This product is part of the RAND Corporation technical report series. Reports may include research findings on a specific topic that is limited in scope; present discussions of the methodology employed in research; provide literature reviews, survey instruments, modeling exercises, guidelines for practitioners and research professionals, and supporting documentation; or deliver preliminary findings. All RAND reports undergo rigorous peer review to ensure that they meet high standards for research quality and objectivity.

TECHNICAL REPORT

New Tools for Assessing State and Local Capabilities for Countermeasure Delivery

Christopher Nelson, Edward W. Chan, Carol Fan, Debra Lotstein, Leah B. Caldarone, Shoshana R. Shelton, Amy L. Maletic, Andrew M. Parker, Alexandria Felton, Amanda Pomeroy, Elizabeth M. Sloss

Prepared for the Department of Health and Human Services

HEALTH

This work was sponsored by the U.S. Department of Health and Human Services Office of the Assistant Secretary for Preparedness and Response. The research was conducted in RAND Health, a division of the RAND Corporation.

Published 2009 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665
RAND URL: http://www.rand.org
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

SUMMARY

If a biological attack occurs or an infectious disease breaks out, medical countermeasures, such as vaccines, antibiotics, or antidotes, may need to be delivered rapidly to the general public. The federal government, through the Centers for Disease Control and Prevention (CDC), maintains a cache of pharmaceutical drugs and other medical materials, known as the Strategic National Stockpile (SNS). Those materials [called materiel in the main text] can be sent to state and local public health departments. State and local governments are responsible for requesting the material, receiving it from the federal government, distributing it to local areas, and, ultimately, dispensing countermeasures to members of the public.

Assessing state and local departments' readiness to perform these functions is a critical element for improving their ability to conduct a countermeasure-delivery campaign in an emergency. The CDC's Division of the Strategic National Stockpile (CDC/DSNS) has been using a checklist-based approach, known as the Technical Assistance Review (TAR), to assess the jurisdictions' countermeasure delivery plans, equipment, and personnel. However, a consensus has emerged within the U.S. Department of Health and Human Services (HHS) and CDC that assessing these *capacity* elements is not sufficient; methods are needed for assessing the jurisdictions' *capabilities* to implement and adapt these plans in real-world situations.

To address this need, the United States Department of Health and Human Services Office of the Assistant Secretary for Preparedness and Response (HHS/ASPR) asked RAND to assist CDC/DSNS in developing a system for assessing the *operational capabilities* to engage in mass countermeasure delivery.

This report describes the assessments that were developed, and it provides an approach for measuring development for public health emergency preparedness that can be applied more broadly. The operations manuals that are discussed in this report provide guidance on assessing a jurisdiction's countermeasure-delivery capabilities. Users can download and print spreadsheet-based data-reporting tools that will facilitate calculation of required metrics and collection of key data. The manuals and the data-reporting tools are available on the CDC website at: http://emergency.cdc.gov/cotper/coopagreement/.

Throughout this report and the operations manuals, we use the term "patient" to refer to the individuals that visit PODs to receive countermeasures. We chose this term to be consistent with the language used in the POD modeling programs referenced in the Dispensing assessment manual. In an actual emergency, PODs would be expected to serve healthy, non-injured persons, who would not be considered patients in the strict medical sense. True patients, i.e., sick or injured individuals, would go to healthcare facilities for treatment instead.

ASSESSMENT TOOLS FOCUS ON "BUILDING-BLOCK" CAPABILITIES

Given the rarity of large-scale public health emergencies, there is little evidence base to guide decisions about what aspects of preparedness are most important to measure. Thus, we conducted engineering-style process analysis to identify core "building-block" capabilities that can be deployed, combined, and adapted in response to a broad spectrum of response scenarios. Focusing assessments around a small number of critical operational capabilities helps keep assessment burdens reasonable and promotes flexibility by encouraging jurisdictions to think of situation-specific plans as containing certain common elements. Assessment tools were ultimately created for five building-block capabilities:

- **Staff call-down**: the capability to notify and mobilize staff to perform emergency response functions.
- **Site call-down**: the operational capability to notify facilities (e.g., points of dispensing [PODs] or warehouses) that are not ordinarily under the health department's control and determine how quickly the sites could be made ready for emergency operations.
- **Warehouse pick-list generation**: a jurisdiction's ability to use its warehouse inventory system to ensure a timely match between the demand for and supply of critical resources.
- **POD set up**: the ability to quickly set up a facility for use as a POD.
- **Mass dispensing through PODs**: the ability to quickly deliver medications or other "countermeasures" to large numbers of people at a POD.

The assessments for each of these capabilities are structured in the form of an operational exercise. Each assessment features a set of clearly defined performance metrics. The Homeland Security Exercise Evaluation Program (HSEEP) and other documents provide a considerable amount of guidance on exercise evaluation; however as yet there are no comparable performance metrics that can support reliable cross-sectional and temporal comparisons of performance of key capabilities during exercises. This work is intended to meet that need.

To reduce assessment burdens, we designed each of the assessments to be used in a small-scale drill that can be conducted individually or combined as components of larger exercises. This versatility increases the range of exercises on which the assessments can be used and decreases the need for additional exercises. The assessments can also be embedded in routine public health functions, thus reducing the need to assemble key staff solely for exercising.

For each assessment, an operations manual was created with a summary of the assessed capability and detailed instructions on drill design, data collection, and calculating and reporting metrics.

FIELD TESTS ASSESSED THE FEASIBILITY, UTILITY, AND RELIABILITY OF ASSESSMENTS

Four of the five assessment tools were field-tested in 2007–2008 (one of the assessments, site call-down, was not field tested separately, because its structure and mechanics are nearly identical to the assessment of staff call-down, which was tested). In nearly all cases, field tests were done in conjunction with already-scheduled drills being conducted by the state or local health department. Following the exercise, RAND staff used a semi-structured interview protocol to guide discussions with members of the health department about the feasibility, reliability, validity, and utility of the assessment tools. Key findings were as follows:

- *Assessments were useful and feasible*. In general, field-test sites reported that the metrics provided useful feedback, that the manuals and spreadsheet-based tools were useful for managing data, and that the level of effort and skill required to use them was reasonable. However, the assessment of the pick-list generation required a higher level of assistance and explanation, largely because the field-test sites were unfamiliar with the concept of a drill devoted solely to exercising the pick-list-generation function.
- *Time studies and computer models presented special challenges.* The mass dispensing assessment requires sites to conduct time studies of POD stations to measure the number of minutes or seconds required to process patients. While collecting the information on timing was feasible for sites during the field tests, staff expressed concern about the number of people needed to do the timing. Moreover, staff at the field-test site were generally not comfortable with computer modeling.
- *Flexibility built into assessments might limit data comparability across sites.* The assessments provide considerable flexibility to sites in specifying exercise scenarios and response processes. However, such variations might significantly limit the ability to draw valid comparisons among performance-metric data across sites and over time. For instance, allowing health departments to choose their own scenario parameters (e.g., time of day for staff call-down, patient mix in the mass-dispensing assessment, start and stop times for POD set up) affected observed performance and made it nearly impossible for us to compare performance on the same drill across pilot sites.
- *More testing is required to assess validity.* Those in the test sites generally found that the results of the assessments had face validity. However, we were unable to validate the results from these assessments against results from real-world responses or other assessments of the same capability.

IMPLICATIONS AND NEXT STEPS

Policymakers must consider several issues when implementing the assessments.

Further Refinements Will Require Decisions About Balancing Standardization and Flexibility

Findings from the field tests suggest that further efforts will be required to control sources of variation captured by the assessments that go beyond the operational capabilities they were designed to measure (also known as *error variance*[1]). Two possible approaches to address these issues are (a) to prescribe more specifically the drill scenarios and tasks, and (b) to collect information on drill context and processes and use this information to make post-hoc adjustments in interpreting assessment results.

Given that a key goal in designing the assessments was to maximize flexibility to jurisdictions, the assessments generally favored the second option (b), collecting extensive information on drill scenarios and procedures. However, the success of any post-hoc adjustment strategy will depend on the number of jurisdictions reporting data and the ability to identify the right contextual variables. As a result, post hoc adjustment might need to be supplemented by greater standardization in scenarios and procedures.

Consider Technical-Assistance Requirements of Performance Assessments

Most of the assessments appear to be quite feasible; however, it is likely that training will be required to ensure consistent application of the procedures and metrics provided in the manuals. Given that expertise with computer modeling varies among health departments, the computer-based dispensing assessment will likely require substantially more technical assistance than the others.

Recommended Next Steps

While making plans to roll out the assessment tools described in this report, HHS/ASPR and CDC/DSNS should consider the following recommendations:

The results of the field testing highlight that the science and art of measuring public health emergency response capabilities remain in their infancy. The assessments were found to be useful and feasible; yet, it is clear that data collected across different sites might not be fully comparable. Even imperfect assessments of capability provide an important complement to the current practice of measuring preparedness by plans, staffing, and equipment. Thus, it seems reasonable to consider rolling out the assessments on a trial basis.

However, decision makers should ensure that processes are in place to capture lessons learned from initial rollout and revise the measures accordingly. Specifically, they must

- **Collect performance-metrics data during the first year of implementation.** Such data will assist in the process of refining the measures. CDC/DSNS may want to consider developing a standardized feedback form to capture lessons from awardees.

---

[1] Variance arising from unrecognized or uncontrolled factors in a study.

- **Analyze data collected during the first year and revise assessments accordingly.** Analyzing data from initial implementation will help assess the quality of the data. Data distributions should be examined for unexpected variation, with the goal of distinguishing true score variance from that due to errors in conducting the assessment. Assessment manuals and metrics should be revised to address problems revealed by this analysis.
- **Develop strategies for ensuring that awardees use the metrics to improve performance.** CDC should clarify how the data from the assessments might fit into the Pandemic and All-Hazards Preparedness Act's (PAHPA's) penalties for nonperformance and pay-for-performance requirements and consider a more focused pilot study to develop best practices in process improvement.

Continue to develop the evidence base behind measures. The data collected for these metrics will provide a valuable tool for more in-depth studies. For example, the data might be used to examine whether capacity measures on the Technical Assistance Review predict actual performance on operational-capability metrics.