipt, Staging, and Storage (RSS) operations). Two focus on crosscutting capabilities: call down (which includes hypothetical assembly) and site activation. The fifth – site setup – is in principle a crosscutting capability but will look quite different

depending on the functional context in which it is implemented. These capabilities are highlighted in Figure 2. Two considerations guided our decision to develop assessments for these five capabilities.

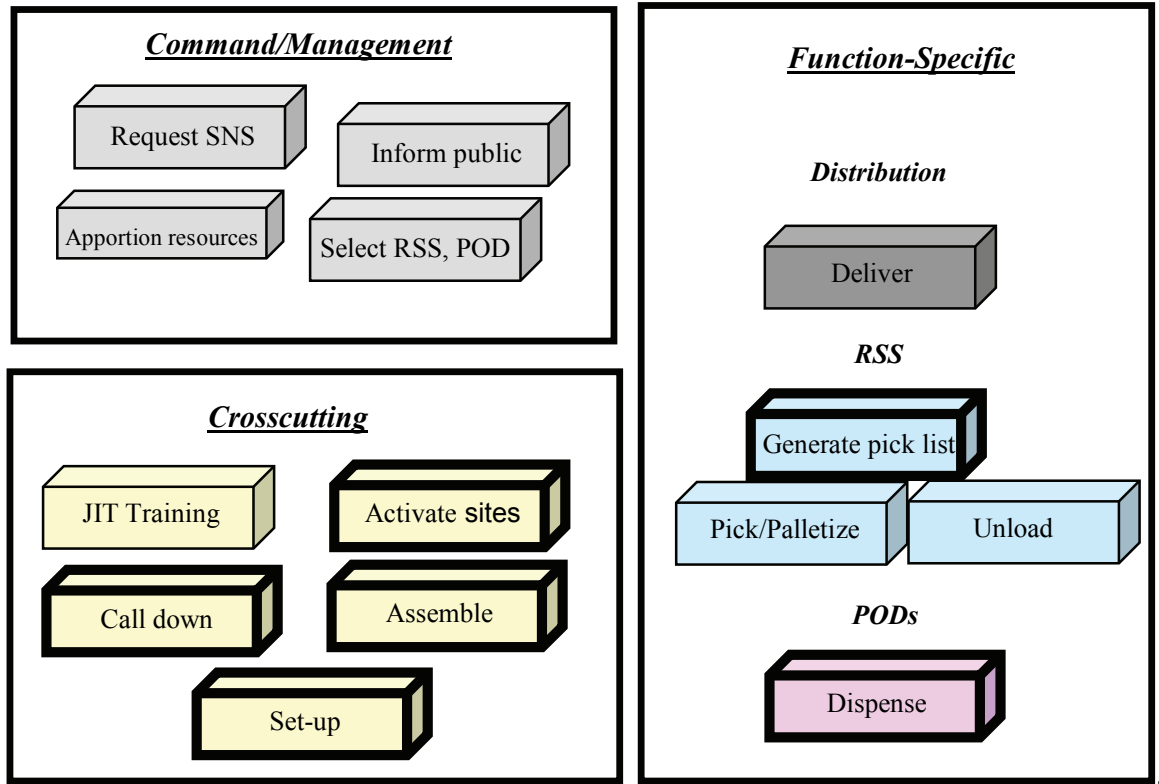


Figure 2: Candidate Capabilities for Assessment Development

Focus on movement of materiel and staff. First, we decided to focus on capabilities that involve the movement of goods and the assembly of staff. We recognize the importance of assessments that test decision-makers' ability to determine when to activate procedures (e.g., requesting SNS), allocate resources (e.g., apportionment of materiel across Points of Dispensing (PODs)), and so on. However, developing assessments for these capabilities requires measurement of highly qualitative and context-specific skills. While we believe that it is possible to assess these capabilities doing so will require more resources than were available for this project. We provide specific ideas about how to approach these assessments in Section 8.

Focus on the most failure-prone capabilities. We also focused on capabilities that seemed most failure-prone. For instance, conversations with subject-matter experts and our own observations of exercises suggested that unloading SNS shipments is a fairly straightforward process as long as an RSS layout plan exists, a suitable facility has been

activated, and enough RSS staff can be called up for service.⁹ These aspects can be verified through other drills and through the checklist-based assessment tool.

Similarly, it did not seem to make sense to hold drills of only the delivery process; if sufficient trucks and trained drivers are present with appropriate directions, it is reasonable to assume that drivers can drive the trucks to the assigned location. Traffic jams, damaged roads, weather, and other factors could complicate distribution, but these factors would be difficult to simulate during a drill.

An argument could be made that picking and palletizing are failure-prone enough to warrant a separate, joint assessment. However, resource constraints forced us to limit the number of assessments included in this initial development process. Moreover, a picking and palletizing drill would require such a large amount of equipment and supplies that a jurisdiction may as well conduct a full RSS drill.

⁹ However, unloading managed inventory on pallets is more difficult in that it requires a trained forklift operator or personnel with pallet jacks.

3. HOW TO ASSESS CAPABILITIES: MAIN FEATURES OF THE PROPOSED SYSTEM

This section describes the main features of the proposed assessment approach and provides a brief overview of the five assessments developed thus far.

Logic of the Proposed System Derives from a Conceptual Model of PHEP

The proposed assessment system reflects the logic of the conceptual model of Public Health Emergency Preparedness (PHEP) presented above. Just as the model envisions preparedness as a matter of developing, deploying, and combining building blocks, the assessment system is based on a set of small-scale, standardized, modular, scalable, extensible, and embeddable assessments of key capabilities. Each of these system characteristics helps address the design requirements outlined in Section 1.

Also, we suggest that the conceptual model – or some version of it – might be useful in communicating the fundamentals of the assessment system to health departments.

Small Scale Drills Test Operational Capabilities without Excessive Burdens

The fact that the drills are small in scale helps the assessment system balance the need to test operational capabilities with the desire to limit burdens on assessed jurisdictions. For instance, the pick list generation drill – the RSS-relevant drill described below – tests jurisdictions’ ability to generate guidance to warehouse staff in moving materiel from the containers or pallets they arrive on to pallets to be shipped to POD locations, hospitals, or other dispensing sites. While pick list generation is a critical and potentially rate-limiting step, the drill can be conducted with just a few staff and requires only a small room, a computer, and minimal office equipment.

Standardized Metrics Promote Comparability

Each drill includes standard metrics and provides enough standardization in the assessed tasks to ensure at least a modicum of comparability across jurisdictions and over time. Most of the metrics measure the amount of time required to complete critical steps in the SNS process. But standard metrics alone are not enough to ensure comparability. Thus, the operations manuals for each drill include detailed instructions on how to structure the assessed task, equip evaluators, and collect/report data.

Modularity, Scalability, and Extensibility Increases Relevance to Quality Improvement

The drills can be run as standalone assessments or combined with other drills or assessment activities to form compound assessments. For instance, assessments of function-specific capabilities such as RSS/pick list generation and POD dispensing require processes for notifying staff and setting up facilities. They could be readily

combined with assessments of crosscutting capabilities such as call down, site activation, and facility setup (illustrated visually in Figure 3 as a process of “snapping” the Lego®-like blocks together).

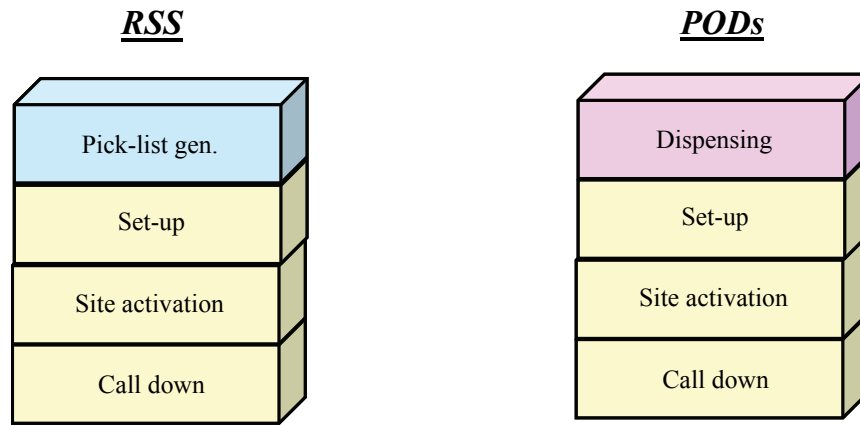


Figure 3: Call Down Drill May Be Assembled Into Compound Drills

Thus, the assessment system can be adapted to meet state and local quality improvement needs without sacrificing policymakers’ need for standardized – and therefore comparable – metrics that can provide an overall summary of performance. Because the standardized drill modules are fairly small, they need not dominate ongoing training and assessment schedules.

The drills are also extensible, allowing jurisdictions to add locally-relevant metrics to the standard set. The operations manuals that accompany the drills include suggestions for add-on metrics.

Embedding Drills in Routine and Small-Scale Events Might Increase Relevance

It might also be possible to embed operational preparedness assessments in more routine public health functions, thus reducing the need to assemble key staff solely for exercising. For example, jurisdictions might use annual refresher training for POD staff as an opportunity to assess the entire call down and assembly process. Instead of telling people to simply assemble at an appointed hour, they could instruct them to not leave their homes/offices until they received a call telling them to do so. Having gathered all applicable staff, a jurisdiction might take the opportunity to walk them through the facility setup process.

Several jurisdictions are already using non-routine practices or small-scale emergencies to test large-scale response capabilities. For example, several jurisdictions used the influenza vaccine shortfall of 2004 as an opportunity to test the public health department’s incident command structure during an effort to mass vaccinate a large number of individuals in a single day (Seid et al., 2006). Other jurisdictions have used

annual flu clinics to test mass dispensing procedures for the SNS program, which alleviates some of the costs associated with recruiting volunteers to act as test patients. Similarly, a local jurisdiction in Virginia recently used a tuberculosis screening process as an opportunity to stand up and exercise incident command structures (Rendin, Welch, & Kaplowitz, 2005).

Format of Operations Manuals Designed to Promote Self-Assessment

We developed short operations manuals for each assessment in order to ensure consistent and easy use of the drills across jurisdictions and over time. The manuals are designed to support self-assessment by most jurisdictions. However, further field testing will be required to ensure that the manuals are usable by a broad spectrum of jurisdictions.

Manuals were drafted by RAND staff after a review of the relevant literature, discussions with subject matter experts and practitioners, and observation of exercises. Drafts were critiqued by CDC staff and a limited number of health department staff. Each of the manuals uses a common outline, summarized in Table 1.

Data Collected From Assessments Could Help Develop SNS Evidence Base

In addition to metric data, each data collection tool asks for a small amount of data on the conduct of the drill and on a jurisdiction's approach to the assessed task. For instance, the tool for the call down drill asks jurisdictions to provide information on communications equipment and protocols. All drills ask for basic staffing information. Such information might provide an important and convenient data source for examinations of whether certain ways of structuring responses are associated with better performance on key metrics and thus help deepen elements of the PHEP evidence base.

Table 1: Content of Operations Manuals

Section of Manual	Description of Content
Overview	An overview of the drill, including a brief description of the assessed capability, a summary of required metrics, and any special issues that bear emphasis.
Planning checklist	A list of key tasks that must be completed before the drill and data collection begins. Typically, this involves recruiting staff and volunteers, ensuring that facilities and equipment are available, and that relevant plans and protocols are ready.
Data collection procedures	Detailed instructions for running the drill and collecting relevant data in a way that will ensure a reasonable degree of consistency across jurisdictions and over time.
Required metrics	Detailed instructions for how to turn the data collected into the metrics we recommend that CDC/DSNS require as part of the SNS assessment process.
Optional metrics	Ideas about additional measurement and assessment activities that might be useful for internal quality improvement purposes. ¹⁰
Linking the drill with other assessments	Ideas about how the drill can be combined with other assessments.
Recommendations for a hot wash and improvement planning	Brief recommendations for running a hot wash after each drill and specific topics and questions that might be covered.
Data reporting checklist	A spreadsheet-based tool that helps users calculate key metrics and provides a convenient place to provide other data.

¹⁰ However, we do not recommend that metrics from these activities be included in the formal SNS assessment system, either because doing so would create excessive burdens or because it is unlikely that the metrics would be sufficiently consistent over time and across jurisdictions to warrant use for accountability purposes.

4. OVERVIEW OF ASSESSMENT MANUALS DEVELOPED THUS FAR

This section describes the main features of each of the five assessment manuals we have developed thus far (see Table 2 below). The operations manuals themselves are included in Appendix B of this report. This section is intended as a concise overview of the manuals, including the logic of drill design and measurement and issues that CDC/DSNS should consider as the assessment system is further developed and rolled out. Readers interested in more detail should turn to the manuals themselves.

Staff Call Down

Staff call down is the ability to contact and mobilize staff to perform emergency response functions. Successful call down requires (a) a contact list or “call down list” that is up to date, readily accessible, and usable; (b) workable methods for contacting staff in a timely fashion; and (c) the ability of staff to report to their assigned stations in a timely manner. The staff call down drill tests the validity of jurisdictions’ call down lists and their ability to contact those staff in a timely manner, and estimates the percentage of staff who could report for duty within a designated time frame. Call down is a crosscutting capability, applicable to a wide variety of SNS functions, including dispensing, warehousing (RSS), distribution, security, and command centers, among others.

Overview of the Assessment Task. An ideal call down drill would be conducted without notice, and involve actual assembly at a site. However, requiring staff to travel might create undue burdens. Moreover, it is not clear that travel times derived from drills would reflect real emergency conditions (e.g., traffic, compromised infrastructure). Therefore the procedures and metrics outlined are based on a no-notice, *hypothetical* assembly, call down drill. That is, staff members are contacted using whatever communications system would be used during a real emergency (e.g., automated calling system, calling tree, etc.) and asked whether they could report to a given location, but are not actually required to assemble.

Since staff requirements will vary by scenario, the manual instructs jurisdictions to attempt to contact *every* staff member on their lists and states that lists should have enough staff to meet the CRI 48-hour standard. Of course, this requirement will remain ambiguous until there are clearer standards about what staffing levels are required to meet the CRI standard. Nevertheless, requiring that all assessed jurisdictions contact every person on the relevant call down list will provide some degree of consistency across jurisdictions. Instructions are provided for both manual and automated calling systems

Table 2: Overview of Current Set of Operational Assessments

Assessment	Overview of Assessment Task	Metrics (required and optional)
Staff call down	Unannounced call down drill; respondents indicate ability to report, but no actual staff movement required (hypothetical assembly).	<u>Required:</u> <ul style="list-style-type: none"> ○ Calling completion time ○ Acknowledgement completion time ○ Contact rate ○ Assembly rate (hypothetical) <u>Optional:</u> <ul style="list-style-type: none"> ○ Actual assembly rate ○ More detailed information about assembly activities
Site activation	Unannounced calls placed to managers/owners of facilities used for SNS operations (e.g., warehouses, POD sites).	<u>Required:</u> <ul style="list-style-type: none"> ○ Calling completion time ○ Contact rate ○ Availability rate (hypothetical) <u>Optional:</u> <ul style="list-style-type: none"> ○ Actual activation time ○ More detailed information about key steps in activation process
Facility setup	Announced actual setup of a facility (e.g., RSS, POD) ; respondents indicate availability of facility to be made ready for SNS use.	<u>Required:</u> <ul style="list-style-type: none"> ○ Setup time ○ % items completed on jurisdiction-generated checklist <u>Optional:</u> <ul style="list-style-type: none"> ○ Time to complete critical subcomponents of setup process
Pick list generation (RSS)	Generate pick list in response to apportionment decision pre-specified by jurisdiction, using CDC/DSNS inventory files.	<u>Required:</u> <ul style="list-style-type: none"> ○ Total elapsed time <u>Optional:</u> <ul style="list-style-type: none"> ○ Data importation time ○ Total time for managed inventory shipment
Dispensing (POD)	Run a set of 100 “express” (non-special-needs) patients through the POD.	<u>Required:</u> <ul style="list-style-type: none"> ○ Patient throughput (per hour) ○ Mean patient flow time ○ Dispensing accuracy <u>Optional:</u> <ul style="list-style-type: none"> ○ Throughput for multiple regimens ○ Throughput, flow time, and accuracy for special needs patients ○ Throughput using additional time posts ○ Station queue length ○ Diagnostic accuracy

Metrics. We recommend that jurisdictions be required to collect and report data for four metrics to CDC/DSNS:

- *Call process completion time.* The amount of time required to complete the call down process
- *Acknowledgement process completion time.* The amount of time needed to receive acknowledgements from staff that they have received the call down message, regardless of whether they would be able to report for duty.
- *Contact rate.* The number of staff successfully contacted as a percentage of all staff on the call down list
- *Assembly rate (hypothetical).* The number of staff who report being able to go to a designated site within a designated time as a percentage of all staff on the list.

The manual also includes suggestions for more in-depth data collection to help validate self-reported assembly information. One way to validate assembly rates and times is to require actual – as opposed to hypothetical – assembly. In this case, the assembly rate would be calculated based on the number of staff actually reporting to the site in the numerator. Recognizing the burdens of actual assembly and that (as noted above) drilled assembly times might not reflect all of the contingencies of real emergencies (e.g., traffic, compromised infrastructure), the manual also suggests more detailed questions assessors might ask to get a more realistic picture of likely assembly times and rates. These questions are not required because they will take considerably more time than a real emergency call down and may be incompatible with automated calling systems.

Site Activation

The site activation drill tests jurisdictions’ ability to quickly contact operators/owners of sites that would house critical SNS facilities (e.g., RSS warehouses, PODs, EOCs) and determine how quickly the sites could be made ready for emergency operations. Site activation is a crosscutting capability, applicable to multiple functions in a mass prophylaxis scenario, such as dispensing (POD), warehousing (RSS) and command centers (EOC).

Overview of the Assessment Task. An ideal site activation drill would be conducted without notice and involve actual site preparation for setup. However, requiring sites, many of which normally function as schools, warehouses, or medical centers, to interrupt their normal non-emergency functions without warning might create an undue burden. While such tests should be conducted periodically, procedures in the manual are based on a no-notice, hypothetical site activation drill to permit more frequent, lower-cost assessments.

As with the staff call down drill, the site activation drill manual instructs jurisdictions to attempt to contact every site on their lists, and states that lists should have enough

facilities to meet the CRI 48-hour standard. Instructions are provided for both manual and automated calling systems

Metrics. We recommend that jurisdictions be required to collect and report data for three metrics to CDC/DSNS:

- *Completion time.* The amount of time needed to contact all sites on the list
- *Contact rate.* The percentage of sites successfully contacted
- *Hypothetical availability rate.* The number of sites that report being able to be prepared within a designated time, as a percentage of all sites on the list.

As with the staff call down drill, the manual also includes suggestions for more in-depth data collection to help validate self-reported assembly information, including actual activation and more in-depth questioning of facility owners/operators.

Facility Setup

The goal of a setup drill is to test the amount of time it takes to completely set up a facility with the materiel, layout, and supplies necessary to perform a given SNS function. Facility setup is a crosscutting capability, applicable to a wide variety of SNS functions, including dispensing (POD), warehousing (RSS), and command and control (EOC), among others. Proper setup is an important precondition of a rapid and effective response, and may in some situations lie on the critical path to successfully prophylaxing a population within 48 hours.

Overview of the Assessment Task. The drill requires players to set up a POD, RSS, or EOC, given the necessary space and materials. The scope of the task is defined by a facility-specific setup checklist, developed ahead of time by the jurisdiction, that identifies all the tasks required to prepare the facility (e.g., moving tables, chairs, and signs into position). The checklists should be congruent with the layouts (e.g., for PODs, RSS, EOC) specified in the jurisdiction's SNS plan.

Metrics. We recommend that jurisdictions be required to report two metrics to CDC/DSNS as part of the SNS assessment process:

- *Setup time.* The time required to complete the tasks listed in the setup checklist's specifications.
- *Percentage of checklist items completed.* The percentage of items completed on a checklist that defines the activities required for site setup.

In addition, jurisdictions are encouraged to measure the time required to complete intermediate steps in the setup process in order to identify bottlenecks.

Pick List Generation (RSS)

The pick list generation (PLG) drill is designed to assess jurisdictions' proficiency in generating pick lists. These lists translate command-level decisions about allocation of resources among PODs into detailed guidance about what specific quantities of countermeasures should go to what PODs. As such, they provide essential guidance for picking, palletizing, and loading materials stored in RSS facilities into trucks destined for POD sites. Discussions with subject matter experts and observation of RSS drills identified pick list generation as a common bottleneck on the critical path to moving SNS materiel to affected populations.

Overview of the Assessment Task. The drill requires players to construct a pick list based on (a) an apportionment decision and (b) the 12-hour Push Pack inventory data file provided ahead of time by CDC/DSNS. The drill does not test jurisdictions' ability to make good apportionment decisions. Nor does it test their ability to generate pick lists for managed inventory or shipments without clear bills of lading. However, jurisdictions are encouraged to test these scenarios and the manual provides guidance in doing so. The drill could be run in an actual RSS warehouse and as part of a larger RSS exercise, but could also be run in a separate room. This facilitates more frequent assessment and quality improvement activities.

Metrics. We recommend that assessed jurisdictions report just one required metric to CDC/DSNS as part of the SNS assessment process: the total amount of time required to generate the pick list. We also suggest that jurisdictions measure the time required to complete intermediate steps (e.g., data importation, where the data are not pre-loaded) in order to identify bottlenecks in the process (as noted below). The operations manual also provides suggestions for testing the ability to generate pick lists for managed inventory and other situations in which a pre-existing inventory data file is not available.

Issues and Options for Future Consideration. As noted above, the drill is designed to test pick list generation using the inventory data file for the 12-hour Push Pack. In the future, CDC/DSNS might wish to add a second stage to the drill where jurisdictions have to generate a pick list for managed inventory or shipments without bills of lading. This would involve manual entry of data and would increase the degree of difficulty and the amount of time required.

Dispensing (POD)

The dispensing drill is designed to assess the speed and accuracy with which jurisdictions can dispense oral prophylaxis in a POD setting. The manual includes an appendix on how the assessment can be applied to vaccinations.

Overview of the Assessment Task. Conversations with subject matter experts, observations of drills, and examination of after action reports (AARs) from drills all suggest that jurisdictions often have a difficult time recruiting enough volunteers to "stress" the POD. Often, POD drills are run with 100 or 200 persons spread out over

several hours, never creating an opportunity to test the POD design and staff's ability to handle large numbers of patients at one time.

Assessment based on 100+ non-special needs persons. Requiring large numbers of persons seems unreasonable, so an alternative discussed with CDC/DSNS is to run a relatively small number of persons through a POD during a compressed time period. Although this does not test the staff's ability to manage the fatigue associated with long-term operations, it seems more important to provide even a short-term test of the POD's surge capacity. Thus, the drill requires jurisdictions to recruit 100 non-special needs patients, with upward adjustments for large PODs. Non-special needs persons are those people that do not require additional consideration for dealing with children, disabilities, language barriers, or complicating medical conditions. In order to increase comparability of the data collected, the manual requires that test persons go through the POD before it is opened up to others. Restricting the test to non-special needs persons will concentrate dispensing activity in certain areas of most PODs (often a "Green" line), which will increase the stress placed on that part of the POD by even a relatively small sample of persons.

Compatible with most POD designs. Another consideration in developing dispensing metrics was to ensure their compatibility with a wide variety of dispensing modalities. Thus, two of the required metrics (throughput and flow time) involve determining the amount of time it takes test persons to get from the beginning to the end of the POD, regardless of how the POD is set up. We believe that this will work for any dispensing modality that requires persons to come to a central location to pick up their medication, whether that is a large clinic or a small dispensing point in a workplace, parking lot, or other location. The drill and metrics will not work, however, for dispensing modalities (e.g., USPS delivery) that require citizens to shelter in place while medications are delivered to them.

Metrics. We recommend that jurisdictions be required to collect and report data for three metrics to CDC/DSNS:

- *Throughput.* The number of persons processed per hour
- *Mean flow time.* The average amount of time it takes a person to move through the POD
- *Dispensing accuracy.* The percentage of persons who received the correct countermeasure prescribed for them.

The metric on dispensing accuracy presumes that more than one countermeasure is being dispensed. Clearly, this metric would be dropped if this is not the case. We have provided room for jurisdictions to note this on the data capture form for this assessment.

The manual also includes optional metrics that could be useful for quality improvement, but that are unlikely to provide consistent and comparable data required for accountability purposes. These include:

- *Throughput for multiple regimens.* The required throughput metric indicates the number of individuals served per hour. However, if jurisdictions use head-of-household dispensing serving one person might result in the prophylaxis of several individuals. Jurisdictions using this strategy might determine the number of *units* dispensed per hour, in addition to the number of individuals processed.
- *Testing with non-special needs persons.* As noted above, the required metric is based on non-special needs persons, which will have few mobility problems or other special conditions. Thus, the metric values produced by the procedures above will not be fully representative of a jurisdiction's ability to respond to real-world conditions. Therefore, the manual encourages jurisdictions to run the procedures described above with special needs persons, such as families with children or those with disabilities, language barriers, or complicating medical conditions. Here, test persons would be given a written sheet containing a patient profile with characteristics and symptoms will be given to each volunteer allowing him/her to play the assigned role. POD workers should treat these persons as provided for in their jurisdiction's dispensing plan.
- *Throughput using additional time posts.* Jurisdictions may add time posts other than at the beginning and end of the POD to identify bottlenecks.
- *Station queue length.* Jurisdictions could measure the average number of people queued at various stations as another way to identify trouble spots in POD designs.
- *Diagnostic accuracy.* Jurisdictions might assess the accuracy of decisions about which persons should be dispensed which medications. This would require a medical professional as the evaluator or in the development of an evaluation tool.

Issues and Options for Future Consideration. There are a number of issues and options that CDC/DSNS might consider as the assessment system continues to be developed and rolled out.

- *Optimal number of patients for a POD test.* As noted above, the required metric calls for 100 patients (with upward adjustments for large PODs) running through the express line of a POD in order to stress the POD – albeit over a short period of time. This number is not based on any systematic evidence and should be reexamined in light of further experience and guidance.
- *Making inferences about large-scale dispensing operations.* Perhaps the most fundamental issue involves the extent to which a test with 100 patients can provide an estimate about a jurisdiction's ability to prophylax thousands of patients in the aftermath of a large-scale terrorist, disease outbreak, or other public health emergency. This should be the subject of modeling efforts and stakeholder discussions. In the meantime, CDC/DSNS should consider the likelihood that throughputs in large-scale operations are likely to be different (probably lower) than throughput in small controlled settings. CDC/DSNS might set slightly higher throughput and flow time standards for a single POD than would be

required for all PODs in order to meet the CRI standard of full community prophylaxis within 48 hours of the decision to do so.

5. WHO SHOULD DO THE ASSESSING?

This section describes options for who should administer the SNS assessments described in this report. Our discussion assumes that the operational assessments described above will be given in addition to – rather than instead of – the current checklist assessment tool.

Expanded Number of SNS Assessments Will Make Site Visit-Based Assessment Difficult

Currently, CDC/DSNS administers SNS assessments to state health departments and CRI-related local health departments during site visits.¹¹ As noted in Section 1, however, our assessment design process proceeded on the assumption that the range of assessed jurisdictions might expand considerably, in part due to the ongoing expansion of the CRI program.

Administering assessments during site visits has clear advantages. For instance, instead of having jurisdictions self-report that an MOU exists, site visitors can inspect the document themselves and judge its adequacy rather than relying on self-reported assessments. Site visits might provide opportunities for evaluators to speak with jurisdiction staff to gauge their awareness and understanding of preparedness plans and resources.

However, the fact that, according to CDC/DSNS, staffing constraints have reduced the number of site visits from twice to once a year makes it unlikely that the agency could continue site visit-based administration with more jurisdictions and with the addition of the capability-based assessments described above. Thus, it will be necessary to consider other approaches for administering the SNS assessments. The remainder of this section reviews options for (a) in-process administration of the assessments and (b) post-hoc review of the assessments. Table 3 summarizes the options discussed below.

Assessments Could Be Self-Administered or Administered by Other Health Departments

As noted above, we designed the operations manuals for the capability assessments under the assumption that they must support self-assessment. This was done to preserve the option of self-assessment. However, some jurisdictions – particularly those with little prior experience running drills and exercises – will probably find the tasks and data collection challenging. Moreover, if CDC assigns consequences to performance on the

¹¹ During such visits the CDC/DSNS rater meets with staff members of the health jurisdiction and uses multiple data sources to complete the checklist. This might include document review, interviews, and observation. Many of the checklist items focus on the specific content of the jurisdiction's SNS plan. For these items, the rater either reviews the written SNS plan or relies on responses to specific questions during interviews with staff members.

metrics, some jurisdictions may be tempted to misreport data or change the tasks to decrease their difficulty.

Table 3: Options for Administration and Review of Assessments

In-Process Administration of Assessments	Post Hoc Review of Assessments
CDC/DSNS (probably not feasible)	CDC/DSNS
State HDs	State HDs
LHDs	LHDs
Self-assessment	Self-review

As an alternative, assessments might be administered by other health departments, thus retaining some of the advantages of an external assessor while spreading the burdens beyond CDC/DSNS. Specifically, state health departments might be assessed by neighboring state health departments. Local Health Departments (LHDs) could be assessed either by their state health department (HD) or by neighboring LHDs. Some attempt would have to be made to ensure that jurisdictions less adept at drilling and exercising are assessed by more adept jurisdictions.

To maintain consistency, it is helpful if assessment teams commit to the process for a significant amount of time (perhaps two years) for the development of shared understandings of assessed functions and evaluative criteria. This could require support from CDC or another organization – perhaps akin to an “evaluation fellowship” – for exemplary public health professionals to serve on the committees on a part-time basis. This approach would also expand the cadre of thought leaders on assessment and process design. Given the costs involved in supporting teams of evaluators, CDC/DSNS and HHS/ASPR might explore the possibility of teams that conduct assessments for both HHS and the Department of Homeland Security (DHS). This would spread costs and foster integration across these two agencies.

Post Hoc Reviews Would Create Smaller Burden But Offer Less Quality Assurance

Given the present demands on health departments, asking them to conduct assessments on other jurisdictions may be unreasonable. A slightly less burdensome approach would be for jurisdictions to provide post hoc review of other departments’ SNS self-assessments. Reviews might include interviews with key personnel and examination of documentation generated by the drills. This would also encourage the development of networks of health officials who could facilitate the diffusion of exemplary practices.

These networks would also help foster coordination during real emergencies. The burden on state and local health departments could be further reduced if CDC/DSNS were to conduct some or all of reviews itself (this, of course, would increase the burden on CDC/DSNS, however).

A key question with any post hoc review process is how to select jurisdictions for review. Reviewing all jurisdictions is probably not feasible. However, it might be more feasible to review all jurisdictions during a 4 or 5 year period. Alternatively, reviews could be targeted to jurisdictions that have (a) received low ratings in the recent past and/or (b) not been reviewed recently. Similarly, jurisdictions could be selected randomly¹² or randomly stratified by measures of risk (e.g., population size, past rating, etc.).

CDC/DSNS could also “audit the auditors” (see, e.g., Klerman, 2004) by reviewing the reviews conducted by state and local health departments.

¹² This is the strategy used in OSHA’s “programmed inspections.” See OSHA (2005) and IOM (2000).

6. ROLLING UP THE ASSESSMENTS INTO AN OVERALL EVALUATION

The previous sections have described the development of five new capabilities-based assessments and discussed options for how these new assessments might be administered. This section discusses how the proposed assessments might be combined – with each other and with the existing checklist tool – to generate an overall evaluation of state and local SNS readiness. In particular, we address the following:

- How to conceptualize “levels of preparedness?”
- Which assessments are required for each level of preparedness?
- How frequently should jurisdictions be assessed?
- How should assessments be administered and reviewed?
- How well must jurisdictions do on each assessment, and how should the results from individual assessments be combined into an overall rating (scoring)?

“Levels of Preparedness” Construct Is Anchored in the Conceptual Model

It is necessary to begin with a conceptual definition of “levels of preparedness” to guide the more technical choices about how to combine assessments mathematically. The conceptual model presented in Section 2 suggests such a construct.

As discussed in Section 2, our model suggests that SNS preparedness (and PHEP more generally) has both static and dynamic components. The static component lies in using planning, training, exercising, leadership and other processes to develop equipment, personnel and other inputs into building block capabilities that are relevant to a broad spectrum of potential threats. The dynamic component lies in the “improvisational” capacity to assess a situation and deploy, adapt, and combine capabilities aimed at reducing morbidity and mortality. Using this definition, we can distinguish three levels of preparedness, which we term “beginning,” “developing,” and “secure”:

- *Beginning*. The jurisdiction has acquired the *inputs* of preparedness (e.g., equipment, personnel) and has made satisfactory progress in implementing key preparedness *processes* (e.g., planning, training, exercising, organization/leadership), but has not demonstrated the ability to implement and adapt key capabilities in operational situations.
- *Developing*. The jurisdiction has acquired the necessary inputs and is engaging in adequate preparedness processes, *and* has demonstrated proficiency in performing separate tasks related to the building-block capabilities. However, the jurisdiction has *not* demonstrated an ability to execute core SNS functions *simultaneously*.
- *Secure*. The jurisdiction has demonstrated that it possesses adequate resources, is engaged in required preparedness processes, has demonstrated separate mastery of

core capabilities, *and* has shown the ability to successfully combine crosscutting capabilities in the execution of the full spectrum core SNS functions.

Figure 4 shows how the levels-of-preparedness construct maps to the PHEP conceptual model.

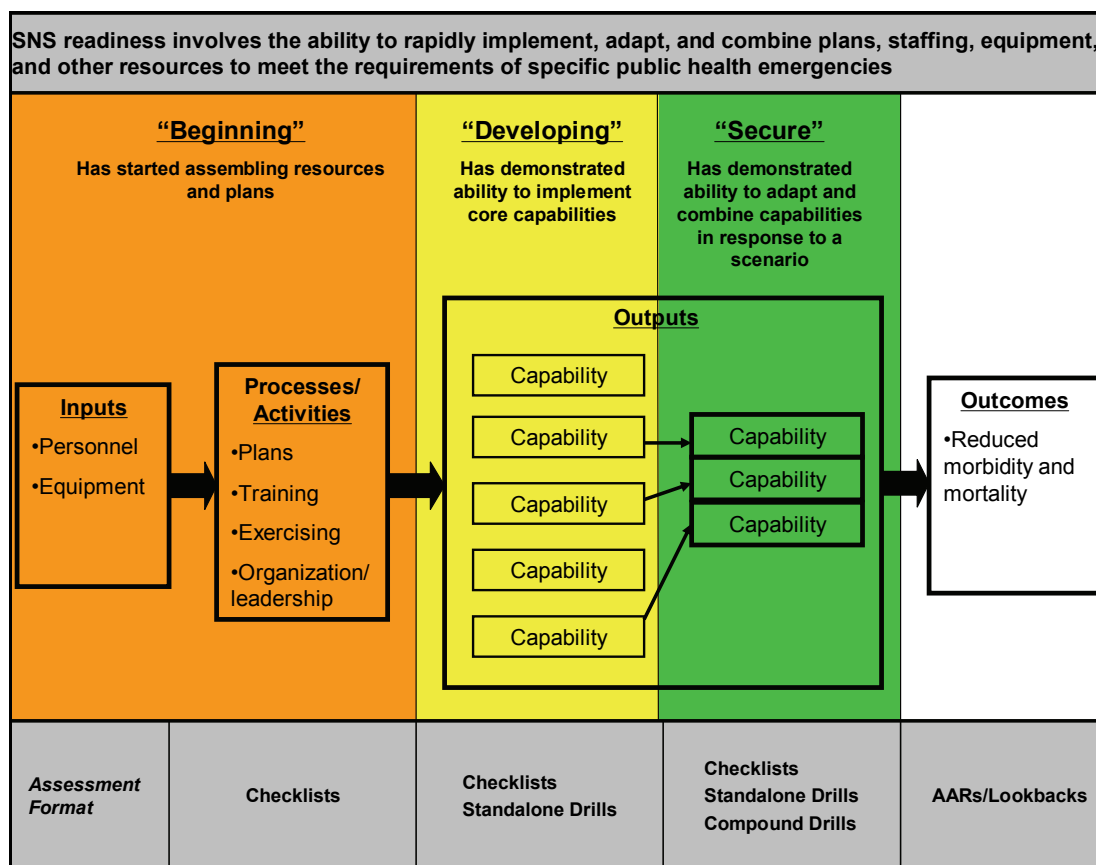


Figure 4: Performance Levels Related to Conceptual Model and Assessment Formats

We have proposed a color coding scheme for the three ratings, with Orange representing “Beginning,” Yellow representing “Developing,” and Green “Secure.” CDC/DSNS, however, might wish to choose other labels and color codings or to eliminate the color codings altogether. We suggest that this decision be based on what is most likely to be salient and constructive when communicating with state and local health departments.¹³

¹³ Note that the levels-of-preparedness construct is consistent with the “crawl-walk-run” philosophy articulated by CDC/DCDC/DSNS and others.

Each Level of Preparedness Would Require Different Assessments

The bottom row in Figure 4 shows which assessments are required to attain each level of preparedness. The current checklist-based tool does a reasonably good job of assessing the presence of key SNS inputs and processes. Thus, only the checklist tool would be required to attain the “Beginning” level. Operational assessments, however, are required to demonstrate that a jurisdiction is able to implement key SNS capabilities.¹⁴ As a result, separate administration of the capability-based assessments described above would be required to attain the “Developing” level. Finally, demonstrating that a jurisdiction is able to simultaneously implement key capabilities would require simultaneous drills or exercises of these capabilities. Thus, jurisdictions would be required to conduct full-scale exercises – what we call “compound drills” – to attain the “Secure” level.

One unresolved issue is whether jurisdictions should be required to administer all assessments on a regular basis, or whether they should have some choice about whether or when they seek a higher performance level. Given the importance of drills as instruments of process improvement the former option is advisable. However, CDC/DSNS may wish to provide flexibility to jurisdictions if there is concern about assessment burdens.

Requirements for Assessment Frequency, Administration, and Review Might Vary By “Jurisdiction Tier”

It is also necessary to determine how many of each assessment type are required to demonstrate level of preparedness, who should administer them, and how they should be reviewed.

We recommend that higher priority “Tier 1” jurisdictions be assessed more frequently than other jurisdictions (see Table 4). For instance, while all jurisdictions might be required to complete the checklist annually, CDC might consider requiring Tier 1 and Tier 2 jurisdictions to conduct all 5 capability drills every 2 years as a condition of attaining “Yellow” status, while requiring only 3 drills every 2 years from Tier 3 jurisdictions. Similarly, Table 4 suggests 2 compound drills (each including at least 2 component drills) per year for Tier 1 jurisdictions but only 1 per year for Tier 2 jurisdictions and 1 every 2 years for Tier 3 jurisdictions as a condition of “Green” status.

Table 4 also suggests tier-related differences in who administers and reviews the assessments. For instance, while state and CRI-related LHDs might be required to complete their checklist assessments during CDC/DSNS site visits, Tier 2 and 3

¹⁴ The current version of the checklist contains numerous items that relate to drills and exercises. For example, Section 2, item 4 is “State conducts call down exercises of all SNS function rosters to test response rates quarterly.” However, the current version does not provide enough guidance to ensure comparability of assessments. Nor does it provide clear performance metrics. Thus, these items should be regarded measures of preparedness processes, not outputs or outcomes.

jurisdictions might self-administer the checklist with peer review and random audits in order to ensure quality (see Section 5).

Table 4: Operational Details on Assessment Frequency

Jurisdiction Type	“Beginning” (Checklists)	“Developing” (Checklists + Standalone Drills)	“Secure” (Checklists + Standalone Drills + Compound Drills*)
Tier 1: States/ CRI sites	<u>Frequency</u> : Annual <u>By whom?</u> : Assessment by CDC during site visit <u>Threshold</u> : Exceed cut score on checklist tool	<u>Frequency</u> : Do all 5 drills every 2 years <u>By whom?</u> : Self assessment; audit during annual CDC site visit <u>Threshold</u> : Exceed cut score on X drills	<u>Frequency</u> : 2 per year <u>By whom?</u> : Self assessment with audit during CDC site visit <u>Threshold</u> : Exceed cut score on X drills
Tier 2: Other metro/suburban HDs	<u>Frequency</u> : Annual <u>By whom</u> : Self-assessment; peer review every 2 years. Subject to random CDC (or state?) audit <u>Threshold</u> : Exceed cut score on checklist tool	<u>Frequency</u> : Do all 5 drills every 2 years <u>By whom</u> : Self-assessment; with peer review every 2 years. Subject to random CDC (or state?) audit <u>Threshold</u> : Exceed cut score on X drills	<u>Frequency</u> : 1 per year <u>By whom</u> : Self-assessment; peer review every 2 years. Subject to random CDC (or state?) audit <u>Threshold</u> : Exceed cut score on X drills
Tier 3: rural, etc	<u>Frequency</u> : Annual <u>By whom</u> : Self-assessment; peer review every 5 years. Subject to random state audit <u>Threshold</u> : Exceed cut score on checklist tool	<u>Frequency</u> : Do any 3 drills every 2 years <u>By whom</u> : Self-assessment; peer review every 5 years. Subject to random CDC (or state?) audit <u>Threshold</u> : Exceed cut score on X drills	<u>Frequency</u> : 1 every 2 years <u>By whom</u> : Self-assessment; peer review every 2 years. Subject to random CDC (or state?) audit <u>Threshold</u> : Exceed cut score on X drills

*Each compound drill must include at least two of the capability drills. However, compound drills can also include elements not tested by the capability drills; that is, they may be embedded within other assessments.

There are a number of possible ways to define tiers. The system proposed in the table is organized loosely around population or risk. State health departments and CRI-related local health departments are in Tier 1, other urban and suburban areas in Tier 2, and rural areas in Tier 3.¹⁵ Another possible division among tiers might involve states and *original*

¹⁵ The urban-suburban-rural classification might draw upon the categories used by the U.S. Census.

CRI sites in Tier 1, non-CRI Urban Area Security Initiative (UASI) cities in Tier 2, and other areas in Tier 3.

A tiered system has the advantage of recognizing differences in jurisdictions' preparedness, prior level of preparedness, and prior experience with drills and exercises. A disadvantage is that such a system is more complex and more amenable to interpretation. Also, it is worth considering whether different preparedness requirements for different jurisdiction types would be viewed as legitimate – and therefore likely to be implemented – by key stakeholders.

Performance Thresholds and Scoring

Simply performing the assessments outlined in Table 4 is not enough to demonstrate competency in the skills assessed. There must be a way to determine whether jurisdictions have performed adequately in all essential categories. It is also desirable to be able to make finer-grained distinctions in the performance of jurisdictions within each broad level-of-performance category.

We discuss two options below:

- Set performance thresholds for each assessment and then determine the number of assessments that must be “passed” in order to achieve a given level of performance
- Use a mathematical algorithm to combine metrics into a smaller set of scores that can be compared to the CRI 48-hour requirement.

Defining Assessment-specific Thresholds. One option would be for CDC/DSNS to set performance thresholds for passing each assessment.¹⁶ CDC/DSNS has already defined thresholds for the checklist tool. Thus, to attain the Beginning level, jurisdictions would simply have to clear the threshold score on the checklist assessment.

However, additional steps are required to operationalize the Developing and Secure performance levels. First, CDC/DSNS would have to set performance thresholds for each of the metrics in the capability-based assessments (see Section 4 for the metrics). For example, a performance threshold for the pick list generation drill might be the demonstrated ability to generate a pick list within 30 minutes of receiving an apportionment decision and data files. Similarly, jurisdictions might be required to attain POD throughput rates of 300 patients per hour on the Dispensing drill.

Second, CDC/DSNS would have to determine how many assessments jurisdictions must “pass” to demonstrate proficiency at each level of preparedness. A Tier 3 jurisdiction could attain “Yellow” status by doing any 3 drills every two years, provided they exceed the threshold on a certain number of those 3 drills. Table 4 provides placeholders for these thresholds.

¹⁶ Setting performance thresholds was outside the scope of work for this project.

CDC/DSNS could also include opportunities for jurisdictions to earn “pluses” or “minuses” (e.g., “Orange-plus,” “Green-minus,” etc.). Such modifiers might be introduced through adding more thresholds to Table 4. For, instance, Tier 1 jurisdictions passing all 5 standalone drills might be awarded Yellow-plus, while those passing 4 might receive a Yellow and those passing 2 a Yellow-minus.

A somewhat different approach would be to award pluses and minuses for all categories based on the quality of after action reports and quality improvement efforts emanating from the assessment process. In other words, a jurisdiction scoring in any of the three performance ranges could increase its score by making extra efforts in quality improvement in the months following the assessment. The advantage of this approach is that it would call attention to the importance of quality improvement and likely increase the quality of AAR-related data available to CDC and others seeking to improve the PHEP evidence base. However, it would also require development of clear requirements and standards for AARs and quality improvement efforts.

Mathematical Algorithm. The scoring option described above has several advantages. It could be implemented fairly quickly and requires only that CDC/DSNS define performance thresholds for each of the assessment metrics. On the other hand, the evidence base on which such standards could be based is weak, and the standards might have a somewhat arbitrary quality. Moreover, the resulting scores – percentage of assessments “passed” for a given category – has little meaning in the SNS context.

Another approach is to use a simple mathematical algorithm to combine scores on the metrics into summary measures with more substantive meaning, such as the number of PODs that could be stood up in a given time period. We provide an example of such an algorithm in Appendix C.

Raters would enter required metrics from the capabilities-based drills (see Table 2) into the model that estimates the following consolidated performance metrics:

- The number of PODs with enough staff to meet the CRI 48-hour standard
- The number of hours required for the RSS to move materials out to PODs
- The number of PODs that could prophylax their assigned part of the entire population in a 36 hour period (assuming that the first 12 hours of the 48-hour timeline is required for CDC/DSNS delivery).

Note that CDC/DSNS would still need to develop techniques for assuring that assessed jurisdictions have enough PODs in their plans and performance thresholds for the number of hours to move materiel out to PODs.

Because each level of performance relies on different combinations of data (checklists, standalone drills, compound drills), it will be necessary to customize it for each of the three levels of performance. The example in Appendix C is designed for the Developing category. However, the model could be readily adapted for the Secure category. Instead

of using measurements taken from individual drills, jurisdictions would collect the metric data in the context of compound drills (e.g., a full RSS drill) or from full-scale exercises.

Adapting the roll up model for use in the Beginning category would be difficult, as jurisdictions in that category are primarily assessed by plans-based checklists rather than operations-based drills that provide metrics. However, it might be possible to substitute checklist-based estimates for some of the drill-based metrics in the current draft.

- In some cases, we could replace metrics achieved by exercises with those achieved “on paper” or by simulation, such as estimates of dispensing throughput capability using Bioterrorism and Epidemic Outbreak Response Model (BERM).
- In other cases we might be able to give credit based on a checklist score. For instance, one might give credit for the ability to unload the trucks in an hour if the jurisdiction has all the equipment and sufficient personnel in place.

These modifications have not been developed. The primary difficulty lies with being able to express confidence on performance ability based only on paper assessments.

It is important to emphasize that the algorithm will require more development before being ready for use and that we present this version as a discussion draft. First, there are a few metrics in the model for which there are not yet drills and metrics (e.g., unloading, picking). Second, the logic of this draft model must be subjected to greater scrutiny before it is used to evaluate jurisdictions. In particular, the model makes a number of assumptions about work flow that need to be assessed by subject-matter experts, particularly with regards to which operations happen in sequence versus happening in parallel, thus affecting the completion time of the overall mass prophylaxis operation.

Finally, it is important to bear in mind that a roll up algorithm based on idealized drills cannot fully capture the challenges that would be faced in a real emergency. Thus, success in these drills cannot guarantee that a jurisdiction is prepared for an actual disaster. But the inability to perform well on these drills does indicate cause for concern and identifies areas in need of improvement.

7. AGGREGATING LOCAL ASSESSMENTS INTO STATE-LEVEL SCORES

Another issue is how to combine ratings of local HDs into aggregated state-level ratings. Once each jurisdiction has completed its required assessments and the scores have been “rolled up” into an overall evaluation, how should those jurisdiction-level evaluations be combined?

Aggregation Options Depend on Degree of Formal Quantification in Roll Up

Options for aggregation will depend on the level of quantification at the roll up stage. That is, the more quantification is used in combining each jurisdiction’s individual assessments into an overall rating, the easier it will be to use straightforward and standard statistical algorithms to combine scores from various jurisdictions (some such options are presented below). We present three options for aggregation, each of which assumes that each jurisdiction’s rating is expressed in terms of some sort of quantitative index.¹⁷

Option 1: Simple (unweighted) Mean

The simplest option is the unweighted mean which is simply the sum of all local scores (LHD_i) divided by the total number of LHDs (N_{LHD}) (see Equation 1).

$$\frac{\sum_{i=1}^n LHD_i}{N_{LHD}} \quad (1)$$

Option 2: Weighted Mean

An unweighted mean, however, assumes that all LHDs are of equal importance to the aggregate score. Some local sites might warrant greater weight than others in the final aggregate score. To calculate a weighted average, multiply each local score (LHD_i) by the local weight (W_i), sum all of the products, and divide by the sum of the weights (see Equation 2).

¹⁷ The first roll up strategy proposed above – the one built around assessment-specific performance thresholds – is deliberately quite free of formal quantification, both for simplicity and to keep the focus on the content of the assessments rather than on small variations in scores, which are less instructive. However, it would be useful to construct a quantitative index *after the fact* to support aggregation of scores across jurisdictions. For instance, a quantitative index could be based on the percent of required metrics on which a jurisdiction cleared the relevant cut score. Alternatively, one could derive a performance index from the mathematical model described in Section 6 and provided in Appendix E.

$$\frac{\sum_{i=1}^n LHD_i * W_i}{\sum_{i=1}^n W_i} \quad (2)$$

If, for instance, population were used as the weighting variable, this would effectively weight the LHD scores by the percentage of the total population in each LHD. Population would provide the most obvious and readily understandable weight. However, CDC might also choose to use other measures of risk or importance as weights, including recent work by RAND, models developed by various insurance companies, and others (see, e.g., Willis, et al., 2006).

Option 3: Using LHD Performance Thresholds

Both the unweighted and weighted means assume that low scoring LHDs in a state are offset by high scoring LHDs, and vice-versa. However, one could argue that even a small number of extremely low performing LHDs is unacceptable. For instance, an LHD with particularly weak surveillance systems might be an easy point of entry for an infectious disease. Similarly, a weak LHD might be unable to provide mutual aid to other LHDs experiencing high surge demands. Thus, CDC might wish to incorporate some minimum performance threshold (a “pass-fail” cutoff point) into the aggregation of LHDs. The simplest way to incorporate a minimum performance threshold is to use the percentage of LHDs achieving a “passing” score in each state.

8. NEXT STEPS

The proposed improvements to the SNS assessment system described above provide a roadmap. In this final section, we describe a number of concrete steps that are required to move down that road.

Step 1: Field Testing

Having drafted the capability assessments, the next critical step is to field test them. This will help spot and correct problems before the assessments are rolled out on a larger scale. Field testing should be viewed as an opportunity to improve the assessments. Field testing should involve the following:

- *Continuing desk audits.* Share the operations manuals with a small number of jurisdictions and ask them to identify ambiguities and obvious problems with the manuals
- *First round of revisions.* Revise the manuals based on the initial feedback from jurisdictions
- *Operational walk-through.* Ask jurisdictions to walk through the assessment procedures described in the manuals (e.g., run a pick list generation drill)
- *Debrief.* Conduct a semi-structured debrief with jurisdictions to identify remaining problems
- *Second round of revisions.* Revise the manuals based on feedback from the operational walk-throughs described above

Throughout, the field testing should seek to assess validity, reliability, and feasibility.

- *Validity.* Validity is the extent to which an assessment adequately captures the characteristics it seeks to measure. If one imagines repeated administrations of an assessment, validity requires that the mean value taken by the assessment converges on the true value. Semi-structured debriefs with those involved in the assessments will produce useful information on face validity. In addition, the process of selecting and developing the assessment –summarized above – should provide some assurance that the assessments contain sufficient content and construct validity – that they measure sample relevant content domains and that they are related to other characteristics in expected ways.
- *Reliability.* If validity is the extent to which an assessment converges on the true value, reliability is the extent to which there is low variance around that value. A reliable assessment will tend to produce the same values when given to similar subjects or groups. One key concern in developing SNS assessments has been identifying metrics that are comparable across jurisdictions and over time. Thus, the pilot process will seek to expose several groups of individuals in a jurisdiction

- through semi-structured discussions – to the same materials to assess whether they form similar understandings of that material.
- *Feasibility.* Feasibility is the extent to which an assessment’s burdens are reasonable given existing skills, resources, and motivations. This will be assessed through semi-structured discussions with participants. In addition, we will seek detailed feedback from key individuals on the estimates of assessment burdens provided below.

Step 2: Initial Roll Out and Ongoing Evaluation

After conducting initial pilot tests, the next step is to roll out the assessments to a larger group of health departments for initial implementation. It might be wise to begin by requiring only Tier 1 jurisdictions to conduct the assessments, making them optional for lower risk (i.e., Tier 2 and 3) jurisdictions. Tier 1 jurisdictions generally have more prior experience running drills and exercises and are more likely to have the requisite skills for conducting the assessments described above. A period of optional participation for lower- risk jurisdictions – along with a program of training and technical assistance from CDC/DSNS – would increase the likelihood that less experienced jurisdictions would be able to conduct the assessments successfully.

During this period it would be wise to engage in continuing evaluation activities, including:

- Monitoring implementation difficulties, including unclear wording, skills deficits, and other barriers to implementation
- Looking for opportunities to assess criterion validity – the extent to which scores on the assessments correlate with the ability to respond to actual public health emergencies. Hopefully, there will be no large scale public health emergencies, in which case these analyses might seek to focus on responses to smaller scale events. We are not aware of any previous attempts to validate operational assessments in this way and it is likely that doing so will require considerably more planning.

Step 3: Set Performance Standards

As noted above, the focus of this project was restricted to developing assessments. Thus, we have sought to develop a set of yardsticks for assessing SNS readiness. An important next step is to determine what levels of performance are deemed “acceptable.” Specifically, decisions will have to be made about:

- Cut-points for scoring the performance metrics described in the manuals (e.g., throughput for the dispensing drill)
- How many of the metrics must jurisdictions exceed the cut-point in order to attain the various performance levels (beginning, developing, secure)

While setting standards is extremely important to the effectiveness of the SNS assessment system, it would be reasonable to delay this step in favor of getting the assessments ready for broader roll out and giving health departments a chance to become accustomed to them.¹⁸

Developing Assessments of Other Operational Capabilities

In addition to these specific steps, CDC/DSNS should keep several other considerations in mind as it moves forward with finalizing and rolling out the assessments described above.

Several other assessments would be desirable for a fully-developed SNS assessment system. Future assessment development efforts might focus on the following capabilities:

- Security operations
- Public information and communication
- Just-in-time training
- Picking and palletizing
- Command and control functions, including SNS request, apportionment, risk assessment, asset management, and coordination.

Command and control issues are worth further discussion because, unlike other capabilities, they require more judgment-laden qualitative indicators. We believe that such qualitative indicators could be developed by examination of approaches and perspectives from cognitive science and group decision making, as well as more practice-oriented literature on team function in aviation, surgery and other operational areas.

Alternatively, command and control capabilities could be assessed through tabletop exercises linked to computer models of disease outbreaks and countermeasure delivery. Although it appears that many important elements of such a model-based assessment system are currently available, considerably more work is required to determine how those pieces should be linked in an integrated system.

Linking Assessment with Quality Improvement

Performance measurement systems should improve preparedness, not just represent it. Measuring key attributes of SNS preparedness does help focus attention on key problems and can provide a common vocabulary for discussions about process improvement, but these attributes are rarely sufficient to drive real improvements (see, e.g., Seid et al., 2006). Thus, CDC/DSNS should promote tools for closing performance gaps revealed by the assessment system.

¹⁸ However, the recently enacted Pandemic and All-Hazards Preparedness Act (S. 3678) requires HHS to develop, and as appropriate adopt, standards within 180 days.

Ideally, such “change packages” are based on rigorous scientific knowledge and extensive experience, as is often in the case in clinical medicine. The evidence and experience base in PHEP, by contrast, is more limited. It is likely that improvement efforts for the foreseeable future will have to rely on the sharing of experience among state and local HDs and consultation with CDC technical advisors. In addition, the planned SNS quality improvement learning collaborative mentioned above should provide a forum in which change packages can be developed and tested, along with the assessments described in this report.

Assigning Consequences to Assessment Results

Process improvements are more likely to come about with careful consideration of what consequences should attach to performance on assessments. Consequences can be both positive and negative. Positive consequences might include some sort of certification that a jurisdiction has achieved a certain level of SNS preparedness or even increase funding. Negative consequences should be used with caution. For instance, cutting funding to low performers might only limit their capacity to improve, thus undercutting the purposes of the assessment system. A less draconian – and more constructive – option might involve increased technical assistance to low-performing jurisdictions. Consistent low-performers, moreover, might be threatened with partial loss of budgetary and programming authority, as is the case for low performing schools under the No Child Left Behind Act. We emphasize, however, that any schedule of consequences should be subject to careful analysis to avoid unintended effects.

Improving the Evidence Base

Ideally, performance assessments and standards should be heavily informed by a solid evidence base. This is especially important in PHEP where the fortunate rarity of large-scale events and the costs associated with large-scale exercises leave few opportunities to observe jurisdiction readiness directly. In such areas, assessments and standards must be based largely on knowledge about the connection between “upstream” structures, resources, and processes and downstream public health outcomes. We believe that the operational assessments provide a valuable opportunity to collect data that deepens our knowledge of these linkages. For instance, collecting data on how jurisdictions conduct staff call downs along with data on call down times during drills could provide an important data source for exploring these relationships. Thus, we strongly recommend that CDC/DSNS take care to collect and maintain these data, use it for its own internal research purposes, and make it available to the larger research community.

One area of research particularly germane to the SNS assessment system is the degree to which the results of relatively small-scale assessments can be used to predict jurisdictions’ ability to mount responses to large-scale public health emergencies. Throughout, we have assumed that the ability to, for instance, run a successful dispensing operation in a single POD is at least some indication – along with the ability to call down a large number of staff and activate a large number of facilities – of the ability to run

multiple PODs simultaneously across a region. At this time, however, this is just an assumption.

Efforts to address this problem of scale in assessment might proceed along two paths. First, the results of small scale assessments in jurisdictions can be compared with performance on full-scale exercises and – better yet – responses to real emergencies. Second, advances in mathematical and computer simulations might illuminate these issues and help to identify drivers of and barriers to effective responses. Such models could eventually form the basis of a more rigorous roll up strategy, where results from small-scale drills could be entered into validated mathematical models in order to generate reasonably good predictions of the ability to respond to larger scale emergencies.

Consider Efforts to Improve Coordination with Other PHEP Assessments

As noted earlier in the report, the conceptual framework and assessment concepts used to develop tools for measuring SNS readiness should be broadly applicable to other areas of PHEP. Utilization of these concepts could help foster a greater degree of coordination between SNS assessment efforts and those directed to pandemic influenza and the CDC's Cooperative Agreement on PHEP (CDC, 2006).

APPENDIX A: ANALYSIS USED TO SELECT SNS ELEMENTS AROUND WHICH TO BUILD ASSESSMENTS

We used critical path analysis (see, for instance, McClain, Thomas, & Mazzola, 1992) to identify the “rate-limiting factors” in the SNS process – those things that, if not done in a timely fashion, could prevent delivery of medications to the public with the 48 hours required by the Cities Readiness Initiative guidance. Examination of these factors helped identify the specific capabilities that are represented in Section 2’s conceptual model (see Figure A.1).

The results of the analysis are summarized in the precedence diagram in Figure A.1. Precedence diagrams show the temporal flow of key tasks in a process and help to illuminate the most critical process elements. Each box in the diagram represents an SNS-related activity, with some carried out in sequence and others in parallel. Arrows indicate which functions are predecessors of other functions. For instance, the initial SNS shipment must arrive before the RSS can unload and repackage the shipment, which must occur before the distribution process can deliver materiel to the PODs, which must occur before PODs can begin dispensing.¹⁹

The diagram is organized into horizontal streams of function-specific activities, including Receipt, Staging, and Storage (RSS); Distribution; Dispensing (POD), and Public Information. We have also included a stream of activities related to inventory management – however, we recognize that some of the activities will also involve command-and-control (e.g., selection of dispensing sites). We have also included prophylaxis of local staff from local caches, which, though not part of the SNS program, are nonetheless important elements of an SNS response.

The precedence diagram is also useful in identifying key SNS capabilities that are candidates for drill development. *Crosscutting* capabilities are more general response elements that can be assembled and combined in order to execute a variety of tasks. For example, setting up a POD requires a solid command and control function, the ability to call-up and mobilize staff, and so on. Visually, *crosscutting* capabilities are those activities that appear in multiple activity streams. These are highlighted in Figure A.1 by the vertical, solid-line,²⁰ ovals (more detail on potential drill operations is provided in the earlier report).²¹

¹⁹ In an ideal SNS response, the total time from request to dispensing would be limited only by the physical time it takes to move the SNS materiel. However, delays in any of the activities that parallel the tasks on the critical path could become rate-limiting factors and delay the response.

²⁰ The dashed oval around “setup” indicates that application of this capability will differ more substantially across RSS, Distribution, and PODs than other crosscutting capabilities.

²¹ Note that the precedence diagrams do not include the re-ordering process. Unfortunately, we have not had time to fully understand this process and have not included it as a candidate for drill

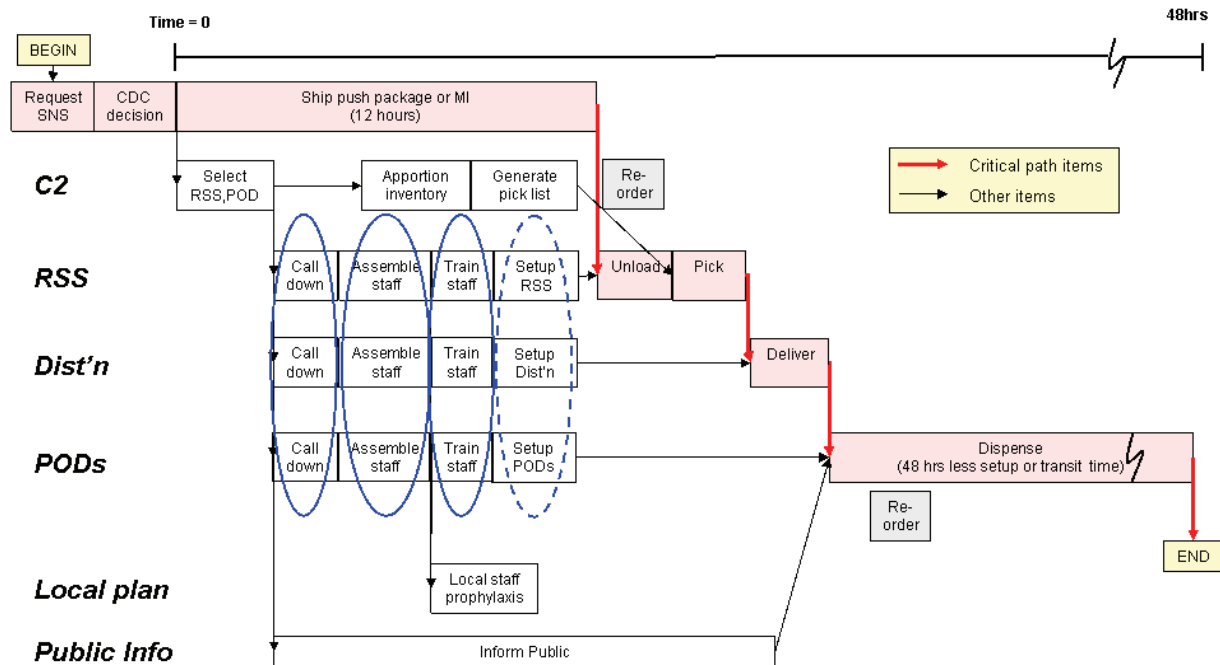


Figure A.1: Precedence Diagram with Crosscutting Capabilities Highlighted

In addition to these generic crosscutting capabilities, the precedence diagram also helps identify *function-specific capabilities*, highlighted in the diagram by the horizontal ovals in Figure A.2. These capabilities are those activities that “directly facilitate the dispensing of drugs and vaccines,”²² including requesting SNS assets, receipt of the SNS cache, distributing the materiel to Point of Dispensing sites (PODs) or other dispensing nodes, actual dispensing of the materiel, and so on.

development. For instance, it is not yet clear to us whether one should test the parts of the reordering process or whether it should be tested as a whole.

²² This phrase is taken from a ConOps statement from AHRQ on the design and operation of dispensing/vaccination clinics (Hupert, et al., 2004). The document distinguishes “core” and “support” functions, which is roughly analogous to our distinction between “core SNS functions” and “crosscutting capabilities.” We believe that the term crosscutting is more apt and serves to highlight the extent to which different core functions draw upon a largely common pool of capabilities. Similar distinctions are also found in the literature on Six Sigma and other business improvement packages (see, e.g., Pande, Neuman, & Cavanaugh, 2000).

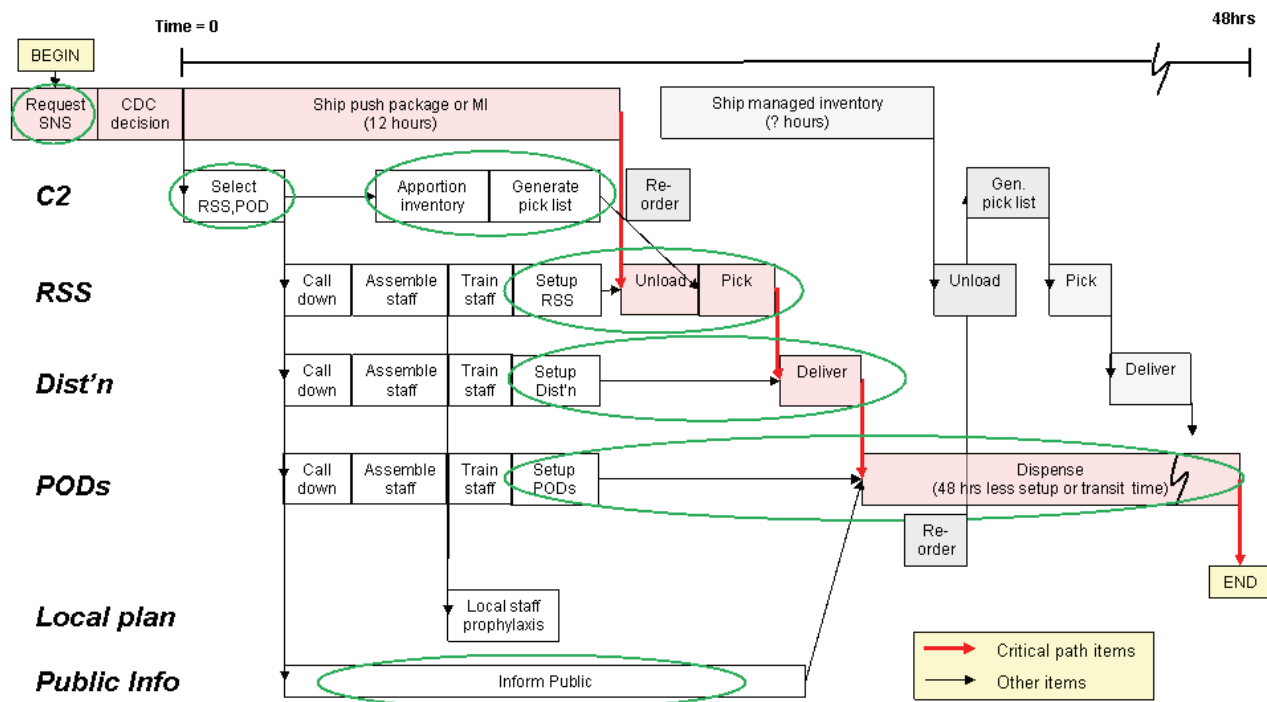


Figure A.2: Precedence Diagram with Function-Specific Capabilities Highlighted

The results of this paper-and-pencil exercise were validated by discussions with SMEs, practitioners in state and local health departments, examination of after action reports, and direct observation of SNS drills and exercises.

APPENDIX B: OPERATIONS MANUALS FOR THE PROPOSED DRILLS

METRICS FOR STAFF CALL DOWN DRILL

Operations Manual

OVERVIEW

Call down is the ability to contact and mobilize staff²³ to perform emergency response functions. It is a cross-cutting capability, applicable to a wide variety of Strategic National Stockpile (SNS) functions, including dispensing, warehousing (RSS), distribution, security, command centers, among others.

Successful call down requires (a) contact lists that are up to date, readily accessible, and usable; (b) workable methods for contacting staff in a timely fashion; and (c) the ability of staff to report to their assigned stations in a timely manner. Thus, the manual describes procedures and metrics for validating call down lists, testing communications, and estimating or documenting staff's ability to report to their stations.

A small set of *required* metrics should be reported to CDC as part of its assessment tool for SNS readiness. These include:

- *Call completion time.* The amount of time required to complete the call down process
- *Acknowledgement completion time.* The amount of time needed to receive acknowledgements from staff who confirm receipt of the call down message, regardless of whether they would be able to report for duty.
- *Contact rate.* The number of staff successfully contacted as a percentage of all staff on the call down list
- *Assembly rate.* The number of staff who report being able to arrive at a designated site within a designated time, as a percentage of all staff on the list

An ideal call down drill would be conducted without notice and involve actual assembly at a site. However, requiring personnel to assemble on short notice might create undue burdens. Moreover, it is not clear that travel times would reflect real emergency conditions (e.g., traffic, compromised infrastructure). It seems most important to preserve the no-notice nature of the drill in order to gauge readiness when staff have not been given advance warning or planning time. Therefore the procedures and metrics outlined below are based on a no-notice, *hypothetical* assembly, call down drill.

This drill is designed to be a standalone assessment but could also be linked with other assessment activities or embedded in larger exercises/drills. Ideas for other uses of the drill format are provided at the end of the manual.

²³ This includes volunteers and others for whom emergency response may not be their day-to-day activity.

PLANNING CHECKLIST

Prior to running the drill, jurisdictions should ensure that the following planning steps are completed:

Recruit drill participants. Required participants include (a) players, (b) a controller, and (c) an evaluator. The evaluator and controller may be one and the same person.

- *Players.* Jurisdictions should run this drill with the actual number of callers – and preferably with the same individual people – that they would expect to use in an actual emergency, as outlined in their SNS plan. (Jurisdictions should be careful not to use callers who would have other responsibilities in an actual emergency.)
- *Controller.* The drill controller will indicate when the drill starts and stops. This drill is designed to evaluate performance, so the controller should not take part as a drill player, and should *not* intervene while the drill is in progress.
- *Evaluator.* The evaluator is responsible for timing the drill. In addition, the evaluator should take notes on the drill in order to identify opportunities for improvement. An evaluator can be a peer from another jurisdiction, a CDC representative, or another individual. In some situations, more than one evaluator may be appropriate.

Ensure that call down list is up to date. As part of the jurisdiction’s development of an emergency plan, a list of personnel required for each function must be developed. The emergency plan should ensure the following:

- *Up-to-date lists.* Keeping the list of available personnel up to date is a continuous process and may be one of the biggest challenges. (One of the benefits of this drill is to audit the accuracy of the list.)
- *Non-overlapping lists.* The call down list for each function should be unique: functions that would be activated simultaneously (including other emergency functions outside of SNS) cannot share the same personnel.
- *Enough personnel for prophylaxis of entire community.* The CRI goal is that the jurisdiction would be able to prophylax 100 percent of the community within 48 hours. In order to meet this goal, ideally the call down list should have enough personnel to staff all the facilities that would be needed to meet the CRI target. For instance, the POD list should include enough staff (including registered volunteers) to cover *all* PODs, and should include enough medically-trained and other staff for handling both oral and injected countermeasure scenarios.
- *Backup personnel.* The list should also contain *extra* personnel to compensate for those who may be unreachable or unavailable to work. (The number of extra personnel needed is a decision for the jurisdiction. The CDC Smallpox

Vaccination Clinic Guide, Annex 3-11, suggests 20 percent extra for no-shows, breaks, surge needs, and other contingencies.)

It is important to stress that the drill should be run whether or not the call-down lists contain enough staff and volunteers to meet the CRI program goals, and whether or not the lists include extra names.

Ensure availability of calling and response-monitoring equipment. Drill players should have access to all calling equipment and any call response monitoring technology that would be used during a real emergency to keep track of the call responses received. This may be as simple as a set of phones and the list of contacts.

Protocol for dealing with non-responses. Jurisdictions, as part of their planning process, should develop their own protocol for how callers or automated systems should deal with voicemail, busy signals, messages left with others, and wrong numbers when placing calls.

Decide which call down lists to test. Before running the drill, jurisdictions will have to decide which call down lists will be tested, such as (a) EOC (b) RSS (c) distribution, (d) POD, (e) security, or (f) public information/communication. Jurisdictions may combine tests of different functions in a single drill or run separate drills by function.

Alternate drills during work hours and non-work hours. Jurisdictions should include, during the course of a year, drills conducted both during normal work-hours and after normal work-hours. The time and day of the drill should be noted on the data capture sheet (see below).

DATA COLLECTION PROCEDURES

Please note that calling procedures will vary somewhat depending upon whether the jurisdiction uses centralized manual calling (i.e., manual calling conducted from a centralized location), automated calling, or a calling tree (where calling responsibilities are distributed).

Test the entire call down list. When testing a staff list, jurisdictions should test the complete list for a given function. For instance, if the POD call down list is being tested, the jurisdiction will seek to contact *all* staff (including registered volunteers) for *all* planned POD sites. Furthermore, if plans call for different staffing at PODs depending on the scenario (e.g. chemoprophylaxis vs. vaccination), the jurisdiction should call staff members for both scenarios.

Keep in mind that the goal of the drill is to provide a semi-standardized measure of call-down capability, not test the response to a particular scenario. Calling all staff on the list will provide some measure of consistency across tests (at least those conducted for the

SNS assessment system) and helps work toward ensuring that each jurisdiction is prepared to respond to high-intensity scenarios.

Steps for centralized manual calling. Although many jurisdictions use some form of automated calling system, some do not. Additionally, jurisdictions should also be prepared to conduct call down procedures manually in case automated means fail, with callers placing calls to individual staff. We begin with steps for manual calling.

- Players should be instructed to work as they would during a real emergency (e.g., with the same sense of urgency) and to begin making calls immediately.
- Timing for the drill **begins** when the first call is made.
- Callers place unannounced calls to all of the staff listed as eligible for work for the facility or function in question. **Callers should clearly state that the call is part of a drill.**
- For each person called, the caller should follow these steps:

Step 1: The caller should record whether each person on the list could be reached and the time at which the call was placed.

Step 2: If the person called cannot be reached, the caller should follow their local protocol as to whether a message should be left, or any further effort made to contact the person. If the person called is declared unreachable, that should be recorded.

Step 3: If the person called is reached, the caller will then ask the following set of questions (see Box 1). The caller should stick close to the script to ensure consistency in the task and to ensure that the drill metrics are comparable over time and across jurisdictions.²⁴

- The timed drill **ends** when an attempt has been made to contact each person on the list, and the last call made has been completed.

Steps for automated calling. Jurisdictions using automated call down systems should instruct all called staff to acknowledge receipt of the message, regardless of whether the staff can report for duty. Additional steps depend on whether the system allows respondents to acknowledge the call and report an assembly time.

- **If system allows for automated response function present.** If the automated system allows for it, jurisdictions should program their systems to require respondents to acknowledge receipt of the call and to indicate the length of time it

²⁴ Jurisdictions may modify these questions, but changes should not substantially change the question content, and all callers should use the same modified script. Most importantly, jurisdictions should not add to or subtract any questions from the list.

would take them to report to a designated location in the event of a real emergency. As with the manual drill, timing the automated calling process begins with the initiation of the first call and ends after calls have been placed to all individuals on the call down list and the last call made has been completed.

Box 1: Call Down Script for Centralized Manual Calling

- “This is a drill being conducted by [jurisdiction]. Your name is on (jurisdiction’s) list of emergency staff. If this were a real emergency, you would be asked to report to (designated site).
- “Are you prepared to report to (designated site) immediately/as soon as possible?” (Yes/No)
- “How soon could you get to (designated site)?” [Record answer.] Callers should not reveal a target assembly time, which might bias responses.
- “Again, this is only a drill. There is no need for you to take any action as a result of this call. Thank you.”

- **If system does not allow for automated response function.** If a jurisdiction’s automated system does not include a function for receiving responses from the called staff, an alternate means of receiving acknowledgements will be required. One possibility would be to stand up the public information call center and use it to receive calls back from the called staff. In this case, there would need to be additional timing of the acknowledgement process, recording the times when each call is received back from each of the called staff.

Steps for calling trees. Jurisdictions using a calling tree (i.e., where calling responsibilities are distributed across many staff and not conducted from a centralized location) should use the following steps:

- Players should be instructed to work as they would during a real emergency (e.g., with the same sense of urgency) and to begin making calls immediately.
- Timing for the drill **begins** when the first calls are made.
- Callers place unannounced calls to all of the staff listed as eligible for work for the facility or function in question. **Callers should clearly state that the call is part of a drill.**
- For each person called, the caller should follow these steps:

Step 1: The caller should record whether each person on the list could be reached and the time at which the call was placed. (This information should be provided to the drill controller or evaluator as soon as possible after calls are completed.)

Step 2: If the person called cannot be reached, the caller should follow their local protocol as to whether a message should be left, or any further effort made to contact the person. If the person called is declared unreachable, that should be recorded.

Step 3: If the person called is reached, the caller will then ask the following set of questions (see Box 1). The caller should stick close to the script to ensure consistency in the task and to ensure that the drill metrics are comparable over time and across jurisdictions.²⁵

Box 2: Call Down Script for Calling Tree

- “This is a drill being conducted by [jurisdiction]. Your name is on (jurisdiction’s) list of emergency staff. If this were a real emergency, you would be asked to report to (designated site).
- “Are you prepared to report to (designated site) immediately/as soon as possible?” (Yes/No)
- “How soon could you get to (designated site)?” [Record answer.] Callers should not reveal a target assembly time, which might bias responses.
- “Again, this is only a drill. There is no need for you to report to any location as a result of this call. However, you should call the people on your designated list (per the jurisdiction’s calling tree plans) using this same script. Thank you.”

- When callers have completed their calls (or attempted calls), they should report their results back to the drill controller. One way to do this is to stand up the public information call center and use it to receive calls back from called staff.
- The drill ends when an attempt has been made to contact each person on the list. The timed portion ends when the last of the members of the phone tree report back to the drill controller or call center.
- Instruct all callers to use the script provided in Box 2. Ideally, this would be incorporated in regular training, with the drill instructions distributed as part of each calling tree member’s set of standard operating procedures. Realistically, jurisdictions may have to remind calling tree members of the drill procedures and contact lists. Jurisdictions should beware that distribution of the script just before an unannounced drill might limit the element of surprise, thus biasing the results of the drill.

²⁵ Jurisdictions may modify these questions, but changes should not substantially change the question content, and all callers should use the same modified script. Most importantly, jurisdictions should not add to or subtract any questions from the list.

REQUIRED METRICS

Call completion time. The amount of time needed to place calls to all staff on the list. This metric should be reported per unit location (i.e., per POD²⁶ or RSS) as well as aggregated across all locations.²⁷ If a calling tree is used in place of a dedicated calling staff, the amount of time required to complete the drill can be reconstructed from the times when call-backs are received from each of the called staff, per the directions provided above.

Acknowledgement completion time. The amount of time needed to receive acknowledgements from staff who confirm receipt of the call down message, regardless of whether they would be able to report for duty.

- In a manual calling situation, the acknowledgement time may be considered the same as the call completion time, except for staff who return calls based on messages left with them.
- Jurisdictions should record the call back time of each individual call.
- To prevent late responders from skewing metrics, percentile times should be reported: (a) the time for 75% of responses to be received, (b) the time for 90% of responses to be received, and (c) the time of the last received response.

Contact rate. This is the percentage of staff the jurisdiction sought to contact who were actually reached by a caller or, in the case of an automated system, who acknowledge receipt of the message, regardless of how much later they respond. This is the number of staff contacted successfully (i.e., with acknowledgement) – regardless of whether they report being able to show up – divided by the total number of staff called.²⁸

$$\text{Contact rate} = \frac{\text{\# Staff successfully contacted}}{\text{Total \# staff on call down list}}$$

²⁶ Note that some jurisdictions might wish to have different target start times for each POD. If so, this should be reflected in the metrics. However, separate start times are not necessary for purposes of the assessment.

²⁷ The completion time is important because in an actual emergency, the call down process would have to start shortly after the decision had been made to activate particular RSS and POD locations, and have to be completed quickly enough to enable staff to assemble, conduct just-in-time training, and perform the facility setup, before the SNS material arrives.

²⁸ The contact rate is important because it lets the jurisdiction get an idea of how many people will be unreachable at a given moment on a given day. This will vary depending not only the time of day, but the time of year (e.g., it will be easier to contact people on a workday morning than late at night on a holiday weekend). Thus, despite the temptation to run the drill during the workday when calling staff are present, jurisdictions may want to run their drills at different times of day, week, and year, in order to get a wider sampling. Contact numbers for their called staff will also vary depending on the time of day they are called.

Assembly rate. This is the number of staff who report being able to assemble at the designated location by the target assembly time divided by the total number of staff on the call down list being tested.

$$\text{Assembly rate} = \frac{\text{\# Staff able to report to designated site by the target assembly time}}{\text{Total \# staff on call down list}}$$

- Callers should not reveal the target assembly time to those being called, in order to avoid biasing the answers. Thus, jurisdictions will need to determine after the drill what percentage of respondents' answers would get them to the site within the target assembly time.
- This metric should be reported per unit location (i.e., per POD or RSS) as well as aggregated across all locations.²⁹

OPTIONAL METRICS

Actual assembly time. One way to validate assembly rates and times is to require actual – as opposed to hypothetical – assembly. Here, the assembly rate would be calculated as above, but with the number of staff actually reporting to the site in the numerator.³⁰

More detailed questions. The assessment described above relies on staff to self-report whether they could report to a designated location and how long it would take to do so. However, respondents might underestimate barriers and overestimate their ability to successfully overcome those barriers.

- *Ask probing questions.* Callers might ask a more probing set of questions designed to elicit more reliable answers about assembly than is possible using automated call down systems, and given the constraints of a timed drill. For instance, instead of simply asking respondents to report their expected assembly time, callers could start with a short set of questions designed to ensure that

²⁹ The assembly rate is similarly important because it lets the jurisdiction get an idea of how many people will be unavailable on a given day. We expect this estimate to be somewhat optimistic; more people may self-report being available in an emergency than would truly report in a real emergency. On the other hand, if a no-notice drill were to require staff to actually drop everything and show up, we might expect the observed assembly rate for the exercise to be lower than the number who would report in an actual emergency.

³⁰ Actual assembly drills, however, can be burdensome. Moreover, it is not clear that assembly rates and times observed during drills provide good estimates of assembly rates and times during real emergencies, when traffic and infrastructure conditions might be quite different than in non-emergency circumstances.

respondents consider the full range of possible constraints in their answers. The call down script might include a question such as:

“Please list any factors that might reasonably prevent or delay you from reporting, such as responsibilities to children, pets, or other family members, traffic, or job responsibilities.”

- *Piece together response times from narrower questions.* Alternatively, the script could seek to “piece together” the various steps most individuals have to accomplish in order to assemble at a designated location:

“How long would it take until you could leave your present location?”

“How much travel time do you estimate will be required?”

The idea is that respondents are less likely to suffer from biases in responding to specific questions than the more general question, “How long until you could assemble?”

Adding these questions, however, could create problems for automated calling systems. Thus, jurisdictions— even those with have automated systems – should periodically run a manual call down drill and include additional questions in the caller script. In addition to providing additional response accuracy, this would also allow jurisdictions using automated systems to ensure that they are prepared for situations where their automated system does not function properly.

LINKING THE DRILL WITH OTHER ASSESSMENTS

As noted above, the call down drill is designed to be run as a stand-alone assessment. However, it could also be linked with other assessments or embedded in larger drills and exercises. Health department officials are encouraged to be creative in developing ideas for linkages with other drills, exercises, and assessments. Here are a few ideas:

Link call down drill with test of call-in center. Jurisdictions using an automated system could program it to request that respondents call a call-in center to report their assembly information. This would also allow the jurisdiction to test their ability to run call-in centers, which is often an important part of an effective public information/communication strategy.

Link with test of decision to activate emergency response plan. During a real emergency, a full staff call down would likely be preceded by a decision by a jurisdiction’s senior command staff about whether the situation warrants a full call down. Thus, jurisdictions might link a call down of Emergency Operations Center staff with a tabletop drill designed to test EOC staff ability to make decisions based on available epidemiological and other information.

HOT WASH AND AFTER ACTION REPORT

Drills should end with a hot wash – a time period reserved for all participants, including controllers and evaluators, to discuss the exercise and how the group responded to it. The hot wash allows participants to receive feedback from evaluators and the controller, as well as from one another. The hot wash reinforces learning by:

- Giving participants time to reflect on their performance during the exercise
- Allowing participants to hear other people’s perspectives and views about the quality of responses developed during the exercise.

Participants can also learn from the drill evaluator, who can:

- Outline the strengths and weaknesses of the responses
- Provide exercise organizers with suggestions for improving the exercise design.

Jurisdictions should develop a brief, written, after action report. After action reports summarize the key activities tested during the drill and key findings and improvement strategies for addressing performance gaps. There is no required format for such a report. However, most include brief summaries of the exercise as well as bulleted lists of strengths and areas for improvement. Readers might consult the Department of Homeland Security’s Homeland Security Exercise Evaluation Program web site for guidance on writing effective after action reports (<http://www.hseep.dhs.gov>). It is usually helpful to include specific corrective actions, timelines, and individuals responsible for carrying out corrective actions, in the written report.

It is useful to circulate the after action report to participants and then meet to discuss it. Such a meeting is useful because, unlike the hot wash session, participants will have had more time to consider their performance during the exercise.

DATA CAPTURE SHEET: CALL DOWN DRILL

Required Metrics

Calling Completion Time minutes

Acknowledgement Process Completion Time minutes

Contact Rate

For first 50% on list

Input these values

of staff called
of staff successfully contacted
Total # of staff on list

Your contact rate* is **#DIV/0!** **percent**

*Contact rate uses total # staff in denominator

For first 90% on list

Input these values

of staff called
of staff successfully contacted
Total # of staff on list

Your contact rate* is **#DIV/0!** **percent**

*Contact rate uses total # staff in denominator

For all staff on list

Input these values

of staff called
of staff successfully contacted
Total # of staff on list

Your contact rate* is **#DIV/0!** **percent**

*Contact rate uses total # of staff in denominator

Hypothetical Assembly Rate

Input these values

of staff able to report by target time
Total # of staff on list

Your assembly rate **#DIV/0!** **percent**

Other Required Information

CDC-DSNS collects this information in order to interpret the metric data and to help develop the knowledge base about exemplary practices. The information below is not used for assessment purposes

Date of drill (mm/dd/yyyy)

Time of drill

 AM / PM

Did you communicate scenario information to staff called? (Yes/No)

If so, briefly describe the scenario:

What target time did you use in calculating assembly rate? (e.g., 2 hours)

Did you use an automated calling system? (Yes/No)

Check all means by which recipients can receive calls

Phone

☐

Email

☐

Pager

☐

Blackberry

☐

Other (describe):

staff used to make calls

Describe where call down lists are kept, and in what format

Other comments:

METRICS FOR SITE ACTIVATION DRILL

Operations Manual

OVERVIEW

Site activation is the ability to contact and ensure that facilities are available for emergency response functions. It is a cross-cutting capability, one that is applicable to multiple functions in a mass prophylaxis scenario, especially Points of Dispensing (PODs) and Receiving, Staging, and Storing (RSS) warehouses.

During an emergency, facilities such as RSS warehouses and PODs may be set up in sites currently in use by other entities for other functions. These sites will have to be activated and prepared for emergency use. It may be helpful to think of site activation as “clearing the space” for subsequent site setup. For example, RSS functions might be conducted in warehouses designed for non-medical materials, and dispensing functions are often situated in schools.

Successful site activation requires (a) site lists that are up-to-date, readily accessible, and usable; (b) workable methods for contacting sites in a timely fashion; and (c) confirmation of sites’ ability to prepare for their assigned functions in a timely manner. This manual describes procedures for validating site lists by contacting those in charge of the listed facilities, testing communications through a facility call down, and asking for estimates of the amount of time required to prepare sites for their assigned functions.

In an actual emergency, the number and selection of sites to be activated is a decision that will have to be made by health department and emergency management officials. Testing that decision-making is vital, but beyond the scope of this drill. However, for the purposes of this assessment you should assume that *all* the RSS or *all* the PODs will be called for activation. This is done to facilitate comparisons of the drill performance over time. In addition it serves to audit the accuracy of contact information for the complete list of facilities.

A small set of *required* metrics should be reported to CDC as part of its assessment tool for Strategic National Stockpile (SNS) readiness. These include:

- *Completion time.* The amount of time needed to contact all sites on the list.
- *Contact percentage.* The number of sites successfully contacted, as a percentage of all sites on the facilities list.
- *Availability percentage.* The number of sites that report being able to be prepared within a designated time, as a percentage of all sites on the list

An ideal site activation drill would be conducted without notice and involve actual site preparation for setup. However, requiring sites, many of which normally function as

schools, warehouses, or other facilities, to interrupt their normal non-emergency functions without warning might create an undue burden. Although such tests should be conducted periodically, the procedures in this manual are based on a no-notice, hypothetical site activation drill to permit more frequent, lower-cost assessments.

This drill is designed to be a stand-alone assessment but, as with other operational drills, it could also be linked with other assessment activities or embedded in larger exercises/drills. Ideas for other uses of the drill format are provided at the end of the manual.

PLANNING CHECKLIST

Prior to running the drill, jurisdictions should ensure that the following planning steps are completed:

Recruit drill participants. Required participants include (a) the drill “players,” (b) a drill controller, and (c) an evaluator. The evaluator and controller may be one and the same person.

- *Players.* The drill players are the calling staff that would be employed in the case of a real emergency site activation. Jurisdictions should run this drill with the same number of callers, and preferably with the same actual callers, that they would expect to use in a real emergency according to their local plan. Jurisdictions should be careful not to use callers who would have other responsibilities in an actual emergency.
- *Controller.* The drill controller will indicate when the drill starts and stops. This drill is designed to evaluate performance, so the controller should not take part as a drill player, and should *not* intervene while the drill is in progress.
- *Evaluator.* The evaluator is responsible for timing the drill. The evaluator should also take notes on the drill in order to identify opportunities for improvement. An evaluator can be a peer from another jurisdiction, a CDC representative, or another individual. In some situations, more than one evaluator may be appropriate.

Ensure that site list is up to date. As part of the jurisdiction’s development of an emergency plan, a list of sites (and contact people) required for each emergency function must be developed, including EOC, RSS, and PODs. The drill planner should ensure the following:

- *Up-to-date lists.* Keeping the list of available personnel up to date is a continuous process and may be one of the biggest challenges. (One of the benefits of this drill is to audit the accuracy of the list.)
- *Non-overlapping lists.* The facility list for each function should be unique: functions that might be activated simultaneously (including for other emergency

non-SNS purposes) cannot share the same space (although in some cases they may be in different parts of the same facility).

- *Enough sites for prophylaxis of entire community.* The CRI goal is that the jurisdiction would be able to prophylax 100 percent of the community within 48 hours. In order to meet this goal, ideally site activation lists should have enough facilities that would be needed to meet the CRI target.
- *Backup facilities.* The site list should also contain extra facilities to compensate for those that may be unreachable or unavailable for timely use.

Ensure availability of calling and response-monitoring equipment. Drill players should have access to whatever calling equipment and call response monitoring technology would be used during a real emergency to keep track of the call responses received (e.g., computers, software, paper materials).

Protocol for dealing with non-responses. Jurisdictions, as part of their planning process, should develop their own protocol for how callers or automated systems should deal with answering machines, voicemail, busy signals, numbers out of service, and wrong numbers when placing calls.

Decide which facility lists to test. Before running the drill, jurisdictions will have to decide which facility lists will be tested: (a) EOC, (b) RSS, or (c) POD. Jurisdictions may combine tests of different functions in a single drill or run separate drills by function. Please note that testing EOC facilities is **optional**.

Alternate drills during work-hours and non-work-hours. Jurisdictions should include drills both during and after normal work-hours. The time and day of the drill should be noted on the data capture sheet (see below).

DATA COLLECTION PROCEDURES

Test the entire call down list. For the purposes of this required drill, jurisdictions should assume that all potentially relevant sites should be called. That is, if the POD site list is being tested, the jurisdiction should seek to contact *all* potential POD sites. This will provide comparability across drills and ensure that each jurisdiction is prepared to respond to high-intensity scenarios.

Steps for manual calling. We assume that most jurisdictions will use manual calling for site activation. However, some might use some form of automated calling system. In any case, jurisdictions should also be prepared to conduct call down procedures manually, with callers placing calls to individual sites, in case automated means fail. We begin with steps for manual calling.

- Players should be instructed to work as they would during a real emergency (e.g., with the same sense of urgency) and to begin making calls immediately.
- Timing **begins** when the first call is made.
- Callers place unannounced calls to all of the sites listed as eligible for use as the facility or function in question. Callers should clearly state that the call is part of a drill.
- For each site contacted, the caller should follow these steps:
 - Step 1: The caller should record whether the site could be reached.
 - Step 2: If the site cannot be reached, the caller should follow their local protocol as to whether a message should be left, or any further effort made to contact the person(s). If the site is declared unreachable, that should be recorded.
 - Step 3: If the site called is reached, the caller will then ask the questions in Box 1 below. The caller should stick close to the script provided to ensure that the drill metrics are comparable over time and across jurisdictions.³¹

Box 1: Site Activation Drill Script

- “This is a drill being conducted by (jurisdiction). Your site, (insert site name), is on (jurisdiction’s) list of emergency facilities. If this were a real emergency, your site would be asked to prepare for activation as a [insert emergency function(s) – POD, RSS, and/or EOC]. These preparations would involve interrupting your facility’s normal functions, clearing out any spaces intended for emergency functions, evacuating unnecessary staff or community members, and preparing equipment such as telephones, computer systems, electricity and water systems for use by emergency teams.” [*Callers should not reveal a target time for site readiness, so as not to bias the answers.*]
- Q2: “If this were a real emergency, how long would it take your site to prepare for activation, meaning being ready for [insert emergency function]?” [Record answer.]
- “Again, this is only a drill. There is no need for you to take any action as a result of this call. Thank you.”

³¹ Jurisdictions may modify these questions, but changes should not substantially change the question content, and all callers should use the same modified script. Most importantly, jurisdictions should not add to or subtract any questions from the list.

- The timed drill **ends** when an attempt has been made to contact all sites on the list, and the last call made has been completed.

Steps for automated calling. Jurisdictions using automated call down systems should instruct all called sites to acknowledge receipt of the message, regardless of whether the staff can report for duty. Additional steps depend on whether the system allows respondents to acknowledge the call and report a site activation time.

- If the automated system allows for it, jurisdictions should program their systems to require respondents to acknowledge receipt of the call and to indicate the amount of time it would take them to ready the site in the event of a real emergency. As with the manual drill, timing the automated calling process begins with the initiation of the first call and ends after calls have been placed to all sites on the call down list and the last call made has been completed.
- If a jurisdiction's automated system does not include a function for receiving responses from the called sites, an alternate means of receiving acknowledgements will be required. One possibility would be to stand up the public information call center and use it to receive calls back from the called staff. In this case, there would be additional timing for the acknowledgement process, recording the times when each call is received back from each of the called site.

REQUIRED METRICS

Completion time. The amount of time needed to contact all sites on the list. This metric should be reported by function (i.e., POD, RSS, or EOC).³²

Contact percentage. The percentage of sites the jurisdiction sought to contact that were actually reached by telephone or other communication medium. This is the number of sites successfully contacted (regardless of whether they report being able to activate), divided by the total number of attempted sites called.³³

³² The completion time is important because in an actual emergency, the site activation process would have to be completed quickly enough to enable staff to assemble, conduct just-in-time training, and perform the site setup, before the SNS material arrives. The jurisdiction should allow enough time for site activation in their local plan. However, with the tight timelines involved, we recommend that the target completion time be 2 hours.

³³ The contact rate is important because it lets the jurisdiction get an idea of how many sites may be unreachable at a given moment on a given day. This will vary depending not only the time of day, but the time of year (e.g., it may be easier to contact many sites on during normal business hours than late at night on a holiday weekend). Thus, despite the temptation to run the drill during the workday when calling staff are present, jurisdictions may want to run their drills at different times of day, week, and year, in order to get a wider sampling. Contact numbers for site contacts may also vary depending on the time of day or time of year that they are called.

$$\text{Contact rate} = \frac{\text{\# Sites successfully contacted}}{\text{\# Sites on call down list}}$$

The contact rate should be reported by function – that is, there should be separate rates for RSS, POD, EOC, etc.

Availability percentage. The percentage of sites that report being able to be prepared for their designated function by the designated time (selected by the jurisdiction according to their planning; known to the jurisdiction but not disclosed to the sites called). This is the number of called sites that report (see Q2 in Box 1) being able to activate within the target time, divided by the total number of attempted site calls, reportable per function (i.e., per POD or RSS).³⁴

$$\text{Availability rate} = \frac{\text{\# Sites able to be ready for SNS function within target time}}{\text{Total \# sites on list}}$$

The availability rate should be reported by function – that is, there should be separate rates for RSS, POD, EOC, etc.

Recall that callers should not reveal the target activation time to those being called, to avoid biasing the answers. Thus, jurisdictions will need to determine after the drill what percentage of respondents' answers indicate the sites would be ready within the target assembly time. This metric should be reported by function (i.e., per POD or RSS).

OPTIONAL METRICS

Actual activation time. One way to validate activation times is to require actual – as opposed to hypothetical – activation. This could include one or many facilities. The activation time would be the time elapsed between when the call was received to activate the site and when the facility was cleared out and ready to be turned over to the health department to begin their setup operations.

More detailed questions. The assessment described above relies on facility staff to self-report how long it would take to get the site ready for emergency SNS use. However, respondents might underestimate barriers and overestimate their ability to successfully overcome those barriers.

³⁴ The availability rate is important because it lets the jurisdiction get an idea of how many sites will be unavailable on a given day. We expect this estimate to be somewhat optimistic; in an actual emergency there may be infrastructure damage.

- *Ask probing questions.* Callers might ask a more probing set of questions designed to elicit more reliable answers about assembly than is possible using automated call down systems, and given the constraints of a timed drill. For instance, instead of simply asking respondents to report their expected assembly time, callers could start with a short set of questions designed to ensure that respondents consider the full range of possible constraints in their answers. For instance, the call down script might include a question such as:

“Please list any factors that might reasonably prevent or delay you from preparing the site for use in an emergency, including lack of space, scheduling constraints, and so on.”

- *Piece together response times from narrower questions.* Alternatively, the script could seek to “piece together” the various steps most sites have to accomplish in order to be ready for emergency use:

How long would it take until you could find/contact any additional staff required to clear the facility?

How much time would be required to gain access to power, keys, or other critical equipment?

Adding these questions, however, could create problems for automated calling systems and will skew times reported for the “completion time” metric discussed above. Thus, these are presented as optional – not required – assessment and training activities and should not be included as part of drills to generate metrics reported as part of the SNS assessment system.

LINKING THE DRILL WITH OTHER ASSESSMENTS

As noted above, the site activation drill is designed to be run as a standalone assessment. However, it could also be linked with other assessments or embedded in larger drills and exercises. Health department officials are encouraged to be creative in developing ideas for linkages with other drills, exercises, and assessments. Generally, site activation drills can be combined with any other assessment or exercise that requires use of facilities. Here are a few specific ideas:

Link site activation drill with RSS exercise. Jurisdictions might run a site activation drill before running a Pick List Generation drill or other RSS exercise.

Link site activation drill with POD drills. Jurisdictions might test the ability to contact all POD sites before standing up and exercising any one POD (perhaps using the dispensing drill metrics described in another manual).

HOT WASH AND AFTER ACTION REPORT

Although it is not required for performance metrics reported to CDC, all exercises should end with a hot wash – a time period reserved for participants, including controllers and evaluators, to discuss the exercise and how the group responded to it. The hot wash gives participants time to reflect on their performance during the exercise, and provide suggestions on how to improve performance in an actual emergency, as well as how to improve the exercise process itself.

Jurisdictions should develop a brief, written, after action report. After action reports summarize the key activities tested during the exercise and key findings and improvement strategies for addressing performance gaps. There is no required format for such a report. However, most include brief summaries of the exercise as well as bulleted lists of strengths and areas for improvement. Readers should consult the Department of Homeland Security's Homeland Security Exercise Evaluation Program web site for guidance on writing effective after action reports (<http://www.hseep.dhs.gov>). It is usually helpful to include specific corrective actions, timelines, and individuals responsible for carrying out corrective actions, in the written report.

DATA CAPTURE SHEET: SITE ACTIVATION DRILL

Required Metrics

Completion Time

minutes

Contact Rate

Input these values

of sites called

of sites successfully contacted

Total # of sites on list

Your contact rate* is #DIV/0! percent

*Contact rate uses total #sites in denominator

Hypothetical Availability Rate

Input these values

of sites ready by target time

Total # of sites on list

Your availability rate #DIV/0! percent

Other Required Information

CDC-DSNS collects this information in order to interpret the metric data and to help develop the knowledge base about exemplary practices. The information below is not used for assessment purposes

Date of drill (mm/dd/yyyy)

Time of drill

AM / PM

Did you communicate scenario information to sites called? (Yes/No)

If so, briefly describe the scenario

What target time did you use in calculating availability rate? (e.g., 2 hours)

Did you use an automated calling system? (Yes/No)

Check all means by which recipients can receive calls

Phone

Email

Pager

Blackberry

Other (describe):

staff used to make calls

Describe where the contact is kept, and in what format

Other comments:

METRICS FOR SETUP DRILL

Operations Manual

OVERVIEW

Setup is the ability to quickly set up materiel and spatial layouts. It is a cross-cutting capability, applicable to a variety of SNS functions, particularly Points of Dispensing (PODs), and Receiving Staging Storing (RSS) warehouses. Proper setup is an important precondition of a rapid and effective response, and may in some situations lie on the critical path to successfully prophylaxing a population within 48 hours.

The following elements are prerequisites for conducting setup: (a) a facility cleared of any extraneous materials and/or personnel that might pertain to the facility's normal operations, but which are unnecessary for the facility's emergency function;³⁵ (b) a setup plan, including the layout, materials, setup crew and facility services required; (c) the materials needed to set the facility up (e.g., signage, furniture, computing and communications equipment); and (d) a setup crew assembled per the setup plan.

The goal of a setup drill is to test the amount of time it takes to completely set up a facility with the materiel, layout, and supplies necessary to perform a given SNS function. A small set of *required* metrics should be reported to CDC as part of its assessment tool for Strategic National Stockpile (SNS) readiness. These include:

- *Setup time.* Time required to complete the setup according to the specification of a setup checklist.
- *Percentage of checklist items completed.* Percentage of items completed on a checklist that defines the activities required for site setup.

This drill is designed to be a standalone assessment but could also be linked with other assessment activities or embedded in larger exercises/drills. Ideas for other uses of the drill format are provided at the end of the manual.

PLANNING CHECKLIST

Jurisdictions should appoint a drill planner to take responsibility for ensuring that the following planning steps are completed:

³⁵ The Facility Activation Drill surveys the availability of potential sites. In the drill, the host facility is asked the amount of time it would take to clear the facility and otherwise make it available to the health department. The time to clear the facility is not considered part of the time necessary for setup.

Recruit participants. Required participants includes (a) the drill “players,” (b) a drill controller, and (c) an evaluator.

- *Players.* Jurisdictions should run this drill with a setup crew assembled according to the SNS plan for each facility, preferably with the same individual crew members that they would expect to use in an actual emergency. (Jurisdictions should also be careful not to use crew members who would have other responsibilities in an actual emergency that would prevent them from helping with setup.)
- *Controller.* The drill controller will indicate when the drill starts and stops. This drill is designed to evaluate performance, so the controller should not take part as a drill player, and should *not* intervene while the drill is in progress.
- *Evaluator(s).* A lead evaluator is responsible for timing the drill. In addition, the lead evaluator should take notes on the drill in order to identify opportunities for improvement. The lead evaluator and controller may be one and the same person. An evaluator can be a peer from another jurisdiction, a CDC representative, or another individual. In some situations, more than one evaluator may be appropriate.

Develop a setup checklist. Each facility should have a setup checklist that describes layout, space requirements, office equipment, computers and communications, utilities and environmental controls, supplies, and security.

Gain access to the actual facility to be used. The drill is most useful when the actual people who would be performing these functions in an emergency are practicing on the actual facilities that would be used. Jurisdictions are encouraged to include in their memoranda of understanding (MOU) with facility owners, provisions for periodic access to the facility for drill purposes. If this is not possible, performing the drill in a substitute location is better than no drill at all, but is far from ideal.

Ensure availability of necessary equipment. Drill players should have access to whatever equipment, materials, and facility services would be used during a real emergency. It is important to remember that this includes any keys, passwords, or other mechanisms that control the facility’s water, power (including backup generators), communications technology, and access to restrooms or other amenities.

DATA COLLECTION PROCEDURES

Starting the drill. The controller should start the drill when all necessary personnel, equipment, and facilities are ready. Timing should begin upon commencement of the drill.

Players should be instructed to work as they would during a real emergency (e.g., with the same sense of urgency) and to begin setting up the facility immediately upon the drill

controller's announcement that the drill has begun; the drill controller should begin timing at this announcement.

Ending the drill. The timed drill ends when the setup crew notifies the drill controller that the facility setup is complete.

REQUIRED METRICS

Setup time. The amount of time elapsed between arrival of all personnel, materiel, and equipment and completion of all setup activities.

Percentage of checklist items completed. Upon completion of the drill, the drill evaluator(s) should complete a Setup Checklist appropriate to the facility and emergency function assessed by the drill. This checklist should address all of the elements required in the facility-specific setup plan. Jurisdictions should report the percentage of items completed.

A generic checklist template is provided at the end of this manual (see Appendix). Any additional elements included in a jurisdiction's plan should be noted on the checklist and assessed as well.

OPTIONAL METRICS

The number of required setup metrics has been minimized to encourage regular and frequent assessments. They represent necessary, but not sufficient, conditions for ensuring readiness. Some additional metrics jurisdictions may want to consider are:

Setup time by component. Jurisdictions are encouraged to consider timing the components of setup as a way of identifying areas in particular need of improvement. These components may include: clearing space, moving furniture, setting up signage and traffic flow (for people, material, or vehicles), and setting up computers and communications.

Additional facility considerations. Having set up the facility, jurisdictions should also go beyond their existing checklist, walking through the operations to look for any problems that had not been previously anticipated, and to revise their site plans accordingly. This might include space and layout requirements, patient or material flow, traffic and parking plans, and other general public needs.

LINKING THE DRILL WITH OTHER ASSESSMENTS

As noted above, the setup drill is designed to be run as a standalone assessment. It could also be linked with other assessments or embedded in larger drills and exercises. Health department officials are encouraged to be creative in developing ideas for linkages with other drills, exercises, and assessments. Here are a few ideas:

Link with facility-specific call-down and assembly drills. Given that the setup team will already be arriving at the facility for the purposes of this drill, the time it takes them to arrive at the facility could be assessed by running call down and assembly specific to the facility. The drill could be (a) announced, (b) announced but unscheduled, or (c) completely unannounced.

Link with a full scale exercise. Most full scale exercises are conducted with the intent of replicating and assessing as many activities related to an actual emergency situation as possible. Therefore, as part of a full scale exercise, it may be helpful to run a setup drill for all facilities that are performing RSS, Command, or POD functions, in conjunction with their other exercise activities.

HOT WASH AND AFTER ACTION REPORTS

Although it is not required for the performance metrics reported to CDC, all exercises should end with a hot wash – a time period reserved for participants, including controllers and evaluators, to discuss the exercise and how the group responded to it. The hot wash gives participants time to reflect on their performance during the exercise and provides suggestions on how to improve performance in an actual emergency, as well as how to improve the exercise process itself.

Jurisdictions should develop a brief, written, after action report. After action reports summarize the key activities tested during the exercise and key findings and improvement strategies for addressing performance gaps. There is no required format for such a report. However, most include brief summaries of the exercise as well as bulleted lists of strengths and areas for improvement. Readers should consult the Department of Homeland Security's Homeland Security Exercise Evaluation Program web site for guidance on writing effective after-action reports (<http://www.hseep.dhs.gov>). It is usually helpful to include specific corrective actions, timelines, and individuals responsible for carrying out corrective actions, in the written report.

It is useful to circulate the after action report to participants and then meet to discuss it. Such a meeting is useful because, unlike the hot wash session, participants will have had more time to consider their performance during the exercise.

GUIDANCE ON DEVELOPING SETUP CHECKLISTS

Jurisdictions should include the setup checklists used to evaluate the drill in their submissions to CDC. A generic template for setup checklists is provided below. Jurisdictions should also refer to the Version 10 SNS guidance for further considerations in developing checklists.

GENERIC CHECKLIST TEMPLATE

- ☐ The layout conforms to the facility layout plan
- ☐ The setup is accessible to: Fire, EMS, Police, Handicapped, and the general public.
- ☐ The setup meets all administrative materials requirements in the facility setup plan (e.g., pens, pencils, forms, notes, clips, scissors)
- ☐ The setup meets all health-related materials requirements in the facility setup plan (e.g., medical supplies, sanitation supplies, and personal protective equipment such as masks and gloves)
- ☐ The setup meets all computer and communications requirements in the facility setup plan (e.g., copy machines, phone lines, phones, computers, internet access or connectivity with department networks, including firewall and security considerations)
- ☐ The setup meets all facility management materials requirements in the facility setup plan (e.g., extension cords, trash bags, batteries, flashlights)
- ☐ The setup meets all crowd management requirements in the facility setup plan (e.g., tape, bullhorns, whistles, vests, cones)
- ☐ Miscellaneous: The setup meets any additional requirements in the facility setup plan. Please specify:

DATA CAPTURE SHEET: SETUP DRILL

Required Metrics

Setup Time

minutes

Checklist Completion Rate

Input these values

of checklist items completed

Total # of checklist items

Your checklist completion rate is #DIV/0! percent

Other Required Information

Date of drill (mm/dd/yyyy)

Time of drill

AM / PM

Facility type tested in drill (check one)

POD

RSS

EOC

Other (describe)

Routine use of facility (e.g., school)

Is relevant Floorplan/layout attached? (Yes/No)

Number of staff used to conduct setup

Use of volunteers

volunteer POD staff (not patients):

health department employees:

other paid staff:

Other comments:

PICK LIST GENERATION DRILL

Operations Manual

OVERVIEW

During an SNS response, medical countermeasures will be shipped to a Receipt, Staging, and Storage (RSS) site designated by your jurisdiction. After these have been unloaded and staged in the RSS warehouse, the inventory must be apportioned, picked, and shipped to the PODs. A pick list tells RSS warehouse workers the items that are to be picked and shipped to particular PODs, including the item type, quantity, lot number, and location within the warehouse. One or more pick lists combine to create an order that will be shipped to the POD sites.

This manual will help users design and conduct a pick list generation drill. The primary objective of this drill is to measure the amount of time required to generate a pick list.

This manual contains instructions for administering a pick list generation drill as a modular, standalone assessment. However, the drill module could also be linked with other assessment activities (such as equipment setup) or embedded in larger exercises/drills. Ideas for other uses of the drill format are provided at the end of the manual.

PLANNING CHECKLIST

Before conducting the drill, jurisdictions should ensure that the following planning steps are completed:

Recruit drill participants. Required participants include (a) the drill “players,” (b) a drill controller, and (c) an evaluator. The controller may be able to serve as an evaluator.

- *Drill Players.* Drill players include all staff that would be involved in generating the pick list during an actual response that would require requesting the SNS.
- *Drill Controller.* The drill controller will indicate when the drill starts and stops. This drill is designed to evaluate performance, so the controller should *not* intervene while the drill is in progress. The controller should not participate in the drill as one of the players.
- *Evaluator.* The evaluator is responsible for timing the drill. The evaluator should also take careful notes on the drill, based on the drill evaluation checklist (presenter later in this manual), in order to identify areas for improvement. The evaluator and drill controller can be the same person. An evaluator can be a peer from another jurisdiction, a CDC representative, or another individual.

Acquire equipment and arrange for a room. The more realistic the drill circumstances are—including facilities and equipment—the more accurate the resulting drill assessment will be.

- *Room.* Ideally, the drill should be run in the same room that would be used during a real emergency. In most instances, this is likely to be a room at the RSS facility. If this is not practical, the drill can also be run in a room that resembles one at the RSS. Special attention should be paid to the availability of office furniture, electrical outlets, and telecommunications support (e.g., phone and computer jacks).
- *Computer equipment and software.* Drill players should use whatever computer equipment would be used during a real emergency. Likewise, the inventory tracking and pick list generation software used in the drill should be the software that would be used according to the jurisdiction's plans: a locally-designed program, a commercial package, or the CDC/DSNS Inventory Management Program (RITS). Jurisdictions might wish to practice or test setting up the computer equipment. However, this should not be part of the timed drill.

Acquire (but not yet import) inventory file. The inventory file represents the assets that the RSS would have in place and available for shipment to POD sites and other locations. For this drill, we will assume that the assets are in the form of the 12-hour Push Package. CDC provides this list in an electronic format. At this point, prior to the actual start of the drill, the jurisdiction should have a copy of that data file. Later, during the drill itself, jurisdictions will be required to import this data, i.e., load the data into their warehouse and inventory management system³⁶. A later section of the manual provides ideas for practicing loading inventory information for managed inventory shipments, which typically will require manual entry.

Develop apportionment instructions. To generate a pick list during the drill, players will need to be told how to apportion the medical countermeasures among POD sites. The “apportionment instructions” tell how to distribute the materiel, including how many boxes should be delivered to each POD. This decision can range from a single simple rule (e.g., divide the inventory equally among PODs) to a longer list of detailed instructions.

DATA COLLECTION PROCEDURES

Ensure that room and equipment are set up. The pick list generation drill begins with all of the necessary equipment (e.g., computers, power cords, etc.) set up in the room where the drill will be run.

Starting the drill. The controller should start the drill when players have received:

- Apportionment instructions
- Inventory data file for the 12-hour Push Pack

Executing the drill. Drill players must complete two steps.

³⁶ Jurisdictions using RITS will find that the inventory data has been pre-loaded by CDC.

1. Import the inventory information into their inventory system. (For jurisdictions using the RITS system, the data file for the 12-hour Push Pack is already loaded into the system, so this step will take no additional time.)
2. Generate a pick list.

Ending the drill. The controller should stop timing when players have produced a printed pick list for *each* of the PODs included in the Apportionment Instructions.

REQUIRED METRIC

Jurisdictions must report one metric from this drill.

Completion time. Jurisdictions must report the length of time it takes players to import the data file (if necessary) and generate the pick list during the drill as described above.

OPTIONAL METRICS

Data importation time. Jurisdictions not using the RITS system (into which the 12-hour Push Pack inventory data are pre-loaded) might also wish to time separately the amount of time required to load the inventory data into their own inventory management software system.

Time to generate pick list for managed inventory. Jurisdictions may also wish to test their ability to handle managed inventory shipments. Drill players could practice inputting sample bills of material corresponding to bulk supplies that might be received from stockpile or vendor managed inventory. This would likely require manual input of data into the computer system.

LINKING THE DRILL WITH OTHER ASSESSMENTS

As noted above, the pick list generation drill is designed to be run as a standalone assessment. However, it could also be linked with other assessments or embedded in larger drills and exercises. Health department officials are encouraged to be creative in developing ideas for linkages with other drills, exercises, and assessments. Here are a few ideas:

Call down. Link the drill with a call down drill to test call down lists for RSSs and the ability to muster RSS staff in a timely manner.

Site activation. Link the drill with a site activation drill to test the ability to ensure that the space required for one or more RSS can be made available in a timely manner

Setup. Link the drill with a setup drill to test the ability to set up the materials and equipment required for a RSS, including the computer systems needed for RSS inventory management, in a timely manner.

Embed within larger RSS drill. The pick list generation drill described in this manual could be embedded within a larger RSS drill that includes actual picking, palletizing, loading, and other movement of materiel.

Practicing apportionment decision-making. Jurisdictions should practice the apportionment of available inventory among the PODs. Given surveillance/epidemiological data, a list of PODs (and perhaps the population expected at each POD), a list of inventory available at each POD, and a list of inventory available at the RSS, the drill would consist of deciding how the inventory should be divided among the PODs. The assumption underlying this activity would be that PODs have either not placed orders for specific quantities (as might be the case at the beginning of the dispensing operation).

HOT WASH AND AFTER ACTION REPORT

Drills should end with a hot wash – a time period reserved for participants to discuss the exercise and how the group responded to it. The hot wash allows participants to receive feedback from evaluators and the controller, as well as from one another. The hot wash reinforces learning by:

- Giving participants time to reflect on their performance during the exercise
- Allowing participants to hear other people’s perspectives and views about the quality of responses developed during the exercise.

Participants can also learn from the drill evaluator, who can:

- Outline the strengths and weaknesses of the responses
- Provide exercise organizers with suggestions for improving the exercise design.

Table G.1 below provides a checklist that can be used to guide a hot wash for a pick list generation drill. However, this checklist is only suggestive and should be adapted to meet each jurisdiction’s particular needs.

Jurisdictions should develop a brief written after action report. After action reports summarize the key activities tested during the drill and key findings and improvement strategies for addressing performance gaps. There is no required format for such a report. However, most include brief summaries of the exercise as well as bulleted lists of strengths and areas for improvement. Readers might consult the Department of Homeland Security’s Homeland Security Exercise Evaluation Program web site for guidance on writing effective after action reports (<http://www.hseep.dhs.gov>). It is usually helpful to include specific corrective actions, timelines, and individuals responsible for carrying out corrective actions, in the written report.

Table G.1: Sample Checklist for Pick List Generation Drill Hot Wash

-
- Were participating staff adequately trained? What training interventions might improve performance?
 - Would including staff with other skills improve performance? If so, what skills are needed?
 - Was the equipment (computers, furniture, etc.) adequate? Was anything missing?
 - Did participants understand their roles adequately? Was JIT training used?
 - Was communication among participants adequate?
 - If problems were encountered during the drill, how well did players adapt and problem solve?
 - How might the response have to be adapted to meet the needs of other scenarios and situations?
-

It might be useful to circulate the after action report to participants and then meet to discuss it. Such a meeting is useful because, unlike the hot wash session, participants will have had more time to consider their performance during the exercise.

DATA CAPTURE SHEET: PICK LIST GENERATION DRILL

Required Metric

Completion Time

minutes

Other Required Information

CDC-DSNS collects this information in order to interpret the metric data and to help develop the knowledge base about exemplary practices. The information below is not used for assessment purposes

Date of drill (mm/dd/yyyy)

Time of drill

Was RITS software used? (Yes/No)

List other software used

AM / PM

Briefly describe apportionment decision used

Number of staff used to conduct pick list generation

Use of volunteers

volunteer POD staff (not patients):

health department employees:

other paid staff:

Other comments:

TIMED METRICS FOR A POD EXERCISE

Operations Manual

OVERVIEW

This manual provides guidance on collecting objective performance data from POD drills/exercises. The metrics are flexible enough to work with virtually any dispensing model where persons receive countermeasures in some sort of central location, including streamlined POD models and dispensing operations in workplaces and other nontraditional venues. They are *not*, however, designed for models involving home delivery. Also, the metrics focus on the dispensing process itself and do not include rate-limiting steps before entry into the PD (e.g., parking, traffic flow).

A small set of *required* metrics should be reported to CDC as part of its assessment tool for Strategic National Stockpile readiness. These include:

- *Throughput*. The number of persons who receive prophylaxis per hour
- *Mean flow time*. The average time it takes a person to go through the POD from start to finish
- *Dispensing accuracy*. The percent of persons given the correct countermeasure (i.e., medication dispensed is the medication “prescribed”)³⁷

The manual also describes a set of *optional* metrics that can help jurisdictions diagnose and develop corrective actions for specific performance gaps.

Data collection for the metrics is based on a standard set of 100 able-bodied (non-special needs) persons going through the POD. There is also an upward adjustment in the number persons for large PODs. The standard number of persons is designed to facilitate consistency over time and across jurisdictions, ensuring that the drills stress the system without the need to recruit an unrealistically large number of persons.

Data on the 100+ person sample can be gathered during a larger exercise, so long as the standard person sample goes through the POD first. Thus, the data collection procedures described in this manual can be conducted as a standalone drill or embedded within a larger exercise.

The manual focuses only on gathering information for the metrics and is not intended as a complete guide to planning and executing a POD drill/exercise. (See the appendix for a list of recommended resources for planning exercises and drills.)

³⁷ As noted below, this metric may be omitted when testing streamlined POD models – especially those that seek to minimize paperwork.

PLANNING CHECKLIST

Jurisdictions should appoint a drill planner to take responsibility for ensuring that the following planning steps are completed:

Recruit 100+ players. Data collection for the metrics is based on a standard sample of 100 players. These persons should be able to walk unassisted and should represent the types of persons that typically would be routed through as “express line.” Note that for POD designs with separate lines for special needs persons, this implies that these lines will be inactive during the timed portion of the drill. Procedures for testing other persons are described in the section on “optional metrics.”

Please note the following additional considerations:

The number of test persons may need to be adjusted upward according to the size of the POD. The size of the set of persons must be at least 100, in order to provide enough persons to test the POD’s throughput. For large PODs, however, 100 persons may not be enough.

Specifically, **to ensure a sufficient number of persons, each POD drill should involve approximately a half-hour’s work at the required throughput rate.** This can be determined as follows:

- Step 1: Divide the total jurisdiction population by the number of PODs provided by the SNS plan for the largest possible prophylaxis operation. This will yield the population per POD. (If the SNS plan calls for some PODs to handle more persons than others then these adjustments should be factored in.)
- Step 2: Divide the result from Step 1 by 36 hours, which will provide the required throughput per hour for each POD.
- Step 3: Divide the result from Step 2 by 2 to calculate the number of persons the POD would be expected to service in a half hour

$$\text{Sample size} = \frac{\text{Jurisdiction Population}}{\text{\# of PODs in SNS plan}} \times \frac{1}{2 \times 36 \text{ hrs}} = \underline{\hspace{2cm}} \text{ persons/half hour}$$

The jurisdiction should use the *greater* of (a) the results of Step 3 or (b) 100 persons.

“Shrink” the POD as a last resort if sample sizes are infeasible. If, despite the best efforts of a jurisdiction, a sufficient test pool of persons still cannot be recruited, then as a last resort, jurisdictions may “shrink” the size of the POD. At each process step,

reduce the staff by half, shutting down half the number of parallel stations.³⁸ This will provide (roughly) a half-sized POD, while attempting to preserve the interaction between stations and steps. The throughput that is achieved should be half of the throughput that could be achieved if the POD were fully staffed. **If used, this strategy should be clearly documented in materials submitted to CDC-DSNS.**

Recruit an exercise controller to coordinate timing of the drill. An exercise controller will indicate when the timed portion of the exercise starts and stops. This exercise is designed to evaluate performance, so the controller should *not* intervene or participate in the exercise as a POD staff member.

Recruit evaluators to do the actual timing. Exercise timers are responsible for timing the progression of the standard persons through the POD during the exercise, as described below. We recommend using at least 1 timer for each time post (described below), whose exclusive responsibility is time keeping. An evaluator can be a peer from another jurisdiction, a CDC representative, or another individual.

Please note the following:

- ***Time stamping methods may vary depending on available technology.*** Each timer should be equipped with a pen and a timing device. Some previously tested time evaluation methods have been time stamping,³⁹ manual time-dating⁴⁰ from a central time source, and microchip (RFID) tracking of individuals as they pass through the POD.^{41,42}

³⁸ However, shrinkage should not result in complete elimination of any function. For instance, if the POD design calls for one staff member at a station, that staff member should be retained.

³⁹ Jeffrey Herrmann of the University of Maryland has conducted several time studies using time stamping devices. These devices are reported to be available at office supply stores, used for noting the receipt of documents. For an example of his work, see: Aaby, Kay, J. Herrmann, C. Jordan, M. Treadwell, and K. Wood, [Improving Mass Vaccination Clinic Operations](#), Proceedings of the International Conference on Health Sciences Simulation, New Orleans, Louisiana, January 23-27, 2005.

⁴⁰ An exercise that used manual time dating by watches was the San Francisco County HD Emergency Drill, June 17, 2003. Report written up by James Bowman Associates, Inc. Accessible at: <http://www.dph.sf.ca.us/Reports/June17Drill/Attach/Attach26y27.pdf#search=%22%20time%20studies%20patient%20flow%22>

⁴¹ L.A. County has tested using microchip tracking in one of its POD exercises. More information on RFID tracking can be found at: UNISYS website, Asset Tracking, http://www.unisys.com/services/global_commerce_visibility/asset_tracking.htm. Last accessed 8/27/06.

⁴² The CDC Public Health Information Network (PHIN) has several archived broadcasts on POD operations, including one on measurement of POD performance.

DATA COLLECTION PROCEDURES

Collect data based on sample persons. A set of recruited sample persons will go through the POD. As noted above, these persons will be assumed to not have children or medical complications.

Collect data on 100+ sample persons before others. To ensure consistency of test results, data on the performance metrics should be collected on the basis of the test pool entering an otherwise empty POD. This may be done at the beginning of the POD exercise, before all other persons, or could be done by holding other persons from entering the POD until the POD has emptied out. The sample of standard persons will then run through the POD. Jurisdictions may allow POD staff to practice by running a few persons before the test so that the POD staff can learn their roles and “warm up”. However, jurisdictions are discouraged from running the test sample through the POD for practice because they may leave after their first run or become so acquainted with POD procedures that they run through unrealistically quickly. All persons should be lined up at the entrance to the POD facility and should be prepared to enter quickly.

Observe time elapsed from entry to exit. While PODs usually involve a number of steps, the most basic required metric requires only that evaluators observe *entry* of persons at the POD intake point (i.e., not including queuing outside the POD or at a remote site) and *exit* of persons at the end of the process. This corresponds with (A) and (D) on Figure D.1 and should apply to even the most streamlined of POD designs. Adding timers at other points, such as (B) or (C), as appropriate to the design of the POD, will return more detailed data, and this is discussed as an option in the optional metrics section.

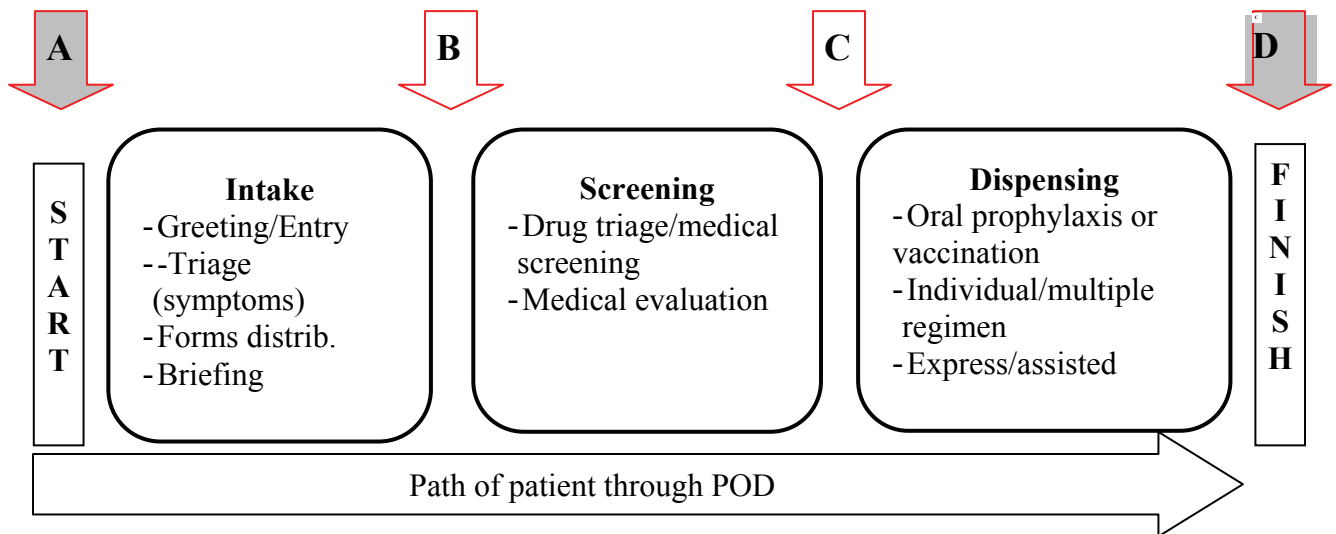


Figure D.1: Functions and Timing Post Locations

If using multiple clocks or time stamps, they must be synchronized. In addition the timer at time post (A) will need a supply of forms to hand out to persons for recording time, which will be carried by persons as they progress through the POD.

Time stamp persons going through the POD. When each person enters the POD at time post (A), the first timer will hand the person a form noting the time (exact to the minute) the person entered. The timer at the exercise finish, time post (D), should mark the time the person exits, and should collect the person’s sheet as they exit the POD facility.

REQUIRED METRICS

Time stamps for entry (A) and exit (D) should be used – after the exercise – to derive two of the three required performance metrics, throughput and mean flow time. The third, dispensing accuracy, will use other methods as described later in this section.

Throughput. Number of persons who exit the POD having received prophylaxis per unit time. The total number of persons (which should be approximately 100) is entered into the numerator. The denominator is the time elapsed from the exercise start until the last person arrives at the Finish point (i.e., the latest Finish time – this can be determined by the time stamps on the last persons’ forms). The resulting quotient is the throughput, reported in persons per hour. These equations for calculating throughput are:

Finish time of last person – Start time of first person = Time elapsed

$$\text{Throughput} = \frac{\text{\# of persons}}{\text{Time elapsed (min.)}} \times 60 \text{ min. /hour} = \underline{\hspace{2cm}} \text{ persons/hour}$$

Calculation of the maximum possible throughput is important because it is one measure of the ability of the local jurisdiction to provide prophylaxis in a timely manner. If PODs share similar designs, then multiplying the maximum throughput by the number of PODs will estimate the rate at which the population can be administered prophylaxis.

Mean flow time.⁴³ The average time it takes a person to go through the POD from start to finish. Calculate this metric by averaging all of the individual start-to-finish (or “flow”) times of all of the persons in the test sample, each of which can be calculated from the time stamps on each of the person’s forms. These equations may help to calculate mean average flow time:

For each person:

$$\text{Person’s finish time} - \text{Person’s Start time} = \text{Person’s flow time}$$

⁴³ Mean is used here instead of median so that the average patient flow time reflects the “outlier” persons that take longer to process.

The average flow time is then given by:

$$\text{Average flow time} = \frac{\text{Person 1's time} + \text{Person 2's time} + \dots + \text{Person n's time}}{\text{\# Persons}} = \underline{\hspace{2cm}} \text{ minutes}$$

The mean flow time is an estimate of the amount of time each person will spend in the POD, including the time spent waiting in line. Long flow times do not necessarily mean that a mass prophylaxis campaign cannot be completed within the designated time period, but they do affect morale and may indicate the need to streamline the process.

Dispensing Accuracy. Jurisdictions should also measure the percent of persons dispensed the accurate pill type. This can be done after the last time post so as not to bias the total flow time metric. After persons pass through the final time post (D), an evaluator would check the pill type received against their prescription sheet from their medical evaluation.

$$\text{Percent of persons with correct pill type} = \frac{\text{\# of persons with correct pill type}}{\text{Total \# of persons}} \times 100 = \underline{\hspace{2cm}} \%$$

This only measures dispensing accuracy – whether the pills received by the person matched the dose and type of medication they were “prescribed” – not diagnostic accuracy, which is treated later in this manual under “optional metrics”. This metric is also not applicable to vaccination PODs. **This metric may be omitted when using streamlined POD models that minimize paperwork.**

Please see the Data Capture Sheet at the end of this manual for help in recording and calculating data that must be reported to CDC-DSNS

OPTIONAL METRICS

In addition to collecting and reporting data on the required metrics, jurisdictions are strongly encouraged to collect other data that can help develop corrective actions to improve future performance. These might include:

Throughput given multiple regimens. Jurisdictions might be interested in knowing not only how many persons can be processed in an hour but also how many regimens can be dispensed. This is relevant for jurisdictions using head-of-household or similar dispensing strategies. This metric is calculated the same as the early throughput metric, with one exception: replace the numerator (total number of persons) with the total number of regimens dispensed across the sample persons.

$$\text{Patient throughput} = \frac{\text{\# of regimens}}{\text{Time elapsed (min.)}} \times 60 \text{ min. /hour} = \text{Regimens/hour}$$

Testing persons with special needs. Express line persons should have few mobility problems or other special conditions. Thus, the metric values produced by the procedures above will not be fully representative of a jurisdiction’s ability to respond to real-world conditions. Jurisdictions should consider running the procedures described above with special-needs persons (e.g., families with children, those requiring assistance, etc.). When doing so, a written sheet containing a profile with characteristics and symptoms will be given to each volunteer allowing him/her to play their assigned role.⁴⁴ POD workers should treat these persons as described in their jurisdiction’s dispensing plan.

Additional time posts. As mentioned previously, jurisdictions that wish to do so may add time posts to obtain function-specific time metrics, such as at (B) and (C) in Figure D.1. For instance, time posts before and after dispensing will show how much of the overall flow time is spent at that function, and if it may be a bottleneck that could operate more efficiently. When deciding where to place time posts, jurisdictions should consider whether to measure queuing time versus actual function time, whether or not to differentiate between different stations for the same function (such as multiple dispensing stations), and how additional time posts may add to total flow time and take away from throughput. The metrics reported would be function flow time, or average time at each function.

Station queue length. The number of people in line at the various stages and stations in the POD and the time they spend in line at those stations. Measuring the number of people in line would require additional evaluators to periodically count the number in queue at each station. The time spent in line could be determined by the time-stamping method described above.

Diagnostic accuracy. As mentioned in the previous section, the required quality measure does not assess the accuracy of a diagnosis, only whether the person received the right pill type for their diagnosis. A more refined metric would be an assessment of how many of the persons were accurately diagnosed given their characteristics and symptoms. This metric would require a medical professional as the evaluator or at least as the designer of the tool used in the measurement.

ADAPTING THE GUIDANCE FOR VACCINATIONS

One method for testing a POD is to stand up a flu clinic in the form of a POD. In POD flu clinics, persons receive a tangible benefit—vaccination—for participating, thus providing a source of willing volunteers.

⁴⁴ The intention is to NOT to measure the ability of the system to detect these individuals, but to measure the ability to “serve” or “prophylax” them. (The assessment is not to depend on volunteers’ acting skills).

The most straightforward way to adapt this manual to vaccination clinics is to dispense a prophylaxis proxy to the 100 person standard sample *before* opening the clinic to vaccination persons.

Jurisdictions could also gather the metric data from the vaccination process. However, they should bear in mind several differences between vaccination and dispensing oral prophylaxis:

- Time and staff training levels required for the physical act of injecting vs. dispensing;
- Numbers of injectors vs. dispensers required (injection process may be slower and injectors must be relieved more often due to fatigue);
- Ratios of special vs. express persons;
- Forms; and
- Dosing processes

In addition, the best order of POD operations leading to vaccination may not be the same as that leading to dispensing. Most importantly, a vaccinating POD must be prepared to accommodate a larger number of persons than a medicating POD because there are no opportunities for a single individual to pick up multiple vaccination regimens.

Considerations for each applicable area of difference should be made and documented in detail. Jurisdictions that test POD capabilities by running a flu vaccination clinic should report their required performance metrics to CDC-DSNS with clear indication that the metrics relate to a vaccinating POD rather than a dispensing POD.

LINKING THE DRILL WITH OTHER ASSESSMENTS

The dispensing drill can be linked with a number of other capability drills.

Call down. Link the drill with a call down drill to test call down lists for PODs and the ability to muster POD staff in a timely manner.

Site activation. Link the drill with a site activation drill to test the ability to ensure that the space required for one or more PODs can be made available in a timely manner.

Setup drill. Link the drill with a setup drill to test the ability to set up the materials and equipment required for a POD in a timely manner.

HOT WASH AND AFTER ACTION REPORT

Although it is not required for the performance metrics reported to CDC, all exercises should end with a hot wash – a time period reserved for participants, including evaluators and

controllers, to discuss the exercise and how the group responded to it. The hot wash gives participants time to reflect on their performance during the exercise, and provide suggestions on how to improve performance in an actual emergency, as well as how to improve the exercise process itself.

Table D.1 provides a checklist that can be used to guide a hot wash for a dispensing exercise.

Table D.1: Sample Checklist for Dispensing Exercise Hot Wash

<ul style="list-style-type: none"> • Where did the longest lines form? Did this differ during busy and non-busy periods? • Were participating staff adequately trained? What training interventions might improve performance? • Would including staff with other skills improve performance? If so, what skills are needed? • Was the POD floor plan effective and efficient? How could the person flow be improved? • Were the facility and materials adequate? Was anything missing? • Did participants understand their roles adequately? Was JIT training used? • Was communication among participants adequate? • If problems were encountered during the exercise, how well did players adapt and problem solve? • Were any persons more difficult to serve than others? How might dispensing be improved for them? • Did the POD respond appropriately to all health conditions? Were all person assessments and dosages correct (per external medical review)?

Jurisdictions should develop a brief, written, after action report. After action reports summarize the key activities tested during the exercise and key findings and improvement strategies for addressing performance gaps. There is no required format for such a report. However, most include brief summaries of the exercise as well as bulleted lists of strengths and areas for improvement. Readers should consult the Department of Homeland Security’s Homeland Security Exercise Evaluation Program web site for guidance on writing effective after action reports (<http://www.hseep.dhs.gov>). It is usually helpful to include specific corrective actions, timelines, and individuals responsible for carrying out corrective actions, in the written report.

RESOURCES ON DISPENSING

Several resources are available for jurisdictions to help guide the design and implementation of SNS Dispensing Drills.

CDC Version 10 guidance: www.cdc.gov

Best guide to helping jurisdictions design exercises.

CDC Dispensing Checklist

CRI Checklist

The Weill/Cornell Bioterrorism and Epidemic Outbreak Response Model:

http://www.hospitalconnect.com/aha/key_issues/disaster_readiness/resources/vaccination.html

An interactive planning tool designed to estimate the number of staff needed to operate a mass prophylaxis center given specific population size and staff limitations.

NACCHO BT ToolBox: <http://bt.naccho.org/Bt-Toolbox/>

Repository of emergency preparedness information by, for, and about local public health systems. Its central feature is a searchable database of public health preparedness materials, such as response plans, emergency protocols, table-top exercises, training curricula and programs, videos, CD-ROMs, outbreak investigation tools, vaccination and treatment implementation guides, etc.

The CDC Public Health Information Network (PHIN) has several archived broadcasts on POD operations – including one on measurement of POD performance.

DATA CAPTURE SHEET: DISPENSING DRILL

Required Metrics

Throughput

Input these values

Total # of persons

Time elapsed (in minutes)

Your throughput is **persons per hour**

Mean Flow Time

Input these values

Sum of all flow times* (min.)

Total # of persons

Your mean flow time is **minutes**

*Flow time is the amount of elapsed from a person's entry to exit

Dispensing Accuracy

Input these values

of persons with correct pill type

Total # of persons

Your dispensing accuracy rate is **percent**

Other Required Information

CDC-DSNS collects this information in order to interpret the metric data and to help develop the knowledge base about exemplary practices. The information below is not used for assessment purposes

Date of drill (mm/dd/yyyy)

Time of drill

AM / PM

Did drill use express line? (Yes/No)

List other lines employed

Is POD Floorplan/layout attached? (Yes/No)

Staff used

medical staff

non-medical staff

other staff

Use of volunteers

volunteer POD staff (not patients)

health department employees

other paid staff

Other comments (esp. documentation of POD shrinkage (see manual):

APPENDIX C: DRAFT MATHEMATICAL ROLL UP MODEL

Each of the drills presented in Section 4 results in one or more metrics related to the performance of a vital function within a mass prophylaxis operation. The spreadsheet model presented in this section provides a means by which those metrics may be rolled up into an estimate of each function's ability to perform at a level that would enable a jurisdiction to meet the CRI 48-hour timeline.

The roll up tool consists of three tabs, presented at the end of this Appendix:

Tab 1: PODstaff. Judges the number of PODs with sufficient staffing

Tab 2: PODtimes. Judges the timeliness of the POD function from site activation and staff call down, through dispensing

Tab 3: RSStimes. Judges the timeliness of the RSS and distribution function from site activation and staff call down, through pick list generation, picking, and distribution

Tab 1: POD Staffing Levels (POD staff)

The assessed jurisdiction starts by inputting the number of staff required to operate one shift of each POD, according to local POD plans (line 1.1). An example with 5 PODs is shown, but the tool could be expanded to accommodate a larger number of PODs. Jurisdictions also input the number of shifts each POD will be operating (line 1.2). Multiplying the two together gives the number of staff needed at each POD (line 1.3).

The jurisdiction is then asked how many people have been recruited to work at each POD (line 1.4). Dividing this number (line 1.4) by the number of people required (line 1.3) gives the percentage staffing level for each POD (line 1.5). Ideally there are more staff recruited than are required (i.e., these percentages should be greater than 100%) because on any given day some percentage of the staff will be unreachable or otherwise unavailable to work.

Availability percentages from staff call down drills are entered on the next line, for each POD (line 1.6). Multiplying this percentage (line 1.6) by the number of people recruited (line 1.4), gives the number of recruits who are available to work, for each POD (line 1.7).

Jurisdictions will want to compare the number of recruits available to work for each POD against the number of people needed to staff each POD, and take note of which PODs will be understaffed. The spreadsheet indicates whether a POD sufficiently staffed (line 1.8), and counts the number of PODs that are sufficiently staffed (line 1.9).

Tab 2: POD Timing (PODtimes)

The POD Timing portion of the roll up model assumes that three functions operate in parallel (i.e., simultaneously), and that all three of these functions must occur before any other POD operations may begin:

- Site activation
- Call down
- Assembly

Site activation. The jurisdiction begins by entering the time needed for the call and acknowledgement process from the site activation drill (line 2.1).⁴⁵ As part of that drill, each site owner reports the amount of time needed to prepare the facility, to clear it and make it ready for health department to begin setup. These numbers are entered, one for each site (line 2.2). Adding the amount of time for calling all the sites (line 1.1) to the preparation time needed at each site (line 2.2) provides an estimate of the number of hours needed to have each site ready for setup (line 2.3).

Call down and Assembly of staff. The staff call down process would normally occur as sites are being activated. Thus, the jurisdiction enters the time needed for the call down and acknowledgement process from the call down drill (line 2.4). If an actual assembly is performed, the amount of time needed for the staff to actually assemble would be reported next. However, in the absence of actual assembly, the roll up tool uses the targeted assembly time chosen by the jurisdiction in their POD planning (line 2.5).⁴⁶

Presumably, a jurisdiction's targeted assembly time will have been set to be well after the call down process has completed. If the call down process is slow, however, then it becomes the constraining factor as to when staff will assemble. The calculation for total time elapsed for Call down and Assembly is the *greater* of the Call down time and the targeted Assembly time.

Setup. Finally the greater of these two subtotals – the time needed for the facility to be available (line 2.3) and the time needed for the call down and assembly (line 2.6) – gives the total time for the facility to be available and staffed, ready for setup (line 2.7).

⁴⁵ As noted in Section 4, the site activation drill involves calling *all* the sites on the list. Since POD sites are generally more numerous than other facility sites, this should provide a close approximation of the time required to activate the POD sites only. Also, the model assumes that all sites will be cleared after the calling process when, in fact, the clearing of some sites may begin while others are still being called.

⁴⁶ This number was used in the call down drill as part of the calculation of the Assembly percentage: Called staff report the amount of time needed to assemble; that number is compared against the targeted assembly time chosen by the jurisdiction, to give the percentage of staff members available to assemble by the deadline.

Next, jurisdictions enter the amount of time required to set up each POD, as calculated in the Setup drill (line 2.8). Adding the time needed for the facility and staff to be ready for setup (line 2.7) to the setup time (line 2.8) gives the total time for each facility to be ready to begin dispensing operations (line 2.9). The spreadsheet tool counts the number of PODs that are ready for dispensing operations *12 hours*⁴⁷ from when site activation and call down began (line 2.10).

Dispensing. The assessed jurisdiction begins by entering the approximate portion of the community population assigned to each POD, according to their plan (line 2.11). They then enter the throughput achieved for each POD from their dispensing drill, expressed as patients per hour (2.12).⁴⁸ Dividing the population assigned to a POD by that POD's throughput gives the amount of time needed for each POD to conduct the dispensing operation for its assigned population, from the time the dispensing operation starts (line 2.13). By adding to that time, the time needed for Site Activation, Call Down, Assembly, and Setup, we obtain the total amount of time needed for prophylaxis from the time sites are first activated (line 2.14).

The spreadsheet calculates the mean (line 2.15), as well as shortest (line 2.16) and longest (line 2.17) dispensing times for each POD. It identifies the number of PODs that can conduct the dispensing operation in *36 hours*,⁴⁹ counted from when dispensing begins (line 2.18). The spreadsheet also identifies the number of PODs that can conduct the dispensing operation in *48 hours*, counted from when site activation begins (line 2.19).

Tab 3: RSS Timing (RSS times)

Site activation, Call down, Assembly. The calculations for these operations occur in a similar fashion to that of the PODs, above. Assessed jurisdictions enter times for site activation (lines 3.1, 3.2, and 3.3), staff call down (line 3.4) and assembly (lines 3.5 and 3.6), which are assumed to occur simultaneously. The maximum of these completion times is the time needed for facility and staff to be ready to begin setup (line 3.7). The spreadsheet counts the number of RSS ready to begin setup within *6 hours* of the beginning of Site activation (line 3.8), which is the standard provided by the Target Capabilities List (TCL) for Medical Supplies Management and Distribution.

⁴⁷ This is *not* an established SNS standard or requirement. Indeed, the Target Capabilities List (TCL) for Mass Prophylaxis only requires that operations be ready to begin 24 hours from the decision to activate. However, this would leave only 24 hours for dispensing if cities are to meet the 48 hour CRI timeline. We believe that cities should strive for a 12 hour target, corresponding to the amount of time needed for a push package to arrive at the RSS, and leaving 36 hours available for dispensing.

⁴⁸ As the drill has been designed for “green” or “express” patients, this is an optimistic estimate.

⁴⁹ One again, this is *not* an established SNS standard. However, it corresponds to the CRI standard of 48 hours, less 12 hours for the push package to arrive at the RSS (during which Site activation, Call down and Assembly of staff, and setup for the PODs should also occur). It is also an optimistic assumption; in reality more time will be taken for the materiel to be processed at the RSS and distributed to the PODs.

Setup. Assessed jurisdictions enter the time needed to perform the setup of the RSS, from the Setup drill (line 3.9). The spreadsheet adds this time to the amount of time until the facility and staff will be ready to begin setup (line 3.7) to obtain the total time for the RSS to be staffed and set up (line 3.10).

Inventory apportionment, Pick list generation. Assessed jurisdictions enter the time needed for to apportion inventory (line 3.11) – note that this drill has not been drafted yet. They also enter the time needed to generate pick lists based on the pick list generation drill (line 3.12). These times are added to the elapsed time through setup (line 3.10) to give the time needed for the RSS to be ready to begin handling materiel (line 3.13). The spreadsheet determines the number of RSS facilities for which all these operations can be performed in 12 hours⁵⁰ from when Site activation begins.

Unload, Pick, Distribute. These are operations that occur after the SNS materiel has arrived at the RSS warehouse. Note that drills for these operations have not been drafted. Most likely these operations will be practiced as part of a full RSS drill, or as part of a full scale exercise (FSE). Jurisdictions would enter the time needed to conduct unloading of the SNS materiel (line 3.15) and picking of the orders for the PODs (line 3.16). They would then enter the amount of time needed for these materials to be distributed to the PODs (line 3.17).⁵¹

These times would be added to the time it took for the RSS to be ready to handle materiel (line 3.13) to give the total time needed to get materiel out to each POD, starting from the activation of the RSS (line 3.18). Subtracting these times from 48 gives the number of hours left for each POD to conduct dispensing operations, assuming their facilities and staff are ready (line 3.19).

⁵⁰This is not an established SNS standard. However, we feel that these operations should be completed by the time the SNS materiel arrives at the RSS (thus the 12-hour window). If not, any further time taken to complete these operations will further delay the movement of materiel.

⁵¹ In the absence of drill data, a combination of checklists and simulations may be used to estimate the time to perform these tasks. This is an area that requires further development.

Tab 1: POD Staffing Levels Roll Up

		Site 1	Site 2	Site 3	Site 4	Site 5
1.1 Number of people required to staff one shift	<i>enter # of staff</i>	87	90	80	85	85
1.2. Number of shifts	<i>enter # of shifts</i>	2	2	2	2	2
1.3. Total number of people required	<i>Calculation: staff per shift * shifts</i>	174	180	160	170	170
1.4. Number of people recruited	<i>enter # of staff</i>	180	200	150	250	230
1.5. Percent of required staff that have been recruited	<i>Calculation: staff recruited / staff required</i>	103%	111%	94%	147%	135%
1.6. Assembly percentage from call down drill	<i>enter % reporting available</i>	85%	90%	87%	70%	75%
1.7. Number of recruits available to report	<i>Calculation: staff recruited * assembly percentage</i>	153	180	130.5	175	172.5
1.8. Sufficient staff?	<i>Yes if recruits available >= staff required</i>	N	Y	N	Y	Y
1.9. Number of PODs with sufficient staff	<i>Count</i>	3				

Tab 2: POD Timing Roll Up

Operations done in parallel, during the same 12-hour window							
Site Activation	2.1: Time needed to complete the call and acknowledgement process	enter # of hours	2				
			Site 1	Site 2	Site 3	Site 4	Site 5
	2.2: Facility owner's reported preparation time (time needed to clear facility)	enter # of hours	2	12	4	8	3
		Calculation: Facility Activation time + Facility preparation time	4	14	6	10	5
2.3: Total time for facility to be ready for setup							
Call Down	2.4: Time needed to complete the call and acknowledgement process	enter # of hours	4				
Assembly	2.5: Target Assembly Time stated in Call Down Drill, stated in terms of # of hours from start of Call Down Drill (E.g., 4 hours)	enter # of hours	6				
	2.6: Total time elapsed for Call Down and Assembly: time for staff to be ready	Calculation: Max of Call Down, Assembly (hours)	6				
			Site 1	Site 2	Site 3	Site 4	Site 5
2.7: Total time for facility AND staff to be ready for setup		Calculation: Max of facility ready time, staff ready time	6	14	6	10	6
Operations done after preceding steps							
Setup			Site 1	Site 2	Site 3	Site 4	Site 5
	2.8: Setup time, from drill, by site	enter # of hours	2	2	4	3	2
2.9: Hours until POD is ready to operate		Calculation: Facility and staff ready time + Setup (hours)	8	16	10	13	8
2.10: # of PODs ready within 12-hour window for SNS delivery		Calculation: Count	3				
Operations done after the 12-hour window							
Dispensing	2.11: Population assigned to site	enter population (people)	10,000	10,000	15,000	5,000	12,000
	2.12: Throughput, from POD Drill, by site	enter drill throughput (people/hour)	400	300	500	200	300
		Calculation: Population / Throughput (hours)	25	33.3	30	25	40
	2.13: Hours to prophylax population, by site (not including:						
	2.14: Hours to conduct entire sequence of operations: Site Activation, Call down, Assembly, Setup, and Prophylaxis	Calculation: Hours until POD is ready + Hours to prophylax population	33	49.3	40	38	48
	2.15: Mean Dispensing Time	Calculation: Mean of site prophylax time (hours)	30.7				
	2.16: Shortest Dispensing Time	Calculation: Min of site prophylax time (hours)	25				
	2.17: Longest Dispensing Time	Calculation: Max of site prophylax time (hours)	40				
	2.18: # of PODs able to prophylax in 36 hours	Calculation: Count	4				
	2.19: # of PODs able to Call Down, Assemble, Set up and Prophylax assigned population in 48 hours	Calculation: Count	4				

Tab 3: RSS Timing Roll Up

Operations done in parallel, during the same 12-hour window

Site Activation	3.1: Time needed to complete the facility call and acknowledgement process	enter # of hours	1
	3.2: Facility owner's reported preparation time (time needed to clear facility)	enter # of hours	2
	3.3: Total time for facility to be ready for setup	Calculation: Facility Activation time + Facility preparation time	3
Call Down	3.4: Time needed to complete the staff call and acknowledgement process	enter # of hours	4
Assembly	3.5: Target Assembly Time stated in Call Down Drill. In # of hours from start of Call Down Drill (E.g., 4 hours)	enter # of hours	6
	3.6: Total time elapsed for Call Down and Assembly: time for staff to be ready	Calculation: Max of Call Down, Assembly (hours)	6
	3.7: Total time for facility AND staff to be ready for setup	Calculation: Max of facility ready time, staff ready time (hours)	6
	3.8: # of RSS staffed within 6-hour window (per TCL)	Calculation: Count	1

Operations to be done after the above two operations are done

Setup	3.9: Setup time, from drill, by site	enter # of hours	2
	3.10: Hours until RSS is staffed and set up	Calculation: Facility and staff ready time + Setup (hours)	8
Apportion	3.11: Inventory apportionment time, from drill	enter # of hours	1
Picklist generation	3.12: Picklist generation time, from drill	enter # of hours	1
	3.13: Total time for RSS to be ready to handle material	Calculation: RSS staffed & set up time + Apportion time + Picklist generation time	10
	3.14: # of RSS ready to begin handling material within 12-hour window	Calculation: Count	1

Operations done after the 12-hour window

Unload	3.15: Unloading time, from RSS drill	enter # of hours	1					
Pick	3.16: Pick time, from RSS drill	enter # of hours	2					
Distribute	3.17: Delivery time, from FSE	enter # of hours		Site 1	Site 2	Site 3	Site 4	Site 5
			1	1	0.5	2	1	1
	3.18: Number of hours until delivery to PODs	Calculation: RSS ready-to-handle time + Unloading time + Pick time + Delivery time	14	14	15	14	14	
	3.19: Number of hours within 48-hour window remaining for POD prophylaxis	Calculation: 48 hours less delivery time	34	35	33	34	34	

REFERENCES

- Centers for Disease Control and Prevention (CDC). 2006. *Continuation Guidance for Cooperative Agreement on Public Health Preparedness and Response for Bioterrorism---Budget Year Six*.
<http://www.bt.cdc.gov/planning/coopagreement/#fy06>
- Davis P. 2002a. *Analytic Architecture for Capabilities-Base Planning, Mission-System Analysis, and Transformation*. MR1513-OSD. Santa Monica: RAND.
- Davis P. 2002b. "Integration Transformation Programs." In *Transforming America's Military*, Hans Binnendijk, ed. Center for Technology and National Security Policy, National Defense University Press.
- Hupert N, J Cuomo, M Callahan, A Mushlin & S Morse. 2004. *Community-Based Mass Prophylaxis A Planning Guide for Public Health Preparedness*. Washington, DC: Agency for Health Research and Quality [AHRQ].
- Institute of Medicine. 2000. *To Err Is Human: Building a Safer Health System*, Washington, D.C.: National Academy Press.
- Kelly J. 1961. Critical Path Scheduling and Planning: Mathematical Basis. *Operations Research* 9(3): 296-320.
- Klerman, J. 2004. Measuring Performance. In Light, P. & R. Klitgaard, eds. *High Performance Government: Structure, Leadership, Incentives*. Santa Monica, CA: RAND.
- McClain J, L Thomas, & J Mazzola. 1992. *Operations Management: Production of Goods and Services*. Englewood Cliffs NJ: Prentice-Hall.
- Nelson C, L Shugarman, N Lurie, & C Hartman. 2006. Improving Performance Accountability Systems for Public Health Emergency Preparedness: An Assessment of CDC's Cooperative Agreement Guidance and a Proposed Strategy for Improvement. PM-1905-DHHS. Santa Monica, CA: RAND.
- Nelson C, E Chan, M Moore, M Shea, E Sloss. 2005. Assessing State and Local Capacity to Receive, Distribute, and Dispense Strategic National Stockpile Assets: Proposed System Architecture and Design Options. DRR-3823-DHHS. Santa Monica, CA: RAND.
- Nelson C & E Sloss. 2005. Proposed Short- and Long-Term Changes to Current SNS Assessment Tool. DRR-3824-DHHS. Santa Monica, CA: RAND.

- Pande P, R Neuman, & R Cavanagh. 2000. *The Six Sigma Way: How GE, Motorola, and Other Top Companies are Honing Their Performance*. New York: McGraw-Hill.
- Rendin R, N Welch, & L Kaplowitz. 2005. Leveraging Bioterrorism Preparedness for Nonbioterrorism Events: A Public Health Example. *Biosecurity and Bioterrorism: Biodefense, Strategy, and Science*. 3: 309-315.
- Seid M, Lotstein D, Williams V, Nelson C, Lurie N, et al. 2006. Quality Improvement: Implications for Public Health Preparedness. TR-316-DHHS, RAND, Santa Monica, CA. http://www.rand.org/pubs/technical_reports/TR316/
- U.S. Department of Homeland Security. 2005. *Target Capabilities List*, Version 2.0. Washington, DC: Government Printing Office.
- U.S. Federal Emergency Management Agency. 1996. *Guide for All-Hazard Emergency Operations Planning (SLG-101)*. Washington, DC: Government Printing Office.
- U.S. Occupational Safety and Health Administration. 2005. <http://www.osha.gov>.