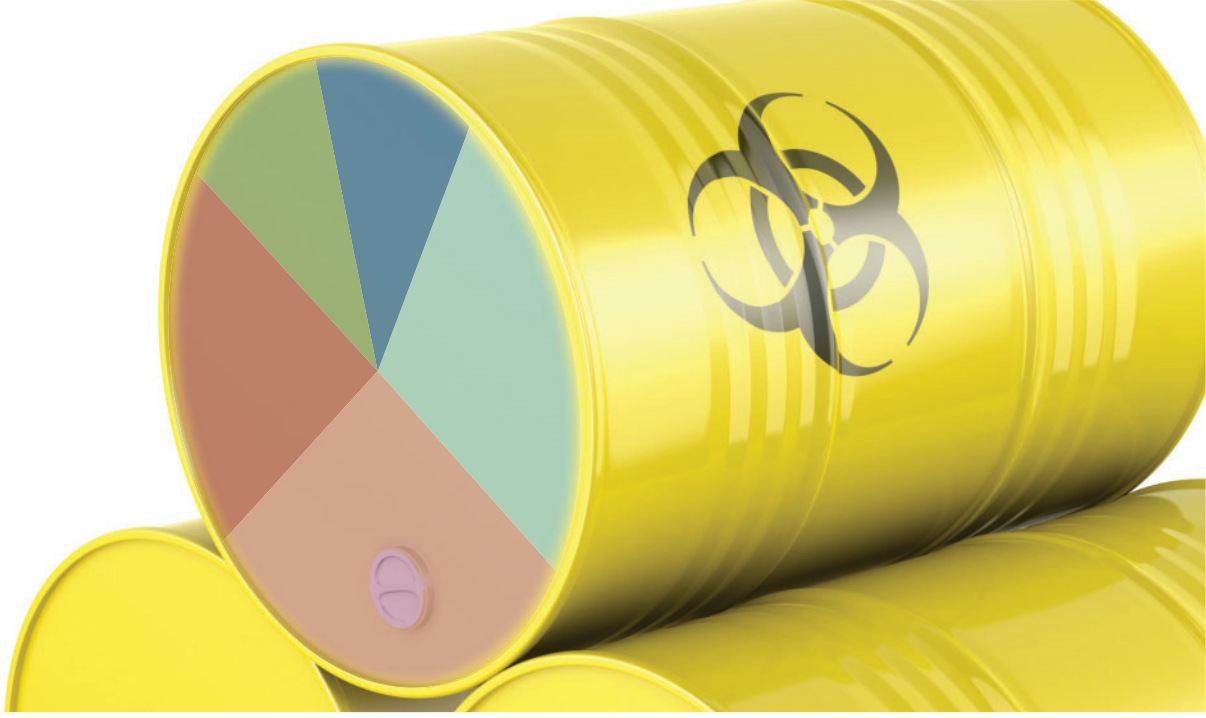
This report is part of the RAND Corporation research report series. RAND reports present research findings and objective analysis that address the challenges facing the public and private sectors. All RAND reports undergo rigorous peer review to ensure high standards for research quality and objectivity.



Measuring Cooperative Biological Engagement Program (CBEP) Performance

Capacities, Capabilities, and Sustainability Enablers for Biorisk Management and Biosurveillance

Stephanie Young, Henry H. Willis, Melinda Moore, Jeffrey Engstrom

RAND CORPORATION

NATIONAL DEFENSE RESEARCH INSTITUTE

Measuring Cooperative Biological Engagement Program (CBEP) Performance

Capacities, Capabilities, and Sustainability Enablers for Biorisk Management and Biosurveillance

Stephanie Young, Henry H. Willis, Melinda Moore, Jeffrey Engstrom

Prepared for Cooperative Biological Engagement Program

For more information on this publication, visit www.rand.org/t/rr660

Library of Congress Cataloging-in-Publication Data
ISBN: 978-0-8330-8693-8

Published by the RAND Corporation, Santa Monica, Calif.

Preface

At the end of the Cold War, remnants of the Soviet Union's biological weapon complex presented a significant threat to U.S. and international security. As part of the U.S. government's broader Cooperative Threat Reduction (CTR) program, the Biological Threat Reduction Program (BTRP) addressed proliferation risks associated with biological agents, related materials, and technical expertise associated with the defunct biological weapon program. Under its authorities to operate in the former Soviet Union, BTRP worked to destroy bioweapon production facilities, consolidate collections of dangerous pathogens, and support peaceful research activities to employ personnel with knowledge. By 2002, the program had also began to work to improve partner biosurveillance capabilities where needed. Today, global health security challenges arise from developments far broader than state-sponsored biological weapon programs, such as that of the Soviet Union, and BTRP has evolved to meet them. Now called the Cooperative Biological Engagement Program (CBEP), the biological threat component of CTR partners with about 20 countries in different regions around the world and works with them to address diverse threats to international security, including terrorist organizations seeking to acquire pathogens of security concern; human, animal, and agricultural facilities operating with inadequate safety and security safeguards; and the spread of diseases with potential security or economic consequences.

As the program has evolved, so too should approaches to monitoring program performance. Recognizing this, Congress requested that the Department of Defense improve metrics for measuring the performance of its CTR efforts and called on the National Academy of Sciences to recommend refinements to CTR performance metrics. In its 2012 report, the academy noted that, relative to other CTR programs, CBEP poses unique challenges for measurement. The research reported here builds on existing work by developing an evaluation framework for CBEP, recommending a set of metrics for assessing CBEP performance, and providing guidance on interpreting and implementing metrics to support reporting on CBEP performance to several audiences, ranging from the program to the strategic level. This work should be of interest to CBEP and CTR leadership as they seek to improve program design, budget allocation, and execution and to communicate effectively about program performance in the areas of biorisk management and biosurveillance to external audiences. For the

broader community of program evaluation practitioners, the work advances innovative approaches, aimed at targeting measurement on desired impacts more precisely.

This research was sponsored by the Cooperative Biological Engagement Program in the Defense Threat Reduction Agency's Cooperative Threat Reduction Program, and conducted within the Acquisition and Technology Policy Center of the RAND National Defense Research Institute, a federally funded research and development center sponsored by the Office of the Secretary of Defense, the Joint Staff, the Unified Combatant Commands, the Navy, the Marine Corps, the defense agencies, and the defense Intelligence Community.

For more information on the RAND Acquisition and Technology Policy Center, see http://www.rand.org/nsrd/ndri/centers/atp.html or contact the director (contact information is provided on the web page).

Contents

Figures and Tables

Figures

Tables

Summary

The Cooperative Biological Engagement Program (CBEP) within the Department of Defense (DoD) is part of the larger Cooperative Threat Reduction (CTR) program that was initiated at the end of the Cold War. While CTR began in the former Soviet Union (FSU), it has now has expanded beyond its Cold War legacy to address current biological threats of mutual interest to the United States and 20 partner countries in different regions around the world.

As the program has evolved since its inception two decades ago, so too have its content and approaches to performance measurement. Congress has shown particular interest in encouraging DoD to improve CTR assessments. Legislation for fiscal year 2010, for example, called on DoD to develop metrics for assessing CTR programs and on the National Academy of Sciences (NAS) to assess DoD's efforts. CBEP leaders approached the RAND National Defense Research Institute to build on existing work related to measurement of program performance, develop a comprehensive evaluation framework, and recommend metrics for assessing and communicating CBEP progress toward program goals. This report details the work we carried out between February and October 2013, including our methods, findings, recommendations, and guidance for implementing and aggregating the recommended metrics.

Using a Conceptual Framework to Identify Metrics for CBEP

At each level of management, the demand for metrics is motivated by the need to explain how actions lead to anticipated results. The development and application of logic models underpinned our efforts. The logic model approach establishes a logical sequence from program inputs to activities, outputs, outcomes, and impacts. Using logic models to identify metrics to recommend required us to address three key questions related to measurement of program performance.

What should CBEP measure to evaluate program performance? This question focused on identifying CBEP program objectives. To do so, we reviewed international, U.S. governmental, DoD, and CBEP guidance documentation related to biological engagement and engaged top stakeholders in discussions to fully capture diverse perspectives related to overall program goals.

We identified two objectives for CBEP to serve as the basis for the development of logic models:

- *Strengthen enduring partner capabilities for biorisk management.* This objective refers to both biosafety and laboratory biosecurity control measures, as well as associated risk assessment and oversight functions. We developed separate logic models for biosafety and biosecurity, which reflect both common elements and the components that make each unique.
- *Strengthen enduring partner capabilities for biosurveillance.*

These objectives introduce two key elements of our conceptual frameworks. First, they specify that CBEP's objective is to focus on building partner *capability*, meaning a demonstrated operational ability. The distinction between partner capabilities and specific program outputs is a key element of the logic models. Second, the objectives specify that the partner capabilities CBEP seeks to strengthen need to be *enduring*. This element drove our focus in the logic model on sustainability enabling factors.

Other important CBEP program activities, most prominently cooperative biological research (CBR), *support* these overall program goals. As we discuss in Chapter Two, CBR is a CBEP program activity with logical relationships to capacities, enduring capabilities, and sustainability enablers in support of biorisk management and biosurveillance objectives; in itself, CBR has a fundamentally different relationship to threat reduction than does biorisk management or biosurveillance. CBR activities resource and exercise partner biorisk management and biosurveillance programs and support the longer-term sustainability of other CBEP investments in partner nations. The metrics we recommend reflect CBR's status as a sustainability enabler.

How should CBEP systematically identify metrics? Next, we developed conceptual frameworks for systematically capturing program goals, objectives, and program activities to serve as the foundation for developing metrics. The development and application of logic models underpinned our efforts, although other approaches would also work.

We developed separate logic models for biosafety, biosecurity, and biosurveillance that consisted of three primary classes of elements—capacities, capabilities, and sustainability enables. Our distinction between these three elements allows CBEP to ensure that measurement focuses on outputs that are closer to desired impacts. Our logic models focus measurement on the level of program outputs and partner-nation operational abilities. At this level of analysis, we emphasized three types of program outputs: capacities, sustainability enablers, and enduring capabilities. *Capacities* are the building blocks required to enable partner capabilities. *Sustainability enablers* are the building blocks required to maintain partner capacities and capabilities in the longer term. *Enduring capabilities* are demonstrated and sustainable partner operational abilities. These are the ultimate outputs of CBEP program activities, which in turn lead

to strategic-level outcomes and mission impacts; CBEP helps to create or strengthen partner capacities and implements partner sustainability enablers to enable enduring partner capabilities.

What metrics do we recommend for CBEP? We identified a set of metrics, grounded in the logic models and selected to best meet relevant criteria, that helps ensure appropriate targeting and robustness in the measurement scheme. Our approach to developing metrics began with an environmental scan of metrics related to biorisk management and biosurveillance that CBEP or others currently recommend or use. We then scored these potential metrics by taking into account their strength along several dimensions:

- Validity: To what extent does the metric capture the concept being addressed?
- Reliability: How consistent is the measurement?
- Feasibility: How easy is it to collect the data needed?
- Utility: How useful is the metric?
- Relevance: To what extent does the metric reflect CBEP programming?

These criteria helped us achieve several objectives with respect to identifying metrics. The validity criterion helped us reduce the possibility that program assessment motivates activities that receive good evaluations but do not in fact achieve program goals. This form of validity is necessary for measurement but is not sufficient. The metrics must also be reliable, such that evaluators can confidently make comparisons across countries and over time as engagement with partner countries progresses. It must also be feasible for program managers to be able to apply the metrics within the constraints of the time, knowledge, and people available to conduct assessments. We also sought to identify metrics that would be most useful for decisionmaking, assuming they were valid and reliable measures, and that data collection to support them was feasible. Finally, CBEP is one of several U.S. government programs with responsibilities for improving partner biorisk management and biosurveillance capabilities. The criterion of "relevance to CBEP" helped ensure that the metric was appropriate to CBEP's threat reduction mission and focused on pathogens of security concern.

The scoring process allowed us to screen potential metrics to select only those that scored relatively highly across all our criteria. We then aligned metrics emerging from our screening process with specific elements of our logic models and assessed coverage across all elements of the logic models. This step helped us develop a list of potential metrics, which we then vetted with CBEP personnel, especially with respect to practicality and completeness of the overall metrics framework. The vetting process allowed for further revision of recommended metrics.

Recommended Metrics

Our screening process allowed us to identify almost 300 metrics in the literature, from which we were able to isolate a subset of the strongest scoring along several dimensions. We recommend 47 metrics for immediate CBEP use for assessing program performance. These include 21 metrics for biorisk management (BRM) and 26 metrics for biosurveillance to be introduced now, and 13 metrics recommended for introduction later. Table S.1 shows that the distribution of the recommended metrics across the three elements of our logic models and two time frames.

Table S.2 lists all 47 recommended metrics and provides additional key information, including the type of metric and an indication of whether the data to support the metric currently exist in CBEP, are planned to exist in CBEP, or will require new data collection. CBEP leaders will need to carefully consider the desirable number of metrics, balancing the desire for robustness of actionable and strategically important information with the burden of data collection. Fortunately, CBEP already collects substantial performance information, which will lessen the additional burden (about one-half of the recommended metrics rely on data already existing or planned for collection within CBEP). CBEP will nonetheless need to make decisions about expanding data collection or making existing data more reliable and robust.

Many existing metrics draw on indicators of partner performance that CBEP tracks using a software and visualization program called the Metrics Assessment Program (MAP). As Table S.2 indicates, MAP inputs are more likely to support metrics than biosurveillance metrics. In some cases, we recommend adapting MAP inputs for use as CBEP metrics because they offer a valid approximation of the capacity, enduring capability, or sustainability enabler in our logic model. In other cases, we recommend aggregating inputs into a single metric or adjusting the language of the MAP question to better align with our desired measurement target. Tables in the appendixes indicate how we recommend translating existing data into metrics aligned with our conceptual frameworks. Implementers on the ground provide MAP inputs; an implementer

Table S.1
Summary of Recommended Metrics

	Metrics Recommended for Use Now		Metrics Recommended for Introduction Later	
Type of Metric	Biorisk Management (no.)	Biosurveillance (no.)	Biorisk Management (no.)	Biosurveillance (no.)
Capacities	6	9	0	2
Capabilities	9	11	2	3
Sustainability Enablers	6	6	4	2
Total	21	26	6	7

Table S.2
Recommended Metrics for Use Now

Number	Metric	Type	Data Existing, Planned, New Within CBEP
Biorisk Management			
1	Percentage completion: BRM assessment, requirements and planning of facilities (MAP checklist)	Capacity	Existing[a]
2	Percentage completion: BRM management personnel (MAP checklist)	Capacity	Existing[a]
3	Number or percentage of facilities with complete sets of relevant standard operating procedures (SOPs) in place (MAP checklist)	Capacity	Existing[a]
4	SME finds that facilities are designed to allow employees to work safely and securely	Capacity	Existing[a]
5	Percentage completion: BS&S equipment (MAP checklist)	Capacity	Existing[a]
6	Percentage of personnel trained in laboratory biosafety and biosecurity relative to requirements	Capacity	Existing[a]
7	Percentage completion: biosafety: work practice and administrative control (MAP Checklist)	Capability	Existing[a]
8	Pathogen consolidation criteria and tiers	Capability	Planned[c]
9	Percentage of disclosed biological weapon–related infrastructure that has been eliminated (FSU only)	Capability	New
10	Percentage completion: maintain control of pathogens (physical security, material control and accountability, information security, transportation security) (MAP checklist)	Capability	Existing[a]
11	SME finds mechanisms are in place sufficient to ensure that personnel are competent and reliable (adapted MAP language)	Capability	Existing[a]
12	Reporting in 1540 matrix affirming existence of a mechanism for penalizing violator of national prohibition on engagement in biological weapon–related activities	Capability	Existing
13	Proficient test scores for CBEP courses related to biorisk assessment and planning	Capability	Existing[a]
14	SME finds that accidents or incidents and nonconformities related to biorisk are correctly managed	Capability	Existing[a]
15	SME finds that the BRM system is reviewed regularly	Capability	Existing[a]
16	SME finds that a policy concerning management of laboratory biorisk (biosafety and biosecurity) been written	Sustainability enabler	Existing[a]
17	SME finds that operational plans include materiel sustainment considerations	Sustainability enabler	New

Table S.2—Continued

Number	Metric	Type	Data Existing, Planned, New Within CBEP
18	Number of national or regional institutions that make BS&S training to international standards available (training centers, universities, professional societies, ministries)	Sustainability enabler	New
19	Number of peer-reviewed publications, number of conference papers (abstract, poster, or oral)	Sustainability enabler	Existing[d]
20	Number or value of internationally competitive research grants won	Sustainability enabler	New
21	SME finds that operational plans include resource sustainment considerations	Sustainability enable	New
Biosurveillance			
22	Number and percentage of priority pathogens for which baseline data have been established (e.g., via CBR) and used to detect anomalies (modified TT tiers)	Capacity	Planned[c]
23	Reporting capability (TT criteria and tiers)	Capacity	Planned[c]
24	Number and percentage of major jurisdictions using electronic reporting system (e.g., EIDSS)	Capacity	New
25	Accuracy and scope: national guidance for reporting, e.g., case definitions—specified animal and human diseases	Capacity	New
26	Number trained of field epidemiologists and number per 200,000 population	Capacity	New
27	Technical capacity (TT criteria and tiers)	Capacity	Planned[c]
28	Number and percentage of laboratories certified	Capacity	New
29	Number and list of pathogens for which (a) the national and (b) each provincial or state laboratory can test	Capacity	New
30	Laboratory networking: written or established protocols for specimen referral (a) within country or (b) to international laboratory	Capacity	New
31	Epidemiological surveillance analysis performance—through exercise or supervisory observation	Capability	Planned[c]
32	Epidemiological investigation performance documented in written report or tested via exercise (tabletop or functional)	Capability	New
33	Number of suspected priority pathogen cases or outbreaks in the past 12 months and the percentage for which investigations were conducted and results documented	Capability	New
34	Number laboratory tests performed for each priority pathogen in the past 12 months	Capability	New

Table S.2—Continued

Number	Metric	Type	Data Existing, Planned, New Within CBEP
35	Number and percentage of laboratories participating in proficiency testing at least once in the past 12 months	Capability	New
36	Number and percentage of laboratories that passed all proficiency tests in the past 12 months, among laboratories that participated in proficiency testing in the past 12 months	Capability	New
37	Specimen collection and transport: number of specimens received by laboratory and the number and percentage that are adequate for testing	Capability	Planned[c]
38	Laboratory referral network: number of specimens sent or received for confirmatory testing in the past 12 months to an (a) in country or (b) international laboratory	Capability	New
39	Biosurveillance (or CBR) directly informed decisions (e.g., policy, clinical practice)	Capability	New
40	Supervised demonstration of electronic reporting system use (e.g., EIDSS)	Capability	New
41	Number of cases of internationally reportable pathogens or diseases detected in the country in past 12 months; percentage reported to the appropriate international authority (e.g., WHO, OIE)	Capability	New
42	Regulatory environment criteria and tiers	Sustainability enabler	Planned[c]
43	Degree of consistency between national and CBEP planning and implementation (e.g., CBEP within national plan or vice versa, joint planning and execution of exercises)	Sustainability enabler	New
44	Number of peer-reviewed publications; number of conference papers (abstract, poster, or oral)	Sustainability enabler	Existing[d]
45	Number of institutions with epidemiology core curriculum that meets international (e.g., CDC FETP) standards; number of trainees per year	Sustainability enabler	New
46	Career track for trained epidemiologists	Sustainability enabler	New
47	Operational laboratory certification program	Sustainability enabler	New

[a] Existing MAP input.

[b] Metric the CBEP BRM team proposed for country-level criteria.

[c] Metric the CBEP-led TT recommended metric for the annual report to Congress.

[d] Existing metric the CBEP science team tracks.

may be a biothreat reduction integrating contractor or another on-the-ground partner. Many of the metrics in Table S.2 require the assessment of a subject-matter expert (SME). In the context of metrics related to MAP, the SME is defined to be the CBEP implementer responsible for reporting.

Planned metrics indicated in Table S.2 will be supported by the outputs of two recent CBEP efforts to strengthen evaluation and reporting on CBEP performance. First, the CBEP biorisk management team is considering adding country-level criteria to MAP, many of which would provide the basis for strong metrics. Second, in fall 2013, CBEP participated in a tiger team (TT in the tables) tasked with strengthening the metrics tracked in CTR's annual report to Congress. The tiger team defined five tiers and the criteria associated with partner maturity toward biorisk management and biosurveillance goals. These planned criteria and tiers provide the foundation for the strong metrics we recommend using to assess performance in several areas (see Table S.2).

MAP proved to be a less useful source for data regarding partner biosurveillance capacities, capabilities, and sustainability enablers than for data on biorisk management. This means that implementing the recommended biosurveillance metrics will require more CBEP time and resources and that CBEP program leaders will want to carefully weigh the benefits of robust metrics data against the burden of data collection. We also recognize that different leaders may reach different conclusions about the relative importance of value and burden for specific metrics. While most of our recommended biosurveillance metrics require data that CBEP does not currently collect, our scoring and screening process allowed us to choose ones for which a source of data exists or for which technical resources exist that might provide guidance or assistance in the collection of new biosurveillance metrics data.

The recommended metrics include both qualitative and quantitative indicators of CBEP program performance. Currently, many MAP metrics take inputs from SMEs in the form of checklist responses marked either yes or no ("Are standard operating procedures in place?") or in the form of questions to support tracking of program deliveries ("How many epidemiologists have been trained?" "How much equipment has been delivered?"). However, other MAP inputs require the SME implementer to make substantial qualitative judgments ("Is the behavior of laboratory personnel safe?"). To support more consistent and stable reporting for qualitative questions, the CBEP biorisk management team is currently working to develop scoring criteria for implementers on the ground. Other qualitative metrics we recommend, such as the existence of a career track for epidemiologists, are well defined by other U.S. government agencies tracking their own program performance, in this case by the Centers for Disease Control and Prevention for the purposes of assessing the Field Epidemiology Training Program.

As discussed above, CBEP supports CBR projects as means of supporting overall program goals. While an explicit evaluation framework and metrics for evaluating

the impact of CBR are beyond the scope of this research, the recommended metrics represent elements of CBR's multifaceted role in CBEP. For example, the numbers of conference presentations or peer-reviewed publications are direct outputs of CBR that provide a sustainability enabling metric related to professional development opportunities for trained personnel. This metric is only a proxy measure, but it provides useful insight into the state of the scientific and technical community in which CBEP has invested. Another example is the biosurveillance capacity metric for priority pathogen baseline data. For some partners, CBR projects are the mechanism for establishing the critical baseline data. While these are relatively clear examples of CBR as reflected in the CBEP metrics framework discussed here, they are a subset of the broader network of complex relationships through which CBR supports CBEP program goals.

Additional Metrics for Introduction at a Later Date

In addition to the 47 metrics in Table S.2, we also identified 13 metrics for later implementation (see Table S.3). These additional metrics provide CBEP with opportunities for phased implementation, which will build the robustness of the evaluation framework over time. Four of the metrics recommended for later introduction are intended

Table S.3
Metrics Recommended for Later Introduction

Number	Addition or Replacement	Metric	Type	Existing, Planned, New
Biorisk Management				
1	Replacement for recommended metric 2	SME finds that collections of dangerous pathogens have been consolidated into a minimum number of facilities	Capability	Planned[b]
2	Replacement for recommended metric 9	SME finds that there are mechanisms for oversight, enforcement, and attribution for biosafety and biosecurity regulations and/or guidelines	Capability	Planned[a]
3	Replacement for recommended metric 10	SME finds that there are regulations and/or guidelines for biosafety and biosecurity	Sustainability enabler	Planned[a]
4	Replacement for recommended metric 10	Are there national incident management systems for naturally occurring or intentional biological events? (concerns policies, frameworks, MOUs for the laboratory, health security and law enforcement, and ER sectors)	Sustainability enabler	Planned[a]
5	Addition	Percentage of required equipment that is domestically or regionally sourced	Sustainability enabler	New
6	Addition	SME finds that there is adequate availability of funding to support BS&S programs and initiatives	Sustainability enabler	Planned[a]

Table S.3—Continued

Number	Addition or Replacement	Metric	Type	Existing, Planned, New
Biosurveillance				
7	Addition	Demonstrated performance in surveillance of at least three of five defined core syndromes (U.S. government IHR measure)	Capability	New
8	Addition	Number and percentage of laboratories that meet specified standards (mutually agreed on by country and CBEP)	Capability	New
9	Addition	Laboratory timeliness: number of hours following receipt of specimens that laboratory testing is (a) initiated and (b) completed (from laboratory logs or exercise)	Capability	New
10	Addition	Trained and funded internal staff or active contracts or vendor agreements for life-cycle equipment maintenance	Sustainability enabler	New
11	Addition	Number and percentage of priority pathogens for which country (a) produces or (b) can sustainably procure needed media and reagents	Sustainability enabler	New
12	Addition	Performance under standardized clinical scenarios or clinical preceptor validation (and TT tiers)	Capacity	Planned[b]
13	Addition	Number of specified level jurisdictions, number and percentage with functional community-based surveillance (submit ≥50 percent reports)	Capacity	New

[a] Metric the CBEP BRM team proposed for country-level criteria.

[b] Metric the CBEP-led TT recommended metric for the annual report to Congress.

to replace existing metrics and therefore do not add to the data collection burden. While we scored the validity of these metrics as higher than the ones they will eventually replace, data collection requirements made them somewhat less feasible in the near term than those in the original set. The other nine metrics in Table S.3 are recommended as additions to the original set. While their introduction would increase the overall data collection burden, they represent different dimensions of overall CBEP program impacts, and their inclusion would make the metrics framework more robust.

CBEP leaders will clearly need to make important decisions regarding the metrics recommend here. For example, they will need to carefully consider the desirable number of metrics, balancing the desire for robustness of actionable and strategically important information against the burden of data collection. If CBEP elects to begin with a subset of the recommended metrics, it will then need to select the desired ones and decide when and how to begin implementing them. Nonetheless, this report offers a foundation for such decisions, including concrete information regarding sources

of additional technical resources as well as guidance for scoring our recommended metrics.

Recommended Metrics Relative to Other Approaches to Evaluating CBEP Impact

The development of recommended metrics described here builds on other efforts aimed at strengthening the evaluation of CBEP performance toward program goals. One example of this is the significant work CBEP has committed to designing, aggregating, and maintaining data in MAP. MAP establishes a strong foundation for data collection that will support metrics for assessing several dimensions of program impact. MAP reflects broader U.S. government interest in evaluating the impact of CTR programs. In 2010, DoD developed a set of metrics for CTR, including CBEP, which NAS subsequently evaluated. To discuss metrics development for biorisk management (Chapter Three) and biosurveillance (Chapter Four), we compare our recommended metrics with the NAS framework. NAS made several recommendations for an improved approach to developing, prioritizing, and utilizing metrics for CTR, including CBEP. Within the scope of our tasking from CBEP, our approach addresses many of the NAS recommendations:

- NAS recommended that DoD include a concise statement of each CTR program's objectives and of how each objective is intended to reduce threat or risk. We identify CBEP objectives and explicitly relate them to the overall threat reduction mission. (NAS, 2012, p. 34.)
- NAS also recommended that the United States and a partner country jointly develop objectives for projects in that partner country. We address this by including a "sustainability enabler" category of activity (and measurement) that takes into account partner participation in development of program goals. Notably, our explicit focus on sustainability reinforces an existing concept in CTR program evaluation. NAS noted that most CTR programs work to "develop a sustaining capability," and the report praised DoD for factoring sustainability into its recommended metrics for assessing CBEP performance. (NAS, 2012, p. 4.)
- NAS also recommended using a consistent framework to prioritize and refine metrics. We accomplish this by developing logic models to serve as the foundation for metrics development. Logic models are not one of the two approaches to meeting this goal that NAS explicitly mentioned, but they serve the desired purpose. Our framework also facilitates prioritization through our explicit scoring criteria and recommendations on prioritization through a time-phased implementation. (NAS, 2012, p. 39.)
- With respect to data, NAS recommended that CTR needed independent evaluation of how the capabilities being built perform in practice. We address this recommendation in part by firmly distinguishing between *capacities*—the building blocks CBEP helps a partner country put in place—and the partner's demon-

> strated ability to operationalize desired *capabilities*. With respect to the need for an independent evaluation, we note that current CBEP reliance on metrics data inputs from CBEP implementers would be more reliable if replaced in the longer term by unambiguously objective information, third-party audits, exercises, or demonstrations. (NAS, 2012, p. 47.)

Ultimately, NAS found that the 49 DoD metrics for CBEP exceeded what was desirable for measurement because the longer list, especially in the absence of an approach for prioritization, muddled the task of assessing and communicating program performance. We recommend a similar number of metrics, but as will be discussed in the next section, we also recommend an approach to aggregation for reporting on program performance at several levels. This both enhances clarity and simplicity for decisionmaking at higher levels and enables unpacking into more granular, actionable metrics for program management at lower levels. The number of metrics we recommend for use in the near term (47) is about the same number of overall metrics recommended in the initial DoD report (49), but our approach maintains a strong conceptual framework while improving potential for implementation and use. We also identify additional metrics (nine) recommended for introduction and use at a later date, should CBEP desire opportunities to make the evaluation framework more robust.

Aggregating Metrics to Support Reporting on CBEP Performance at Several Levels

The 47 recommended metrics tell a comprehensive story about the state of biorisk management and biosurveillance in CBEP partner countries. While the set of recommended metrics represents a concise picture selected from a larger set of nearly 300, making sense of a set of nearly 50 metrics is challenging, as NAS noted. The utility of a set of metrics is in part determined by whether it helps answer the question, "What is CBEP accomplishing?" Of course, there is no one answer to this relatively simple question. How it is answered depends on who is asking it and why. When DoD or congressional leadership asks the question, the answer must explain how CBEP contributes to the overall goals of using CTR to make the world a safer place. For CTR program leadership, the answer must explain how activities are changing the threats nuclear, chemical, and biological weapons pose across the globe. For CBEP, the answer must explain what progress is being made on the objectives of biorisk management and biosurveillance. Finally, for country managers, the answer must help them understand where work remains to build capacities, capabilities, and future sustainability.

In this report, we present a formal approach employing a clear measurement hierarchy that can transparently aggregate individual metrics progressively up to more strategic-level assessments. While the framework we describe relies on notional values

we assigned for program goals and thresholds and on notional choices regarding aggregation rules, actual implementation would require policy guidance and program expertise to determine actual (rather than notional) values for these. To the extent to which the proposed metrics can tell all parts of this story, the effort required to collect and organize the metrics will be offset by the benefits from using them to guide decisionmaking at all levels of CTR and CBEP program management.

Conclusions and Recommendations

We developed an evaluation framework, including metrics, for assessing the progress of CBEP activities in partner countries toward achievement of two main program objectives. This effort suggests conclusions in the specific case of CBEP, as well as broader conclusions related to using logic models as a foundation for measurement. We provide recommendations on a phased approach to implementing the framework and recommendations on additional analyses that could help increase the impact of CBEP programming.

Conclusions

This work builds on substantial existing efforts related to measurement that will greatly facilitate implementation. Work on evaluation within CBEP and CTR and by experts outside the U.S. government, such as NAS, on assessments has established a strong foundation for measurement. Because of the data it gathers for MAP, for example, CBEP already collects a significant fraction of the data required for the recommended metrics. To date, however, the existing data and bases for evaluation have been better developed for biorisk management than for biosurveillance. Despite such strides, implementation of a new evaluation framework will require additional time, effort, and resources. CBEP can mitigate this burden by phasing in the implementation or by establishing priorities; our recommendations on such approaches appear in the appendix.

Developing an evaluation framework for CBEP also allowed us to make some broader observations about logic models as a foundation for management. The framework we recommend facilitates the use of the assessments and communications about the program to decisionmakers at several levels, from in-country project managers to central program managers and up to strategic-level managers within and beyond DoD. The proposed evaluation framework and associated metrics put CBEP at the leading edge of DoD's efforts to monitor program performance. This framework can provide CBEP with new and valuable tools for program monitoring. Some observations about the approach include the following:

- **Logic models provide a framework for evaluating and communicating program performance.** Logic models provide CBEP with a systematic means of capturing program activities and outputs in support of enduring and demonstrable partner capabilities. A framework like this supports measurement and reporting on program performance and for communicating a story about CBEP activities and impacts.
- **The functional frameworks for each program objective provide another approach to selecting and communicating metrics information.** As compared to the relatively abstract representation of the logic models, we found that an alternative, functional representation of metrics for each CBEP objective provided a more intuitive means of "telling the story" and facilitating choices about implementation.
- **The key distinction between *capacities* and *capabilities* allows CBEP to ensure that measurement is focused on outputs that are closer to stated goals.** We use these terms to distinguish demonstrated partner operational abilities (capabilities) from the building blocks that enable them (capacities). Distinguishing between capacities and capabilities was one of the elements of the logic models that most resonated with CBEP personnel.
- **Identification of *sustainability enablers* allows CBEP to focus CBEP programming *now* concretely on *future* sustainability.** In addition to distinct metrics for capacities and capabilities, we propose metrics that specifically look at progress toward the longer-term sustainability of CBEP investments in partner nations (sustainability enablers). Sustainability is vital to program success. By specifically highlighting enablers, the framework better equips CBEP to address sustainability explicitly in program activities and track progress, rather than relegating sustainability to the status of a hoped-for byproduct of program activities, to be measured only at the end of the association with a partner nation.
- **Metrics can be used to support management and reporting at several levels, but doing so requires CBEP deliberation and analysis.** One significant virtue of the recommended framework is its adaptability to support reporting to and the use of several audiences. Implementation of an approach, such as the maturity model described in Chapter Six, can help CBEP tailor reporting to the specific needs of a range of decisionmakers, with transparency and accountability. Doing so, however, requires formalizing the maturity model approach. This will require programmatic and policy expertise, deliberation, and analysis to specify the desired standards and appropriate rules for defining hierarchical relationships.

Recommendations

We recommend that CBEP take steps to refine and implement the metrics framework to support internal evaluations and external reporting on program impacts. To refine

the framework by identifying implementation challenges and mitigation strategies, we recommend pilot testing the recommended metrics in a small set of CBEP partner countries. This approach would allow CBEP to identify challenges and opportunities for refinement in advance of widespread implementation. An important element of pilot testing will also be formalizing the structure for aggregating metrics to support strategic-level reporting. The output of this pilot effort will be evaluations of the selected partners with respect to capacities, capabilities, and sustainability enablers based on the recommended evaluation framework and metrics. Another output for CBEP would be having a tried and tested framework and metrics that CBEP can use broadly to assess performance toward program goals.

We also recommend that CBEP consider moving evaluation efforts forward by focusing on assessing program outcomes in areas key to ultimate program success. As CBEP expands its program activities to new geographic regions, it should build on lessons learned from more than a decade of experience partnering with countries in the FSU. This history positions CBEP well to evaluate the effectiveness of programming in, for example, CBR and other elements of health diplomacy activities, as well as the ultimate sustainability of CBEP investments. Evaluations of lessons learned and best practices would position CBEP to support effectiveness by improving alignment of program activities with program objectives.

Acknowledgments

We would like to thank the staff of the Cooperative Biological Engagement Program (CBEP) in the Cooperative Threat Reduction (CTR) Program of the Defense Threat Reduction Agency (DTRA), especially Lance Brooks, CBEP Division Chief, and Jeffrey Fields, International Projects Manager, for their guidance and support. We also extend special thanks to CAPT Randolph Pierson, Will Chapman, Pete Pesenti, Carl Newman, and LCDR Adam Samuels, who generously shared their time and insights throughout the course of our study. Our products are much improved as a result of their contributions. Finally, the authors would like to thank Rick Eden and Christopher Nelson for their careful and thoughtful reviews and Director Cynthia Cook and Associate Director Paul Deluca of the RAND Acquisition Policy Center for their guidance, feedback, and support.

The views expressed herein are our own and do not necessarily represent the policy of CBEP.

Abbreviations

AFHSC	Armed Forces Health Surveillance Center
APHL	Association of Public Health Laboratories
BMBL	*Biosafety in Microbiological and Biomedical Laboratories*
BRM	biorisk management (in tables only)
BS&S	biosafety and biosecurity
BTRIC	biological threat reduction integrating contractor
CBEP	Cooperative Biological Engagement Program
CBR	cooperative biological research
CDC	Centers for Disease Control and Prevention
CEN	European Committee for Standardization
CSTE	Council of State and Territorial Epidemiologists
CTR	Cooperative Threat Reduction
DoD	Department of Defense
DTRA	Defense Threat Reduction Agency
EDP	especially dangerous pathogents
EIDSS	Electronic Integrated Disease Surveillance System
ER	emergency response
FETP	Field Epidemiology Training Programs (in tables only)
FOC	full operational capability
FSU	former Soviet Union
FY	fiscal year

GAO	Government Accountability Office
IHR	International Health Regulations
IOC	initial operational capability
ISO	International Organization for Standardization
JCI	Joint Commission International
MAP	Metrics Assessment Program
MOH	Ministry of Health
MOU	memorandum of understanding
NAS	National Academy of Sciences
NDAA	National Defense Authorization Act
NDRI	RAND National Defense Research Institute
NRC	National Research Council
OIE	World Organization for Animal Health
OSD-ATL	Office for the Secretary of Defense for Acquisition Technology and Logistics
OSD-P	Office for the Secretary of Defense for Policy
SME	subject-matter expert
SOP	standard operating procedure
TEPHINET	Training Programs in Epidemiology and Public Health Interventions Network
TT	tiger team (in tables only)
WHO	World Health Organization

CHAPTER ONE

Introduction and Background

The Cooperative Biological Engagement Program (CBEP) within the Department of Defense (DoD) is part of the larger Cooperative Threat Reduction (CTR) program that was initiated at the end of the Cold War in the former Soviet Union (FSU) but now has expanded beyond that legacy to address current biological threats of mutual interest to the United States and 20 partner countries in different regions around the world. In recent years, the biological mission has become an increasingly significant component of DoD's CTR program, and according to DoD's plans for CTR, this emphasis is projected to grow. As CBEP has evolved since its inception two decades ago, so too have the content and approaches for performance measurement. Congress has shown particular interest in encouraging DoD to improve assessments for CTR. Legislation for fiscal year (FY) 2010, for example, called on DoD to develop metrics for assessing CTR programs and on the National Academy of Sciences (NAS) to assess DoD's efforts. CBEP leaders approached the RAND National Defense Research Institute (NDRI) to build on existing work related to measurement of program performance by developing a evaluation framework and recommending metrics for measuring CBEP performance and progress toward program objectives. This report details the work we carried out between February and October 2013, including our methods, findings, recommendations, and guidance for implementing and aggregating the recommended metrics.

Background

CBEP initially focused narrowly on partnering with countries in the FSU to reduce the significant threat that existing biological weapon–related expertise, materials, and infrastructure posed. Today, CBEP partners with a broader range of countries in different regions around the world to implement biosafety and biosecurity measures for facilities housing pathogens of security concern and to enhance partner ability to quickly detect, diagnose, and report diseases with implications for international security.

Despite being a signatory to the Biological and Toxin Weapon Convention in 1972, the Soviet Union covertly developed the largest offensive biological weapon program in the world (Government Accountability Office [GAO], 2000a). In 1992, Russian President Boris Yeltsin acknowledged the existence of the biological weapon program and vowed to end it (Cook and Woolf, 2002). Under CTR authorities, DoD has supported a variety of efforts related to biological threat reduction. In 1995, the Defense Special Weapons Agency, predecessor to the Defense Threat Reduction Agency (DTRA), began to work to dismantle a biological weapon production facility in Kazakhstan. In 1998, the newly stood up DTRA established the Biological Weapons Proliferation Prevention, with authorities to carry out biosafety, biosecurity, and cooperative biological research (CBR) projects in the FSU. In 2002, Biological Weapons Proliferation Prevention began to support activities aimed at enhancing partner biosurveillance capabilities through Threat Agent Detection and Response Program projects. In 2006, Biological Weapons Proliferation Prevention became the Biological Threat Reduction Program, responsible for partnering with about six countries in the FSU on biosafety, biosecurity, CBR, and biosurveillance projects (DTRA). In 2010, the program was renamed CBEP to reflect a shift toward new approaches to partnering, as well as the broader range of security challenges the program sought to address.

Biosafety and Biosecurity

Notwithstanding the evolution of CBEP since its origins in the 1990s, several features of some partner countries remain relevant today. For example, the inherently dual-use nature of many biological-related activities in former Soviet countries means that not all facilities housing dangerous pathogens are slated for elimination. Many facilities support important national functions, including research, diagnostic testing, and production of agricultural products. At such facilities, CBEP activities seek to manage biological risks more effectively by reducing collections of pathogens to a minimum and by enhancing facility safety and security at the facilities that remain. CTR's biosafety and biosecurity activities have, for example, enhanced the physical security of bacterial and viral research institutes and animal vaccine production facilities and enhanced personnel access control by supporting installation of alarm systems and guards (Voronova-Abrams, 2011).

Cooperative Biological Research

The threats associated with biological materials include not just the physical materials and infrastructure but potentially also the expertise of scientists and technical personnel. CTR addresses the dual-use nature of biological expertise by supporting opportunities for collaborative research between U.S. and partner-nation scientists and technical personnel. The National Research Council (NRC), 2007, notes that "[i]t is the human dimension of a nation's infrastructure . . . that is the critical determinant in a nation's effort to control dual-use assets" and calls for increased levels of engage-

ment with partner-nation scientists. The intent of CBR funding for research opportunities and peaceful applications of technical expertise includes removing incentives for intentional misuse of biological materials, infrastructure, or expertise and helping resource partner-nation scientists and laboratories, supporting the longer-term sustainability of CBEP investments. However, as the has program expanded its activities to regions facing different security challenges and begun partnering with countries with varying levels of existing capabilities, CBR's role in CBEP has also evolved. In 2012, the CBEP science team released a research strategy to guide how CBR will be planned, coordinated, and executed in this new context. It detailed four distinct goals for CBR, relevant across all partner countries (Pesenti, 2013):

- support biosurveillance and biosafety and biosecurity capacity-building efforts
- execute local projects of regional importance
- promote the One Health initiative (emphasize the nexus between human and animal health)
- foster an international culture of responsible and ethical conduct in biological research.

Biosurveillance

In addition to the control of dangerous materials, CBEP also supports efforts to enhance partner capabilities to conduct effective biosurveillance, i.e., to detect, diagnose, and respond to naturally occurring or intentionally caused outbreaks, pandemics, and other biological threats. CBEP supports training in field epidemiology and equipment and training related to diagnostic laboratory testing and electronic reporting mechanisms. The international framework underlying commitments to global health security is the revised International Heath Regulations (IHR), which were adopted in 2005 and entered into force in 2007 (World Health Organization [WHO], 2005; Miller and Dowell, 2012). In 2010, DoD reported in its annual report to Congress that one desired goal for CBEP was to have partners capable of complying with international requirements, standards, and guidelines, such as the IHR (DoD, 2010a). The purpose and scope of the IHR (WHO, 2005) is to

> prevent, protect against, control and provide a public health response to the international spread of disease in ways that are commensurate with and restricted to public health risks, and which avoid unnecessary interference with international traffic and trade. (WHO, 2005)

The 2005 IHR requires member states to detect, assess, and notify WHO of all events that may constitute public health emergencies of international concern, as defined in the document. The IHR also requires signatories to develop minimum core capacities to meet its obligations to prevent, protect against, control, respond to, and notify WHO of emergencies (WHO, 2005). The requirements under the IHR have provided

a significant input, shaping CBEP program activities, especially in support of enhancing partner biosurveillance capabilities. Chapter Four will consider the consequences of IHR for CBEP program activities and program assessment in more detail.

Expansion of CBEP Beyond the FSU

In the FY 2002 National Defense Authorization Act (NDAA) (Public Law 107-107, 2001), Congress asked DoD to consider development of "cooperative threat reduction programs with India and Pakistan" (NAS, 2009). In 2004, Congress explicitly authorized DoD to use CTR funds (up to $50 million) outside the FSU but made no specific suggestions on how or where. With language in the 2008 NDAA (Public Law 110-181, 2008), however, Congress made the path to engagement beyond the FSU clearer, stating that "the CTR model should be expanded to address threats beyond the states of the former Soviet Union" (Committee on Armed Services United States Senate, 2007; Nikitin and Woolf, 2013). CBEP has since expanded to new regions that are defined by very different threat environments and have different existing capabilities and economic environments. For each major command, the following are the countries in which CBEP has current or planned engagements in the near term (CBEP, 2013c):

- U.S. Pacific Command
  - Vietnam
  - India
  - Malaysia
  - Laos
  - Cambodia
  - Philippines
  - Indonesia
- U.S. Africa Command
  - Djibouti
  - Kenya, South Africa
  - Tanzania
  - Uganda
- U.S. Central Command
  - Afghanistan
  - Iraq
  - Pakistan
  - Kazakhstan
  - Uzbekistan
- U.S. European Command
  - Armenia
  - Azerbaijan
  - Georgia

– Russia
– Ukraine.

The partners with the longest history of engagement with CBEP in the FSU include Ukraine, Russia, Kazakhstan, Uzbekistan, Georgia, and Azerbaijan; the "newer engagement" countries include partners in U.S. Pacific Command, U.S. Africa Command, those other than Kazakhstan in U.S. Central Command, and Armenia in U.S. European Command. CBEP projects that, by 2015, the overall commitment of program funding for partners outside the FSU will exceed that for partners in it. The future of CBEP will be in regions outside the FSU. Moreover, the focus of CTR activities is likely to be less on reducing the threat state-sponsored weapon programs pose and more on engagement to address security concerns associated with threats with "smaller footprints, less distinct signatures, and more closely associated with industrial activities related to energy, biology, health, or chemistry" (NAS, 2009).

Recent Interest in Improved Metrics for Cooperative Threat Reduction

In recent years, decisionmakers at many levels have focused on the need to improve assessments of CTR programs, especially as they expand in scope and geographic reach. In 2009, the NAS noted that new approaches for cooperative threat reduction would require new approaches for evaluating performance. The metrics used to assess CTR in the FSU would not necessarily be the right approach for measuring the impact of the program in new partner countries.

The NDAA for FY 2010 (Public Law 111-84, 2009) called on DoD to "develop and implement metrics to measure the impact and effectiveness of activities of the Cooperative Threat Reduction Program," and it also called on the NAS to provide an assessment of the DoD metrics development effort. The resulting DoD metrics report (NAS, 2012, Appendix B) and NAS assessment (NAS, 2012) broke new ground for assessing progress toward the biological threat reduction mission, and significantly shaped our approach. DoD recommended 49 metrics for assessing CBEP, derived from a systematic approach which defined program objectives, program activities and desired partner capabilities, measures, metrics, and evaluation criteria (NAS, 2012, Appendix B). NAS praised DoD for the strength of the CBEP metrics, as having "nearly all of the elements that the committee thinks are needed for development of useful metrics" (NAS, 2012). The NAS praise was notable because NAS also noted that, relative to other CTR efforts, CBEP faces unique challenges for measurement. "CBEP's work resides in a gray mission space that has elements that overlap with public health activities," the NAS noted. It further noted that

> [i]mpact, effectiveness, and success are difficult to measure in such efforts—there are no simple parallel metrics to counting the number of delivery vehicles destroyed

> or the fraction of weapon-useable nuclear material secured or eliminated that can effectively measure the impact of any complex capacity building program, such as CBEP. (NAS, 2012)

Traditional CTR metrics focusing on quantifiable indicators, such as warheads deactivated, missile silos eliminated, and mobile missile launchers destroyed, would be insufficient for capturing the impact of "relationships and processes" central to new approaches to engaging partners (NAS, 2009).

The NAS identified several opportunities for improving the consistency, clarity, and utility of DoD's metrics framework. For example, the NAS noted that the objectives were not related consistently to the threat reduction mission; program activities were sometimes confused with desired partner capabilities; the metrics framework lacked a mechanism for implementation that took into account prioritization; and there were too many metrics. Within the scope of our tasking, our approach and recommendations specifically address many of the NAS recommendations, which will be detailed in Chapters Four and Five and summarized in Chapter Seven.

What RAND Was Tasked to Do

CBEP tasked NDRI with addressing three key questions related to measurement of program performance.

- **What should CBEP measure to evaluate program performance?** This question focused on identifying CBEP program objectives. For this, we reviewed international, U.S. governmental, DoD, and CBEP guidance documentation related to biological engagement, and we engaged top stakeholders in discussions related to CBEP program objectives. This included representatives from CTR leadership, CBEP leadership, the Office for the Secretary of Defense (OSD) for Policy (OSD-P), and the OSD for Acquisition Technology and Logistics (OSD-ATL) Nuclear, Chemical, and Biological Defense Programs to fully capture diverse perspectives related to overall program goals.
- **How should CBEP systematically identify metrics?** CBEP asked us to develop conceptual frameworks for systematically capturing program goals, objectives, and program activities, which would serve as the foundation for metrics development.
- **What metrics do we recommend for CBEP?** CBEP asked us to recommend a set of metrics, grounded in the conceptual frameworks and selected to best meet the criteria that will help ensure appropriate targeting and robustness in the measurement scheme. CBEP called on RAND to draw on and leverage as appropriate previous efforts within DoD and from nongovernmental organizations, such as NAS, that have contributed to the development of a significant body of existing work relating to metrics for assessing CBEP.

Report Overview

This report describes our methodology and approach, findings, recommendations, and suggestions on the use and interpretation of metrics. We used several approaches to collect information related to the project tasks, primarily document review and semi-structured discussions with subject-matter experts (SMEs) to develop, iterate on, and vet interim products. The chapters that follow described the use of these methods to support project tasks.

Chapter Two details our overall study approach, including the development of logic models and their employment as a foundation for measurement.

Chapters Three and Four are parallel discussions of the specific application of our study approach to biorisk management (Chapter Three) and biosurveillance (Chapter Four). These chapters discuss definitions and terms, the development of logic models, and the identification of recommended metrics in detail. We present the recommended metrics in different ways and recommend which could be used now or phased in later. In each chapter, we also compare the metrics we recommend with those from other current or previously recommended measurement frameworks.

Chapter Five discusses issues related to implementation of the recommended metrics framework. This discussion is tightly linked to the implementation information appearing in Appendixes A and B.

Chapter Six discusses an approach to using the logic model frameworks and metrics to support program-level management progressively up the chain to strategic-level reporting on CBEP performance. Using an existing RAND tool, the Portfolio Assessment Tool, we demonstrate a rules-based hierarchical approach for using recommended metrics, combined with the logic models, to support aggregation for reporting on program performance from country-level to progressively higher-level decisionmakers at the program, DoD strategic, and non-DoD federal levels.

Finally, Chapter Seven offers concluding comments for CBEP as it seeks to build on this research to improve measurement.

CHAPTER TWO

Using a Conceptual Framework to Identify Metrics for CBEP

This chapter describes our approach to using a conceptual framework to develop metrics to monitor CBEP program performance. At each level of management, the demand for metrics is motivated by the need to explain how actions lead to anticipated results. The development and application of logic models underpinned our efforts, although alternative approaches could also have achieved this goal. The logic model approach establishes a logical sequence from program inputs to activities; outputs; outcomes; and, ultimately, impacts. The logic model approach also allows the framework to be used to support aggregation of metrics as needed for decisionmaking at several levels, from country to central program management to more-strategic levels within and beyond DoD. Our approach to using logic models to identify metrics for CBEP included three key steps. First, we identified CBEP program objectives. These objectives represent the impact that CBEP programming seeks to support. Second, we described our development of logic models to systematically link program objectives to specific program outputs and partner operational abilities. This framework provides a conceptual framework for measuring progress toward these objectives. Third, we used the logic models to identify a set of recommended metrics that allow monitoring of CBEP performance. This chapter details the first step and our general approach to the second and third steps; Chapters Three and Four address the latter two steps in more detail for the two CBEP program objectives.

Identifying Program Objectives

We identified CBEP program objectives by reviewing national-level guidance and by holding structured discussions with top stakeholders in CBEP, OSD-ATL, OSD-P, and CTR. For these discussions, we employed an interview protocol that included questions related to formal and informal CBEP program objectives; objectives that may be distinct, complementary, or redundant with other U.S. government programs; regional variations in program objectives; the role of the partner nation in shaping CBEP programmatic activities or objectives; and sustainability as it relates to program objectives. Engagement with a diverse set of stakeholders with varied responsibilities

for CBEP facilitated identification of CBEP objectives and highlighted key issues for resolution.

We framed our evaluation frameworks around two objectives for CBEP. Detailed discussions of the definitions of these objectives will appear in Chapters Four and Five; here, we describe the process of framing the CBEP objectives. The objectives we identified include strengthening enduring partner capabilities in two areas[1]:

- **biorisk management**: Identify, consolidate, and secure collection of pathogens and diseases of security concern to prevent the sale, theft, diversion, or accidental release of pathogens and diseases. Biorisk management includes biosafety and biosecurity control measures, as well as risk assessment, oversight, and review activities. We consider two distinct subobjectives:
  - **biosafety**: Prevent the unintentional exposure to biological agents and toxins, or their accidental release. Our definition for this subobjective includes partner capability to control unintentional releases into communities, animal populations and the environment (Centers for Disease Control and Prevention [CDC], 2009).
  - **laboratory biosecurity**: Prevent loss, theft, diversion, or intentional misuse of microbial agents (CDC, 2009). Our definition for this subobjective includes the partner's ability to "[d]ismantle, destroy, and prevent the sale, theft, diversion, or use of stockpiles of biological weapons, means of delivery, and [biological-weapon]–related equipment, technology, and infrastructure" (Creedon and Weber, 2013, p. 3).
- **biosurveillance**: "Enhance partner country/region's capability to rapidly and accurately survey, detect, diagnose, and report biological terrorism and outbreaks of pathogens and diseases of security concern in accordance with international reporting requirements" (Creedon and Weber, 2013, p. 3).

The objectives we identified share certain important characteristics. First, both objectives describe strengthening partner *capability* as the program objective. This explicit choice highlights the key distinction we maintain between building *capacities* and building *capabilities*. We will discuss this distinction in greater detail in our discussion of logic models, but in short, we define *capacities* to be program outputs that provide the building blocks of partner biorisk management and biosurveillance programs (e.g., having in place plans, trained people, equipment, and facilities), while a *capability* is the ability to operationalize the programs (e.g., respond to exposures in the laboratory, successfully conduct a diagnostic test).

Second, both objectives make explicit reference to *sustainability* as an integral component of program objectives. Our recommendation that supporting partner sus-

[1] Our framing is consistent with that of Creedon and Weber, 2013, p. 3.

tainment of CBEP investments be considered a key component of program objectives, based on examination of relevant literature and our experience-based insights, diverged from several other evaluation frameworks. A 2013 CBEP program overview highlighted the importance of sustainability by including it as an independent program objective: "Ensure the developed capabilities are designed to be sustainable" (CBEP, 2013a). Sustainability is critical for investments in health system development in countries around the world but remains more of an aspiration than an established operational agenda. Sustainability is important to both donors and the recipients of their investments and is relevant across a wide range of sectors, including health. The world is now experiencing both the challenges and opportunities of globalization and the realities of an austere worldwide fiscal environment. The U.S. government supports health system development and international cooperation through programming spread across multiple federal agencies, including DoD. Sustainability is an explicit priority of the U.S. Global Health Initiative. It is more important than ever to invest intelligently in both current and future U.S. initiatives to produce results that endure. DoD can strongly support U.S. government efforts to advance a sustainability agenda through both empirical analysis and action. Specifically, through this new element of CBEP's evaluation framework, CBEP can assess progress toward the sustainability of its efforts while also testing and refining new metrics that reflect presumed sustainability enablers. We determined that embedding sustainability within other objectives, rather than recommending it as a separate objective, better supported its integration into both programming and an evaluation framework for measurement.

Our framing of CBEP objectives diverges in some notable ways from the framing of objectives in other existing documentation. For example, we adopted the term *biorisk management*, with two subobjectives, as the primary CBEP objective, rather than commonly used *biosafety and biosecurity*. The more inclusive term, *biorisk management*, includes biosafety and biosecurity control measures, as well as risk assessment, oversight, and review activities (Stroot and Jenal, 2011b). Our definition also includes under "biosecurity" several earlier objectives closely associated with bioweapons infrastructure. Other constructs include stand-alone objectives such as secure and consolidate collections of pathogens of security concern and into a minimum number of secure facilities, and elimination of bioweapon-related infrastructure and technologies (NAS, 2012; DoD, 2010a). CBEP's strategic policy guidance (Creedon and Weber, 2013) includes "[d]ismantle, destroy, and prevent the sale, theft, diversion, or use of stockpiles of biological weapons, means of delivery, and biological weapon–related equipment, technology, and infrastructure" as a separate objective. We capture this important CBEP objective under our framework for biosecurity.

Cooperative Biological Research

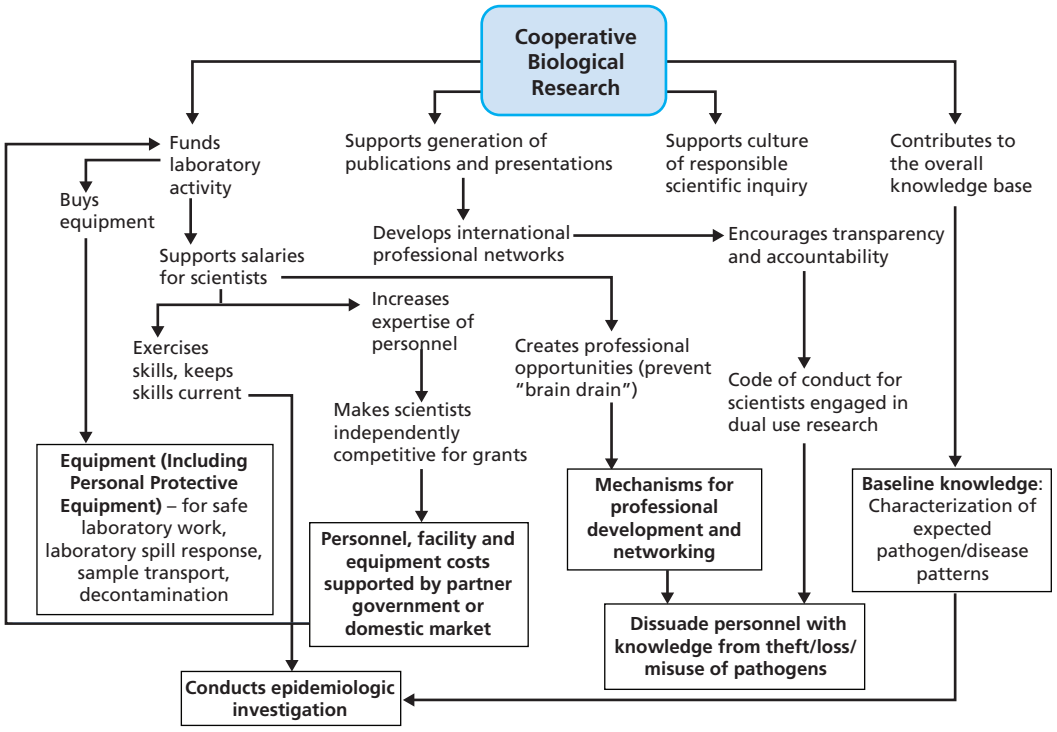
Perhaps the most widely cited CBEP effort not directly included in our list of CBEP objectives is CBR. For example, existing documentation refers to CBR as a category of

activities (NRC, 2007), a product line, a supplemental objective (Nacht, 2009; NAS, 2012), or a CBEP objective (CBEP, 2013a). The varied descriptions of CBR's role in CBEP reflects its long history as a flexible tool for biological engagement and its continuing central role in CBEP programming and longer-term sustainability.

We assessed CBR to be an activity that supports both CBEP program objectives but that is not appropriate to consider a program objective in itself. By supporting research activities in partner nations, CBR supports broader program objectives through such mechanisms as generating resources for partner-nation scientists and institutions, supporting the development of a culture of responsible scientific practice in partner nations, and generating useful knowledge that supports other programmatic objectives. Specific examples of relationships between CBR activities and CBEP program outputs appear in Figure 2.1.

Because we consider CBR to be an activity that supports CBEP objectives and not an objective in itself, we did not develop a conceptual model for evaluating CBR activities. However, CBR does appear indirectly in our models for biorisk management

Figure 2.1
Examples of the Relationship Between CBR and CBEP Program Activities, Program Outputs, and Partner-Nation Activities



RAND *RR660-2.1*

and biosurveillance: The six boxed elements at the bottom of Figure 2.1 are elements of our CBEP logic models, as we will discuss in Chapters Three and Four.

Developing Logic Models for Systematically Identifying Metrics

Decades of research in program evaluation have provided alternative approaches for designing systematic measurement frameworks and selecting metrics derived from them (Rogers et al., 2000). For example,

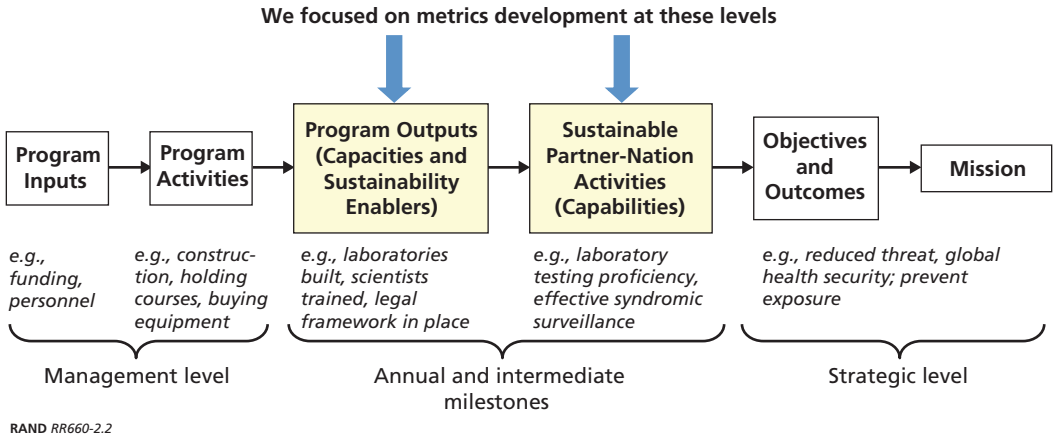
- Describe a "chain of objectives" when evaluating public service programs (Suchman, 1967).
- Use *outcome hierarchies* to evaluate education programs (Bennett, 1975).
- Use *theory of action* to evaluate interventional programs (Schön, 1997).
- Use *program theory* as a framework for program evaluation (Bickman, 1987).
- Use *theory-based evaluation* to evaluate public policy (Weiss, 1997).
- Discuss *program logic* to evaluate the effectiveness of public services (Lenne and Cleland, 1987).

While none of these approaches is uniquely superior to another, they share a common perspective on what constitutes a good framework for selecting metrics. When done well, the metrics will tell a coherent and complete story about what is being accomplished. Even better, if managers responsible for collecting metrics can use the data to make decisions themselves, the effort to put toward measurement may be seen as valuable above and beyond the value of reporting the metrics to others.

Consistent with these criteria and the existing literature on program evaluation, we adapted the approach Greenfield, Williams, and Eiseman (2006) used for using logic models to guide strategic planning and evaluation in response to the Government Performance and Results Act. This approach for using logic models organizes metrics in the context of how activities and decisionmaking at the management level support meeting goals defined in annual program planning, which in turn supports meeting strategic-level objectives and outcomes. Figure 2.2 presents how this logic model framework applies to CBEP.

At the most basic level, CBEP provides people, funding, and equipment. These program inputs are used to carry out a variety of program activities. For example, CBEP projects construct new facilities, hold courses to train scientists, or buy equipment to be used in laboratories. Logically, the *program outputs* for CBEP are connected to these activities and could be tracked, for example, such in terms as laboratories constructed or scientists trained. However, program outputs are not an end to themselves. Rather, program outputs are chosen based on how they build partner-nation operational capabilities related to CBEP objectives. Examples of operational capabilities could include

Figure 2.2
Logic Model Elements



proficient laboratory testing capabilities or effective syndromic surveillance capabilities. In turn, CBEP objectives are chosen to help the nation achieve national strategies such as reducing threats from weapons of mass destruction or improving global health security.

Elements of Our Logic Models: Enduring Capabilities, Capacities, Sustainability Enablers

Figure 2.2 provides a simplified schematic for how logic models can relate resource inputs on the left to national-level objectives and missions on the right. Such models can describe relationships at several levels of analysis. In accordance with sponsor priorities for measurement, our logic models focus on the level of program outputs and partner-nation operational abilities. CBEP leadership felt that the existing Metrics Assessment Program (MAP) does a good job of tracking progress in terms of program outputs and activities (tracking training events, equipment deliveries) and budget justification (tracking dollars and manpower) and that assessments of overall threat reduction or global health security are the purview of other U.S. government efforts.

At this level of analysis, we built on the Greenfield, Williams, and Eiseman (2006) approach to emphasize three types of program outputs that other logic model frameworks did not capture clearly: capacities, sustainability enablers, and enduring capabilities.

- *Capacities* are the building blocks required to enable partner capabilities.
- *Sustainability enablers* are building blocks required to maintain partner capacities and capabilities in the longer term.
- *Enduring capabilities* are demonstrated and sustainable partner operational abilities.

These are the ultimate outputs of CBEP program activities, which in turn lead to strategic-level outcomes and mission impacts; CBEP helps create or strengthen partner capacities and implements partner sustainability enablers to enable enduring partner capabilities.

Our distinction between these three elements in our evaluation framework allows CBEP to focus measurement on outputs that are closer to desired impacts. While CBEP currently has strong mechanisms for collecting data on partner capacities, such as people trained, equipment delivered, or standard operating procedures in place—especially with respect to biorisk management—program leadership knows that such measurement should not be taken as a proxy for partner operational ability (capabilities). Distinguishing between capacities and capabilities allows CBEP to monitor both these critical program outputs. Likewise, identification of specific sustainability enablers gives CBEP a framework to actively support the longer-term sustainability of investments in partner nations, rather than assuming that a partner will be able and willing to sustain capacities and capabilities once CBEP scales back engagement activities.

Identifying Recommended Metrics

Our approach to metrics development consisted of six steps. First, we conducted an environmental scan of existing metrics used or suggested by CBEP or others. Second, we mapped these metrics to the relevant element of the logic model (aiming to map each to just one logic model element, but occasionally mapping to more than one). Second, we grouped the metrics according to logic model element, so that we could more easily compare different metrics addressing the same logic model component. This process also helped identify logic model elements for which there were few, or even no, relevant metrics. Third, we scored each metric based on five criteria. Fourth, based on the scoring, we proposed a set of metrics and identified critical gaps. Fifth, we consulted with CBEP to review potential metrics, particularly with respect to completeness and practicality. Sixth, based on these consultations, we revised and documented our metrics recommendations. These six steps are described in more detail in the three sections that follow.

Environmental Scan and Alignment to Logic Models

We reviewed a broad body of literature related to measurement for biorisk management and biosurveillance. The review included documents from national-level guidance, non–U.S. government guidance, and CBEP-specific approaches to measurement. Several of these documents provided actual metrics, but in most cases, they discussed only the elements of a sound biorisk management and/or biosurveillance program. In such cases, we developed our own language to capture a metric the literature sug-

gested. Our list of recommended metrics clearly distinguishes between metrics derived directly from existing literature and those for which we created the wording. This environmental scan yielded 289 candidate CBEP metrics related to biorisk management and biosurveillance.

After translating an environmental scan into potential metrics, we mapped all metrics to the relevant logic models.

Scoring Metrics and Identification of Gaps

Our scoring took into account the strength of a potential metric along several dimensions. We scored metrics by applying five criteria:

- Validity: To what extent does metric capture the concept being addressed?
- Feasibility: How easy is it to collect the data needed?
- Utility: How useful is the metric, for practical purposes?
- Reliability: How consistent would measurement be?
- Relevance: To what extent does the metric reflect CBEP programming?

These criteria helped us achieve several objectives with respect to identifying recommended metrics. The validity criterion helped us reduce the possibility that program assessment motivates activities that are evaluated well but do not in fact achieve program goals. This form of validity is necessary for measurement but is not sufficient. The metrics must also be reliable, so that evaluators can confidently make comparisons across countries and over time as engagement with partner countries progresses. Program managers must also be able to apply the metrics within the constraints of the time, knowledge, and people available to conduct assessments. Thus, the recommended set of metrics will reflect a review of available data and reactions of CBEP program managers to the feasibility of their collection and the practicality of their use.

We developed a five-point scale from low to high (low, medium-low, medium, medium-high, and high) and applied it to each of our 289 candidate metrics. Three of the five scoring levels we employed appear in Table 2.1. We color-coded each rating cell and then visually organized the potential metrics from highest to lowest within each logic model element (i.e., within each specific capability, capacity, and sustainability enabler). We then selected all metrics that scored at least medium on all five criteria, as well as the highest-scoring metric for the logic model components for which no high-scoring metric emerged. This process resulted in an array of 73 potential metrics grouped by logic model element and loosely organized from highest to lowest rating.

This scoring method and four of the criteria we used are consistent with the guidance from GAO on implementing performance measurement.[2] The fifth criterion, "relevance to CBEP" was perhaps the one most specifically tailored to evaluating CBEP.

[2] See, for example, GAO, 1998.

Table 2.1
Scoring Criteria

Criterion	Definition	High	Medium	Low
Validity	To what extent does the measure capture the concept being addressed?	Direct measure of concept being addressed	Proxy or closely related to concept being addressed	Proxy relationship to concept; indirectly related to concept being addressed
Feasibility	How feasible would it be to collect the data needed?	Data currently collected in all CBEP countries by CBEP	Data not currently collected by CBEP, but a practical mechanism exists for collection	Data not currently collected by CBEP, and the mechanism for collection is unclear
Utility	How useful is the measure for decisionmaking?	Extremely useful for decisionmaking at the strategic, program management, and/or country management level	Somewhat useful for decisionmaking	Not at all useful for decisionmaking
Reliability	How consistent would the measure be?	Quantitative, well defined, stable	Qualitative, well defined, stable or Quantitative, poorly defined, not stable	Qualitative, judgment dependent, anecdotal
Relevance	To what extend does the measure reflect CBEP-specific programming?	Directly relevant to an explicit CBEP program element or activity	Closely but indirectly relevant to an explicit CBEP program element or activity	Not directly or closely but indirectly relevant to CBEP

CBEP is not the only U.S. government program with objectives related to biorisk management or biosurveillance. It is, however, a program uniquely oriented toward DoD's threat reduction mission. The relevance criterion allowed us to identify metrics aligned with CBEP's unique role relative to the other U.S. government programs.

While the scoring of potential metrics was done largely independently, rather than in concert with multiple team members, we took certain steps to improve the consistency and reliability of scoring. For example, at the outset of the scoring exercise, we did, as a team, apply our scoring criteria to a sample of metrics. This exercise allowed us to identify differences in interpretation and application of the criteria and to discuss options for improved consistency. A second check on the application of our scoring was the process of vetting potential metrics through discussion with CBEP personnel, as described below. These discussions allowed CBEP to challenge us to justify our scoring and provided us additional information and insights to help us refine the initial scoring.

Vetting Metrics Emerging from Initial Screen and Final Recommendations

After our initial scoring and screening, we next met with key personnel to vet potential metrics, especially with respect to practicality and completeness. We held interim discussions on metrics with CBEP principals, country managers, science leads, the biorisk management team, and the CTR Advisory and Assistance Services team. These SMEs provided valuable insights into proposed approaches to measurement that did not emerge from our document review. For example, the experts detailed initial operational capability (IOC) and full operational capability (FOC) partner demonstrations, allowing us to refine our understanding of the opportunities and limitations of data currently available to CBEP to support metrics. We also developed a stronger understanding of the process for reporting and validating inputs to MAP, which was emerging as a key data source for biorisk management metrics. While our initial scoring scored MAP inputs that required significant discretion as low on reliability, the discussion with CBEP personnel provided additional information regarding the process, allowing us to recommend such metrics. We accompanied such recommendations with suggestions on how to further strengthen the metrics. The discussions were more than just an opportunity to glean the insights of experts knowledgeable about CBEP. The country teams and science team also provided insights as intended consumers of the recommended metrics. The evaluation framework and recommended metrics are also intended to help CBEP implementers (including country and science teams) track CBEP programming they oversee in their regional areas of responsibility.

The initial list we brought to these discussions consisted of potentially promising metrics—ones that scored relatively high across all criteria or that were strong enough in some areas to warrant follow-on conversations with SMEs. As noted earlier, we considered a metric potentially promising if it scored at least medium across all five criteria.

After these discussions, we refined the wordings for certain metrics and made choices among others to further pare down the total. We used the validity and feasibility criteria, in particular, to further winnow the list. This resulted in a final list of 50 recommend metrics, 24 for biorisk management and 26 for biosurveillance. However, a large number of recommended metrics rely on data CBEP already collects (18 for biorisk management, seven for biosurveillance).

We also considered which metrics could (and perhaps should) be collected beginning in the near term and which could be phased in later. We identified 14 metrics for introduction later as part of a phased implementation. Chapters Four and Five walk through the metric development process for each CBEP objective in greater detail. They also present alternative approaches to visualizing metrics and for implementing a metric framework (i.e., prioritization, phased implementation).

CHAPTER THREE

Conceptual Model and Recommended Metrics for Biorisk Management

Overview

A primary objective of CBEP is to support the safety and security of facilities that house pathogens of security concern by enhancing sustainable partner capability for biorisk management.[1] CBEP defines *biorisk management* as the management of risks of harm arising from biological agents. A biorisk is a combination of the probability of occurrence of harm and the severity of that harm where the source of harm is a biological agent or toxin. The source of harm may be an unintentional exposure, accidental release or "loss, theft, misuse, diversion of, unauthorized access or intentional unauthorized release" (CBEP, undated). Biorisks of the former type, threats due to unintentional release of pathogens, are addressed through biosafety practices. Laboratory biosecurity activities address threats of the latter type—pathogens released intentionally.[2] *Biorisk management* has in recent years been favored over *biosafety and biosecurity* to describe not only the "implementation of biosafety and biosecurity control measures, but also [to] address their oversight and review, as well as the appropriate allocation of responsibilities and resources" (Stroot and Jenal, 2011a). CBEP defines an organization's biorisk management system as

> the part of an organization's management system used to develop and implement its biorisk policy and manage its biorisks. A management system is a set of inter-

[1] CBEP's biorisk management activities are guided primarily by WHO, 2006a; DoD Instruction 5210.89, 2006; and CBEP, 2011.

[2] CBEP adaptation of CEN Workshop Agreement 15793: "**Biosafety:** laboratory biosafety describes the containment principles (facilities), technologies (equipment) and practices that are implemented to prevent the unintentional exposure to biological agents and toxins, or their accidental release"; "**Biosecurity:** laboratory biosecurity describes the protection, control and accountability for biological agents and toxins within laboratories, to prevent their loss, theft, misuse, diversion of, unauthorized access or intentional unauthorized release. In this context, biosecurity is restricted to laboratory biosecurity; laboratory includes bioscience facilities involved in diagnostic testing, animal research, vaccine production, academic research, and manufacturing/industry, and does not include all aspects of biosecurity in the sense of national or regional control measures to prevent the dissemination of alien species and pathogens" (CBEP, undated).

> related elements used to establish policy and objectives and to achieve those objectives and includes organizational structure, planning activities (e.g., risk assessment and the setting of objectives), responsibilities, practices, procedures, processes and resources. (CBEP, undated)

We developed separate logic models for biosafety and biosecurity, based on existing guidance and international standards and discussions with SMEs, especially the CBEP biorisk management team. We developed these logic models with an eye toward developing metrics. Developing the logic models to serve as the foundation for measurement, we placed special emphasis on ensuring that they capture key elements of CBEP programming. To support this, we vetted the models with CBEP personnel responsible for carrying out implementing program activities in partner nations. Importantly, however, we also intend the models to represent the elements of sound biosafety, biosecurity, and biosurveillance capacities, capabilities, and sustainability enablers, irrespective of our ability to identify a recommended metric for each element. The metrics we recommend align, first, with our logic models and, second, with a functional construct.

Logic Models for Biorisk Management

We developed separate logic models for biosafety and biosecurity, while integrating the risk assessment and management elements of biorisk management into each. We consulted a broad range of guidance regarding sound biosafety and biosecurity programs to develop our logic models. This includes international standards, such as the widely cited guidance *Laboratory Biorisk Management* produced by the European Committee for Standardization (2011). We also drew on guidance from U.S. institutions, such as the reference book *Biosafety in Microbial and Biomedical Laboratories* (BMBL) (CDC, 2009). IHR 2005 makes limited mention to partner-nation capabilities for biosafety and biosecurity. Only IHR Core Capacity 8 refers to laboratory services:

> States Parties need to establish mechanisms that assure the reliable and timely laboratory identification of infectious agents and other hazards likely to cause public health emergencies of national and international concern, including shipment of specimens to the appropriate laboratories if necessary. (WHO, 2011)

WHO's *Core Capacity Monitoring Framework* recommended that partner biosafety and biosecurity "Laboratory biosafety and laboratory biosecurity (biorisk management) practices are in place" (WHO, 2011). This provides little guidance for determination of the elements of an effective biorisk management program, but does suggest that biosafety and biosecurity are elements of IHR compliance for CBEP partners.

Our logic models for biosafety and biosecurity share many common elements, especially at the level of capacities and sustainability enablers. This suggests that an effective biorisk management program includes measures that support both laboratory biosafety and biosecurity. The BMBL notes that many laboratory biosecurity measures are common with good laboratory practice (CDC, 2009). For example, even though most laboratories do not have pathogens of security concern, good laboratory practice dictates that personnel maintain control and inventory materials, protect sensitive information, and maintain appropriate access controls. Such measures represent good laboratory practice for addressing biorisks, whether originating from intentional or unintentional actions.

Biosafety—Enduring Capabilities

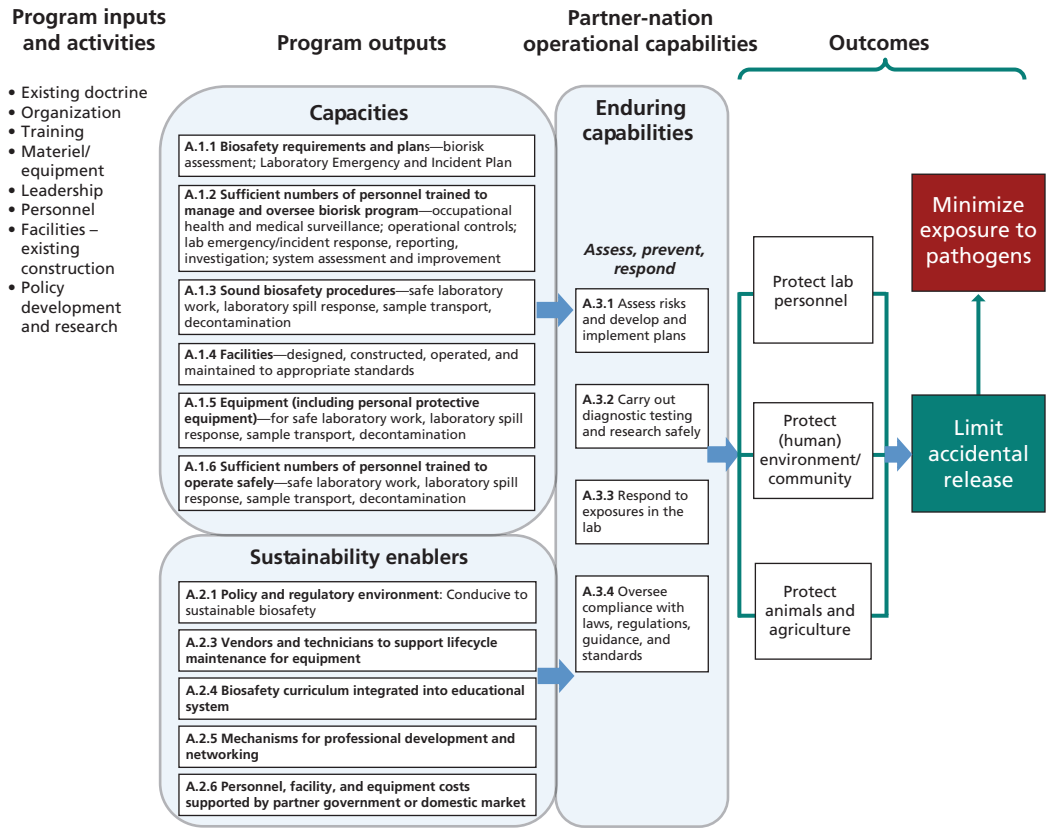
Figure 3.1 shows our logic model for enduring partner capability for biosafety. The numbering scheme includes information that specifies the CBEP objective and logic model type. In accordance with CBEP's measurement priorities, the other levels of analysis are not fleshed out to the same level of detail (program inputs on the left side of the diagram; outcomes at the right side of the diagram). However, brief consideration of these levels (as indicated on the left and right of the diagram) can help tell a clear story about how CBEP programming depends on resource inputs and relates to national goals.

Limiting populations' exposure to pathogens requires four partner capabilities to assess, prevent, and respond to laboratory biorisks with respect to biosafety. First (A.3.1 in the figure), the partner must be able to appropriately assess risks and develop and implement plans. The BMBL defines the elements of effective risk assessment as identifying hazards associated with biological materials, which guides the selection of plans to implement appropriate practices, equipment, and facility safeguards to help prevent and limit exposure to pathogens (CDC, 2009). Second, the partner must be able to carry out the required laboratory tasks or tasks associated with laboratory work. We designate this as a capability to carry out diagnostic research and testing safely, which includes associated biosafety activities, such as safe sample transport. This capability prepares the partner to prevent a biosafety incident or event. However, if an incident or event does occur, the partner must also be prepared to take appropriate action. This third capability requires the partner to be prepared to respond to exposures in the laboratory and extends to include laboratory spill response and decontamination. Finally, the partner needs to demonstrate an ability to oversee compliance to ensure that it abides by existing laws, regulations, guidance, and standards.

Biosecurity—Enduring Capabilities

While the widely cited BMBL focuses primarily on biosafety, the fifth edition was the first to include a section in the main document specifically addressing "Principles of Laboratory Biosecurity" (CDC, 2009). Since publication of the previous edition 1999,

Figure 3.1
Logic Model for Biosafety



NOTE: The numbering scheme includes information that specifies the CBEP objective (a first character A signifies biosafety; B signifies biosecurity; and C signifies biosurveillance) and logic model type (a second character 1 signifies capacity; 2 signifies sustainability enabler; 3 signifies capability).

RAND *RR660-3.1*

> significant events have brought national and international scrutiny to the area of laboratory security. These events, including the anthrax attacks on U.S. citizens in October 2001 and the subsequent expansion of the United States Select Agent regulations in December 2003, have led scientists, laboratory managers, security specialists, biosafety professionals, and other scientific and institutional leaders to consider the need for developing, implementing and/or improving the security of biological agents and toxins within their facilities. (CDC, 2009)

Our logic model for biosecurity is informed by the principles outlined in the fifth edition of the BMBL (CDC, 2009).

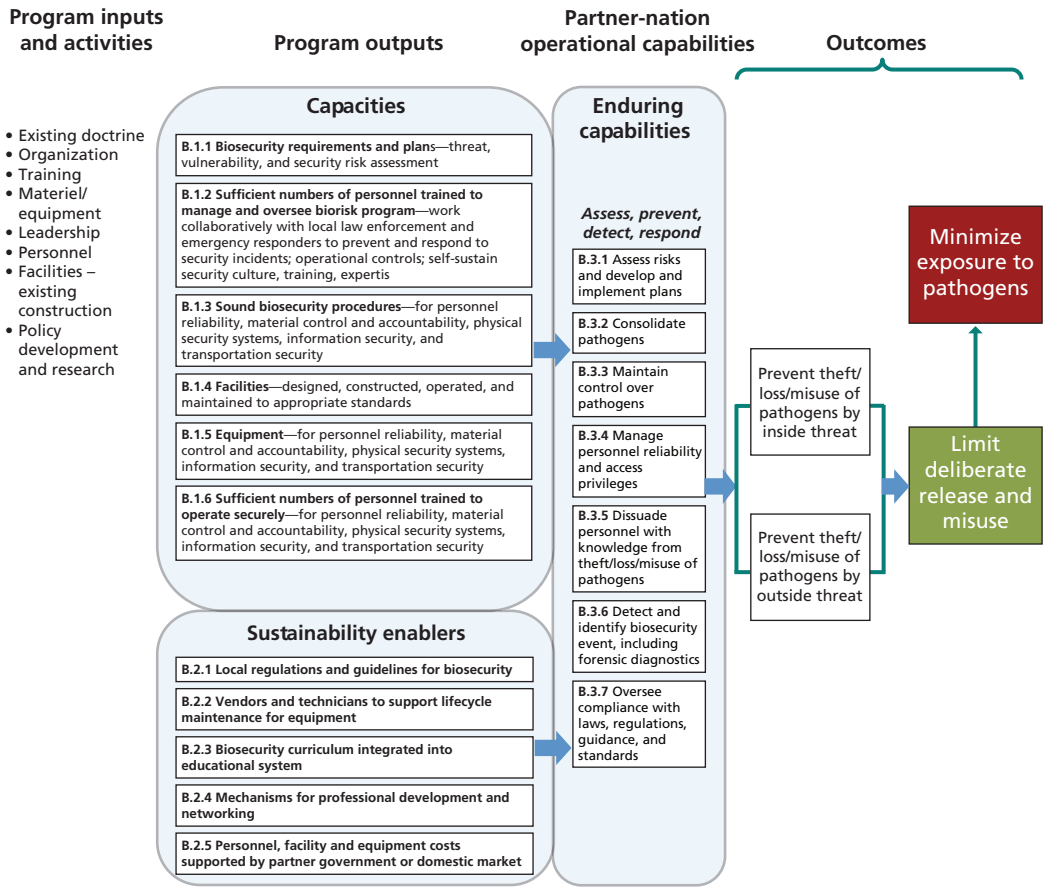
In addition to the literature on sound laboratory practices, our biosecurity logic model is also informed by international agreements related to weapons of mass destruc-

tion. In April 2004, the United Nations (UN) Security Council adopted Resolution 1540, establishing legally binding obligations for member states to develop and enforce effective measures against the proliferation of weapons of mass destruction, their means of delivery, and related materials (Bakanidze, Imnadze, and Perkins, 2010). Member states are required to report on implementation to the 1540 Committee, responsible in turn for reporting on compliance to the UN Security Council. In support of implementation goals, the 1540 Committee developed matrices for the status of measures, procedures, and legislation for securing biological weapons and related materials. Another important component of international efforts to counter biological weapons is the Biological Weapons Convention. The convention entered into force in 1975 as the first multilateral disarmament treaty that banned the production of an entire category of weapons (Bakanidze, Imnadze, and Perkins, 2010). Since 1986, the convention has defined voluntary, nonbinding confidence-building measures involving an exchange of information between member states regarding a wide range of biological weapon–related activities. While the confidence-building measures are nonbinding, consideration of the Biological Weapons Convention framework helped us identify elements of effective biosecurity.

Figure 3.2 shows our logic model for enduring partner capability for biosecurity. At the strategic level (looking at the right side of the figure), effective laboratory biosecurity requires that a partner be capable of limiting deliberate release or misuse of dangerous pathogens by both inside personnel with authorized access to the facility and outside personnel (CDC, 2009).

Limiting these populations' deliberate release or misuse requires seven enduring capabilities to assess, prevent, and respond to laboratory biorisks with respect to biosecurity. First, the partner must be able to appropriately assess risks and develop and implement plans. The BMBL defines the elements of effective risk assessment as identifying hazards associated with biological materials, which guides the selection of plans to implement appropriate practices, equipment, and facility safeguards to help prevent and limit exposure to pathogens (CDC, 2009). The next four capabilities seek to prevent a biosecurity incident through various mechanisms intended to keep dangerous materials away from personnel with the means to do harm. The second and third capabilities address the need to control dangerous materials by consolidating collections of pathogens to a minimum number of facilities and by maintaining control over pathogens through such functions as material control and accountability, physical security, information security, and transportation security. The fourth and fifth capabilities address the human dimension—managing personnel reliability and access privileges and dissuading personnel with knowledge from acting badly—through, for example, cooperative research, research grants, or international engagement. The sixth capability is the operational ability to detect and identify a biosecurity event. Finally, the seventh capability is the partner country's demonstrated ability to oversee compliance to ensure that it abides by existing laws, regulations, guidance, and standards.

Figure 3.2
Logic Model for Biosecurity



RAND *RR660-3.2*

Biorisk Management—Common Capacities and Sustainability Enablers

We identified numerous common elements for partner capacities and sustainability enablers with respect to biosafety and biosecurity. While recognizing the distinct focuses of biosafety and biosecurity programs, our logic models include common building blocks. Two capacities relate to management functions: having sufficient numbers of biorisk management personnel and having biorisk plans, including determination of requirements, in place. The BMBL next recommends adoption of specified procedures designed to reduce the risks these hazards pose (CDC, 2009). Strict adherence to procedures requires trained personnel, appropriate laboratory equipment, and appropriate physical infrastructure. These three constitute the remaining capacities for biosafety and biosecurity.

Notably, for a biosafety program, a facility's biosafety level describes the present threats from biological agents and provides guidance on appropriate safeguards to

mitigate the risks from such threats. The biosafety level takes into account the specific pathogens present at the facility, their mode of transmission, their association with disease in humans, and the resulting severity of the diseases in humans (CDC, 2009). Meeting the safeguard standards appropriate to the biosafety level of a facility requires capacity building with respect to laboratory practices and techniques, appropriate equipment, and appropriate facility design. If executed to standards, the biosafety level represents the condition under which the agents in a laboratory can ordinarily be handled safely (CDC, 2009). The BMBL provides guidance on criteria and essential elements of biosafety program at different biosafety levels (CDC, 2009).

While we emphasized commonalities in the building blocks, we also emphasized elements that make a biosafety program distinct from a biosecurity program. For biosafety, we aligned capacities with biosafety functions, including safe laboratory work, laboratory spill response, sample transport, and decontamination. For biosecurity, we aligned capacities with the five pillars of a biosecurity program: personnel reliability, material control and accountability, physical security, information security, and transportation security. While both biosafety and biosecurity require biorisk management personnel to manage and oversee the biorisk program, their emphases were distinct. Personnel who manage and oversee a *biosafety* program focus on occupational health and medical surveillance and laboratory emergency and incident response, reporting, and investigation, while personnel who manage and oversee a *biosecurity* program focus more on working collaboratively with local law enforcement and emergency responders to prevent and respond to security incidents and developing a sustaining culture, training, and expertise for biosecurity.

CBEP can enable the sustainability of its investments in biorisk management capacities and enduring capabilities by putting five sustainability enablers in place. CBEP can support the institutionalization of a partner's biorisk management program by putting a policy, legal, or regulatory framework in place that is conducive to sustainability. Such a framework would be aligned with a partner's national priorities and international standards. CBEP can enable the sustainability of the equipment delivered by ensuring that the partner or region has vendors or technicians who can support life-cycle maintenance, which should be understood to also include modernization and recapitalization of existing stock. CBEP can enable the sustainability of its investments in a partner's human capital by ensuring that the partner or regional educational system has a biorisk management curriculum and has mechanisms for professional development for trained personnel. Opportunities for partner scientists to participate in CBR is one example of a CBEP activity that supports professional development. A final sustainability enabler CBEP can support would be developing goals for projects in a partner country or region that are ultimately within the ability of a partner nation to resource. Planning program activities with these sustainability enablers in mind will work to ensure that the partner *capabilities* CBEP helps strengthen are *enduring capabilities*.

CBEP's Metrics Assessment Program

CBEP tracks many indicators of partner performance toward biorisk management goals using a software and visualization program called MAP. The CBEP biorisk management team determined the data fields to be collected in MAP in consultation with the CTR Advisory and Assistance Services team and in accordance with national and international guidance. Biological threat reduction integrating contractors (BTRICs) or other on-the-ground CBEP implementers in FSU countries collect data quarterly. One intended use for the MAP data is to support decisions about when a partner may be ready for an IOC or FOC demonstration. These detailed demonstrations of how a partner's biosurveillance system responds to the simulated release of a dangerous pathogen test the partner's biosurveillance system at several nodes—a local mobile-response laboratory, a regional diagnostic laboratory, and a central reference laboratory. In an IOC demonstration, the partner is alerted ahead of time of the location and type of outbreak to be detected and diagnosed; in an FOC demonstration, the partner is not advised of the nature of the outbreak ahead of time. Both demonstrations require substantial investments of time and resources. As CBEP evolves its partnerships in New Engagement countries, BTRICs and IOC/FOC events may not be the chosen tools for partner engagement.[3] Needs for engagement and data collection on the ground may be met through smaller, more-flexible tools. CBEP indicates that the intent, however, is for program implementers operating at the country or regional levels to continue to collect MAP data.[4]

MAP is a rich data repository well aligned with national and international guidance on biorisk management. The repository maintains data on partner capacities, certain capabilities, and sustainability enablers through 28 specific biorisk management questions. Implementers answer *yes*, *no*, or *in progress* for each question. MAP also maintains data on other key capacities, such as standard operating procedures adopted, personnel trained in biorisk management related areas, and performance on proficiency exams after training. While we think there are opportunities to improve the fidelity of MAP inputs, the program provided a rich repository of data to support metrics development.

Recommended Metrics for Biorisk Management

Much of our work on recommended metrics focused on finding the best ways to leverage the rich existing MAP data. In some cases, we used a direct input from MAP that

[3] New Engagement countries are Vietnam, India, Malaysia, Laos, Cambodia, Philippines, Indonesia, Djibouti, Kenya, South Africa, Tanzania, Uganda, Afghanistan, Iraq, and Pakistan.

[4] Conversations with CBEP and CTR Advisory and Assistance Services personnel, August 2013.

was a valid approximation of the capacity, capability, or sustainability enabler in our logic model. In other cases, we recommended adjusting the language of the MAP question to better align with our desired target for measurement. For example, the language of the question "Is there mechanism/s to ensure that personnel are competent and reliable?" is framed as a capacity metric, but with language adjustments (supported by scoring criteria for CBEP implementers, as will be discussed below), it could measure partner capability to maintain personnel reliability safeguards. We recommended the following rewording: "SME finds that mechanisms are in place sufficient to ensure that personnel are competent and reliable." In still other cases, we recommended aggregating MAP inputs. The appendixes list potential MAP elements to aggregate, although specific elements to include and aggregation rules would require additional input from CBEP.

As a result of our assessments and subsequent discussions with CBEP stakeholders, we recommend a total of 21 biorisk management metrics that can be implemented now (Table 3.1) and six that can be phased in later (Table 3.2). Figure 3.3 shows the recommended metrics for biosafety capabilities, and Figure 3.4 shows the ten metrics for biosecurity capabilities. Figure 3.5 shows the capacity metrics for both biosafety and biosecurity, and Figure 3.6 shows the sustainability enabling metrics for both. The figures map each metric to the corresponding logic model element. In some instances, several metrics map to the same element; in other instances, no metrics that we could recommend mapped to an element of the logic model. Multiple metrics for a given element offer opportunities for more robust understanding of performance within the element, or for reducing the number of metrics. The figures also show the color-coded rating of each promising metric against the five criteria described in Chapter Two. The metric column in each figure indicates the proposed timetable for introduction of the metric (now or later); metrics shaded in light yellow reflect those for which we formulated the wording of the metric, and the "TT" designation indicates those that correspond to metrics developed in parallel by the tiger team.[5] MAP inputs scored very high on feasibility because the program is already collecting these data. We were less certain, however, on how to score the *reliability* of metrics derived from MAP, especially for questions that required substantial judgment from BTRICs (or other CBEP implementers). For example, one MAP input—"Is the behavior of personnel safe?"—is a highly valid metric for biosafety capabilities, but to the extent that assessment requires significant judgment, its reliability score could be lower. Questions related to the reliability of MAP inputs were major topics for conversation with CBEP country managers and science leads as we revised recommended metrics. On the basis of SME inputs, we scored the reliability of metrics based on MAP questions that were basically

[5] In August 2013, CBEP took the lead on an effort to update the framework for reporting on CTR in the Annual Report to Congress. CBEP, OSD-P, and OSD-ATL Nuclear Chemical Biological Defense personnel assembled two "tiger teams" (TT in the tables) for revamping the document for reporting on biorisk management and biosurveillance. The RAND project team participated in tiger team discussions.

Table 3.1
Recommended Biorisk Management Metrics for Immediate Use

Number	Metric	Type	Data Existing, Planned, New Within CBEP
1	Percentage completion: BRM assessment, requirements and planning of facilities (MAP checklist)	Capacity	Existing[a]
2	Percentage completion: BRM management personnel (MAP checklist)	Capacity	Existing[a]
3	Number or percentage of facilities with complete sets of relevant standard operating procedures (SOPs) in place (MAP checklist)	Capacity	Existing[a]
4	SME finds that facilities are designed to allow employees to work safely and securely	Capacity	Existing[a]
5	Percentage completion: BS&S equipment (MAP checklist)	Capacity	Existing[a]
6	Percentage of personnel trained in laboratory biosafety and biosecurity relative to requirements	Capacity	Existing[a]
7	Percentage completion: biosafety: work practice and administrative control (MAP checklist)	Capability	Existing[a]
8	Pathogen consolidation criteria and tiers	Capability	Planned[b]
9	Percentage of disclosed biological weapon–related infrastructure that has been eliminated (FSU only)	Capability	New
10	Percentage completion: maintain control of pathogens (physical security, material control and accountability, information security, transportation security) (MAP checklist)	Capability	Existing[a]
11	SME finds that mechanisms are in place sufficient to ensure that personnel are competent and reliable (adapted MAP language)	Capability	Existing[a]
12	Reporting in 1540 matrix affirming existence of a mechanism for penalizing violator of national prohibition on engagement in biological weapon–related activities	Capability	Existing
13	Proficient test scores for CBEP courses related to biorisk assessment and planning	Capability	Existing[a]
14	SME finds that accidents or incidents and nonconformities related to biorisk are correctly managed	Capability	Existing[a]
15	SME finds that the BRM system is reviewed regularly	Capability	Existing[a]
16	SME finds that a policy concerning management of laboratory biorisk (BS&S) been written	Sustainability enabler	Existing[a]
17	SME finds that operational plans include materiel sustainment considerations	Sustainability enabler	New

Table 3.1—Continued

Number	Metric	Type	Data Existing, Planned, New Within CBEP
18	Number of national or regional institutions that make BS&S training to international standards available (training centers, universities, professional societies, ministries)	Sustainability enabler	New
19	Number of peer-reviewed publications, number of conference papers (abstract, poster, or oral)	Sustainability enabler	Existing[c]
20	Number or value of internationally competitive research grants won	Sustainability enabler	New
21	SME finds that operational plans include resource sustainment considerations	Sustainability enable	New

[a] Existing MAP input.

[b] Metric the CBEP-led TT recommended metric for the annual report to Congress.

[c] Existing metric the CBEP science team tracks.

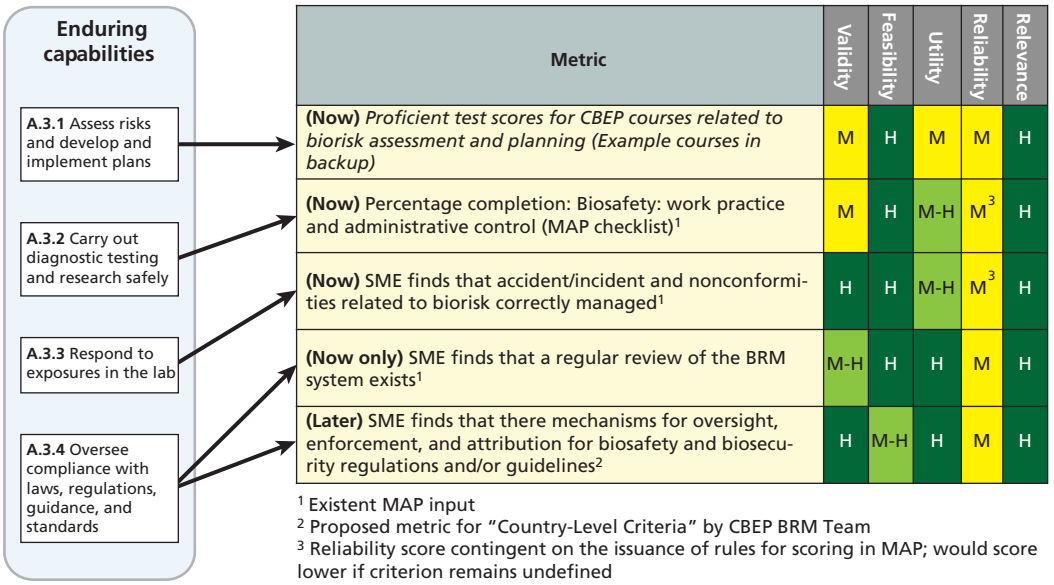
Table 3.2
Recommended Biorisk Management Metrics for Later Introduction

Number	Addition or Replacement	Metric	Type	Existing, Planned, New
1	Replacement for recommended metric 2	SME finds that collections of dangerous pathogens have been consolidated into a minimum number of facilities	Capability	Planned[a]
2	Replacement for recommended metric 9	SME finds that there are mechanisms for oversight, enforcement, and attribution for biosafety and biosecurity regulations and/or guidelines	Capability	Planned[b]
3	Replacement for recommended metric 10	SME finds there are regulations and/or guidelines for biosafety and biosecurity	Sustainability enabler	Planned[b]
4	Replacement for recommended metric 10	Are there national incident management systems for naturally occurring or intentional biological events? (concerns policies, frameworks, MOUs for the laboratory, health security and law enforcement, and ER sectors)	Sustainability enabler	Planned[b]
5	Addition	Percentage of required equipment that is domestically or regionally sourced	Sustainability enabler	New
6	Addition	SME finds that there adequate availability of funding to support BS&S programs and initiatives	Sustainability enabler	Planned[b]

[a] Metric the CBEP-led TT recommended metric for the annual report to Congress.

[b] Metric the CBEP BRM team proposed for country-level criteria.

Figure 3.3
Recommended Biosafety Metrics for Enduring Capabilities



Metric	Validity	Feasibility	Utility	Reliability	Relevance
(Now) *Proficient test scores for CBEP courses related to biorisk assessment and planning (Example courses in backup)*	M	H	M	M	H
(Now) Percentage completion: Biosafety: work practice and administrative control (MAP checklist)[1]	M	H	M-H	M[3]	H
(Now) SME finds that accident/incident and nonconformities related to biorisk correctly managed[1]	H	H	M-H	M[3]	H
(Now only) SME finds that a regular review of the BRM system exists[1]	M-H	H	H	M	H
(Later) SME finds that there mechanisms for oversight, enforcement, and attribution for biosafety and biosecurity regulations and/or guidelines[2]	H	M-H	H	M	H

[1] Existent MAP input
[2] Proposed metric for "Country-Level Criteria" by CBEP BRM Team
[3] Reliability score contingent on the issuance of rules for scoring in MAP; would score lower if criterion remains undefined

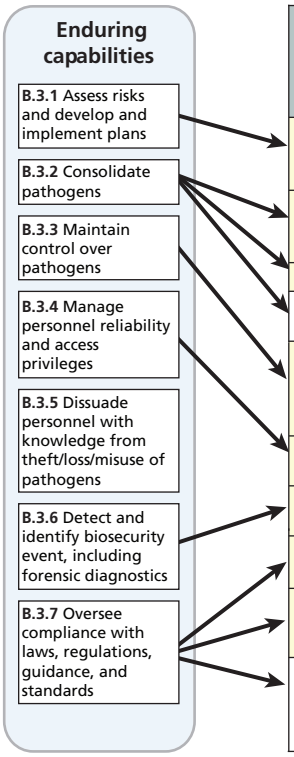
NOTE: M = medium; M-H = medium–high; H = high. Italic font indicates metric language developed by RAND; yellow shading indicates metrics derived from or inspired by CBEP program documentation.

RAND *RR660-3.3*

checklists for the presence or absence of certain elements (generally capacities and sustainability enablers) as *medium*. We scored questions that required significant discretion (generally capabilities) as *low*. However, the CBEP biorisk management team indicated that it is currently developing scoring criteria for CBEP implementers responsible for submitting MAP data. This would allow CBEP to ensure that certain standards were maintained even when implementers had to apply a certain level of discretion. Assuming that CBEP efforts to provide criteria to implementers continue to completion, we increased the score of MAP metrics requiring discretion to *medium*. Efforts to improve the quality of MAP data, such as the identification and implementation of scoring criteria, could allow CBEP to develop more-robust metrics with relatively little additional work.

While we considered MAP inputs to be strong enough metrics to recommend their use, self-reporting by CBEP implementers is not as reliable a foundation for metrics as other sources. If CBEP wants to improve data, especially for capability metrics, it could phase in third-party audits, demonstrations, or smaller-scale exercises. Conversations with CBEP indicate that current CBEP functional exercises and IOC/FOC demonstrations provide stronger measures of partner biosurveillance capability than they do of partner biorisk management capabilities. CBEP could leverage these existing mechanisms to demonstrate partner capabilities more reliably than MAP by tailor-

Figure 3.4
Recommended Biosecurity Metrics for Enduring Capabilities



Metric	Validity	Feasibility	Utility	Reliability	Relevance
(Now) *Proficient test scores for CBEP courses related to biorisk assessment and planning (Example courses in backup)*	M	H	M	M	H
(Later) SME finds that collections of dangerous pathogens have been consolidated into a minimum number of facilities[2]	M-H	M-H	H	M	H
(Now only) Pathogen consolidation (TT)	H	H	H	M-H	H
(Now/FSU only) *Percentage of disclosed biological weapons–related infrastructure that has been eliminated*	H	M-H	H	M	H
(Now) Percentage completion: Maintain control of pathogens (physical security, material control and accountability, information security, transportation security) (MAP checklist)	H	H	M-H	M*	H
(Now) SME finds that mechanisms are in place sufficient to ensure that personnel are competent and reliable*	H	H	M-H	M	H
(Now) SME finds that accident/incident and nonconformities related to biorisk correctly managed	M	H	M-H	M*	M-H
(Now only) SME finds that a regular review of the BRM system exists	M-H	H	H	M	H
(Later) SME finds that there mechanisms for oversight, enforcement, and attribution for biosafety and biosecurity regulations and/or guidelines[2]	H	M-H	H	M	H
(Now/FSU only) Reporting in 1540 matrix affirming existence of a mechanism for penalizing violator of national prohibition on engagement in biological weapons–related activities	H	M-H	H	M	M

*Adaptation of existent MAP language to reference a capability rather than a capacity

NOTE: M = medium; M-H = medium–high; H = high. Italic font indicates metric language developed by RAND; yellow shading indicates metrics derived from or inspired by CBEP program documentation.

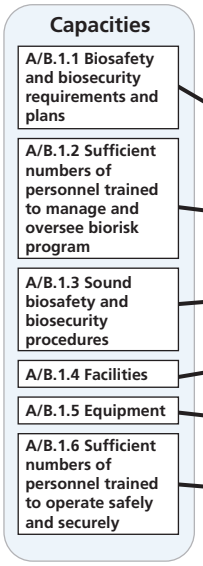
RAND *RR660-3.4*

ing the mechanisms to address biorisk management capabilities to the same extent they currently demonstrate BSV capabilities.

Metrics Within a Functional Framework

Conversations with CBEP SMEs indicated that it might be more intuitive to describe metrics as aligning with functional areas in addition to aligning with logic models. Therefore, we aligned the metrics for each CBEP objective with functional areas as defined by relevant literature and SME input. As Figures 3.7 and 3.8 show, we aligned the biosafety and biosecurity metrics to the framework for biosafety outlined in the BMBL (CDC, 2009).

Figure 3.5
Recommended Metrics for Biorisk Management Metrics for Capacities, Including Both Biosafety and Biosecurity



Metric	Validity	Feasibility	Utility	Reliability	Relevance
(Now) Percent complete for BRM Assessment, Requirements and Planning of facilities (MAP checklist)[1]	M-H	H	M-H	M	M-H
(Now) Percent complete of BRM management personnel (MAP checklist)[1]	M-H	H	M-H	M	M-H
(Now) Number or percentage of facilities with complete set of relevant SOPs in place (MAP checklist)[1]	H	H	M	M	H
(Now) SME finds that facilities are designed to allow employees to work in a safe and secure way[1]	H	H	H	M	H
(Now) Percent complete BS&S Equipment (MAP checklist)[1]	M-H	H	M-H	M	M-H
(Now) Percentage of personnel trained in laboratory biosafety and biosecurity relative to requirements[1]	M-H	M-H	H	M	H

NOTE: M = medium; M-H = medium–high; H = high. Yellow shading indicates metrics derived from or inspired by CBEP program documentation.

RAND *RR660-3.5*

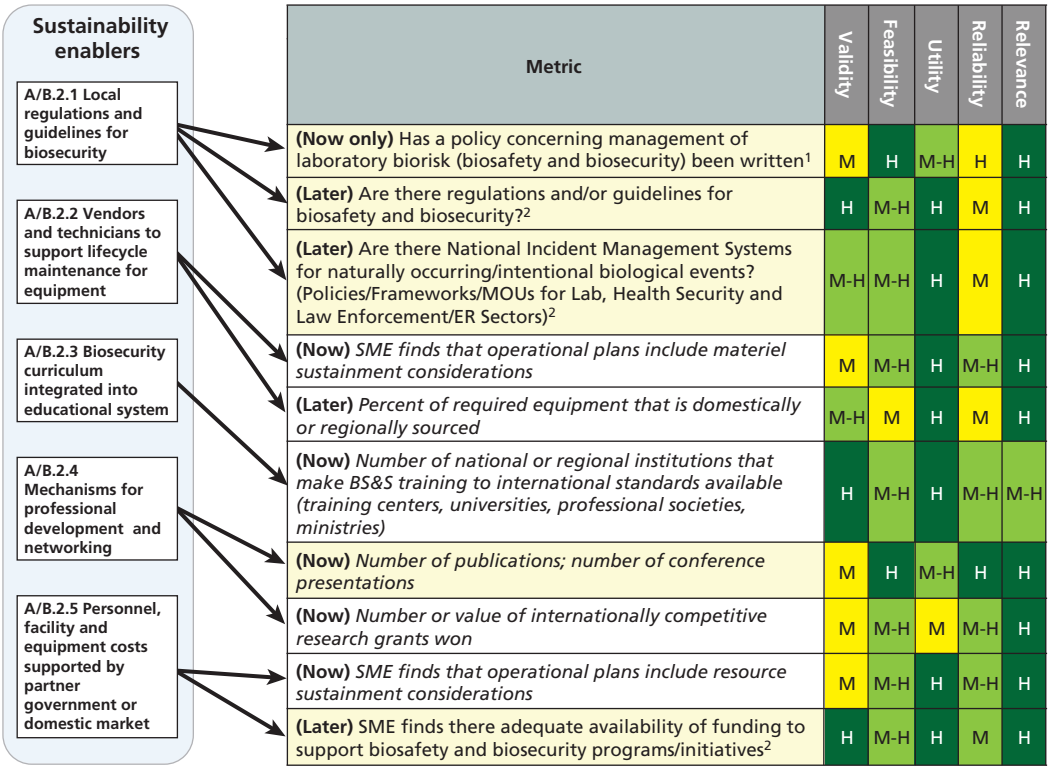
Primary functional areas for biosafety include biorisk assessment, primary barriers (safety equipment), and secondary barriers (facility design and construction). Activities in these areas enable safe practices and emergency or incident response. To this framework, we added "human capital" to capture the range of metrics aimed at measuring the development and sustainment of the biosafety workforce. Figure 3.7 shows this functional alignment.

Primary functional areas for biosecurity metrics are structured around the "five pillars" of biosecurity: physical security, transportation security, information security, material control and accountability, and personnel management and reliability. The pillars flow from a threat or security risk assessment and support the detection and response to a security incident. In this case, we again added "human capital" to support measurement of CBEP activities to develop and sustain a biosecurity workforce. Figure 3.8 shows the functional alignment for biosecurity.

Comparison of RAND Recommendations with Other Relevant Biorisk Management Metrics

A comparison of the metrics recommended here with the metrics recommended in other recent assessments illustrates the value of taking a systematic approach to devel-

Figure 3.6
Recommended Metrics for Biorisk Management Sustainability Enablers, Including Both Biosafety and Biosecurity



Metric	Validity	Feasibility	Utility	Reliability	Relevance
(Now only) Has a policy concerning management of laboratory biorisk (biosafety and biosecurity) been written[1]	M	H	M-H	H	H
(Later) Are there regulations and/or guidelines for biosafety and biosecurity?[2]	H	M-H	H	M	H
(Later) Are there National Incident Management Systems for naturally occurring/intentional biological events? (Policies/Frameworks/MOUs for Lab, Health Security and Law Enforcement/ER Sectors)[2]	M-H	M-H	H	M	H
(Now) *SME finds that operational plans include materiel sustainment considerations*	M	M-H	H	M-H	H
(Later) *Percent of required equipment that is domestically or regionally sourced*	M-H	M	H	M	H
(Now) *Number of national or regional institutions that make BS&S training to international standards available (training centers, universities, professional societies, ministries)*	H	M-H	H	M-H	M-H
(Now) *Number of publications; number of conference presentations*	M	H	M-H	H	H
(Now) *Number or value of internationally competitive research grants won*	M	M-H	M	M-H	H
(Now) *SME finds that operational plans include resource sustainment considerations*	M	M-H	H	M-H	H
(Later) SME finds there adequate availability of funding to support biosafety and biosecurity programs/initiatives[2]	H	M-H	H	M	H

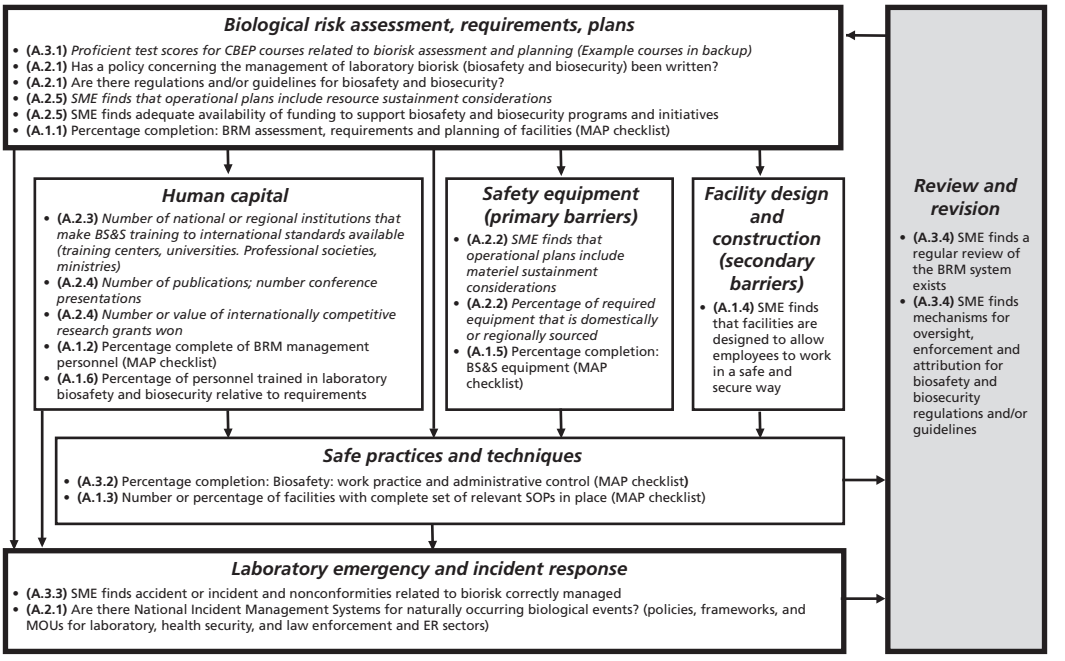
NOTE: M = medium; M-H = medium–high; H = high. Italic font indicates metric language developed by RAND; yellow shading indicates metrics derived from or inspired by CBEP program documentation.

RAND *RR660-3.6*

oping and prioritizing metrics. Both the DoD metrics report NAS assessed and the MAP framework made strides to advance CBEP's goals for program evaluation. While these contributions shaped our approach, our framework also addresses some gaps and imbalances that have limited the utility of existing frameworks.

The DoD metrics report includes a systematic development of metrics for CBEP that links program objectives to desired outcomes and metrics for measurement. NAS rightfully praised DoD for including in the CBEP framework "nearly all of the elements that the committee thinks are needed for development of useful metrics" (NAS, 2012, p. 24). DoD-recommended metrics for CBEP, however, include a larger number of metrics for capacities and certain sustainability enablers. These include, for example, having "facility specific" plans or a "legal framework" in place (see NAS, 2012, pp. 81–82). The latter, which we include as a sustainability enabler, DoD includes as a metric for *sustainability* as a program objective separate from parallel objectives for

Figure 3.7
Recommended Biosafety Metrics, Aligned Functionally



NOTE: Italic font indicates metric language developed by RAND.

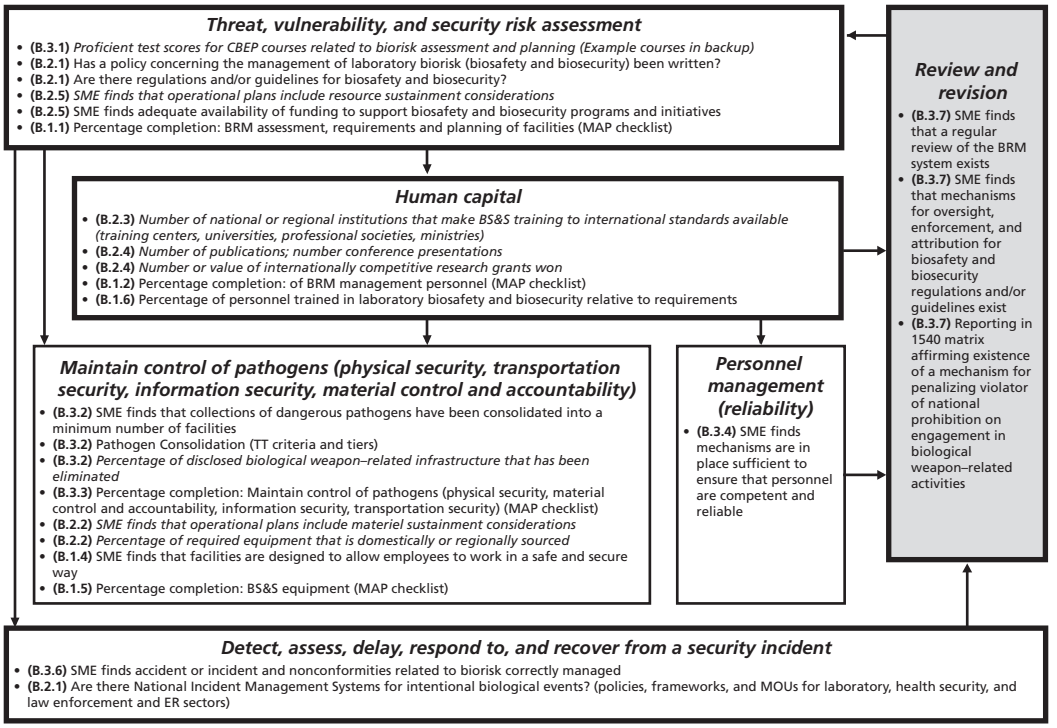
RAND *RR660-3.7*

biorisk management and biosurveillance. We think that integrating sustainability into other program objectives clarifies the relationship of these objectives to overall threat reduction, a central NAS recommendation (NAS, 2012, p. 5). While the DoD report includes several metrics for partner capabilities, the metrics it does include are less robust. For example, DoD includes a metric for immediate notification of a biosecurity event but offers little guidance on the mechanism for assessing this.

To demonstrate the relative emphases of the two metrics frameworks, Tables 3.3 and 3.4 summarize the mapping of the DoD recommended metrics to RAND recommended metrics.

The MAP framework is a rich resource, but it was developed to meet specific and distinct program goals. It especially supports decisionmaking about program management at the level of implementation, by tracking detailed information about equipment delivered, individuals receiving training, and the status of legal agreements with partners. This information supports important monitoring of CBEP inputs and can help CBEP make decisions about partner readiness for visits from CBEP personnel. Our framework leverages much of this valuable data to support metrics at higher levels of analysis and reporting on program performance to internal and external audiences.

Figure 3.8
Recommended Biosecurity Metrics, Aligned Functionally



NOTE: Italic font indicates metric language developed by RAND.

RAND *RR660-3.8*

Given the purpose for which it was developed, MAP's strength is in tracking partner capacities. It includes few metrics for capabilities, and those it does include are less robust. For example a biosafety capability metric of whether or not laboratory personnel operate safely requires substantial judgment from a CBEP implementer on the ground. Our framework recommends making such metrics more robust by either implementing scoring criteria for CBEP implementers or introducing improved mechanisms for third-party observations of partner capabilities to improve consistency and reliability of capability metrics.

Biorisk Management Conclusions

CBEP leaders will clearly need to make important decisions about the metrics we recommend here. For example, they will need to carefully consider the desirable number of metrics, balancing the desire for robustness of actionable and strategically important information with the burden of data collection. Fortunately, CBEP already collects

substantial information on performance toward biorisk management goals, especially in MAP, which will lessen the additional burden. CBEP will nonetheless need to make decisions about expanding data collection or making existing data more robust.

Table 3.3
DoD Biorisk Management Metrics Included in NAS 2012 Report

DoD Metric	Logic Model Category	Functional Category
1.1: Number of EDP collections	Capability	Maintain control
2.1: Biosecurity compliance	Capability	Maintain control
3.1: Established legal framework	Sustainability enabler	Review and revision
3.2: Level of regulation	Capability	Review and revision
4.1: Biosafety guidelines	Capacity	Biological risk assessment, requirements, and plans
4.2: Facility-specific biosafety plans exist	Capacity	Biological risk assessment, requirements, and plans
5.1: Biosecurity standards	Capacity	Threat, vulnerability, and security risk assessment
5.2: Facility-specific biosecurity plans exist	Capacity	Threat, vulnerability, and security risk assessment
6.1: BS&S standards are available	Capacity	Threat, vulnerability, and security risk assessment
6.2: Biosecurity event notification	Capability	Detect, assess, delay, respond to, and recover from a security incident
6.3: Biosafety event notification	Capability	Laboratory emergency incident response
1.1: Sustainability plan to physically maintain collection	Sustainability enablers	Maintain control
2.1: Sustainment cost of application and enforcement of BS&S guidelines and regulations	Sustainability enablers	Review and revision
3.1: Sustainment cost	Sustainability enablers	
4.1: Trainee test results for ability to retrain	Sustainability enablers	Human capital

Table 3.4
Comparison of RAND and DoD Biorisk Management Metrics

Type of Metric	Number of Metrics	
	RAND (n = 31)*	DoD 2010 (n = 20)
Logic model framework		
Inputs and/or activities	0	0
Capacities	6	5
Sustainability enablers	10	5
Capabilities	15	5
Functional framework		
Biological risk assessment, requirements, plans	6	2
Human capital	5	1
Safety equipment	3	0
Facility design and construction	1	0
Safe practices and techniques	2	0
Laboratory emergency and incident response	2	1
Review and revision	3	3
Threat, vulnerability, and security risk assessment	6	3
Maintain control of pathogens	8	3
Personnel management (reliability)	1	0
Detect, assess, delay, respond to, and recover from a security incident	2	1
Other	0	1

NOTE: Numbers do not add because several metrics overlap between biosafety and biosecurity functional frameworks.

CHAPTER FOUR

Conceptual Model and Recommended Metrics for Biosurveillance

Overview

Biosurveillance has become an integral component of CBEP programming. It brings together priorities for capacity building and collection of disease-related data (including CBR) within the context of the WHO IHR, an internationally binding treaty to which nearly all countries are signatories (WHO, 2005). It calls for each country to strengthen and maintain a set of specified "core capacities" (Table 4.1) to detect, diagnose, report, and respond to public health emergencies potentially of international concern and for countries in a position to do so to help others strengthen their core capacities. CBEP works with partner countries to help them build their core public health capacities, particularly as these relate to dangerous pathogens in humans or animals. CBEP does not carry out operational biosurveillance or provide direct support for responses, e.g., outbreak investigations.

Public health professionals consider biosurveillance to be a cornerstone of public health (Moore et al., 2012). While DoD has been carrying out biosurveillance activities for many years, a memorandum from the Deputy Secretary of Defense in June 2013 was the first official issuance using this term (Carter, 2013). That interim guidance specifies that DoD adopts the definition of the term from the 2012 *National Strategy for Biosurveillance* (White House, 2012b) and calls for development of a DoD directive for biosurveillance within 12 months. That definition of biosurveillance is

> the process of gathering, integrating, interpreting, and communicating essential information related to all-hazards threats or disease activity affecting human, animal, or plant health to achieve early detection and warning, contribute to overall situational awareness of the health aspects of an incident, and to enable better decision making at all levels (White House, 2012b).

Table 4.1
International Health Regulations Core Capacities

Core Capacity	Description
National legislation, policy, and financing	State parties need to have an adequate legal framework to support and enable implementation of all their obligations and rights.
Coordination and national focal point communications	Coordination of nationwide resources, including the designation of an IHR national focal point as a national center for IHR communications, a key requisite for IHR implementation.
Surveillance	The IHR require the rapid detection of public health risks and prompt risk assessment, notification, and response to these risks. To this end, a sensitive and flexible surveillance system with an early warning function is necessary.
Response	Command, communications, and control mechanisms are required to facilitate the coordination and management of outbreak operations and other public health events.
Preparedness	Preparedness includes the development of national, intermediate, and community and/or primary response public health emergency response plans for relevant biological, chemical, radiological, and nuclear hazards.
Risk communication	Risk communications should be multilevel and multifaceted, aiming to help stakeholders define risks, identify hazards, assess vulnerabilities, and promote community resilience.
Human resources	Strengthening the skills and competencies of public health personnel is critical to the sustainment of public health surveillance and response at all levels of the health system and the effective implementation of the IHR.
Laboratory	Laboratory services are part of every phase of alert and response, including detection, investigation, and response, with laboratory analysis of samples performed either domestically or through collaborating centers.

SOURCE: WHO, 2011.

This definition is consistent with earlier widely cited ones from one of CDC's most respected epidemiologists and from the Institute of Medicine:

> Public health surveillance is the systematic collection, analysis, interpretation, and dissemination of data regarding a health-related event for use in public health action to reduce morbidity and mortality and to improve health. (Thacker, 2000)

> Public health surveillance is the ongoing systematic collection, analysis, and interpretation of health data, essential to the planning, implementation, and evaluation of public health practice, closely integrated to the dissemination of these data to those who need to know and linked to prevention and control. (Institute of Medicine, 2002)

Simply put, biosurveillance is "systematic information for public health action" (Moore et al., 2008).

Our process for developing biosurveillance metrics to recommend to CBEP was comparable to the one for biosafety and biosecurity. It is important to note that CBEP has already been collecting a substantial number of metrics relevant to biosafety and biosecurity but has a significantly less-robust measurement system in place for biosurveillance. Thus, much of the search for biosurveillance metrics necessarily extended well beyond CBEP's current monitoring protocols. The sections below describe the key elements of biosurveillance and the progressive steps toward the systematic development of recommended biosurveillance metrics.

Logic Model for Biosurveillance

As described in *National Strategy for Biosurveillance* (White House, 2012b), the core functions of biosurveillance—which can also be considered desired outcomes—are to scan and discern the environment, identify and integrate essential information, alert and inform decisionmakers, and forecast and advise on impacts. Surveillance systems require epidemiology, laboratory, information technology, and other communications capacities and capabilities:

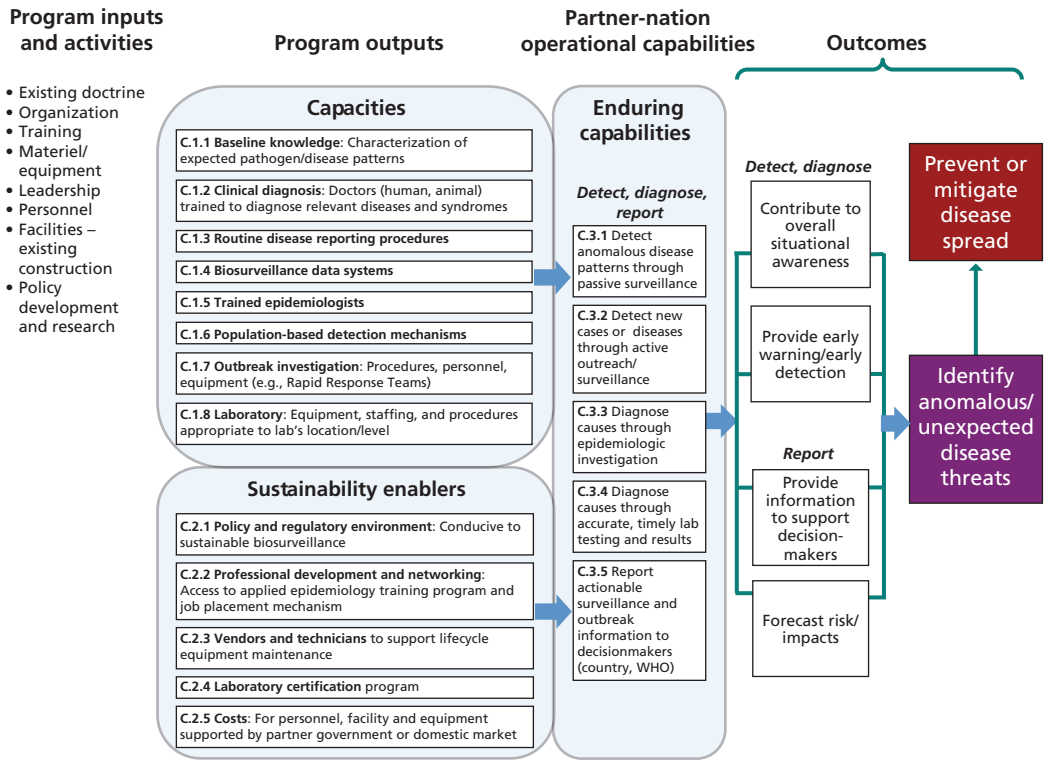
- *Epidemiologists* must be able to collect, analyze, and interpret surveillance data to monitor trends and discern anomalies and must be able to communicate timely information effectively to decisionmakers. Most practicing epidemiologists receive academic and/or applied training in epidemiologic methods and practices. For example, CBEP supports CDC's Field Epidemiology Training Program in several partner countries. This program is modeled after CDC's domestic epidemiology training program (in operation since the early 1950s) and has been operational overseas since 1980. Surveillance data can be reported routinely (passively) or collected actively when health officials contact health facilities to inquire about the occurrence of cases of a suspected outbreak. Surveillance data may reflect individual cases or aggregate numbers of cases of diseases or conditions, including risk factors. Reporting may be of lay diagnoses (chief complaints or reports from nonprofessional community members), clinical diagnoses (from a doctor or nurse), or laboratory-confirmed diagnoses. Enduring epidemiology capabilities will require ongoing training (or access to training) in applied epidemiology and supervisory oversight of surveillance analysis and outbreak investigation by practicing epidemiologists.
- *Laboratory capabilities* are also critical to biosurveillance. Such capabilities rely on appropriate equipment, supplies, and procedures and trained personnel. Enduring laboratory capabilities will require systems for maintenance of equipment, reliable supplies of necessary reagents, and testing proficiency.

- *Communications capabilities* are also essential to surveillance, to facilitate both routine reporting and any additional reporting when outbreaks occur. Electronic reporting systems, such as the Electronic Integrated Disease Surveillance System CBEP supports, facilitate timely and standardized surveillance reporting. Enduring communications capabilities depend on reliable and modern equipment and trained personnel to both use and maintain such equipment.

The first step in the process to develop CBEP biosurveillance metrics was to establish a logical flow of biosurveillance elements, beginning with resource inputs and program activities and flowing logically to program outputs (biosurveillance capacities and sustainability enablers), and then to enduring host-nation biosurveillance capabilities, which in turn enable near-term and longer-term outcomes (Figure 4.1).

In principle, one can begin with outcomes and work backward, or begin with the early elements and proceed forward. In practice, however, the development of the biosurveillance logic model was an iterative process that flowed in both directions and drew from our own extensive professional experiences in public health surveillance and

Figure 4.1
Logic Model for Biosurveillance



RAND *RR660-4.1*

epidemiology, key published documents, and vetting with key stakeholders, including key CBEP leaders and those working in parallel to develop higher-level metrics. The logic model outcomes drew from both the IHR (WHO, 2005) and the definition of the term specified the *National Strategy for Biosurveillance* (White House, 2012b). Interestingly, IHR uses the term *core capacities*, but the recommended four-point rating scale to assess these "capacities" includes two levels that are clearly capacities and two that are clearly capabilities, as we have distinguished them (WHO, 2011).

CBEP programming reinforces the focus on detection, diagnosis and reporting, which is also explicitly reflected in the logic model.

Working backward from the desired outcomes, key *operational capabilities* for biosurveillance include the demonstrated ability to detect anomalous disease patterns through routine (passive) surveillance, detect new cases or emerging diseases through active surveillance, diagnose the problem through epidemiologic investigation and laboratory testing, and report actionable surveillance information to relevant authorities—local, national, and/or international, as appropriate. These capabilities reflect the typically linear progression from suspicion of a problem (detection of anomalous patterns or new threats) to diagnosis and reporting of the problem.

The *capacities* needed to enable these capabilities relate to baseline information, clinical diagnostic skills (which reasonably could also be considered capabilities); disease reporting procedures and systems; trained personnel; population-based reporting mechanisms; outbreak investigation skills (including the required personnel, equipment, and protocols); and laboratories with the equipment, trained staff, and procedures appropriate for the laboratory's location or level. All these capacities can be mapped to relevant elements within the DoD "DOTMLPF-P" construct,[1] as can most of the inputs and program activities.

Factors that *enable the sustainability* of biosurveillance capacities and enduring capabilities include a conducive policy and regulatory environment, access (in country or nearby) to ongoing training, (e.g., applied epidemiology), a mechanism for job placement of trainees, professional networking (e.g., among clinicians, among epidemiologists), mechanisms to ensure maintenance of key equipment (e.g., laboratory, computer networks), laboratory certification programs to ensure that laboratories maintain the high standards intended, and the ability of the host nation or domestic market to support the ongoing costs of personnel, facilities, and equipment. Because explicit attention to facilitating and measuring progress toward sustainability is new, relatively little has been published about sustainability enablers. Thus, these factors and the associated metrics are more developmental than rigorously empirically tested. However, principles likely to contribute to sustainability have been advanced in recent literature (Moore et al., 2012) and current national policy related to global health (White House, 2009):

[1] This construct refers to the combination of doctrine, organization, training, materiel, leadership, personnel, facilities, and policy.

- owner-driven agendas
- explicit attention to transforming "data" (numbers) into "information" (what the numbers mean); "messages" (what to do based on the information); and, finally, action
- accountability
- leveraging strengths across key partners
- effective multisector engagement; and trusting relationships.

These factors are described in the context of sustainability of regional disease surveillance networks and federal policy related to global health, but they are also already recognized as CBEP priorities and principles. Thus, explicit attention to activities that facilitate sustainability and metrics to measure progress are the natural next steps for CBEP.

Recommended Biosurveillance Metrics

The next step in the process to develop CBEP biosurveillance metrics was to identify existing (or suggested) metrics from various relevant sources, including CBEP. As noted earlier, CBEP does not presently collect extensive data on biosurveillance performance in partner countries. In contrast to rich data for biorisk management metrics, MAP collects relatively limited data related specifically to biosurveillance. The system does track data on laboratory tests in place in a country, the mechanism a partner uses for electronic reporting (e.g., the Electronic Integrated Disease Surveillance System), and numbers of personnel training to carry out biosurveillance functions. However, other CBEP-related documentation uncovered or suggested a number of additional items to be measured (e.g., in the Georgia Country Test Plan [DTRA, undated] and NAS assessment of CBEP metrics [NAS, 2012]). Other sources of potentially relevant biosurveillance metrics included a published report from a U.S. government interagency effort to develop measures for selected IHR core capacities (Ijaz et al., 2012), an internal government working document with metrics to reflect different aspects of global health security agenda (some of which are relevant to biosurveillance), and CDC's Public Health Preparedness Capabilities and Public Health Infrastructure metrics from HHS's Healthy People 2020 initiative (Department of Health and Human Services, 2010b; CDC, 2011). We extracted a total of 94 items from these sources and mapped them to specific elements within the logic model.

We next assessed the 94 items using the five criteria described earlier (in Chapter Three) and selected all those meeting at least "medium" on all criteria or the highest-rated metric that mapped to a logic model element not otherwise addressed, for vetting with CBEP stakeholders. As a result of our assessments and subsequent discussions with CBEP stakeholders, we recommend a total of 26 biosurveillance metrics that

can be implemented now (Table 4.2) and seven additional ones that can be phased in later (Table 4.3). Figure 4.2 presents the recommended capability metrics, Figure 4.3 presents the capacity metrics, and Figure 4.4 presents the sustainability-enabling metrics. The figures show the mapping of each metric to the corresponding logic model element. In some instances, several metrics map to the same element. Multiple metrics for a given element offer opportunities either for gaining a more robust understanding of performance within the element or for reducing the number of metrics. The figures also show the color-coded rating of each promising metric against the five criteria. The "Metric" column in each figure indicates the proposed timetable for introduction of the metric (now or later); metrics shaded in light yellow are those for which we formulated the wording of the metric, and the "TT" designation indicates those that correspond to metrics developed in parallel by the tiger team for reporting 2013 progress to Congress. As indicated in the figures, our recommendations include seven metrics that the tiger team also suggested: two capability metrics, four capacity metrics, and one sustainability-enabling metric.

Table 4.2
Recommended Biosurveillance Metrics for Immediate Use

Number	Metric	Type	Data Existing, Planned, New Within CBEP
22	Number and percentage of priority pathogens for which baseline data have been established (e.g., via CBR) and used to detect anomalies (modified TT tiers)	Capacity	Planned[a]
23	Reporting capability (TT criteria and tiers)	Capacity	Planned[a]
24	Number and percentage of major jurisdictions using electronic reporting system (e.g., EIDSS)	Capacity	New
25	Accuracy and scope: national guidance for reporting, e.g., case definitions—specified animal and human diseases	Capacity	New
26	Number trained of field epidemiologists and number per 200,000 population	Capacity	New
27	Technical capacity (TT criteria and tiers)	Capacity	Planned[a]
28	Number and percentage of laboratories certified	Capacity	New
29	Number and list of pathogens for which (a) the national and (b) each provincial or state laboratory can test	Capacity	New
30	Laboratory networking: written or established protocols for specimen referral (a) within country or (b) to international laboratory	Capacity	New
31	Epidemiological surveillance analysis performance—through exercise or supervisory observation	Capability	Planned[a]
32	Epidemiological investigation performance documented in written report or tested via exercise (tabletop or functional)	Capability	New

Table 4.2—Continued

Number	Metric	Type	Data Existing, Planned, New Within CBEP
33	Number of suspected priority pathogen cases or outbreaks in the past 12 months and the percentage for which investigations were conducted and results documented	Capability	New
34	Number laboratory tests performed for each priority pathogen in the past 12 months	Capability	New
35	Number and percentage of laboratories participating in proficiency testing at least once in the past 12 months	Capability	New
36	Number and percentage of laboratories that passed all proficiency tests in the past 12 months, among laboratories that participated in proficiency testing in the past 12 months	Capability	New
37	Specimen collection and transport: number of specimens received by laboratory and the number and percentage that are adequate for testing	Capability	Planned[a]
38	Laboratory referral network: number of specimens sent or received for confirmatory testing in the past 12 months to an (a) in country or (b) international laboratory	Capability	New
39	Biosurveillance (or CBR) directly informed decisions (e.g., policy, clinical practice)	Capability	New
40	Supervised demonstration of electronic reporting system use (e.g., EIDSS)	Capability	New
41	Number of cases of internationally reportable pathogens or diseases detected in the country in past 12 months; percentage reported to the appropriate international authority (e.g., WHO, OIE)	Capability	New
42	Regulatory environment criteria and tiers	Sustainability enabler	Planned[a]
43	Degree of consistency between national and CBEP planning and implementation (e.g., CBEP within national plan or vice versa, joint planning and execution of exercises)	Sustainability enabler	New
44	Number of peer-reviewed publications; number of conference papers (abstract, poster, or oral)	Sustainability enabler	Existing[b]
45	Number of institutions with epidemiology core curriculum that meets international (e.g., CDC FETP) standards; number of trainees per year	Sustainability enabler	New
46	Career track for trained epidemiologists	Sustainability enabler	New
47	Operational laboratory certification program	Sustainability enabler	New

[a] Metric the CBEP-led TT recommended for the annual report to Congress.

[b] Existing metric the CBEP science team tracks.

Table 4.3
Recommended Biosurveillance Metrics for Later Introduction

Number	Addition or Replacement	Metric	Type	Existing, Planned, New
7	Addition	Demonstrated performance in surveillance of at least three of five defined core syndromes (U.S. government IHR measure)	Capability	New
8	Addition	Number and percentage of laboratories that meet specified standards (mutually agreed on by country and CBEP)	Capability	New
9	Addition	Laboratory timeliness: number of hours following receipt of specimens that laboratory testing is (a) initiated and (b) completed (from laboratory logs or exercise)	Capability	New
10	Addition	Trained and funded internal staff or active contracts or vendor agreements for life-cycle equipment maintenance	Sustainability enabler	New
11	Addition	Number and percentage of priority pathogens for which country (a) produces or (b) can sustainably procure needed media and reagents	Sustainability enabler	New
12	Addition	Performance under standardized clinical scenarios or clinical preceptor validation (and TT tiers)	Capacity	Planned[a]
13	Addition	Number of specified level jurisdictions, number and percentage with functional community-based surveillance (submit ≥50 percent reports)	Capacity	New

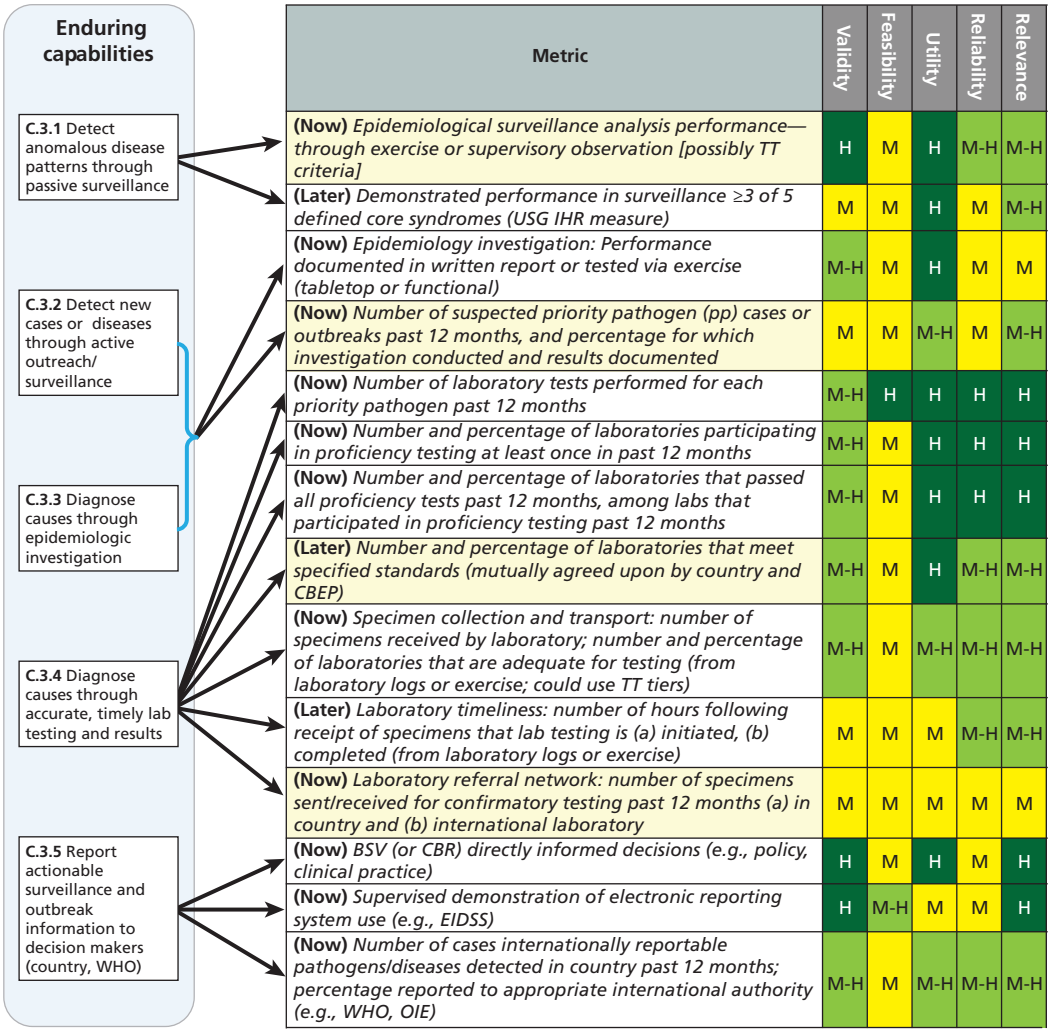
[a] Metric the CBEP-led TT recommended for the annual report to Congress.

The highest priority for metric development was for biosurveillance *capabilities,* since these reflect operational abilities that lead more directly to the desired outcomes. We identified 14 such metrics. As shown in Figure 4.2, these 14 metrics address all five capabilities, with at least two metrics for each one.

Biosurveillance *capacities* are necessary but not sufficient precursors to operational capabilities. As Figure 4.3 shows, the 11 recommended capacity metrics address seven of the eight capacities, with more than one metric for only two capacities, one each for four capacities, and no metric for one capacity.

Active efforts to lay the groundwork for the sustainability of biosurveillance within partner countries is an important priority for CBEP that is also consistent with priorities specified in President Barack Obama's Global Health Initiative Strategy, which calls for public health surveillance and health systems strengthening as well as country ownership, accountability and sustainability (White House, 2009). However, measurement of sustainability-enabling factors is more developmental and exploratory at this point than measurement of capacities and capabilities, for which there is consid-

Figure 4.2
Recommended Biosurveillance Capability Metrics



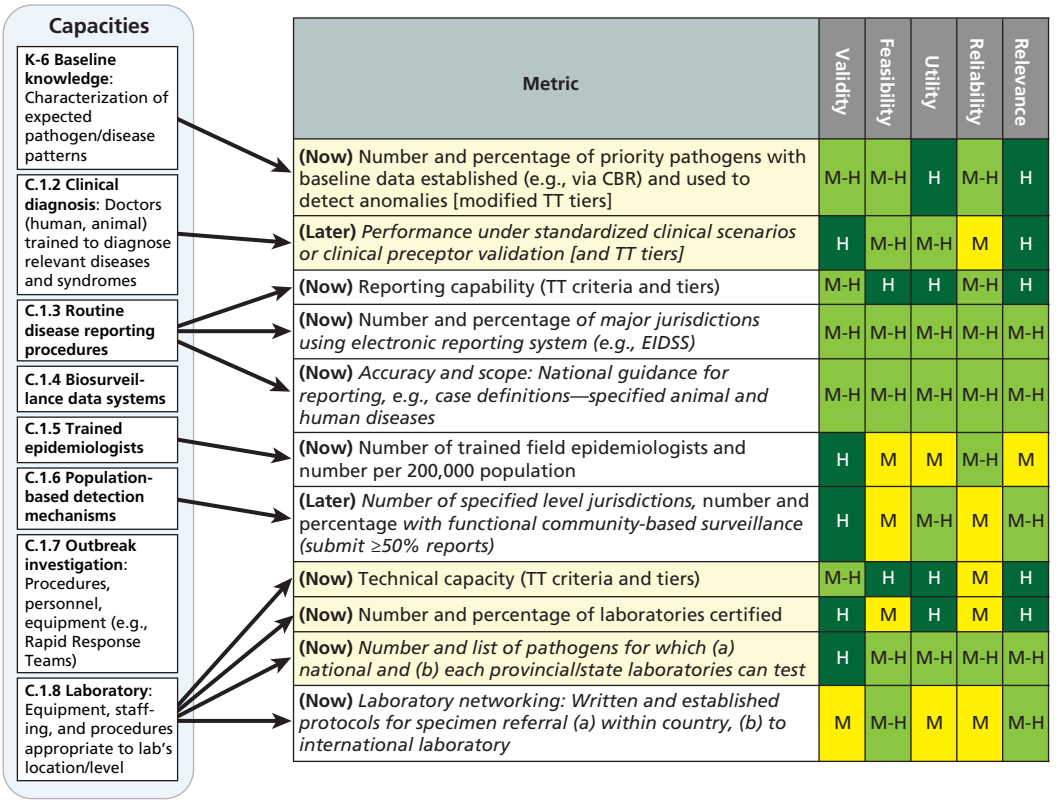
Metric	Validity	Feasibility	Utility	Reliability	Relevance
(Now) *Epidemiological surveillance analysis performance—through exercise or supervisory observation [possibly TT criteria]*	H	M	H	M-H	M-H
(Later) *Demonstrated performance in surveillance ≥3 of 5 defined core syndromes (USG IHR measure)*	M	M	H	M	M-H
(Now) *Epidemiology investigation: Performance documented in written report or tested via exercise (tabletop or functional)*	M-H	M	H	M	M
(Now) *Number of suspected priority pathogen (pp) cases or outbreaks past 12 months, and percentage for which investigation conducted and results documented*	M	M	M-H	M	M-H
(Now) *Number of laboratory tests performed for each priority pathogen past 12 months*	M-H	H	H	H	H
(Now) *Number and percentage of laboratories participating in proficiency testing at least once in past 12 months*	M-H	M	H	H	H
(Now) *Number and percentage of laboratories that passed all proficiency tests past 12 months, among labs that participated in proficiency testing past 12 months*	M-H	M	H	H	H
(Later) *Number and percentage of laboratories that meet specified standards (mutually agreed upon by country and CBEP)*	M-H	M	H	M-H	M-H
(Now) *Specimen collection and transport: number of specimens received by laboratory; number and percentage of laboratories that are adequate for testing (from laboratory logs or exercise; could use TT tiers)*	M-H	M	M-H	M-H	M-H
(Later) *Laboratory timeliness: number of hours following receipt of specimens that lab testing is (a) initiated, (b) completed (from laboratory logs or exercise)*	M	M	M	M-H	M-H
(Now) *Laboratory referral network: number of specimens sent/received for confirmatory testing past 12 months (a) in country and (b) international laboratory*	M	M	M	M	M
(Now) *BSV (or CBR) directly informed decisions (e.g., policy, clinical practice)*	H	M	H	M	H
(Now) *Supervised demonstration of electronic reporting system use (e.g., EIDSS)*	H	M-H	M	M	H
(Now) *Number of cases internationally reportable pathogens/diseases detected in country past 12 months; percentage reported to appropriate international authority (e.g., WHO, OIE)*	M-H	M	M-H	M-H	M-H

NOTE: M = medium; M-H = medium–high; H = high. Italic font indicates metric language developed by RAND; yellow shading indicates metrics derived from or inspired by CBEP program documentation.

RAND *RR660-4.2*

erably more experience. DoD is thus on the leading edge of using and refining metrics related to sustainability. RAND identified eight metrics for the five sustainability-enabling components of the biosurveillance logic model (Figure 4.4).

Figure 4.3
Recommended Biosurveillance Capacity Metrics



Metric	Validity	Feasibility	Utility	Reliability	Relevance
(Now) Number and percentage of priority pathogens with baseline data established (e.g., via CBR) and used to detect anomalies [modified TT tiers]	M-H	M-H	H	M-H	H
(Later) *Performance under standardized clinical scenarios or clinical preceptor validation [and TT tiers]*	H	M-H	M-H	M	H
(Now) Reporting capability (TT criteria and tiers)	M-H	H	H	M-H	H
(Now) Number and percentage *of major jurisdictions using electronic reporting system (e.g., EIDSS)*	M-H	M-H	M-H	M-H	M-H
(Now) *Accuracy and scope: National guidance for reporting, e.g., case definitions—specified animal and human diseases*	M-H	M-H	M-H	M-H	M-H
(Now) Number of trained field epidemiologists and number per 200,000 population	H	M	M	M-H	M
(Later) *Number of specified level jurisdictions,* number and percentage *with functional community-based surveillance (submit ≥50% reports)*	H	M	M-H	M	M-H
(Now) Technical capacity (TT criteria and tiers)	M-H	H	H	M	H
(Now) Number and percentage of laboratories certified	H	M	H	M	H
(Now) *Number and list of pathogens for which (a) national and (b) each provincial/state laboratories can test*	H	M-H	M-H	M-H	M-H
(Now) *Laboratory networking: Written and established protocols for specimen referral (a) within country, (b) to international laboratory*	M	M-H	M	M	M-H

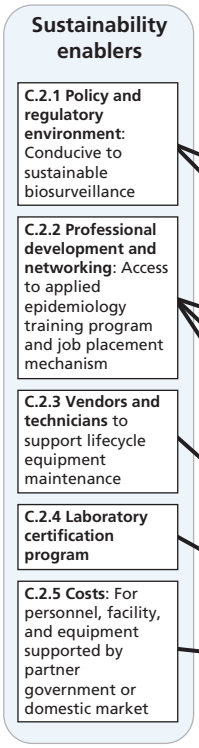
NOTE: M = medium; M-H = medium–high; H = high. Italic font indicates metric language developed by RAND; yellow shading indicates metrics derived from or inspired by CBEP program documentation.

RAND *RR660-4.3*

Biosurveillance Metrics Within a Functional Framework

Another systematic, and presumably also useful, way to think about metrics is by biosurveillance function. We mapped the 33 metrics by broad biosurveillance functions, as shown in Figure 4.5. The functions capture the typical chronological course of biosurveillance from detection to diagnosis, reporting, and action. They also reflect the four desired outcomes from the logic model, detection, epidemiologic and laboratory diagnosis, and reporting. This presentation includes all metrics in one common picture, arrayed functionally. We derived an overall rating for each metric based, in particular, on its validity and feasibility. These ratings can be found in Tables A.4–A.6 and B.4–B.6 in the appendixes. The alphanumeric labels identify a metric as referring to a capability (C.3 series), a capacity (C.1 series), or sustainability enabler (C.2 series).

Figure 4.4
Recommended Biosurveillance Sustainability Enabling Metrics



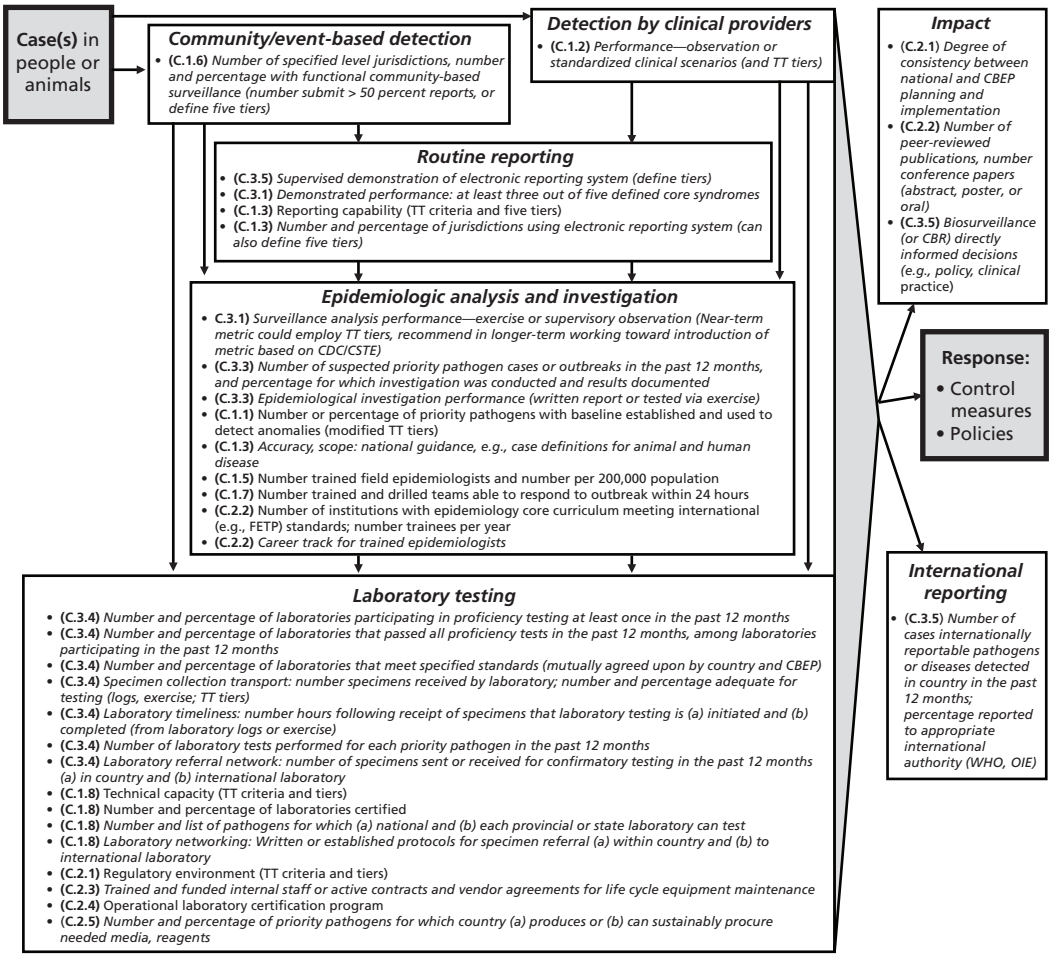
Metric	Validity	Feasibility	Utility	Reliability	Relevance
(Now) Regulatory environment (TT criteria and tiers)	M-H	H	H	M-H	H
(Now) *Degree of consistency between national and CBEP planning and implementation (e.g., CBEP within national plan or vice versa, joint planning and execution of exercises)*	H	H	H	M-H	H
(Now) *Number of peer-reviewed publications, number of conference papers (abstract, poster, or oral)*	H	H	H	H	H
(Now) Number of institutions with epidemiology core curriculum that meets international (e.g., CDC FETP) standards and number of trainees per year	H	M	H	M	H
(Now) *Career track for trained epidemiologists*	H	M-H	H	M-H	H
(Later) *Trained and funded internal staff or active contracts/vendor agreements for lifecycle equipment maintenance*	H	L-M	H	L-M	H
(Now) Operational laboratory certification program	H	M-H	H	M	H
(Later) *Number and percentage of priority pathogens for which country (a) produces or (b) can sustainably procure needed media, reagents*	M-H	M-H	H	M	M

NOTE: L-M = low–medium; M = medium; M-H = medium–high; H = high. Italic font indicates metric language developed by RAND; yellow shading indicates metrics derived from or inspired by CBEP program documentation.

RAND *RR660-4.4*

Arraying metrics in this way provides another way to consider the completeness and robustness of metrics from a functional perspective and another way to consider which, if any, might be deferred or deleted from the package of CBEP biosurveillance metrics. For example, one might be willing to sacrifice robustness in the interest parsimony and thus select just one or a very few capability metrics from each functional area or select just the highest-rated metrics—whether capability, capacity, or sustainability enablers—from each functional area. However, this does create trade-offs in sacrificing information that is useful for management purposes. The relatively large number of metrics associated with the epidemiologic analysis and laboratory testing functions indicates the richness of activity and dimensions of measurement associated with these vital biosurveillance functions.

Figure 4.5
Recommended Biosurveillance Metrics



NOTE: Italic font indicates metric language developed by RAND.

RAND *RR660-4.5*

Comparison of RAND Recommendations with Other Relevant Biosurveillance Metrics

Comparison of the metrics recommended here with the biosurveillance metrics in the DoD metrics report that NAS assessed illustrates the value of a systematic approach to developing and prioritizing metrics because we must assume that careful thought was invested in creating the 25 DoD metrics. Table 4.4 lists the 25 metrics from the NAS report and classifies them according to our logic model and functional framework. Table 4.5 summarizes the distribution of our metrics and these DoD metrics.

Table 4.4
DoD Biosurveillance Metrics Included in NAS 2012 Report

DoD Metric	Logic Model Category	Functional Category	Comment
1.1: Biosafety guidelines	NA	NA	BRM
2.1: Biosecurity standards	NA	NA	BRM
3.1: National pandemic influenza preparedness and response plan	Capacity	Other	
3.2: Bioterrorism preparedness and response plan	Capacity	Other	
3.3: Multihazard national public health emergency preparedness and response plan	Capacity	Other	
3.4: National emergency preparedness and response plan for animal diseases	Capacity	Other	
4.1: National disease surveillance plan	Capacity	Other	
5.1: Disease surveillance system capable of detecting EDP cases	Capability	Epidemiological analysis	Metric is not precisely defined
5.2: Sharing of epidemiological data from EDP case investigations with those responsible for human and animal health	Capability	Routine reporting	
5.3: Provision of laboratory test results to those responsible for human and animal health	Capability	Routine reporting	
5.4: Reporting of all human epidemiological events constituting a public health emergency of international concern to WHO	Capability	International reporting	
5.5: Reporting of all reportable animal diseases to OIE	Capability	International reporting	
5.6: Sharing of case data via appropriate reporting systems	Capability	Routine reporting	
6.1: Prompt reporting of epidemiological data from EDP case investigations to those responsible for human and animal health	Capability	Routine reporting	
6.2: Prompt provision of laboratory test results to those responsible for human and animal health	Capability	Routine reporting	
6.3: Prompt reporting of all human epidemiological events constituting a public health emergency of international concern to WHO	NA	NA	Duplicates 5.4
6.4: Prompt reporting of all reportable animal diseases to OIE	NA	NA	Duplicates 5.5
7.1: Investigation and documentation of suspect EDP cases by those responsible for human or animal health	Capability	Epidemiological analysis	

Table 4.4—Continued

DoD Metric	Logic Model Category	Functional Category	Comment
7.2: Collection and prompt transport of appropriate samples under optimum conditions for laboratory confirmation of the diagnosis	Capability	Laboratory	
7.3: Ability of the partner country to diagnose endemic EDPs	Capability	Laboratory	
7.4: Ability of the partner country to utilize international reference laboratories when there is no country diagnostic capability	Capacity	Laboratory	
8.1: Prompt investigation and assessment of suspect cases by those responsible for human or animal health	NA	NA	Duplicates 7.1
8.2: Collection and prompt transport of appropriate samples under optimum conditions for laboratory confirmation of the diagnosis	Capability	Laboratory	
8.3: Partner country promptly initiates diagnostic testing of endemic EDPs	Capability	Laboratory	
8.4: Ability of the partner country to rapidly utilize international reference laboratories when there is no in-country diagnostic capability	Capability	Laboratory	

As shown in the tables, the DoD metrics leave gaps and imbalances in both the logic model and functional frameworks. Our metrics cover the entire range of biosurveillance functions, while the DoD metrics do not.

Biosurveillance Conclusions

CBEP leaders will clearly need to make important decisions about the metrics we recommend. For example, they will need to carefully consider the desirable number of metrics, balancing the desire for robustness of actionable and strategically important information against the burden of data collection. If they elect to begin with a subset of our recommended metrics, they will then need to select the desired ones and decide when and how to begin implementing them.

Table 4.5
Comparison of RAND and DoD Biosurveillance Metrics

Type of Metric	Number of Metrics	
	RAND (n = 33)	DoD 2010 (n = 20)[a]
Logic model framework		
Inputs and/or activities	0	0
Capacities	11	6
Sustainability enablers	8	5
Capabilities	14	14
Functional framework		
Community-based detection	1	0
Clinical detection	1	0
Routine reporting	4	5
Epidemiological analysis, investigation	9	2
Laboratory testing	15	6
International reporting	1	2
Impact	3	0
Other	0	5

[a] As noted in Table 4.4, of the 25 DoD biosurveillance metrics, two relate to BRM rather than biosurveillance, and three are duplicates.

CHAPTER FIVE

Implementing the Recommended Metrics Framework

This work builds on substantial existing efforts related to measurement that will greatly facilitate implementation. Data collected for MAP, for example, means that CBEP already collects a significant fraction of the data required for recommended metrics. Despite such strides, implementation of a new evaluation framework will require additional time, effort, and resources. For example, existing CBEP data and bases for evaluation are better developed to date for biorisk management than for biosurveillance. This means that implementation of the biorisk metrics will require focusing more attention on scoring and testing the metrics than on data collection. This chapter discusses issues related to implementation of the recommended framework and metrics and provides a user's guide to the extensive material related to implementation that appears in Appendixes A and B.

Data Availability

Operationalizing the recommended metrics first entails knowing where and how to collect the data. In each appendix, the "Sources and Descriptions" section provides the data sources or approaches to data development for each recommended metric. Appendix A aligns metrics with the logic model framework, and Appendix B aligns the same recommended metrics with a functional framework. The tables in these appendixes indicate one or more potential sources of data when such sources were identified and outline approaches for developing data if specific sources have not yet been identified. Given current CBEP approaches to tracking program progress, data availability differs significantly between metrics related to biorisk management and metrics for biosurveillance. Data availability for biorisk management means that metrics development in this area will, to a high degree, focus on leveraging existing CBEP data most effectively, while the task of implementing recommended biosurveillance metrics will focus instead on identifying data sources not yet available in CBEP.

Biorisk Management

As the biosafety and biosecurity tables under "Sources and Descriptions" in Appendix A suggest, the recommended metrics heavily leverage existing MAP inputs. MAP data are generally used in one of three ways. First, in several instances, we pulled the metric directly from the MAP checklist. Second, some of our metrics adapted existing MAP language to strengthen alignment to the evaluation framework. For example, as described below, in certain cases, we adapted existing language originally framed as a capacity, framing it as a capability metric. Third, in numerous cases, the metrics aggregated existing MAP data. While the specific aggregation rules applied would likely require additional analysis and deliberation to score, our metrics are, in many cases, defined as the "percentage complete" across several MAP inputs. We further discuss scoring later.

While MAP is a rich resource, CBEP has opportunities to improve existing data in some important ways, and indeed, the biorisk management team is already leading the way in several areas. One example of such an opportunity would be to revise the language of select MAP questions to better align them with specific outcomes CBEP wants to measure. For example, one biosecurity metric we recommend, "SME finds that mechanisms are in place sufficient to ensure that personnel are competent and reliable," is framed as a capability metric. This is an adaptation of an existing MAP input, "Are mechanisms in place to ensure that personnel are competent and reliable," which is framed as a capacity. CBEP could review the language of MAP inputs to strengthen alignment with elements they seek to measure. However, the recommended language that frames the input as a capability metric also requires substantial discretion for the SME responsible for reporting in MAP. CBEP could improve the reliability of reporting in MAP if it also completed the scoring criteria to guide CBEP implementer reporting. The CBEP biorisk management team is already advancing an effort to accompany MAP checklists with criteria to guide reporting; we think this is an effective way to bolster the reliability of recommended metrics with a minimal additional implementation burden. Finally, in addition to MAP, CBEP has other existing assessment processes and tools that may be leveraged more effectively to provide data for biorisk metrics. For example, functional exercises and IOC/FOC demonstrations reportedly provide stronger inputs for assessing biosurveillance capabilities than for assessing partner capabilities for biorisk management. Opportunities may exist to more directly target such activities on elements of enduring partner capabilities.

Biosurveillance Metrics

Operationalizing the metrics we recommend entails knowing where and how to collect the data and balancing the value of the metrics data against the burden of data collection. This is especially important for the biosurveillance metrics because most of them would be new to CBEP's monitoring framework. It will be important to identify sources of biosurveillance performance data and/or resources for guidance or assis-

tance to collect the data. In each appendix, the biosurveillance tables under "Sources and Descriptions" provide extensive information on existing data for biosurveillance metrics (when identified) and approaches to data development in the case existing data have not yet been identified.

The appendixes are a resource for guidance on specific data sources and information related to data development. In several cases, potential data sources for the metrics point CBEP toward similar existing programs of other U.S. government and international organizations. For example, data related to a metric for partner capability to conduct an epidemiologic investigation may draw on existing extensive CDC documentation or a WHO checklist on conducting such an investigation. Similarly, data related to institutionalization of appropriate curricula in partner institutions may be gleaned from the global network of Field Epidemiology Training Programs, which includes the one CDC operates. Identification of similar, established data collection efforts of other organizations can provide a clear and specific path for CBEP as it seeks to develop new data sources to meet its priorities for measurement.

We recognize that there may be differing views regarding the feasibility of some of the recommended metrics and that different program leaders may reach different conclusions about how the value of data balances against the burden of data collection. Laboratory proficiency testing is a prime example of a set of metrics that is very important to CBEP programming but that some highly experienced program leaders view as more difficult to implement. In such instances, CBEP leaders may wish to draw from the technical expertise and experiences of other organizations that conduct laboratory proficiency testing internationally to further examine its practical feasibility for CBEP. In Appendix A, for example, the table on sources and descriptions for biosurveillance points CBEP toward several organizations with known proficiency testing programs, including such U.S. government entities as the CDC, such laboratory groups as the Association of Public Health Laboratories (APHL), such international organizations as WHO Collaborating Centers, and private firms.

Score and Target Development

Identifying available data is the first step toward implementing the metrics framework, but the important next step is defining scoring criteria and a target value for each metric. Guidance on approaches to scoring appears in the "Scoring and Metrics" section of each appendix, again arrayed using the logic model and functional framework, respectively. The scoring for biorisk management metrics again relies more heavily on leveraging the existing MAP framework than on developing new approaches than do recommended approaches to scoring for biosurveillance metrics. To align as much as possible with the approach the tiger team recommended in its parallel effort to develop

high-level metrics for reporting progress to Congress, we used a similar approach in developing criteria to classify each relevant biosurveillance metric into one of five tiers.

Approaches to scoring will depend on the nature of the tool used to document partner capacities, capabilities, and sustainability enablers. Many of our metrics could be scored using a checklist tool. For example, scoring for MAP inputs primarily involves *yes*, *no*, or *in progress* checklists. For example, in the table on scoring and metrics for biosafety in Appendix A, a metric related to pathogen consolidation could be scored as a *yes* or *no* answer to the proposed MAP input, "Have collections of dangerous pathogens been consolidated into a minimum number of facilities?" As indicated above, CBEP is currently developing scoring criteria to guide the yes-or-no assessments; we think this is an important contribution, especially for scoring MAP inputs that require significant discretion. Development of such criteria could also facilitate scoring at tiers more detailed than those binary assessments allow. Checklist tools could also be used to assess the completeness of desired characteristics for a given metric. For example, as shown in the table on scoring and metrics for biosurveillance in Appendix A, partner performance on an epidemiologic investigation could be scored (by quintile or simple percentage) using a knowledge and skills checklist, such as those the U.S. Agency for International Development uses. Other metrics could be scored as a simple count (e.g., counts of partner research publications, disaggregated by institution, scientist, and research area, or count of trained personnel relative to requirements).

Testing

Finally, and as described in Chapter Seven, testing the new evaluation framework on a small scale before introduction more broadly could help CBEP identify challenges and refine the framework as needed. Pilot testing in a small number of countries would allow CBEP to identify data collection or scoring challenges early and to develop mitigation strategies. This experience could also help CBEP develop approaches for prioritization or a phased approach to implementation. In addition to providing CBEP with a tried and tested set of metrics, pilot testing would also allow socializing proposed metrics with partner countries. This becomes part of the owner-driven agenda that contributes to sustainability.

CHAPTER SIX

Aggregating Proposed Metrics to Report on CBEP Performance to Support Decisionmaking at Several Levels

The preceding chapters of this report described a set of 47 metrics that tell a comprehensive story about the state of sustainable capacities and capabilities for biorisk management and biosurveillance in CBEP partner countries. Despite the fact that this set represents a concise picture selected from a larger set of nearly 300, making sense of a set of 47 metrics is challenging. The utility of this set of metrics is in part determined by whether it helps answer the following question: "What is CBEP accomplishing?"

Of course, there is no one answer to this relatively simple question. How it is answered depends on who is asking it. When DoD or congressional leadership asks it, the answer must explain how CBEP contributes to the overall goals of using CTR to make the world a safer place. When CTR program leadership asks it, the answer must explain how activities are changing the threats across the globe from nuclear, chemical, and biological weapons. When CBEP leadership asks it, the answer must explain what progress is being made on the objectives of biorisk management and biosurveillance. Finally, when country managers ask it, the answer must help them understand where work remains to build capacities, capabilities, and future sustainability.

A formal approach, such as the one this report describes, can transparently tie strategic-level assessments to metrics for specific capabilities, capacities, and sustainability enablers. The framework in this chapter relies on notional values we assigned for program goals and thresholds and on notional choices of aggregation rules. Implementation of an approach to support strategic-level reporting would require policy guidance and program expertise to determine actual (rather than notional) values for thresholds and aggregation rules.

Ideally, the metrics in this report would be part of a story that answers all these questions. The metrics can thus be used to explain

- how project resources and activities support program outputs
- how these outputs lead to outcomes of building sustainable capacities and capabilities in partner countries across CBEP objectives
- how CBEP programming contributes to broader threat reduction goals
- how CTR programs together support U.S. strategic objectives.

To the extent to which the proposed metrics can tell all parts of this story, the effort required to collect and organize the metrics will be offset by the benefits from using them to guide decisionmaking at all levels of CTR and CBEP program management.

Drawing Insights from a Large Number of Metrics

The logic modeling approach we used to identify valid metrics provided a starting point for making sense of the 47 recommended metrics. However, distilling insight from the full set of metrics requires addressing three analytic challenges:

1. developing a clear measurement hierarchy (from the country to program to strategic level)
2. aggregating assessment data at each level of the hierarchy
3. developing ways to present assessments.

Developing a Clear Measurement Hierarchy

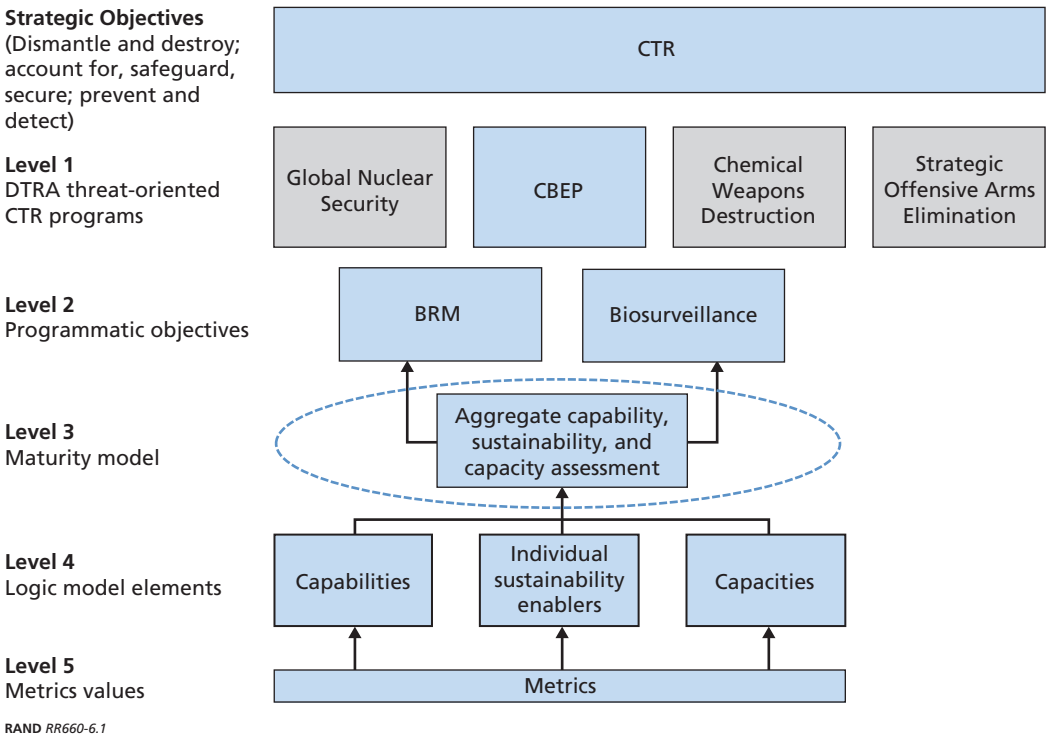
Using metrics to tell the complete story of how program resources and activities ultimately support strategic objectives requires several levels of metrics. Figure 6.1 hierarchically depicts how the proposed metrics can fit into this story.

As a starting point at the bottom of the figure, the proposed individual metrics collected at the country level (level 5) provide the most disaggregated indication of the current state of capabilities, capacity, and sustainability enablers. Looking across a set of metrics, country managers or program managers can assess the state of individual enablers in a partner country (level 4), and it should also be possible to determine an aggregated score indicating in how many of these enablers a partner nation has demonstrated adequate progress (level 3). In turn, the number of enablers that have been demonstrated can help indicate the extent to which a partner nation has achieved CBEP objectives in the areas of biorisk management and biosurveillance (level 2). When similar analysis is done for other threats, CTR leadership can explore how its component programs are contributing to the overall goals of CTR (level 1) and thus how well the overall program supports U.S. strategic objectives.

Combining Assessments at Each Level of the Hierarchy

At each level of reporting in Figure 6.1, evaluation involves making a holistic judgment drawn from metrics reported at the level below. For example, level 4 metrics report the state of specific capacities, capabilities, and sustainability enablers based on data from one or more of the metrics proposed at level 5. Similarly, an assessment of CBEP overall, at level 1, combines judgments across assessment of the objectives reported at level 2.

Figure 6.1
Measurement Hierarchy Depicting How Proposed Metrics Support Strategic Objectives



Judgments can be combined in many ways to provide an overall evaluation (Davis, Shaver, and Beck, 2008; Keeney, 1992). Examples include reporting

- an average or additive score across metrics in a level, for example, when reporting performance on a number of equally important capabilities that compensate for each other
- the number or percentage of goals met across metrics in a level, for example, when judging performance against well-agreed-on standards
- the weakest performance across metrics in a level, for example, when overall performance is constrained by the weakest of several capabilities.

These are only three of many possible ways metrics can be aggregated to provide an overall judgment. The appropriate approach to use is entirely context-specific. The approach to aggregation will depend on the interactions among functions in the system being evaluated, the existence of standards that define acceptable performance (or lack thereof), and leadership judgment of what outcomes are most important or threats most concerning (Davis and Dreyer, 2009)

Similarly, it is important to communicate whether performance is adequate. This can be expressed in terms of whether minimum acceptable performance thresholds or goals are met. These thresholds and goals may be different in each partner country. As with any standard-setting exercise, thresholds and goals must be selected in the context of the current state of affairs in the country being evaluated and the available evidence against which to judge adequacy of performance. When evidence suggests a clearly desirable outcome, all countries might be held to the same goal. Absent such evidence, variable goals may be set across countries that encourage consistent improvement from the current state of affairs. However, if countries are held to different standards, it is important to be clear that assessments are relative to country-specific goals and transparent about what those goals are to avoid confusing intercountry comparisons.

In the end, the choice of how to make an overall assessment from component evaluations is one of judgment. When viewed as such, there is no objectively correct approach to aggregating or setting goals and thresholds. Instead, it is critically important to be transparent about how such judgments are being made (Davis and Dreyer, 2009).

Developing Ways to Present Assessments

Drawing insight from a large set of metrics is challenging even after overcoming the challenges of defining a measurement hierarchy, aggregation rules, and thresholds. A measurement hierarchy provides the framework for insights about overall performance. Aggregation rules codify how to make summary judgments of performance. Goals and thresholds establish the yardsticks for measuring performance. The remaining analytic challenge is developing ways to present these results clearly. Prior work on communicating results of hierarchical analysis points to several principles for doing so effectively (Davis and Dreyer, 2009).

First, organize the measurement hierarchy based on the decision to be made. In some contexts, CTR leadership may be most interested in balancing progress against nuclear, chemical, and biological threats. In other contexts, the focus might be geographic, to show the balance of efforts between ongoing work with states of the FSU and newer engagements in other geographic regions. The displays used to communicate analysis should be organized to most easily communicate the most salient balances and trade-offs.

Second, use the hierarchy to limit the number of comparisons. As the number of metrics grows, so do the combinations that must be compared; minds struggle to parse large numbers of combinations. The measurement hierarchy can be used to reduce the number of combinations to compare at any one level without reducing the number of factors in an evaluation.

Third, ensure that all judgments are transparent. Wherever judgments enter into the analysis, the approaches used to display results should reveal them. Common approaches include:

- including summary assessment from higher levels on subordinate-level views
- describing aggregation rules used when making summary judgments
- using labels consistently to increase traceability between levels
- using colors to indicate acceptable and unacceptable conditions
- defining clearly the thresholds and goals used for assigning these colors

Implementing an assessment framework that incorporates these approaches requires deliberative analysis to ensure that judgments about aggregation and thresholds are valid, that data management is practical, and that reporting is transparent. While the approach has been considered by others in the context of national health security and public health emergency preparedness, implementation is not trivial (Dreyer et al., 2010). An alternative, which could be used instead of this approach or while implementing it, is to use a smaller subset of metrics. For example, Chapters Three and Four, as well as the appendixes, provide priority designations for recommended metrics. This subset could be used instead of aggregating a larger number of metrics. The trade-off between measurement approaches is that it may be easier to implement a subset of the metrics but that doing so may not provide as valid or complete an assessment of CBEP performance.

Demonstrating a Measurement Hierarchy for CBEP Metrics

To illustrate how the proposed metrics for CBEP could be used at multiple levels, from country-level management to strategic-level reporting, we developed a measurement hierarchy and developed a notional program assessment for the CTR program. This section describes the notional assessment, beginning with the strategic perspective, and shows how it is ultimately fed by the proposed metrics. This illustrative example builds on the hierarchical relationships in Figure 6.1 and demonstrates how to support a strategic assessment (level 1) with a logical narrative based on metrics collected at the country level (level 5).

At the strategic level (level 1), the CTR program addresses threats from biological, nuclear, and chemical weapons. Three of the programs that constitute CTR focus specifically on these threats: CBEP, Global Nuclear Security, and Chemical Weapons Destruction. Thus, one strategic assessment would describe the extent to which these three programs are reducing threats from biological, nuclear, and chemical weapons across a set of partner countries.

Table 6.1 illustrates a level 1 program assessment for the CTR program. In this figure, the extent to which a CTR program is reducing the global threat is indicated by a scale ranging from green (very much so) to red (very little). In this notional example, it is apparent that the Chemical Weapons Destruction program is most effective in country 1 and that CBEP is moderately more effective than either of the other pro-

Table 6.1
Notional Level 1 Program Assessment for the Cooperative Threat Reduction Program

	Program		
Countries	Cooperative Biological Engagement	Global Nuclear Security	Destruction of Chemical Weapon
Country 1			
Country 2			
Country 3			

NOTE: Illustrative assessment based solely on notional data.

grams assessed. This type of top-level view may be a useful summary for a report to DoD or congressional leadership but only if accompanied by details that explain why and how each of the assessments was made.

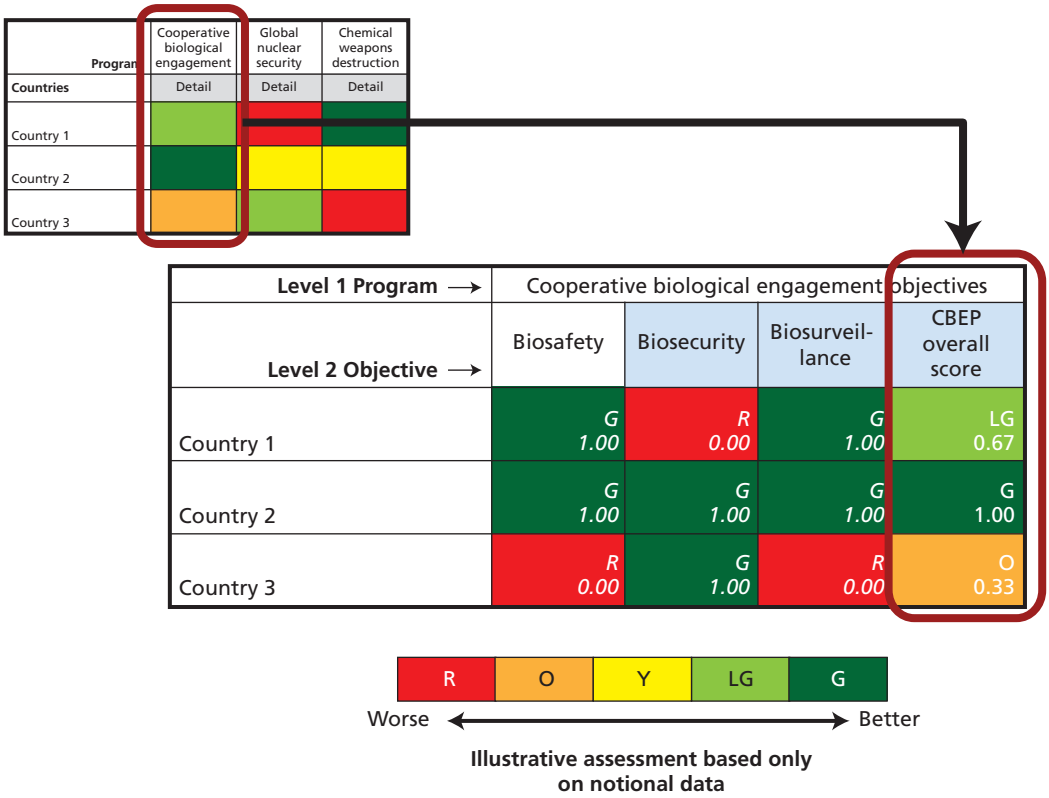
To provide the next level of detail, we use the measurement hierarchy to illustrate the basis for the level 1 assessments were made. Figure 6.2 illustrates the level 2 assessment of CBEP objectives. In this figure, the level 1 assessment is shown on level 2 as the *CBEP overall score*. This overall score represents an aggregated judgment of whether a country has demonstrated acceptable sustainable capabilities across CBEP biorisk management (i.e., biosafety and biosecurity) and biosurveillance objectives. The numbers in the column labeled *CBEP overall score* indicate that the score represents a percentage of goals met in a country across objectives. For example, country 1 was judged to have met the goals for biosafety and biosurveillance, but not for biosecurity.

The level 2 assessment in Figure 6.2 also incorporates judgments about how to aggregate scores and what constitutes acceptable performance. Thus, the hierarchy can be used once again to provide transparency into how these choices were made.

Figure 6.3 illustrates a notional level 3 assessment for biosafety in CBEP. In this case, the assessment describes how the overall assessment of a partner nation's abilities on CBEP objectives was derived from assessments of the overall state of the nation's capacities, capabilities, and sustainability enablers for an objective—in this case, biosafety.

In this example, we adopted a scoring system that attempts to reflect a capability maturity model similar to that used during a recent OSD-P and OSD-AT&L exercise to develop strategic metrics for CTR. This exercise identified criteria to rate the maturity of country's abilities for biosafety, biosecurity, and biosurveillance on a five-point scale. An immature country demonstrates little capacity, let alone capability or sustainability. As the country matures, it demonstrates more capacities and, eventually, capabilities. A fully mature country demonstrates many capabilities and sustainability. The

Figure 6.2
Notional Level 2 Program Assessment for CBEP



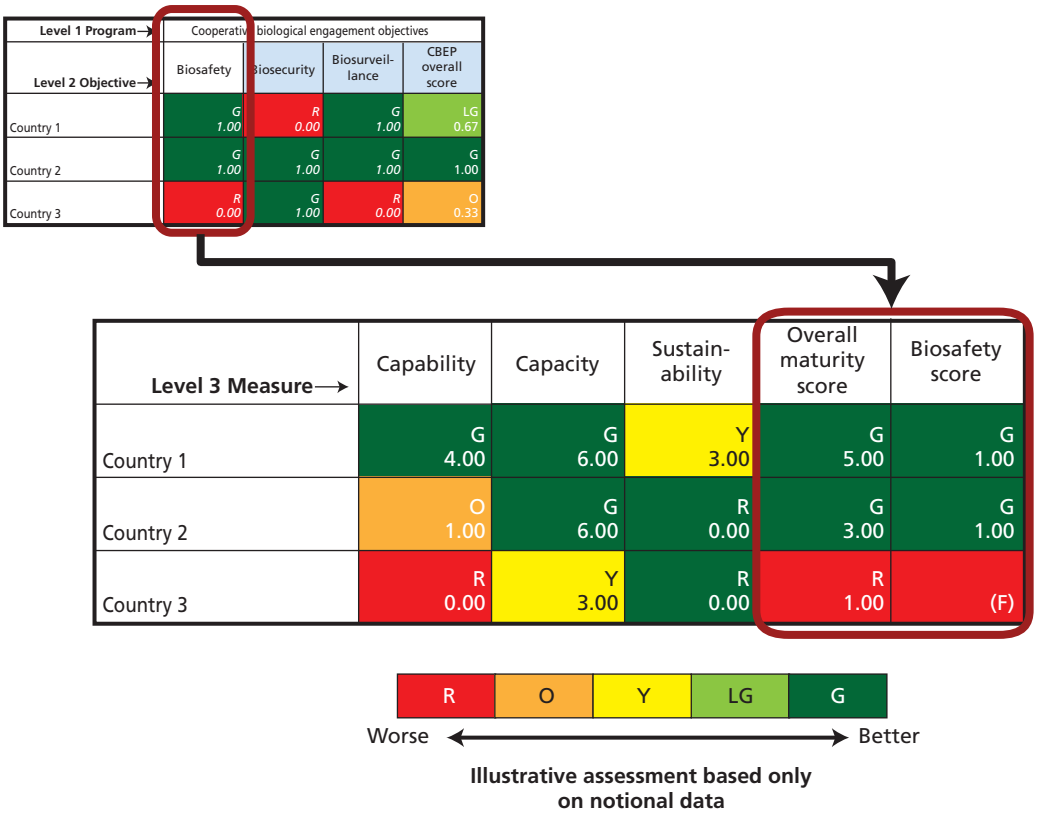
RAND *RR660-6.2*

overall level 3 maturity score reflects a notional application of this logic to the biosafety component CBEP's biorisk management objective.

The proposed biosafety logic model includes six capacities, four capabilities, and five sustainability enablers. The columns for Capacity, Capability, and Sustainability in Figure 6.3 indicate how many of each of these have been demonstrated in a country. The overall score was calculated using the logic shown in Table 6.2. To achieve a score of 3, a country must have demonstrated all the biosafety-relevant capacities. As the country develops more capabilities and implements sustainability enablers, it achieves higher maturity scores. In the notional analysis presented in Figure 6.3, we established a goal for countries to at least demonstrate all capacities and thus achieve a minimum of maturity level 3.

The obvious next question is how to determine whether a country has demonstrated a capacity, capability, or sustainability enabler. Once again, another level of hierarchy is used reveal the judgments made in the assessment. Figure 6.4 illustrates how the notional analysis of the level 3 assessment of biosafety capability is supported through a level 4 assessment of these capabilities and, ultimately, through level 5.

Figure 6.3
Notional Level 3 Program Assessment for the CBEP Biosafety Objective



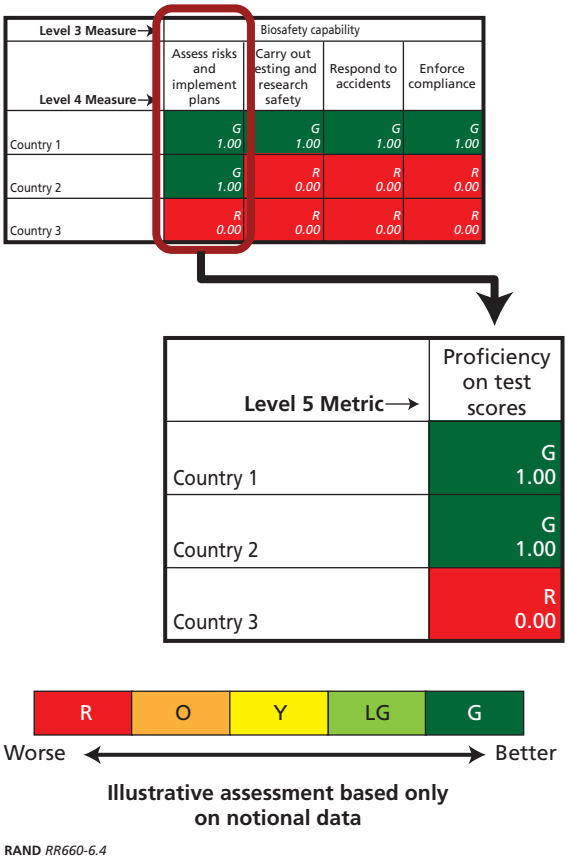
RAND *RR660-6.3*

Table 6.2
Notional Scoring Method for a CBEP Capability Maturity Model Assessment

Capability Maturity Level	Capabilities Demonstrated (no. out of 4)	Capacities Documented (no. out of 6)	Sustainability Enablers Implemented (no. out of 5)
5	4	6	≥3
4	4	6	0–2
3	1–3	6	NR
2	NR	4–5	NR
1	NR	0–3	NR

NOTE: NR = Not relevant when the maturity level is determined strictly by demonstration of capacities or capabilities.

Figure 6.4
Notional Level 4 and Level 5 Program Assessments for CBEP



In this case, the overall assessment of level 3 is simply a count of how many of the capabilities shown in level 4 were demonstrated adequately. This assessment is in turn made by comparing specific metrics (level 5) to thresholds and goals established for the country and the capability. In the notional example, this could be interpreted as the percentage of laboratories in a country that demonstrated proficiency on tests that evaluated laboratory management's ability to assess risks and implement biosafety management plans. In this notional example, countries 1 and 2 have demonstrated 100-percent performance, but no laboratories in country 3 have demonstrated this capability (Figure 6.4).

While this notional example illustrates the potential for using the proposed set of metrics to tell a story of how CBEP resources and activities ultimately support CTR goals, it also highlights some additional steps, such as establishing appropriate goals and aggregation rules, that must be taken to draw insight from this set of metrics. Chapter Six discusses these points in more detail.

CHAPTER SEVEN

Conclusions and Next Steps

We developed an evaluation framework, including metrics, for assessing the progress of CBEP activities in partner countries toward achievement of two main program objectives. This effort suggested conclusions in the specific case of CBEP, as well as broader conclusions related to using logic models as a foundation for measurement. We have provided recommendations toward implementing the framework in phases and suggested additional analyses that could help increase the effectiveness of CBEP programming.

This work has built on substantial existing efforts related to measurement that will greatly facilitate implementation. The work CBEP, CTR, and experts outside the U.S. government, such as NAS, have already done on evaluation and assessments has established a strong foundation for measurement. Because of the data it collects for MAP, for example, CBEP already collects a significant fraction of the data required for the metrics we propose here. Existing data and bases for evaluation, however, have been better developed to date for biorisk management than for biosurveillance. This means that implementing the biorisk metrics will require focusing more attention on scoring and testing the metrics than on data collection. Despite such strides, implementing a new evaluation framework will require additional time, effort, and resources. CBEP can mitigate this burden by phasing in the implementation or by establishing priorities, as we discuss in Chapter Six and further develop in the appendixes.

Development of this evaluation framework for CBEP also allowed us to make some broader observations about logic models as a foundation for management. The recommended framework facilitates the use of the assessments of and communications on the program to decisionmakers at several levels, from in-country project managers to central program managers and up to strategic-level managers inside and beyond DoD. The proposed evaluation framework and associated metrics put CBEP at the leading edge of DoD's efforts to monitor program performance. The recommended framework can provide CBEP with new and valuable tools for program monitoring. Some observations about the approach include the following:

- **Logic models provide a framework for evaluating and communicating program performance.** Logic models provide CBEP a systematic means of captur-

ing program activities and outcomes in support of enduring and demonstrable partner capabilities. Such a framework supports measurement and reporting on program performance and communicating CBEP activities and impacts.

- **The functional frameworks for each program objective provide another approach to selecting and communicating metrics information.** As compared to the relatively abstract representation of the logic models, an alternative, functional representation of metrics for each CBEP objective proved to be a more intuitive way to "tell the story" and facilitated choices about implementation.
- **The key distinction between *capacities* and *capabilities* allows CBEP to ensure that it is focusing measurement on outputs that are closer to the desired outcomes.** We use these terms to distinguish demonstrated partner operational abilities (capabilities) from the building blocks that enable them (capacities). Distinguishing between capacities and capabilities was one of the elements of the logic models that most resonated with CBEP personnel.
- **Identification of *sustainability enablers* allows CBEP to focus CBEP programming *now* concretely on *future* sustainability.** In addition to distinct metrics for capacities and capabilities, we have proposed metrics that specifically look at progress toward the longer-term sustainability of CBEP investments in partner nations (sustainability enablers). By specifically highlighting sustainability enablers, the framework better equips CBEP to address explicitly in program activities and track progress in an area vital to ultimate program success, rather than leaving sustainability as a wishful byproduct of program activities to be measured only at the end of the engagement.

Recommendations

We recommend that CBEP take steps to refine and implement the metrics framework to support internal evaluations and external reporting on program impacts. CBEP is currently advancing efforts to improve the validity of its biorisk management metrics. We think this is a valuable way to strengthen existing data while minimizing additional time and resources. We recommend that CBEP continue to develop and field scoring criteria for MAP implementers and work to identify other opportunities to improve existing data.

To refine the framework, by identifying implementation challenges and mitigation strategies, we recommend pilot testing the recommended metrics in a small set of CBEP partner countries. This would test that the data sources and collection approaches outlined in the appendix actually do support the recommended metrics and that approaches to scoring operate as intended. The pilot effort will yield an evaluation of the selected partners with respect to capacities, capabilities, and sustainability enablers based on the evaluation framework and metrics described here. For CBEP,

this would result in a tried and tested framework and metrics to use broadly to assess performance toward program goals. This approach would allow CBEP to identify challenges and opportunities for refinement in advance of widespread implementation.

A second important element of pilot testing will be formalizing the structure for aggregating metrics to support strategic-level reporting. As discussed in Chapter Six, metrics can be used to support management and reporting at several levels, but doing so requires deliberation and analysis. Implementation of an approach such as the maturity model in Chapter Six can help CBEP tailor reporting to the specific needs of a range of decisionmakers, with transparency and accountability. Doing so, however, will require programmatic and policy expertise, deliberation, and analysis to specify the desired standards and appropriate rules for defining hierarchical relationships.

We also recommend that CBEP consider moving evaluation efforts forward by focusing on assessing program outcomes in areas key to ultimate program success. As CBEP expands its program activities to new geographic regions, it should build on lessons learned from more than a decade of experience partnering with countries in the FSU. This history positions CBEP well to evaluate the outcomes of programming in for example, CBR and other elements of health diplomacy activities, as well as the ultimate sustainability of CBEP investments. Evaluations of lessons learned and best practices would position CBEP to support desired outcomes by improving alignment of program activities with program objectives.

APPENDIX A

Implementation Information for Proposed Metrics—Aligned to Logic Model Framework

The tables in this appendix provide information to support the implementation of the metrics framework, aligning each metric with the logic model structure. There are two sets of tables, one for sources and descriptions and one for scoring and metrics. Each set contains individual tables for biosafety, biosecurity, and biosurveillance.

Note that, in the tables, italics indicate that the language was developed by RAND.

Sources and Descriptions

Tables A.1 through A.3 provide practical information necessary for utilizing existing data and collecting additional data to implement the evaluation framework.

Scoring and Metrics

Tables A.4 through 4.6 provide practical information on using the proposed metrics to evaluate CBEP performance.

Table A.1
Sources and Descriptions of Biosafety Metrics—Logic Model Framework

Metric	Potential Source	Comments
Capabilities		
(A.3.1) *Proficiency test scores for CBEP courses related to biorisk assessment and planning (example courses in backup)*	CBEP implementer responsible for MAP reporting	Test scores are current MAP inputs (before the course begins and after it ends, by individual with specified position title)
(A.3.2) Percentage completion: Biosafety: work practice and administrative control (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs
(A.3.4) SME finds that a regular review of the BRM system exists	CBEP implementer responsible for MAP reporting	Current MAP input
(A.3.4) SME finds that mechanisms for oversight, enforcement, and attribution for biosafety and biosecurity regulations and/or guidelines exist	CBEP implementer responsible for MAP reporting	Proposed MAP input (CBEP BRM team–proposed a country-level criterion)
(A.3.3) SME finds that accident or incident and nonconformities related to biorisk correctly managed	CBEP implementer responsible for MAP reporting	Current MAP input
Sustainability enablers		
(A.2.1) Has a policy concerning the management of laboratory biorisk (biosafety and biosecurity) been written?	CBEP implementer responsible for MAP reporting	Current MAP input
(A.2.1) Are there regulations and/or guidelines for biosafety and biosecurity?	CBEP implementer responsible for MAP reporting	Current MAP input
(A.2.1) Are there national incident management systems for naturally occurring biological events? (These include policies, frameworks, MOUs for laboratory, health security and law enforcement, ER sectors.)	CBEP implementer responsible for MAP reporting	Proposed MAP input (CBEP BRM team–proposed a country-level criterion)

Table A.1—Continued

Metric	Potential Source	Comments
(A.2.2) *SME finds that operational plans include materiel sustainment considerations*	CBEP	CBEP team responsible for developing implementation plan or conducting tabletop exercises with partner government personnel
(A.2.2) *Percentage of required equipment that is domestically or regionally sourced*	Relevant partner governments	
(A.2.3) *Number of national or regional institutions that make BS&S training to international standards available (training centers, universities, professional societies, ministries)*	Relevant partner governments	
(A.2.4) *Number of publications; number of conference presentations*	CBEP science team	Count of CBEP abstracts, posters, presentations, and publications resulting from CBR projects (CBEP science team currently tracks these)
(A.2.4) *Number or value of internationally competitive research grants won*	CBEP science team; relevant partner governments	
(A.2.5) *SME finds that operational plans include resource sustainment considerations*	CBEP	CBEP team responsible for developing implementation plan or conducting tabletop exercises with partner government personnel
(A.2.5) SME finds adequate availability of funding to support biosafety and biosecurity programs and initiatives	CBEP implementer responsible for MAP reporting	Proposed MAP input (CBEP BRM team–proposed a country-level criterion)
Capacities		
(A.1.1) Percentage completion: BRM assessment, requirements, and planning of facilities (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs
(A.1.2) Percentage completion: BRM management personnel in place (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs

Table A.1—Continued

Metric	Potential Source	Comments
(A.1.6) Percentage of personnel trained in laboratory biosafety and biosecurity relative to requirements	CBEP	Required trained personnel: CBEP team responsible for developing implementation plan or conducting tabletop exercises with partner government personnel Actual trained personnel: CBEP implementer responsible for MAP reporting
(A.1.5) Percentage complete BS&S equipment (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs
(A.1.4) SME finds that facilities are designed to allow employees to work in a safe and secure way	CBEP implementer responsible for MAP reporting	Current MAP input
(A.1.3) Number or percentage of facilities with complete set of relevant SOPs in place (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs

Table A.2
Sources and Descriptions of Biosecurity Metrics—Logic Model Framework

Metric	Potential Source	Comments
Capabilities		
(B.3.1) *Proficient test scores for CBEP courses related to biorisk assessment and planning (example courses in backup)*	CBEP implementer responsible for MAP reporting	Test scores are current MAP inputs (before the course begins and after it ends, by individual with specified position title).
(B.3.2) SME finds that collections of dangerous pathogens have been consolidated into a minimum number of facilities	CBEP implementer responsible for MAP reporting	Proposed MAP input (CBEP BRM team proposed a country-level criterion).
(B.3.2) Pathogen consolidation (TT criteria and tiers)	TT	Metric straightforward as written and seems useful.
(B.3.2) *Percentage of disclosed biological weapon–related infrastructure that has been eliminated*	Relevant partner governments	
(B.3.3) Percentage completion: Maintain control of pathogens (physical security, material control and accountability, information security, transportation security) (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs.
(B.3.4) SME finds that mechanisms are in place sufficient to ensure that personnel are competent and reliable	CBEP implementer responsible for MAP reporting	Current MAP input.
(B.3.6) SME finds that accident or incident and nonconformities related to biorisk correctly managed	CBEP implementer responsible for MAP reporting	Current MAP input.
(B.3.7) SME finds that a regular review of the BRM system exists	CBEP implementer responsible for MAP reporting	Current MAP input.
(B.3.7) SME finds that mechanisms for oversight, enforcement, and attribution for biosafety and biosecurity regulations and/or guidelines exist	CBEP implementer responsible for MAP reporting	Proposed MAP input (CBEP BRM team proposed a country-level criterion).

Table A.2—Continued

Metric	Potential Source	Comments
(B.3.7) Reporting in 1540 matrix affirming existence of a mechanism for penalizing violator of national prohibition on engagement in biological weapon–related activities	Relevant partner governments; United Nations Security Council	1540 compliance is reported in 1540 Matrices (UN Security Council, 2005).
Sustainability enablers		
(B.2.1) Has a policy concerning the management of laboratory biorisk (biosafety and biosecurity) been written?	CBEP implementer responsible for MAP reporting	Current MAP input.
(B.2.1) Are there regulations and/or guidelines for biosafety and biosecurity?	CBEP implementer responsible for MAP reporting	Current MAP input.
(B.2.1) Are there National Incident Management Systems for intentional biological events? (policies, frameworks, and MOUs for laboratory, health security, and law enforcement and emergency response sectors)	CBEP implementer responsible for MAP reporting	Proposed MAP input (CBEP BRM team proposed a country-level criterion).
(B.2.2) *SME finds that operational plans include materiel sustainment considerations*	CBEP	CBEP team responsible for developing implementation plan or conducting tabletop exercises with partner government personnel.
(B.2.2) *Percentage of required equipment that is domestically or regionally sourced*	Relevant partner governments	
(B.2.3) *Number of national or regional institutions that make BS&S training to international standards available (training centers, universities, professional societies, ministries)*	Relevant partner governments	
(B.2.4) *Number of publications; number conference presentations*	CBEP science team	Count of CBEP abstracts, posters, presentations, and publications resulting from CBR projects (CBEP science team currently tracks these).
(B.2.4) *Number or value of internationally competitive research grants won*	CBEP science team; relevant partner governments	

Table A.2—Continued

Metric	Potential Source	Comments
(B.2.5) *SME finds that operational plans include resource sustainment considerations*	CBEP	CBEP team responsible for developing implementation plan or conducting tabletop exercises with partner government personnel.
(B.2.5) SME finds adequate availability of funding to support biosafety and biosecurity programs and initiatives	CBEP implementer responsible for MAP reporting	Proposed MAP input (CBEP BRM team proposed a country-level criterion).
Capacities		
(B.1.1) Percentage completion: BRM assessment, requirements, and planning of facilities (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs.
(B.1.2) Percentage completion: BRM management personnel (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs.
(B.1.4) SME finds that facilities are designed to allow employees to work in a safe and secure way	CBEP implementer responsible for MAP reporting	Current MAP input.
(B.1.5) Percentage completion: BS&S equipment (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs.
(B.1.6) Percentage of personnel trained in laboratory biosafety and biosecurity relative to requirements	CBEP	Required trained personnel: CBEP team responsible for developing implementation plan or conducting tabletop exercises with partner government personnel. Actual trained personnel: CBEP implementer responsible for MAP reporting.

NOTES: While most of these are quantitative, all can include additional compelling (qualitative) anecdotes. TT = tiger team (indicators, criteria, and/or tiers defined for biorisk management).

Table A.3
Sources and Descriptions of Biosurveillance Metrics—Logic Model Framework

Metric	Potential Source	Comments
Enduring capabilities		
(C.3.1) *Demonstrated performance: at least three out of five defined core syndromes*	CDC	This is U.S. government interagency indicator; could consult with CDC regarding details
(C.3.1) *Surveillance analysis performance—exercise or supervisory observation (Near-term metric could employ TT tiers, recommend in longer-term working toward introduction of metric based on CDC/CSTE assessment forms)*	CDC FETP, CSTE Applied Epidemiology Competencies	CSTE Applied Epidemiology Competencies documentation provides competencies, job descriptions and assessment forms for four tiers of epidemiologists, from basic- or entry-level to midlevel and senior supervisory; relevance of different tiers may vary by CBEP country and over time Assessment form (self-assessment and supervisory) for Tier 1 (entry-level epidemiologists) (CDC, undated a) and Assessment form for Tier 2 (midlevel epidemiologists) (CDC, undated b).
(C.3.3) *Number of suspected priority pathogen cases or outbreaks in the past 12 months; percentage for which investigation was conducted and results documented*	MOH and laboratory log	Note that denominator (suspected cases) is critical to assessing this indicator, since the metric is proportion (not simply absolute number) of cases or outbreaks that were properly investigated and documented. Laboratory logs (central, provincial or state) should provide denominator data, and MOH provides numerator data (number of cases investigated and documented).
(C.3.3) *Epidemiological investigation performance (written report or tested via exercise)*	CDC FETP	CDC has extensive documentation on the steps involved in epidemiological investigation (e.g., CDC, 2004); consult with CDC as needed. WHO also has at least one checklist that can be used as a tool (e.g., WHO, 2006b, p. 18).

Table A.3—Continued

Metric	Potential Source	Comments
(C.3.4) *Number and percentage of laboratories participating in proficiency testing at least once in the past 12 months*	APHL, CDC, WHO	We believe there are several laboratory proficiency testing programs around the world, including government agencies (e.g., CDC), laboratory groups (e.g., APHL), perhaps the Joint Commission International, and relevant WHO collaborating centers (of which CDC has a number). There are also commercial firms that conduct proficiency testing.
(C.3.4) *Number and percentage of laboratories that passed all proficiency tests in the past 12 months, among laboratories participating in the past 12 months*	APHL, CDC, WHO	See cell immediately above.
(C.3.4) *Number and percentage of laboratories that meet specified standards (mutually agreed on by country and CBEP)*	APHL, CDC, DoD	Consultations with DoD laboratory specialists and/or APHL or CDC to help develop reasonable standards appropriate for various levels of laboratory and country development.
(C.3.4) *Specimen collection transport: number of specimens received by laboratory; number and percentage adequate for testing (logs, exercise; TT tiers)*	Laboratory log	Laboratory logs should be maintained to track specimens received and tested; logs should include specimens that were not suitable for testing, which reflects inadequate collection or transport.
(C.3.4) *Laboratory timeliness: number of hours following receipt of specimens that laboratory testing is (a) initiated, (b) completed (from laboratory logs or exercise)*	Laboratory log or exercise	Laboratory specimens are typically accompanied with brief clinical information (clinical diagnosis, date of onset), which can be captured for this metric.
(C.3.4) *Number of laboratory tests performed for each priority pathogen in the past 12 months*	Laboratory log	
(C.3.4) *Laboratory referral network: number of specimens sent or received for confirmatory testing in the past 12 months (a) in country and (b) international laboratory*	Laboratory log	
(C.3.5) *Supervised demonstration of electronic reporting system (define tiers)*	To be determined	Unless CBEP has already developed a test for demonstrating electronic reporting (e.g., for EIDSS), a test would need to be developed, with five tiers for scoring or some other scoring schema.

Table A.3—Continued

Metric	Potential Source	Comments
(C.3.5) *Number of cases of internationally reportable pathogens or diseases detected in country in the past 12 months; percentage reported to appropriate international authority (WHO, OIE)*	MOH, laboratory log	The needed denominator data (number of reportable cases) may be elusive in countries where transparency is an issue (as CBEP staff noted for FSU), but denominator is critical to really assess this metric, which reflects degree of compliance rather than just number of reports per se.
(C.3.5) Biosurveillance (or CBR) directly informed decisions (e.g., policy, clinical practice)	MOH	Qualitative. This metric is very important but does not lend itself to strictly quantitative measurement.
Sustainability enablers		
(C.2.1) Regulatory environment (TT criteria and tiers)	TT	Metric is straightforward as written and seems useful.
(C.2.1) *Degree of consistency between national and CBEP planning and implementation*		This reflects a critical sustainability-enabling factor; external programming aligns with national planning, rather than vice versa (described for some new engagement CBEP countries, for example).
(C.2.2) *Number of peer-reviewed publications, number of conference papers (abstract, poster, or oral)*	Web of Science, Scopus, or other bibliometric database	These sources have different capabilities and reach (they do not identify the same items, although there is significant overlap); the Web of Science includes very impressive bibliometric analysis capabilities (e.g., can compare countries, including by research area).
(C.2.2) Number of institutions with epidemiology core curriculum meeting international (e.g., FETP) standards; number of trainees per year	TEPHINET, CDC FETP, relevant partner countries	Global network of "FETPs" = TEPHINET. TEPHINET, CDC/FETP, and/or FETP countries undoubtedly have and/or have defined basic core curricula.
(C.2.2) *Career track for trained epidemiologists*	CDC FETP	This is a requirement before establishing CDC FETP in a country. Consult with CDC regarding how to make this happen and what "counts" for tracking purposes.
(C.2.3) *Trained and funded internal staff or active contracts or vendor agreements for life cycle equipment maintenance*	TT or CBEP	Consider a TT-like five-tier scoring scheme for this metric.

Table A.3—Continued

Metric	Potential Source	Comments
(C.2.4) Operational laboratory certification program	MOH; DoD; ISO; JCI	As above for C.1.8 capacity metric, consult with relevant experts to define metric in more detail.
(C.2.5) *Number and percentage of priority pathogens for which country (a) produces or (b) can sustainably procure needed media, reagents*		See cell immediately above.
Capacities		
(C.1.1) Number and percentage of priority pathogens with (a) established baseline and (b) a baseline that is used to detect anomalies (modified TT tiers)	MOH, laboratory, CBEP (CBR)	Quantitative: Most CBEP CBR projects appear to contribute to baseline information for relevant pathogens; CBEP program staff probably can identify pathogens for which baseline information has been collected. Qualitative: It might be more difficult, but is very important, to capture how biosurveillance and related CBR data are actually used (e.g., as a baseline trend for assessment of potential outbreaks; to guide clinical or public health policy).
(C.1.2) *Performance—observation or standardized clinical scenarios (and TT tiers)*	Clinical board examinations	U.S. clinical board examinations involve descriptions and photos of clinical cases for differential diagnostic and other purposes; these could be used to develop appropriate scenarios for testing.
(C.1.3) Reporting capability (TT criteria and five tiers)	TT	Straightforward as written; not highly precise, so should be supplemented by other metrics.
(C.1.3) *Number and percentage of jurisdictions using electronic reporting system (can also define five tiers)*	MOH	
(C.1.3) *Accuracy, scope: national guidance, e.g., case definitions for animal and human disease*	MOH	There are several sources for relevant case definitions (e.g., AFHSC for DoD reportable medical events); scoring could use TT-like five tiers from "no case definitions" to "clear ones for all relevant pathogens."

Table A.3—Continued

Metric	Potential Source	Comments
(C.1.5) Number of trained field epidemiologists and number per 200,000 population	MOH	Set criteria for what "counts"—probably at least a one-year certificate or degree program; consult with CDC FETP. Numbers should increase over time. U.S. government interagency standard of one per 200,000 might be somewhat arbitrary, but metric is useful to track growing capacity.
(C.1.6) *Number of specified level jurisdictions; number and percentage number and percentage with functional community-based surveillance*	MOH or TBD	
(C.1.7) Number of trained and drilled teams able to respond to outbreak within 24 hours	MOH	Suggest using (or developing) standard definition such as that used by WHO and others for rapid-response teams (e.g., Mekong Basin Disease Surveillance group has a definition they use for its rapid-response teams).
(C.1.8) Technical capacity (TT criteria and tiers)	TT	Straightforward as written; not highly precise, so should be supplemented by other metrics.
(C.1.8) Number and percentage of laboratories certified	MOH; DoD; ISO; JCI	In consultation with relevant laboratory specialists (DoD or other), identify relevant certification standards and body; could vary by country.
(C.1.8) *Number and list of pathogens for which (a) national and (b) each provincial or state laboratory can test*	MOH	Metric reflects "reach" of relevant laboratory testing throughout the country.
(C.1.8) *Laboratory networking: Written or established protocols for specimen referral (a) within country and (b) to international laboratory*	MOH	Metric reflects national guidance and should be a precursor to capabilities related to specimen collection and transport (e.g., capability C.3.4).

NOTES: While most of these are quantitative, all can include additional compelling (qualitative) anecdotes. TT = tiger team (indicators, criteria, and/or tiers defined for biosurveillance).

Table A.4
Proposed Scoring and Standards for Biosafety Metrics—Logic Model Framework

Metric	Overall Priority	Scoring and Standard or Target
Capabilities		
(A.3.1) *Proficient test scores for CBEP courses related to biorisk assessment and planning (example courses in backup)*	Medium to high	Number evaluated and percentage scoring at least 70 percent on (specified) CBEP tests related to biorisk assessment and planning (MAP input for posttest score).
(A.3.2) Percentage completion: Biosafety: work practice and administrative control (MAP checklist)	High	Facility level—Number evaluated and percentage reporting *yes* to (specified) MAP inputs related to biosafety: work practice and administrative control. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.
(A.3.3) SME finds that accident or incident and nonconformities related to biorisk correctly managed	High	Facility level—Yes-or-no answer to MAP input: "Are accident or incident and nonconformities related to biorisk correctly managed (i.e., reported, recorded, investigated, and leading to preventive or corrective actions)?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.
(A.3.4) SME finds that a regular review of the BRM system exists	High	Facility level—Yes-or-no answer to MAP input: "Is there a regular review of the biorisk management system?" National level—Percentage of CBEP-engaged facilities answering *yes* to MAP input: "Is there a regular review of the biorisk management system?"
(A.3.4) SME finds that mechanisms for oversight, enforcement, and attribution for biosafety and biosecurity regulations and/or guidelines exist	High	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Are there mechanisms for oversight, enforcement and attribution for biosafety and biosecurity regulations and/or guidelines?"

Table A.4—Continued

Metric	Overall Priority	Scoring and Standard or Target
Sustainability enablers		
(A.2.1) has a policy concerning the management of laboratory biorisk (biosafety and biosecurity) been written?	High	Facility level—Yes-or-no answer to MAP input: "Has a policy concerning the management of laboratory biorisk (biosafety and biosecurity) been written?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.
(A.2.1) Are there regulations and/or guidelines for biosafety and biosecurity?	Medium	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Are there regulations and/or guidelines for biosafety and biosecurity?"
(A.2.1) Are there national incident management systems for naturally occurring biological events? (policies, frameworks, and MOUs for laboratory, health security, and law enforcement and emergency response sectors)	Medium to high	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Are there national incident management systems for naturally occurring biological events? (policies, frameworks, and MOUs for laboratory, health security and law enforcement and ER sectors)?"
(A.2.5) *SME finds that operational plans include resource sustainment considerations*	Medium to high	Percentage of operational plans describing planned delivery of CBEP-furnished facilities, equipment, and training that include plans for partner resourcing sustainment of infrastructure, materiel, and human capital.
(A.2.5) SME finds adequate availability of funding to support biosafety and biosecurity programs and initiatives	High	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Is there adequate availability of funding to support biosafety and biosecurity programs and initiatives?"
(A.2.3) *Number of national or regional institutions that make BS&S training to international standards available (training centers, universities, professional societies, ministries)*	Medium to high	Five potential tiers: (1) no access to BS&S training; (2) access to short-term BS&S training in another country; (3) access to long-term (2-year) training in another country; (4) country has at least one institution that offers short-term training; (5) country has at least one institution or program that provides long-term training.

Table A.4—Continued

Metric	Overall Priority	Scoring and Standard or Target
(A.2.4) *Number of publications; number of conference presentations*	Medium	Simple counts; can be disaggregated by institution, scientist, research area; indexed scores also available via Web of Science.
(A.2.4) *Number or value of internationally competitive research grants won*	Medium	Simple counts; can be disaggregated by institution, scientist, research area.
(A.2.2) *SME finds that operational plans include materiel sustainment considerations*	Medium to high	Percentage of operational plans describing planned delivery of CBEP-furnished facilities and equipment that include plans for partner sourcing of material and expertise required for sustainment of infrastructure and material.
(A.2.2) *Percentage of required equipment that is domestically or regionally sourced*	Medium	(1) (Specified) percentage of equipment used to support partner's BRM program that comes from a domestic or regional source. (2) Percentage remaining that needs to be acquired from nondomestic and nonregional sources is within partner's ability or willingness to resource.
Capacities		
(A.1.2) Percentage complete of BRM management personnel (MAP checklist)	Medium to high	Facility level—Number evaluated and percent reporting *yes* to (specified) MAP inputs related to BRM management personnel. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.
(A.1.6) Percentage of personnel trained in laboratory biosafety and biosecurity relative to requirements	High	(1) Number of personnel trained required, by position and subject area, based on CBEP and partner operational plans. (2) Number of individuals trained in each position and each subject area (MAP inputs for student identification number, position, date course completed).

Table A.4—Continued

Metric	Overall Priority	Scoring and Standard or Target
(A.1.1) Percentage completion: BRM assessment, requirements, and planning of facilities (MAP checklist)	High	Facility level—Number evaluated and percentage reporting *yes* to (specified) MAP inputs related to BRM assessment, requirements, planning. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.
(A.1.5) Percentage completion: BS&S equipment (MAP checklist)	High	Facility level—Number evaluated and percentage reporting *yes* to (specified) MAP inputs related to BS&S equipment. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.
(A.1.4) SME finds that facilities are designed to allow employees to work in a safe and secure way	High	Facility level—Yes-or-no answer to MAP input: "Are the facilities designed to allow to work in a safe and secure way?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.
(A.1.3) Number and percentage of facilities with complete set of relevant SOPs in place (MAP checklist)	High	Facility level—Number evaluated and percentage reporting *yes* to (specified) MAP inputs related to relevant SOPs in place. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.

Table A.5
Proposed Scoring and Standards for Biosecurity Metrics—Logic Model Framework

Metric	Overall Priority	Scoring and Standard or Target
Capabilities		
(B.3.2) Pathogen consolidation (TT criteria and tiers)	High	TT five tiers for "Assessment (Consolidation)" for annual report to Congress.
(B.3.2) *Percentage of disclosed biological weapon–related infrastructure that has been eliminated*	Medium to high	(1) Count of biological weapon–related infrastructure targeted for elimination. (2) Count of above that has been eliminated.
(B.3.3) Percentage completion: Maintain control of pathogens (physical security, material control and accountability, information security, transportation security) (MAP checklist)	High	Facility level—Number evaluated and percent reporting *yes* to (specified) MAP inputs related to capabilities for maintenance of control of pathogens. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.
(B.3.1) *Proficient test scores for CBEP courses related to biorisk assessment and planning (example courses in backup)*	Medium to high	Number evaluated and percent scoring at least 70 percent on (specified) CBEP tests related to biorisk assessment and planning (MAP input for "post-test score").
(B.3.2) SME finds that collections of dangerous pathogens have been consolidated into a minimum number of facilities	Medium to high	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Have collections of dangerous pathogens been consolidated into a minimum number of facilities?"
(B.3.4) SME finds that mechanisms are in place sufficient to ensure that personnel are competent and reliable	High	Facility level—Yes-or-no answer to MAP input: "Is there mechanism/s to ensure that personnel is competent and reliable (e.g., successful completion of training, ability to perform tasks under supervision)?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.

Table A.5—Continued

Metric	Overall Priority	Scoring and Standard or Target
(B.3.6) SME finds that accident or incident and nonconformities related to biorisk correctly managed	High	Facility level—Yes-or-no answer to MAP input: "Are accident or incident and nonconformities related to biorisk correctly managed (i.e., reported, recorded, investigated, and leading to preventive or corrective actions)?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.
(B.3.7) SME finds that a regular review of the BRM system exists	High	Facility level—Yes-or-no answer to MAP input: "Is there a regular review of the biorisk management system?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.
(B.3.7) SME finds that mechanisms for oversight, enforcement, and attribution for biosafety and biosecurity regulations and/or guidelines exist	High	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Are there mechanisms for oversight, enforcement and attribution for biosafety and biosecurity regulations and/or guidelines?"
(B.3.7) Reporting in 1540 matrix affirming existence of a mechanism for penalizing violator of national prohibition on engagement in biological weapon–related activities	High	Percentage of relevant areas (e.g., manufacture, acquire, possess, stockpile, transport) for which a country reports *yes* to question "Does national legislation exist which prohibits persons or entities to engage in one of the following activities? Can violators be penalized?" in the 1540 matrix.
Sustainability enablers		
(B.2.1) Are there national incident management systems for intentional biological events? (policies, frameworks, and MOUs for laboratory, health security, and law enforcement and ER sectors)	Medium to high	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Are there national incident management systems for intentional biological events? (policies, frameworks, and MOUs for laboratory, health security, and law enforcement and ER sectors)?"

Table A.5—Continued

Metric	Overall Priority	Scoring and Standard or Target
(B.2.3) *Number of national or regional institutions that make BS&S training to international standards available (training centers, universities, professional societies, ministries)*	Medium to high	Five potential tiers: (1) no access to BS&S training; (2) access to short-term BS&S training in another country; (3) access to long-term (2-year) training in another country; (4) country has at least one institution that offers short-term training; (5) country has at least one institution or program that provides long-term training.
(B.2.1) Has a policy concerning the management of laboratory biorisk (biosafety and biosecurity) been written?	High	Facility level—Yes-or-no answer to MAP input: "Has a policy concerning the management of laboratory biorisk (biosafety and biosecurity) been written?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.
(B.2.1) Are there regulations and/or guidelines for biosafety and biosecurity?	Medium	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Are there regulations and/or guidelines for biosafety and biosecurity?"
(B.2.5) *SME finds that operational plans include resource sustainment considerations*	Medium to high	Percentage of operational plans describing planned delivery of CBEP-furnished facilities, equipment, and training that include plans for partner resourcing sustainment of infrastructure, materiel, and human capital.
(B.2.5) SME finds adequate availability of funding to support biosafety and biosecurity programs and initiatives	High	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Is there adequate availability of funding to support biosafety and biosecurity programs and initiatives?"
(B.2.4) *Number of publications; number of conference presentations*	Medium	Simple counts; can be disaggregated by institution, scientist, research area; indexed scores also available via Web of Science.
(B.2.4) *Number or value of internationally competitive research grants won*	Medium	Simple counts; can be disaggregated by institution, scientist, research area.

Table A.5—Continued

Metric	Overall Priority	Scoring and Standard or Target
(B.2.2) *SME finds that operational plans include materiel sustainment considerations*	Medium to high	Percentage of operational plans describing planned delivery of CBEP-furnished facilities and equipment that include plans for partner sourcing of material and expertise required for sustainment of infrastructure and material.
(B.2.2) *Percentage of required equipment that is domestically or regionally sourced*	Medium	(1) (specified) Percentage of equipment used to support partner's BRM program that comes from a domestic or regional source. (2) Percentage remaining that needs to be acquired from nondomestic and nonregional sources is within partner's ability or willingness to resource.
Capacities		
(B.1.1) Percentage completion: BRM assessment, requirements, and planning of facilities (MAP checklist)	High	Facility level—Number evaluated and percent reporting *yes* to (specified) MAP inputs related to BRM assessment, requirements, planning National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.
(B.1.2) Percentage completion: BRM management personnel (MAP checklist)	Medium to high	Facility level—Number evaluated and percent reporting *yes* to (specified) MAP inputs related to BRM management personnel. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.
(B.1.6) Percentage of personnel trained in laboratory biosafety and biosecurity relative to requirements	High	(1) Number of personnel trained required, by position and subject area, based on CBEP/partner operational plans. (2) Number of individuals trained in each position and each subject area (MAP inputs for student identification number, position, date course completed).

Table A.5—Continued

Metric	Overall Priority	Scoring and Standard or Target
(B.1.4) SME finds that facilities are designed to allow employees to work in a safe and secure way	High	Facility level—Yes-or-no answer to MAP input: "Are the facilities designed to allow to work in a safe and secure way?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.
(B.1.5) Percentage completion: BS&S equipment (MAP checklist)	High	Facility level—Number evaluated and percent reporting *yes* to (specified) MAP inputs related to BS&S equipment. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.

NOTE: TT = tiger team (indicators, criteria, and/or tiers defined for biorisk management).

Table A.6
Proposed Scoring and Standards for Biosurveillance Metrics—Logic Model Framework

Metric	Overall Priority	Scoring and Standard or Target
Enduring capabilities		
(C.3.1) *Demonstrated performance: at least three out of five defined core syndromes*	Medium	Number evaluated and percentage of (a) clinicians and (b) veterinarians who can identify at least three of the five syndromes on a written test; overall assessment—five tiers, in quintiles. TT five tiers for "identify relevant diseases and syndromes."
(C.3.1) *Surveillance analysis performance—exercise or supervisory observation (Near-term metric could employ TT tiers, recommend in longer-term working toward introduction of metric based on CDC/CSTE)*	High	Initially, TT five tiers for "capability to analyze surveillance data to inform public health decisions." à More detailed CDC/CSTE assessment forms; overall scoring: number assessed by level of epidemiologist, percentage passing, five tiers, in quintiles.
(C.3.3) *Number of suspected priority pathogen cases or outbreaks in the past 12 months, and percentage for which investigation was conducted, results documented*	Medium	Scoring can be simple percentage (of detected cases or outbreaks that were investigated and documented), or in five tiers, in quintiles.
(C.3.3) *Epidemiological investigation performance (written report or tested via exercise)*	Medium to high	Depends on assessment tool (e.g., checklist of investigation steps followed and documented in written report), but can score in five tiers by quintile or simple percentage of those scoring 85–100 percent on skills and knowledge checklist (e.g., as used by USAID).
(C.3.4) *Number and percentage of laboratories participating in proficiency testing at least once in the past 12 months*	Medium to high	Five potential tiers: (1) No laboratories participate in proficiency testing and no intention to do so; (2) no laboratories participate, but country plans to do so; (3) central (or other) laboratory participated in the past 12 months; (4) central and at least one subnational laboratory participated; (5) central and most or all major subnational jurisdictional laboratories participated.
(C.3.4) *Number and percentage of laboratories that passed all proficiency tests in the past 12 months, among laboratories participating in the past 12 months*	Medium to high	Five potential tiers, in quintiles.

Table A.6—Continued

Metric	Overall Priority	Scoring and Standard or Target
(C.3.4) *Number and percentage of laboratories that meet specified standards (mutually agreed on by country and CBEP)*	Medium to high	Five potential tiers, in quintiles.
(C.3.4) *Specimen collection, transport: number specimens received by laboratory; number and percentage adequate for testing*	Medium to high	TT five tiers for "capability to collect and analyze samples" Overall percent of specimens received at (a) central and (b) provincial laboratories that were adequate for testing (logs or exercise); five tiers, in quintiles.
(C.3.4) *Number of hours following receipt of specimens that laboratory testing is (a) initiated, (b) completed (from laboratory logs or exercise)*	Medium	
(C.3.4) Number of laboratory tests performed for each priority pathogen in the past 12 months	High	Five potential tiers: (1) no test for any priority pathogen; (2) at least one test for one or more priority pathogens, performed by international laboratory; (3) at least one test for more than one priority pathogen, performed by country laboratory; (4) at least one test for all or nearly all priority pathogens, performed by country laboratory; (5) more than one test for all or nearly all priority pathogens, performed by country laboratory. *NOTE: CBEP may wish to modify tiers if program does not intend to build country's own diagnostic testing capacity for priority pathogens.*
(C.3.4) *Laboratory referral network: number of specimens sent or received for confirmatory testing in the past 12 months (a) in country and (b) international laboratory*	Medium	Five potential tiers: (1) no specimens received for confirmatory testing; (2) at least one specimen received and sent to international laboratory for confirmatory testing; (3) at least one specimen received and confirmatory testing performed (at least initially) by country laboratory; (4) country performed confirmatory testing on multiple specimens for multiple pathogens; (5) country performed confirmatory testing on all or nearly all specimens received.
(C.3.5) *Supervised demonstration of electronic reporting system*	Medium to high	Number assessed and percentage who successfully demonstrated all relevant reporting actions; overall scoring: five tiers, in quintiles.

Table A.6—Continued

Metric	Overall Priority	Scoring and Standard or Target
(C.3.5) Number of cases internationally reportable pathogens / diseases detected in country in the past 12 months; percentage reported to appropriate international authority (WHO, OIE)	Medium to high	*Standard or target: 100 percent of all cases or outbreaks that meet reporting requirements by the respective organization.*
(C.3.5) *Biosurveillance (or CBR) directly informed decisions (e.g., policy, clinical practice)*	Medium to high	Qualitative but powerful message to country, program, and U.S. Congress.
Sustainability enablers		
(C.2.1) Regulatory environment (TT)	High	TT five tiers.
(C.2.1) *Degree of consistency between national and CBEP planning and implementation*	High	Five potential tiers: (1) country has no plan or strategy relevant to CBEP programming; (2) country has relevant plan or strategy, but it does not guide CBEP planning; (3) CBEP planning and programming aligns somewhat with country plan or strategy; (4) CBEP planning and programming aligns directly and intentionally with country plan or strategy; (5) exercises and/or evaluation of CBEP programming are carried out at least in part by host nation.
(C.2.2) *Number of peer-reviewed publications; number of conference papers (abstract, poster, or oral)*	High	Simple counts; can be disaggregated by institution, scientist, research area; indexed scores also available via Web of Science.
(C.2.2) Number of institutions with epidemiology core curriculum meeting international (e.g., FETP) standards; number of trainees per year	Medium to high	Five potential tiers: (1) no access to epidemiology training; (2) access to short-term epidemiology training in another country; (3) access to long-term (two-year) FETP-like training in another country; (4) country has at least one institution that offers short-term epidemiology training; (5) country has at least one institution or program that provides applied epidemiology training.

Table A.6—Continued

Metric	Overall Priority	Scoring and Standard or Target
(C.2.2) *Career track for trained epidemiologists*	Medium to high	Five potential tiers: (1) No career track or opportunities for trained epidemiologists and no intention to develop these; (2) sporadic, informal job or career opportunities for some epidemiologists; (3) national policy not formalized, but career opportunities for most epidemiologists; (4) national policy established to formalize epidemiology career track; (5) national policy successfully implemented for epidemiology career track.
(C.2.3) *Trained and funded internal staff or active contracts or vendor agreements for life cycle equipment maintenance*	Medium	Five potential tiers: (1) none identified or planned; (2) country plans to establish this; (3) some staff or limited contract for equipment maintenance; (4) well-established staff or contract for central laboratory; (5) well-established staff or contract for country's major laboratory network.
(C.2.4) Operational laboratory certification program	Medium to high	Five potential tiers: (1) none established or planned; (2) country plans to have or use such a program; (3) external certification program for national level laboratory; (4) national or international certification program established and at least one additional laboratory certified; (5) certification program operational throughout country's major laboratory network.
(C.2.5) *Number and percentage of priority pathogens for which country (a) produces or (b) can sustainably procure needed media, reagents*	Medium to high	Simple number and percentage, potentially scored in five tiers, in quintiles. or five potential tiers: (1) none; (2) for at least one priority pathogen; (3) for more than one priority pathogen, mainly for central laboratory; (4) for more than one priority pathogen, for central and major subnational laboratories; (5) for all or nearly all priority pathogens, for central and all major subnational laboratories.

Table A.6—Continued

Metric	Overall Priority	Scoring and Standard or Target
Capacities		
(C.1.1) Number and percentage of priority pathogens with baseline established and used to detect anomalies (modified TT)	High	Modified TT five tiers for "capability to establish and use disease baselines to detect anomalous disease patterns": (1) no baseline established or intent to establish one; (2) baseline assessment for at least one priority pathogen planned; (3) baseline assessment for at least one priority pathogen has been completed; (4) baseline has been established for all endemic pathogens and diseases; (5) baseline established for all such diseases and used for routine epidemiological analysis to detect anomalies.
(C.1.2) *Performance—observation or standardized clinical scenarios (and TT)*	Medium to high	Number evaluated and average (written) test score among (a) clinicians, (b) veterinarians; overall assessment—five tiers, in quintiles. Number evaluated and percentage scoring at least 70 percent on written tests among (a) clinicians, (b) veterinarians; overall assessment—five tiers, in quintiles. TT five tiers for "identify relevant diseases and syndromes."
(C.1.3) Reporting capability (TT)	High	TT five tiers for "reporting capability.
(C.1.3) *Number and percentage of jurisdictions using electronic reporting system*	Medium to high	Five tiers, in quintiles.
(C.1.3) *Accuracy, scope: national guidance, e.g., case definitions for animal and human disease*	Medium to high	Five potential tiers: (1) no case definitions for human or animal diseases developed or intended; (2) case definitions planned or intended for human or animal diseases; (3) case definitions developed, disseminated, and used to report some human or animal diseases; (4) case definitions developed, disseminated, and used for reporting most human or animal diseases; (5) case definitions developed, disseminated, and used for reporting all or nearly all human and animal diseases.

Table A.6—Continued

Metric	Overall Priority	Scoring and Standard or Target
(C.1.5) Number of trained field epidemiologists and number per 200,000 population	Medium	Standard (U.S. government interagency): 1 per 200,000 in each major jurisdiction (note that CBEP managers may select a different standard, and standard may vary by country).
(C.1.6) Number of specified level jurisdictions, Number and percentage with functional community-based surveillance (number submit > 50 percent reports; or define five tiers)	Medium	Number and percentage of communities with CBS/EBS that submit at least 50 percent of reports (including zero reporting) on time. and/or Five tiers—overall percentage of reports (including zero reporting) submitted on time by all participating communities, in quintiles.
(C.1.7) Number of trained and drilled teams able to respond to outbreak within 24 hours	Medium	Five potential tiers: (1) No rapid-response team trained or intended; (2) at least one trained team in country; (3) at least one trained team drilled or responded appropriately to outbreak in the past 12 months; (4) more than one trained team drilled or responded to actual outbreak in the past 12 months; (5) at least one trained team for each major jurisdiction (e.g., province) drilled or responded to outbreak within 24 hours in the past 12 months.
(C.1.8) Technical capacity (TT)	High	TT five tiers
(C.1.8) Number and percentage of laboratories certified	Medium to high	Simple percentage, five tiers, in quintiles. or Five potential tiers: (1) none; (2) central laboratory certified based on national standards; (3) central laboratory certified based on international standards; (4) at least one subnational laboratory also certified based on national standards; (5) all or nearly all major subnational laboratories certified based on national or international standards.
(C.1.8) *Number and list of pathogens for which (a) national and (b) each provincial or state laboratory can test*	Medium to high	Simple lists, by level (national, subnational).

Table A.6—Continued

Metric	Overall Priority	Scoring and Standard or Target
(C.1.8) *Laboratory networking: Written or established protocols for specimen referral (a) within country or (b) to international laboratory*	Medium	Five potential tiers: (1) no formal protocol established or intended; (2) informal or formal protocol for specimen referral to subnational or national reference laboratory established but not implemented; (3) formal protocol established and disseminated at least to subnational laboratories; (4) formal protocol disseminated beyond subnational laboratories (e.g., toward points of care); (5) established protocols implemented throughout country.

NOTES: Most standards/targets are to be determined by CBEP and could vary by country. TT = tiger team (indicators, criteria, and/or tiers defined for biosurveillance).

APPENDIX B

Implementation Information for Proposed Metrics—Aligned to Functional Framework

This appendix provides an alternative way to think about implementation priorities and phasing, in this case looking at the elements of CBEP programming from a functional viewpoint. The structure of this appendix parallels that of Appendix A. There are again two sets of tables, one for sources and descriptions and one for scoring and metrics. Each set contains individual tables for biosafety, biosecurity, and biosurveillance.

Note that, in the tables, italics indicate that the language was developed by RAND.

Sources and Descriptions

Tables B.1 through B.3 present and describe the data sources.

Scoring and Metrics

Tables B.4 through B.6 provide information on scoring and metrics.

Table B.1
Sources and Descriptions for Biosafety Metrics—Functional Framework

Metric	Potential Source	Comments
Biological risk assessment, requirements, plans		
(A.3.1) *Proficient test scores for CBEP courses related to biorisk assessment and planning (Example courses in backup)*	CBEP implementer responsible for MAP reporting	Test scores are current MAP inputs (pre- and post–course completion, by individual with specified position title).
(A.2.1) Has a policy concerning the management of laboratory biorisk (biosafety and biosecurity) been written?	CBEP implementer responsible for MAP reporting	Current MAP input.
(A.2.1) Are there regulations and/or guidelines for biosafety and biosecurity?	CBEP implementer responsible for MAP reporting	Current MAP input.
(A.2.5) *SME finds that operational plans include resource sustainment considerations*	CBEP	CBEP team responsible for developing implementation plan or conducting tabletop exercises with partner government personnel.
(A.2.5) SME finds adequate availability of funding to support biosafety and biosecurity programs and initiatives	CBEP implementer responsible for MAP reporting	Proposed MAP input (CBEP BRM team proposed a country-level criterion).
(A.1.1) Percentage completion: BRM assessment, requirements, and planning of facilities (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs.
Human capital		
(A.2.3) *Number of national or regional institutions that make BS&S training to international standards available (training centers, universities. Professional societies, ministries)*	Relevant partner governments	
(A.2.4) *Number of publications; number conference presentations*	CBEP science team	Count of CBEP abstracts, posters, presentations, and publications resulting from CBR projects (currently tracked by CBEP science team).
(A.2.4) *Number or value of internationally competitive research grants won*	CBEP science team; relevant partner governments	

Table B.1—Continued

Metric	Potential Source	Comments
(A.1.2) Percentage completion: BRM management personnel (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs.
(A.1.6) Percentage of personnel trained in laboratory biosafety and biosecurity relative to requirements	CBEP	Trained personnel, required: CBEP team responsible for developing implementation plan or conducting tabletop exercises with partner government personnel; trained personnel, actual: CBEP implementer responsible for MAP reporting.
Safety equipment (primary barriers)		
(A.2.2) *SME finds that operational plans include materiel sustainment considerations*	CBEP	CBEP team responsible for developing implementation plan or conducting tabletop exercises with partner government personnel.
(A.2.2) *Percentage of required equipment that is domestically or regionally sourced*	Relevant partner governments	
(A.1.5) Percentage completion: BS&S equipment (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs.
Facility design and construction (secondary barriers)		
(A.1.4) SME finds that facilities are designed to allow employees to work in a safe and secure way	CBEP implementer responsible for MAP reporting	Current MAP input.
Safe practices and techniques		
(A.3.2) Percentage completion: Biosafety: work practice and administrative control (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs.
(A.1.3) Number or percentage of facilities with complete set of relevant SOPs in place (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs.

Table B.1—Continued

Metric	Potential Source	Comments
Laboratory emergency and incident response		
(A.3.3) SME finds that accident or incident and nonconformities related to biorisk correctly managed	CBEP implementer responsible for MAP reporting	Current MAP input.
(A.2.1) Are there National Incident Management Systems for naturally occurring biological events? (policies, frameworks, and MOUs for laboratory, health security, and law enforcement and ER sectors)	CBEP implementer responsible for MAP reporting	Proposed MAP input (CBEP BRM team proposed a country-level criterion).
Review and revision		
(A.3.4) SME finds that a regular review of the BRM system exists	CBEP implementer responsible for MAP reporting	Current MAP input.
(A.3.4) SME finds that mechanisms for oversight, enforcement, and attribution for biosafety and biosecurity regulations and/or guidelines exist	CBEP implementer responsible for MAP reporting	Proposed MAP input (CBEP BRM team proposed a country-level criterion).

Table B.2
Sources and Descriptions for Biosecurity Metrics—Functional Framework

Metric	Potential Source	Comments
Threat, vulnerability, and security risk assessment		
(B.3.1) *Proficient test scores for CBEP courses related to biorisk assessment and planning (Example courses in backup)*	CBEP implementer responsible for MAP reporting	Test scores are current MAP inputs (Pre- and post-course completion, by individual with specified position title).
(B.2.1) Has a policy concerning the management of laboratory biorisk (biosafety and biosecurity) been written?	CBEP implementer responsible for MAP reporting	Current MAP input.
(B.2.1) Are there regulations and/or guidelines for biosafety and biosecurity?	CBEP implementer responsible for MAP reporting	Current MAP input.
(B.2.5) *SME finds that operational plans include resource sustainment considerations*	CBEP	CBEP team responsible for developing implementation plan or conducting tabletop exercises with partner government personnel.
(B.2.5) SME finds adequate availability of funding to support biosafety and biosecurity programs and initiatives	CBEP implementer responsible for MAP reporting	Proposed MAP input (CBEP BRM team proposed country-level criterion).
(B.1.1) Percentage completion: BRM assessment, requirements, and planning of facilities (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs.
Human capital		
(B.2.3) *Number of national or regional institutions that make BS&S training to international standards available (training centers, universities. Professional societies, ministries)*	Relevant partner governments	
(B.2.4) *Number of publications; number conference presentations*	CBEP science team	Count of CBEP abstracts, posters, presentations, and publications resulting from CBR projects (currently tracked by CBEP science team).
(B.2.4) *Number or value of internationally competitive research grants won*	CBEP science team; relevant partner governments	

Table B.2—Continued

Metric	Potential Source	Comments
(B.1.2) Percentage completion: BRM management personnel (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs.
(B.1.6) Percentage of personnel trained in laboratory biosafety and biosecurity relative to requirements	CBEP	Trained personnel, required: CBEP team responsible for developing implementation plan or conducting tabletop exercises with partner government personnel; trained personnel, actual: CBEP implementer responsible for MAP reporting.
Maintain control of pathogens (physical security, transportation security, information security, material control and accountability		
(B.3.2) SME finds that collections of dangerous pathogens have been consolidated into a minimum number of facilities	CBEP implementer responsible for MAP reporting	Proposed MAP input (CBEP BRM team proposed country-level criterion).
(B.3.2) Pathogen consolidation (TT criteria and tiers)	TT	Metric is straightforward as written and seems useful.
(B.3.2) *Percentage of disclosed biological weapon–related infrastructure that has been eliminated*	Relevant partner governments	
(B.3.3) Percentage completion: Maintain control of pathogens (physical security, material control and accountability, information security, transportation security) (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs.
(B.2.2) *SME finds that operational plans include materiel sustainment considerations*	CBEP	CBEP team responsible for developing implementation plan or conducting tabletop exercises with partner government personnel.
(B.2.2) *Percentage of required equipment that is domestically or regionally sourced*	Relevant partner governments	
(B.1.4) SME finds that facilities are designed to allow employees to work in a safe and secure way	CBEP implementer responsible for MAP reporting	Current MAP input.
(B.1.5) Percentage completion: BS&S equipment (MAP checklist)	CBEP implementer responsible for MAP reporting	Checklist items are all current MAP inputs.

Table B.2—Continued

Metric	Potential Source	Comments
Personnel management (reliability)		
(B.3.4) SME finds mechanisms are in place sufficient to ensure that personnel are competent and reliable	CBEP implementer responsible for MAP reporting	Current MAP input.
Detect, assess, delay, respond to, and recover from a security incident		
(B.3.6) SME finds that accident or incident and nonconformities related to biorisk correctly managed	CBEP implementer responsible for MAP reporting	Current MAP input.
(B.2.1) Are there national incident management systems for intentional biological events? (policies, frameworks, and MOUs for laboratory, health security, and law enforcement and ER sectors)	CBEP implementer responsible for MAP reporting	Proposed MAP input (CBEP BRM team proposed country-level criterion).
Review and revision		
(B.3.7) SME finds that a regular review of the BRM system exists	CBEP implementer responsible for MAP reporting	Current MAP input.
(B.3.7) SME finds that mechanisms for oversight, enforcement, and attribution for biosafety and biosecurity regulations and/or guidelines exist	CBEP implementer responsible for MAP reporting	Proposed MAP input (CBEP BRM team proposed country-level criterion).
(B.3.7) Reporting in 1540 matrix affirming existence of a mechanism for penalizing violator of national prohibition on engagement in biological weapon–related activities	Relevant partner governments; United Nations Security Council	1540 compliance is reported in 1540 matrices (CDC, 2005).

Table B.3
Sources and Descriptions of Biosurveillance Metrics—Functional Framework

Metric	Potential Source	Comments
Detect: community-based		
(C.1.6) *Number of specified level jurisdictions, number and percentage with functional community-based surveillance (number submit > 50 percent reports, or define five tiers)*	MOH or TBD	
Detect: clinical providers		
(C.1.2) *Performance—observation or standardized clinical scenarios (and TT tiers)*	Clinical board examinations	U.S. clinical board examinations involve description and photos of clinical cases for differential diagnostic and other purposes; these could be used to develop appropriate scenarios for testing.
Report and detect: routine reporting		
(C.3.5) *Supervised demonstration of electronic reporting system (define tiers)*	TBD	Unless CBEP has already developed a test for demonstrating electronic reporting (e.g., for EIDSS), test would need to be developed, with five tiers for scoring or some other scoring schema.
(C.3.1) *Demonstrated performance: at least three out of five defined core syndromes*	CDC	This is U.S. government interagency indicator; could consult with CDC regarding details.
(C.1.3) *Reporting capability (TT criteria and five tiers)*	TT	Straightforward as written; not highly precise, so should be supplemented by other metrics.
(C.1.3) *Number and percentage of jurisdictions using electronic reporting system (can also define five tiers)*	MOH	

Table B.3—Continued

Metric	Potential Source	Comments
Detect and respond: epidemiological analysis and investigation		
(C.3.1) *Surveillance analysis performance—exercise or supervisory observation (Near-term metric could employ TT tiers, recommend in longer-term working toward introduction of metric based on CDC/CSTE)*	CDC FETP, CSTE Applied Epidemiology Competencies	CSTE Applied Epidemiology Competencies documentation provides competencies, job descriptions and assessment forms for four tiers of epidemiologists, from basic- or entry-level to midlevel and senior supervisory; relevance of different tiers may vary by CBEP country and over time. Assessment form (self-assessment and supervisory) for Tier 1 (entry-level epidemiologists) (CDC, undated a) and Assessment form for Tier 2 (midlevel epidemiologists) (CDC, undated b).
(C.3.3) *Number of suspected priority pathogen cases or outbreaks in the past 12 months, and percentage for which investigation was conducted and results documented*	MOH and laboratory log	Note that denominator (suspected cases) is critical to assessing this indicator, since metric is proportion (not simply absolute number) of cases or outbreaks that were properly investigated and documented. Laboratory logs (central, provincial or state) should provide denominator data, and MOH should provide numerator data.
(C.3.3) *Epidemiological investigation performance (written report or tested via exercise)*	CDC FETP	CDC has extensive documentation on steps involved in epidemiologic investigation (e.g., CDC, 2004); consult with CDC as needed. WHO also has at least one checklist that can be used as a tool (e.g., WHO, 2006b, p. 18)
(C.1.1) Number or percentage of priority pathogens with baseline (a) established and used to detect anomalies (modified TT tiers)	MOH, laboratory, CBEP (CBR)	Quantitative: Most CBEP CBR projects appear to contribute to baseline information for relevant pathogens; CBEP program staff probably can identify pathogens for which baseline information has been collected. Qualitative: Might be more difficult, but is very important, to capture how biosurveillance and related CBR data are actually used (e.g., as baseline trend for assessment of potential outbreaks, to guide clinical or public health policy).

Table B.3—Continued

Metric	Potential Source	Comments
(C.1.3) *Accuracy, scope: national guidance, e.g., case definitions for animal and human disease.*	MOH	There are several sources for relevant case definitions (e.g., AFHSC for DoD reportable medical events); scoring could use TT-like five tiers ranging from no case definitions to clear ones for all relevant pathogens.
(C.1.5) Number trained field epidemiologists and number per 200,000 population	MOH	Set criteria for what "counts"—probably at least a one-year certificate or degree program; consult with CDC FETP. Numbers should increase over time. U.S. government interagency standard of one per 200,000 might be somewhat arbitrary, but metric is useful to track growing capacity.
(C.1.7) Number trained and drilled teams able to respond to outbreak within 24 hours	MOH	Suggest using (or developing) standard definition such as that used by WHO and others for rapid-response teams (e.g., Mekong Basin Disease Surveillance group has a definition they use for their rapid-response teams).
(C.2.2) Number of institutions with epidemiology core curriculum meeting international (e.g., FETP) standards; number trainees per year	TEPHINET, CDC FETP, relevant partner countries	Global network of "FETPs" = TEPHINET. TEPHINET, CDC/FETP and/or FETP countries undoubtedly have and/or have defined basic core curricula.
(C.2.2) *Career track for trained epidemiologists*	CDC FETP	This is a requirement before CDC FETP is established in a country. Consult with CDC regarding how to make this happen and what "counts" for tracking purposes.
Diagnose: laboratory testing		
(C.3.4) *Number and percentage of laboratories participating in proficiency testing at least once in the past 12 months*	APHL, CDC, WHO	We believe there are several laboratory proficiency testing programs around the world, from government agencies such as CDC as well as laboratory groups (APHL), perhaps Joint Commission International, and relevant WHO collaborating centers (of which CDC has a number). There are also commercial firms that conduct proficiency testing.
(C.3.4) *Number and percentage of laboratories that passed all proficiency tests in the past 12 months, among laboratories participating in the past 12 months*	APHL, CDC, WHO	See cell immediately above.

Table B.3—Continued

Metric	Potential Source	Comments
(C.3.4) *Number and percentage of laboratories that meet specified standards (mutually agreed upon by country and CBEP)*	APHL, CDC, DoD	Might consult with DoD laboratory specialists and/or APHL or CDC to help develop reasonable standards appropriate for various levels of laboratory and country development.
(C.3.4) *Specimen collection transport: number specimens received by laboratory; number and percentage adequate for testing (logs, exercise; TT tiers)*	Laboratory log	Laboratory logs should be maintained to track specimens received and tested; log should include specimens that were not suitable for testing, which reflects inadequate collection or transport.
(C.3.4) *Laboratory timeliness: number hours following receipt of specimens that laboratory testing is (a) initiated and (b) completed (from laboratory logs or exercise)*	Laboratory log or exercise	Laboratory specimens are typically accompanied with brief clinical information (clinical diagnosis, date of onset), which can be captured for this metric.
(C.3.4) *Number of laboratory tests performed for each priority pathogen in the past 12 months*	Laboratory log	
(C.3.4) *Laboratory referral network: number of specimens sent or received for confirmatory testing in the past 12 months (a) in country and (b) international laboratory*	Laboratory log	
(C.1.8) Technical capacity (TT criteria and tiers)	TT	Straightforward as written; not highly precise, so should be supplemented by other metrics.
(C.1.8) Number and percentage of laboratories certified	MOH; DoD; ISO; JCI	In consultation with relevant laboratory specialists (DoD or other), identify relevant certification standards and body; could vary by country.
(C.1.8) *Number and list of pathogens for which (a) national and (b) each provincial or state laboratory can test*	MOH	Metric reflects "reach" of relevant laboratory testing throughout the country.
(C.1.8) *Laboratory networking: Written or established protocols for specimen referral (a) within country and (b) to international laboratory*	MOH	Metric reflects national guidance and should be a precursor to capabilities related to specimen collection and transport (e.g., capability C.3.4).
(C.2.1) Regulatory environment (TT criteria and tiers)	TT	Metric is straightforward as written and seems useful.
(C.2.3) T*rained and funded internal staff or active contracts and vendor agreements for life cycle equipment maintenance*	TT or CBEP	Consider a TT-like five-tier scoring scheme for this metric.

Table B.3—Continued

Metric	Potential Source	Comments
(C.2.4) Operational laboratory certification program	MOH; DoD; ISO; JCI	As above for C.1.8 capacity metric, consult with relevant experts to define metric in more detail.
(C.2.5) *Number and percentage of priority pathogens for which country (a) produces or (b) can sustainably procure needed media, reagents*		See cell immediately above.
Report: international		
(C.3.5) *Number of cases of internationally reportable pathogens or diseases detected in country in the past 12 months; percentage reported to appropriate international authority (WHO, OIE)*	MOH, laboratory log	The needed denominator data (number of reportable cases) may be elusive in countries where transparency is an issue (noted by CBEP staff for FSU), but denominator is critical to really assess this metric, which reflects degree of compliance rather than just number of reports per se.
Impact		
(C.2.1) *Degree of consistency between national and CBEP planning and implementation*		This reflects a critical sustainability-enabling factor: External programming aligns with national planning rather than vice versa (described for some new engagement CBEP countries, for example).
(C.2.2) *Number of peer-reviewed publications, number conference papers (abstract, poster, or oral)*	Web of Science, Scopus, or other bibliometric database	These sources have different capabilities and reach (they do not identify the same items, though there is significant overlap); Web of Science includes very impressive bibliometric analysis capabilities (e.g., can compare countries, including by research area).
(C.3.5) *Biosurveillance (or CBR) directly informed decisions (e.g., policy, clinical practice)*	MOH	Qualitative. This is one metric that is very important but does not lend itself to strictly quantitative measurement.

NOTE: TT = tiger team (indicators, criteria, and/or tiers defined for biosurveillance).

Table B.4
Proposed Scoring and Standards for Biosafety Metrics—Functional Framework

Metric	Overall Priority	Scoring and Standard or Target
Biological risk assessment, requirements, plans		
(A.3.1) *Proficient test scores for CBEP courses related to biorisk assessment and planning (Example courses in backup)*	Medium to high	Number evaluated and percent scoring at least 70 percent on (specified) CBEP tests related to biorisk assessment and planning (MAP input for "post-test score").
(A.2.1) Has a policy concerning the management of laboratory biorisk (biosafety and biosecurity) been written?	High	Facility level—Yes-or-no answer to MAP input: "Has a policy concerning the management of laboratory biorisk (biosafety and biosecurity) been written?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.
(A.2.1) Are there regulations and/or guidelines for biosafety and biosecurity?	Medium	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Are there regulations and/or guidelines for biosafety and biosecurity?"
(A.2.5) *SME finds that operational plans include resource sustainment considerations*	Medium to high	Percentage of operational plans describing planned delivery of CBEP-furnished facilities, equipment, and training that include plans for partner resourcing sustainment of infrastructure, materiel, and human capital.
(A.2.5) SME finds adequate availability of funding to support biosafety and biosecurity programs and initiatives	High	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Is there adequate availability of funding to support biosafety and biosecurity programs and initiatives?"
(A.1.1) Percentage completion: BRM assessment, requirements, and planning of facilities (MAP checklist)	High	Facility level—Number evaluated and percent reporting *yes* to (specified) MAP inputs related to BRM assessment, requirements, planning. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.

Table B.4—Continued

Metric	Overall Priority	Scoring and Standard or Target
Human capital		
(A.2.3) *Number of national or regional institutions that make BS&S training to international standards available (training centers, universities. Professional societies, ministries)*	Medium to high	Five potential tiers: (1) no access to BS&S training; (2) access to short-term BS&S training in another country; (3) access to long-term (two-year) training in another country; (4) country has at least one institution that offers short-term training; (5) country has at least one institution or program that provides long-term training.
(A.2.4) *Number of publications; number of conference presentations*	Medium	Simple counts; can be disaggregated by institution, scientist, research area; indexed scores also available via Web of Science.
(A.2.4) *Number or value of internationally competitive research grants won*	Medium	Simple counts; can be disaggregated by institution, scientist, research area.
(A.1.2) Percentage completion: BRM management personnel (MAP checklist)	Medium to high	Facility level—Number evaluated and percent reporting *yes* to (specified) MAP inputs related to BRM management personnel. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.
(A.1.6) Percentage of personnel trained in laboratory biosafety and biosecurity relative to requirements	High	(1) Number of personnel trained required, by position and subject area, based on CBEP and partner operational plans. (2) Number of individuals trained in each position and each subject area (MAP inputs for student identification number, position, date course completed).
Safety equipment (primary barriers)		
(A.2.2) *SME finds that operational plans include materiel sustainment considerations*	Medium to high	Percentage of operational plans describing planned delivery of CBEP-furnished facilities and equipment that include plans for partner sourcing of material and expertise required for sustainment of infrastructure and material.

Table B.4—Continued

Metric	Overall Priority	Scoring and Standard or Target
(A.2.2) *Percentage of required equipment that is domestically or regionally sourced*	Medium	(1) (specified) Percentage of equipment used to support partner's BRM program that comes from a domestic or regional source. (2) Percentage remaining that needs to be acquired from nondomestic and nonregional sources is within partner's ability or willingness to resource.
(A.1.5) Percentage completion: BS&S equipment (MAP checklist)	High	Facility level—Number evaluated and percent reporting *yes* to (specified) MAP inputs related to BS&S equipment. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.
Facility design and construction (secondary barriers)		
(A.1.4) SME finds that facilities are designed to allow employees to work in a safe and secure way	High	Facility level—Yes-or-no answer to MAP input: "Are the facilities designed to allow to work in a safe and secure way?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.
Safe practices and techniques		
(A.3.2) Percentage completion: Biosafety: work practice and administrative control (MAP checklist)	High	Facility level—Number evaluated and percentage reporting *yes* to (specified) MAP inputs related to biosafety: work practice and administrative control National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.
(A.1.3) Number or percentage of facilities with complete set of relevant SOPs in place (MAP checklist)	High	Facility level—Number evaluated and percentage reporting *yes* to (specified) MAP inputs related to relevant SOPs in place. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.

Table B.4—Continued

Metric	Overall Priority	Scoring and Standard or Target
Laboratory emergency and incident response		
(A.3.3) SME finds that accident or incident and nonconformities related to biorisk correctly managed	High	Facility level—Yes-or-no answer to MAP input: "Are accident or incident and nonconformities related to biorisk correctly managed (i.e., reported, recorded, investigated, and leading to preventive or corrective actions)?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.
(A.2.1) Are there National Incident Management Systems for naturally occurring biological events? (policies, frameworks, and MOUs for laboratory, health security, and law enforcement and ER sectors)	Medium to high	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Are there national incident management systems for naturally occurring biological events? (policies, frameworks, and MOUs for laboratory, health security, and law enforcement and ER sectors)?"
Review and revision		
(A.3.4) SME finds that a regular review of the BRM system exists	High	Facility level—Yes-or-no answer to MAP input: "Is there a regular review of the biorisk management system?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.
(A.3.4) SME finds that mechanisms for oversight, enforcement, and attribution for biosafety and biosecurity regulations and/or guidelines exist	High	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Are there mechanisms for oversight, enforcement and attribution for biosafety and biosecurity regulations and/or guidelines?"

Table B.5
Proposed Scoring and Standards for Biosecurity Metrics—Functional Framework

Metric	Type	Overall Priority	Scoring and Standard or Target
Threat, vulnerability, and security risk assessment			
(B.3.1) *Proficient test scores for CBEP courses related to biorisk assessment and planning (Example courses in backup)*	CAPAB	Medium to high	Number evaluated and percentage scoring at least 70 percent on (specified) CBEP tests related to biorisk assessment and planning (MAP input for "post-test score").
(B.2.1) Has a policy concerning the management of laboratory biorisk (biosafety and biosecurity) been written?	SUSTN	High	Facility level—Yes-or-no answer to MAP input: "Has a policy concerning the management of laboratory biorisk (biosafety and biosecurity) been written?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.
(B.2.1) Are there regulations and/or guidelines for biosafety and biosecurity?	SUSTN	Medium	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Are there regulations and/or guidelines for biosafety and biosecurity?"
(B.2.5) *SME finds that operational plans include resource sustainment considerations*	SUSTN	Medium to high	Percentage of operational plans describing planned delivery of CBEP-furnished facilities, equipment, and training that include plans for partner resourcing sustainment of infrastructure, materiel, and human capital.
(B.2.5) SME finds adequate availability of funding to support biosafety and biosecurity programs and initiatives	SUSTN	High	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Is there adequate availability of funding to support biosafety and biosecurity programs and initiatives?"
(B.1.1) Percentage completion: BRM assessment, requirements, and planning of facilities (MAP checklist)	CAPAC	High	Facility level—Number evaluated and percent reporting *yes* to (specified) MAP inputs related to BRM assessment, requirements, planning. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.

Table B.5—Continued

Metric	Type	Overall Priority	Scoring and Standard or Target
Human capital			
(B.2.3) *Number of national or regional institutions that make BS&S training to international standards available (training centers, universities. Professional societies, ministries)*	CAPAC	Medium to high	Five potential tiers: (1) no access to BS&S training; (2) access to short-term BS&S training in another country; (3) access to long-term (two-year) training in another country; (4) country has at least 1 institution that offers short-term training; (5) country has at least one institution or program that provides long-term training.
(B.2.4) *Number of publications; number conference presentations*	SUSTN	Medium	Simple counts; can be disaggregated by institution, scientist, research area; indexed scores also available via Web of Science.
(B.2.4) *Number or value of internationally competitive research grants won*	SUSTN	Medium	Simple counts; can be disaggregated by institution, scientist, research area.
(B.1.2) Percentage completion: BRM management personnel (MAP checklist)	CAPAC	Medium to high	Facility level—Number evaluated and percent reporting *yes* to (specified) MAP inputs related to BRM management personnel. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.
(B.1.6) Percentage of personnel trained in laboratory biosafety and biosecurity relative to requirements	CAPAC	High	(1) Number of personnel trained required, by position and subject area, based on CBEP/partner operational plans. (2) Number of individuals trained in each position and each subject area (MAP inputs for student identification number, position, date course completed).
Maintain control of pathogens (physical security, transportation security, information security, material control and accountability			
(B.3.2) SME finds that collections of dangerous pathogens have been consolidated into a minimum number of facilities	CAPAB	Medium to high	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Have collections of dangerous pathogens been consolidated into a minimum number of facilities?"

Table B.5—Continued

Metric	Type	Overall Priority	Scoring and Standard or Target
(B.3.2) Pathogen consolidation (TT criteria and tiers)	CAPAB	High	TT five tiers for "Assessment (Consolidation)" for Annual Report to Congress
(B.3.2) *Percentage of disclosed biological weapon–related infrastructure that has been eliminated*	CAPAB	Medium to high	(1) Count of biological weapon–related infrastructure targeted for elimination (2) Count of above that has been eliminated.
(B.3.3) Percentage completion: Maintain control of pathogens (physical security, material control and accountability, information security, transportation security) (MAP checklist)	CAPAB	High	Facility level— Number evaluated and percent reporting *yes* to (specified) MAP inputs related to capabilities for maintenance of control of pathogens. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.
(B.2.2) *SME finds that operational plans include materiel sustainment considerations*	SUSTN	Medium to high	Percentage of operational plans describing planned delivery of CBEP-furnished facilities and equipment that include plans for partner sourcing of material and expertise required for sustainment of infrastructure and material.
(B.2.2) *Percentage of required equipment that is domestically or regionally sourced*	SUSTN	Medium	(1) (specified) Percentage of equipment used to support partner's BRM program that comes from a domestic or regional source. (2) Percentage remaining that needs to be acquired from nondomestic and nonregional sources is within partner's ability or willingness to resource.
(B.1.4) SME finds that facilities are designed to allow employees to work in a safe and secure way	CAPAC	High	Facility level—Yes-or-no answer to MAP input: "Are the facilities designed to allow to work in a safe and secure way?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.

Table B.5—Continued

Metric	Type	Overall Priority	Scoring and Standard or Target
(B.1.5) Percentage completion: BS&S equipment (MAP checklist)	CAPAC	High	Facility level—Number evaluated and percent reporting *yes* to (specified) MAP inputs related to BS&S equipment. National level—Percentage of CBEP-engaged facilities answering *yes* to (specified number) of above MAP inputs.
Personnel management (reliability)			
(B.3.4) SME finds that mechanisms are in place sufficient to ensure that personnel are competent and reliable	CAPAB	High	Facility level—Yes-or-no answer to MAP input: "Is there mechanism/s to ensure that personnel is competent and reliable (e.g., successful completion of training, ability to perform tasks under supervision)?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.
Detect, assess, delay, respond to, and recover from a security incident			
(B.3.6) SME finds that accident or incident and nonconformities related to biorisk correctly managed	CAPAB	High	Facility level—Yes-or-no answer to MAP input: "Are accident or incident and nonconformities related to biorisk correctly managed (i.e., reported, recorded, investigated, and leading to preventive or corrective actions)?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.
(B.2.1) Are there National Incident Management Systems for intentional biological events? (policies, frameworks, and MOUs for laboratory, health security, and law enforcement and ER sectors)	SUSTN	Medium to high	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Are there National Incident Management Systems for intentional biological events? (policies, frameworks, and MOUs for laboratory, health security, and law enforcement and ER sectors)?"

Table B.5—Continued

Metric	Type	Overall Priority	Scoring and Standard or Target
Review and revision			
(B.3.7) SME finds that a regular review of the BRM system exists	CAPAB	High	Facility level—Yes-or-no answer to MAP input: "Is there a regular review of the biorisk management system?" National level—Percentage of CBEP-engaged facilities answering *yes* to above MAP input.
(B.3.7) SME finds that mechanisms for oversight, enforcement, and attribution for biosafety and biosecurity regulations and/or guidelines exist	CAPAB	High	Yes-or-no answer to CBEP BRM team–proposed input for country-level criteria: "Are there mechanisms for oversight, enforcement and attribution for biosafety and biosecurity regulations and/or guidelines?"
(B.3.7) Reporting in 1540 matrix affirming existence of a mechanism for penalizing violator of national prohibition on engagement in biological weapon–related activities	CAPAB	High	(1) Percentage of relevant areas (e.g., manufacture, acquire, possess, stockpile, transport) for which a country reports *yes* to question "Does national legislation exist which prohibits persons or entities to engage in one of the following activities? Can violators be penalized?" in the 1540 matrix.

NOTE: Most standards or targets are to be determined by CBEP and could vary by country. CAPAB = capability; SUSTN = sustainability enabler; CAPAC = capacity.

Table B.6
Proposed Scoring and Standards for Biosurveillance Metrics—Functional Framework

Metric	Type	Overall Priority	Scoring and Standard or Target
Detect: community-based			
(C.1.6) *Number of specified level jurisdictions, number and percentage with functional community-based surveillance (number submit > 50 percent reports; or define five tiers)*	CAPAC	Medium	Number and percentage of communities with CBS/EBS that submit at least 50 percent of reports (including zero reporting) on time and/or Five tiers—overall percentage of reports (including zero reporting) submitted on time by all participating communities, in quintiles
Detect: clinical providers			
(C.1.2) *Performance—observation or standardized clinical scenarios (and TT)*	CAPAC	Medium to high	Number evaluated and average (written) test score among (a) clinicians, (b) veterinarians; overall assessment—five tiers, in quintiles Number evaluated and percentage scoring at least 70 percent on written tests among (a) clinicians, (b) veterinarians; overall assessment—five tiers, in quintiles TT five tiers for "identify relevant diseases and syndromes"
Report and detect: routine reporting			
(C.3.5) *Supervised demonstration of electronic reporting system*	CAPAB	Medium to high	Number assessed and percentage who successfully demonstrated all relevant reporting actions; overall scoring: five tiers, in quintiles
(C.3.1) *Demonstrated performance: at least three out of five defined core syndromes*	CAPAB	Medium	Number evaluated and percentage of (a) clinicians and (b) veterinarians who can identify at least 3 of the 5 syndromes on a written test; overall assessment—five tiers, in quintiles TT five tiers for "identify relevant diseases and syndromes"
(C.1.3) Reporting capability (TT)	CAPAC	High	TT five tiers for "reporting capability"

Table B.6—Continued

Metric	Type	Overall Priority	Scoring and Standard or Target
(C.1.3) *Number and percentage of jurisdictions using electronic reporting system*	CAPAC	Medium to high	Five tiers, in quintiles
Detect and respond: epidemiological analysis and investigation			
(C.3.1) *Surveillance analysis performance—exercise or supervisory observation (Near-term metric could employ TT tiers, recommend in longer-term working toward introduction of metric based on CDC/CSTE)*	CAPAB	High	Initially, TT five tiers for "capability to analyze surveillance data to inform public health decisions" à More detailed CDC/CSTE assessment forms; overall scoring: number assessed by level of epidemiologist, percentage passing, five tiers—by quintiles.
(C.3.3) *Number of suspected priority pathogen cases or outbreaks in the past 12 months, and percentage for which investigation was conducted, results documented*	CAPAB	Medium	Scoring can be simple percentage (of detected cases or outbreaks that were investigated and documented), or in five tiers, by quintile.
(C.3.3) *Epidemiological investigation performance (written report or tested via exercise)*	CAPAB	Medium to high	Depends on assessment tool (e.g., checklist of investigation steps followed and documented in written report), but can score in five tiers by quintile or simple percentage of those scoring 85–100 percent on skills and knowledge checklist (e.g., as used by USAID).
(C.1.1) Number and percentage of priority pathogens with baseline established and used to detect anomalies (modified TT)	CAPAC	High	Modified TT five tiers for "capability to establish and use disease baselines to detect anomalous disease patterns": (1) no baseline established or intent to establish one; (2) baseline assessment for at least one priority pathogen planned; (3) baseline assessment for at least one priority pathogen has been completed; (4) baseline has been established for all endemic pathogens and diseases; (5) baseline established for all such diseases and used for routine epidemiological analysis to detect anomalies.

Table B.6—Continued

Metric	Type	Overall Priority	Scoring and Standard or Target
(C.1.3) *Accuracy, scope: national guidance, e.g., case definitions for animal and human disease.*	CAPAC	Medium to high	Five potential tiers: (1) no case definitions for human or animal diseases developed or intended; (2) case definitions planned or intended for human or animal diseases; (3) case definitions developed, disseminated and used to report some human or animal diseases; (4) case definitions developed, disseminated and used for reporting most human or animal diseases; (5) case definitions are developed, disseminated, and used for reporting all or nearly all human and animal diseases.
(C.1.5) Number of trained field epidemiologists and number per 200,000 population	CAPAC	Medium	Standard (U.S. government interagency): One per 200,000 in each major jurisdiction (note that CBEP managers may select a different standard, and standard may vary by country).
(C.1.7) Number of trained and drilled teams able to respond to outbreak within 24 hours	CAPAC	Medium	Five potential tiers: (1) No rapid-response team trained or intended; (2) at least one trained team in country; (3) at least one trained team, drilled or responded appropriately to outbreak in the past 12 months; (4) more than one trained team drilled or responded to actual outbreak in the past 12 months; (5) at least one trained team for each major jurisdiction (e.g., province), drilled or responded to outbreak within 24 hours in the past 12 months.
(C.2.2) Number of institutions with epidemiology core curriculum meeting international (e.g., FETP) standards; number of trainees per year	SUSTN	Medium to high	Five potential tiers: (1) no access to epidemiology training; (2) access to short-term epidemiology training in another country; (3) access to long-term (two-year) FETP-like training in another country; (4) country has at least one institution that offers short-term epidemiology training; (5) country has at least one institution or program that provides applied epidemiology training.

Table B.6—Continued

Metric	Type	Overall Priority	Scoring and Standard or Target
(C.2.2) *Career track for trained epidemiologists*	SUSTN	Medium to high	Five potential tiers: (1) No career track or opportunities for trained epidemiologists and no intention to develop these; (2) sporadic, informal job and career opportunities for some epidemiologists; (3) national policy not formalized, but career opportunities for most epidemiologists; (4) national policy established to formalize epidemiology career track; (5) national policy successfully implemented for epidemiology career track.
Diagnose: laboratory testing			
(C.3.4) *Number and percentage of laboratories participating in proficiency testing at least once during in the past 12 months*	CAPAB	Medium to high	Five potential tiers: (1) No laboratories participate in proficiency testing and no intention to do so; (2) no laboratories participate, but country plans to do so; (3) central (or other) laboratory participated in the past 12 months; (4) central and at least one subnational laboratory participated; (5) central and most or all major subnational jurisdictional laboratories participated.
(C.3.4) *Number and percentage of laboratories that passed all proficiency tests in the past 12 months, among laboratories participating in the past 12 months*	CAPAB	Medium to high	Five potential tiers, in quintiles.
(C.3.4) *Number and percentage of laboratories that meet specified standards (mutually agreed upon by country and CBEP)*	CAPAB	Medium to high	Five potential tiers, in quintiles
(C.3.4) *Specimen collection, transport: number of specimens received by laboratory; number and percentage adequate for testing*	CAPAB	Medium to high	TT five tiers for "capability to collect and analyze samples" Overall percentage of specimens received at (a) central and (b) provincial laboratories that were adequate for testing (logs or exercise); five tiers, in quintiles.
(C.3.4) *Number of hours following receipt of specimens that laboratory testing is (a) initiated, (b) completed (from laboratory logs or exercise)*	CAPAB	Medium	

Table B.6—Continued

Metric	Type	Overall Priority	Scoring and Standard or Target
(C.3.4) *Number of laboratory tests performed for each priority pathogen in the past 12 months*	CAPAB	High	Five potential tiers: (1) no test for any priority pathogen; (2) at least one test for one or more priority pathogens, performed by international laboratory; (3) at least one test for more than one priority pathogen, performed by country laboratory; (4) at least one test for all or nearly all priority pathogens, performed by country laboratory; (5) more than one test for all or nearly all priority pathogens, performed by country laboratory. NOTE: CBEP may wish to modify tiers if program does not intend to build country's own diagnostic testing capacity for priority pathogens
(C.3.4) *Laboratory referral network: number specimens sent or received for confirmatory testing in the past 12 months (a) in country and (b) international laboratory*	CAPAB	Medium	Five potential tiers: (1) no specimens received for confirmatory testing; (2) at least one specimen received and sent to international laboratory for confirmatory testing; (3) at least one specimen received and confirmatory testing performed (at least initially) by country laboratory; (4) country performed confirmatory testing on multiple specimens for multiple pathogens; (5) country performed confirmatory testing on all or nearly all specimens received
(C.1.8) Technical capacity (TT)	CAPAC	High	TT five tiers
(C.1.8) Number and percentage of laboratories certified	CAPAC	Medium to high	Simple percentage, five tiers, quintiles. or Five potential tiers: (1) none; (2) central laboratory certified based on national standards; (3) central laboratory certified based on international standards; (4) at least one subnational laboratory also certified based on national standards; (5) all or nearly all major subnational laboratories certified based on national or international standards.
(C.1.8) *Number and list of pathogens for which (a) national and (b) each provincial or state laboratory can test*	CAPAC	Medium to high	Simple lists, by level (national, subnational).

Table B.6—Continued

Metric	Type	Overall Priority	Scoring and Standard or Target
(C.1.8) *Laboratory networking: Written or established protocols for specimen referral (a) within country, (b) to international laboratory*	CAPAC	Medium	Five potential tiers: (1) no formal protocol established or intended; (2) informal or formal protocol for specimen referral to subnational or national reference laboratory established but not implemented; (3) formal protocol established and disseminated at least to subnational laboratories; (4) formal protocol disseminated beyond subnational laboratories (e.g., toward points of care); (5) established protocols implemented throughout country.
(C.2.1) Regulatory environment (TT)	SUSTN	High	TT five tiers.
(C.2.3) *Trained and funded internal staff or active contracts or vendor agreements for life cycle equipment maintenance*	SUSTN	Medium	Five potential tiers: (1) none identified or planned; (2) country plans to establish this; (3) some staff or limited contract for equipment maintenance; (4) well-established staff or contract for central laboratory; (5) well-established staff or contract for country's major laboratory network.
(C.2.4) Operational laboratory certification program	SUSTN	Medium to high	Five potential tiers: (1) none established or planned; (2) country plans to have or use such a program; (3) external certification program for national level laboratory; (4) national or international certification program established and at least one additional laboratory certified; (5) certification program operational throughout country's major laboratory network.
(C.2.5) *Number and percentage of priority pathogens for which country (a) produces or (b) can sustainably procure needed media, reagents*	SUSTN	Medium to high	Simple number and percentage, potentially scored in five tiers by quintile. or Five potential tiers: (1) none; (2) for at least one priority pathogen; (3) for more than one priority pathogen, mainly for central laboratory; (4) for more than one priority pathogen, for central and major subnational laboratories; (5) for all or nearly all priority pathogens, for central and all major subnational laboratories.

Table B.6—Continued

Metric	Type	Overall Priority	Scoring and Standard or Target
Report: international			
(C.3.5) *Number of cases internationally reportable pathogens or diseases detected in country in the past 12 months; percentage reported to appropriate international authority (WHO, OIE)*	CAPAB	Medium to high	Standard or target: 100 percent of all cases or outbreaks that meet reporting requirements by the respective organization.
Impact			
(C.2.1) *Degree of consistency between national and CBEP planning and implementation*	SUSTN	High	Five potential tiers: (1) country has no plan or strategy relevant to CBEP programming; (2) country has relevant plan or strategy, but it does not guide CBEP planning; (3) CBEP planning and programming aligns somewhat with country plan or strategy; (4) CBEP planning and programming aligns directly and intentionally with country plan or strategy; (5) exercises and/or evaluation of CBEP programming are carried out at least in part by host nation.
(C.2.2) *Number of peer-reviewed publications, number conference papers (abstract, poster, or oral)*	SUSTN	High	Simple counts; can be disaggregated by institution, scientist, research area; indexed scores also available via Web of Science.
(C.3.5) *Biosurveillance (or CBR) directly informed decisions (e.g., policy, clinical practice)*	SUSTN	Medium to high	Qualitative but powerful message to country, program, and U.S. Congress.

NOTE: Most standards or targets are to be determined by CBEP and could vary by country. TT = tiger team (indicators, criteria, and/or tiers defined for biosurveillance); CAPAC = capacity; CAPAB = capability; SUSTN = sustainability enabler.

Bibliography

Amorim, Celso, Philippe Douste-Blazy, Hasan Wirayuda, Jonas Gahr Støre, Cheikh Tidiane Gadio, Nkosazana Dlamini-Zuma, and Nitya Pibulsonggram, "Oslo Ministerial Declaration—Global Health: A Pressing Foreign Policy Issue of Our Time," *Lancet*, Vol. 369, April 2, 2007, pp. 1373–1378.

Bakanidze, Lela, Paata Imnadze, and Dana Perkins, "Biosafety and Biosecurity as Essential Pillars of International Health Security and Cross-Cutting Elements of Biological Nonproliferation," *BMC Public Health*, Vol. 10, Suppl. 1, 2010.

Benkert, Joseph A., "Statement for the Record, Joseph A. Benkert Principal Deputy Assistant Secretary of Defense for Global Security Affairs," before the Senate Committee on Armed Services Subcommittee on Emerging Threats and Capabilities, April 11, 2007.

Bennett, Claude, "Up the Hierarchy," *Journal of Extension*, Vol. 13, No. 2, 1975, pp. 7–12.

Bickman, Leonard, ed., *Using Program Theory in Evaluation*, New Directions for Program Evaluation No. 33, San Francisco: Jossey-Bass, 1987.

"Black & Veatch to Coordinate Armenia's National Biological Threat Reduction Program," press release, *BusinessWire*, June 2, 2011.

Bonventre, Eugene V., Kathleen H. Hicks, and Stacy M. Okutani, *U.S. National Security and Global Health: An Analysis of Global Health Engagement by the U.S. Department of Defense*, Washington, D.C.: Center for Strategic and International Studies, 2009.

Cali, Shawn, Sara Mayer, and Roger Breeze, "The United States Department of Defense Biological Threat Reduction Program: Threat Agent Detection and Response and Cooperative Biological Research," briefing, Defense Threat Reduction Agency, February 23, 2007.

Carnegie Endowment for International Peace and the Russian-American Nuclear Security Advisory Council, *Reshaping U.S.–Russian Threat Reduction: New Approaches for the Second Decade*, Washington, D.C., 2002.

Carter, Ashton B., Deputy Secretary of Defense, "Interim Guidance for Implementing the National Strategy for Biosurveillance," memorandum, Washington, D.C., June 13, 2013.

Carter, Ashton B., Robert G. Joseph, et al., *Review Panel on Future Directions for Defense Threat Reduction Agency Missions and Capabilities to Combat Weapons of Mass Destruction*, March 2008.

CBEP—*See* Cooperative Biological Engagement Program.

CDC—*See* Centers for Disease Control and Prevention.

Centers for Disease Control and Prevention, "Epidemiology Competency Assessment Form—Tier 1 Epidemiologist," undated a. As of June 9, 2014:
http://www.cdc.gov/AppliedEpiCompetencies/downloads/AEC_Assessment_Tier1.pdf

———, "Epidemiology Competency Assessment Form—Tier 2 Epidemiologist," undated b. As of June 9, 2014:
http://www.cdc.gov/AppliedEpiCompetencies/downloads/AEC_Assessment_Tier2.pdf

———, "Steps of an Outbreak Investigation," web page, November 17, 2004. As of June 9, 2014:
http://www.cdc.gov/excite/classroom/outbreak/steps.htm

———, *Biosafety in Microbiological and Biomedical Laboratories*, 5th ed., Atlanta, Ga., December 2009.

———, *Public Health Preparedness Capabilities: National Standards for State and Local Planning*, Atlanta, Ga., March 2011.

Coker, Richard J., Benjamin M. Hunter, James W. Rudge, Marco Liverani, and Piya Hanvoravongchai, "Emerging Infectious Disease in Southeast Asia: Regional Challenges to Control," *Lancet*, Vol. 377, February 12, 2011, pp. 599–609.

Committee on Armed Services United States Senate, *National Defense Authorization Act for FY 2008: Report to Accompany S. 1547*, Washington, D.C.: U.S. Government Printing Office, Report 110-77, 2007.

Convention on the Prohibition of the Development, Production, and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction, in force March 26, 1975.

Cook, Michelle Stem, and Amy F. Woolf, *Preventing Proliferation of Biological Weapons: U.S. Assistance to the Former Soviet States*, Washington, D.C.: Congressional Research Service, April 10, 2002.

Cooperative Biological Engagement Program, Cooperative Biological Engagement Program Biorisk Management Memo of Instruction, undated.

———, "Biorisk Management Implementation Guidance, Cooperative Biological Engagement Program Design Basis Threat (Attachment 4)," CTB-11-094, December 21, 2011.

———, *Research Strategic Plan: Addressing Biological Threat Reduction Through Research*, December 2012.

———, *CBEP Program Overview*, 2013a.

———, "Training Workshop: Participant Introductions and Priorities," February 27–28, 2013b.

———, "Program Update," April 2, 2013c.

Creedon, Madelyn R., and Andrew C. Weber, *Strategic Policy Guidance for the Cooperative Biological Engagement Program*, Washington, D.C.: U.S. Department of Defense, August 21, 2013.

Davis, Paul K., and Paul Dreyer, *RAND's Portfolio Analysis Tool (PAT): Theory, Methods, and Reference Manual*, Santa Monica, Calif.: RAND Corporation, TR-756-OSD, 2009. As of May 21, 2014:
http://www.rand.org/pubs/technical_reports/TR756.html

Davis, Paul K., Russell D. Shaver, and Justin Beck, *Portfolio-Analysis Methods for Assessing Capability Options*, Santa Monica, Calif.: RAND Corporation, MG-662-OSD, 2008. As of May 21, 2014:
http://www.rand.org/pubs/monographs/MG662.html

de Sa, Joia, Sandra Mounier-Jack, Chau Darapheak, Ly Khunbun Narann, Rattanaxay Phetsouvanh, Nyphonh Chanthakoummane, Sok Touch, Bounlay Phommasack, and Richard Coker, "Responding to Pandemic Influenza in Cambodia and Lao PDR: Challenges in Moving from Strategy to Operation," *Southeast Asian Journal of Tropical Medicine and Public Health*, Vol. 41, No. 5, September 2010, pp. 1104–1115.

Defense Threat Reduction Agency, "Overview of BTRP Cooperative Biological Research (CBR) Project," briefing, January 19, 2007.

———, "Electronic Integrated Disease Surveillance System (EIDSS)," briefing, July 14, 2011.

———, "DTRA, SCC-WMD and SJFHQ-E," fact sheet, 2012.

Defense Threat Reduction Agency, Cooperative Biological Engagement Program, *Full Operational Capability Country Test Plan (Georgia)*, undated.

Department of Defense Directive 2060.02, *Department of Defense (DoD) Combating Weapons of Mass Destruction (WMD) Policy*, April 19, 2007.

Department of Defense Instruction 5210.89, *Minimum Security Standards for Safeguarding Biological Select Agents and Toxins*, April 18, 2006.

Department of Health and Human Services, *National Biosurveillance Strategy for Human Health*, February 2010a.

———, *Healthy People 2020: Public Health Infrastructure Objectives*, 2010b. As of January 2014: http://www.healthypeople.gov/2020/topicsobjectives2020/objectiveslist.aspx?topicId=35

DoD—*See* U.S. Department of Defense.

Dreyer, Paul, Rick Eden, Matthew W. Lewis, and Christopher D. Nelson, "Toward a Dashboard Capability for Public Health Security: Minimal Requirements for Initial Implementation," unpublished RAND research, 2010.

DTRA, Threat Agent Detection and Response (TADR) Program.

European Committee for Standardization (CEN), *CEN Workshop Agreement (CWA) 15793: Laboratory Biorisk Management*, September 2011.

Feldbaum, Harley, *U.S. Global Health and National Security Policy*, Washington, D.C.: Center for Strategic and International Studies, 2009.

Government Accountability Office, *The Results Act: An Evaluator's Guide to Assessing Agency Annual Performance Plans*, Washington, D.C., GAO/CGD-10.1.20, April 1, 1998.

———, *Biological Weapons: Effort to Reduce Former Soviet Threat Offers Benefits, Poses New Risks*, Washington, D.C., NSIAD-00-138, April 28, 2000a.

———, *Weapons of Mass Destruction: DOD's Actions to Combat Weapons Use Should Be More Integrated and Focused*, Washington, D.C., NSIAD-00-97, May 26, 2000b.

Greenfield, Victoria A., Valerie L. Williams, and Elisa Eiseman, *Using Logic Models for Strategic Planning and Evaluation: Application to the National Center for Injury Prevention and Control*, Santa Monica, Calif.: RAND Corporation, TR-370-NCIPC, 2006. As of May 21, 2014: http://www.rand.org/pubs/technical_reports/TR370.html

Handelman, Kenneth, "Statement of Mr. Kenneth Handelman Acting Assistant Secretary of Defense for Global Strategic Affairs," before the Senate Armed Services Committee Subcommittee on Emerging Threats and Capabilities, Washington, D.C., May 10, 2011.

Hanvoravongchai, Piya, Wiku Adisasmito, Pham Ngoc Chau, Alexandra Conseil, Joia de Sa, Ralf Krumkamp, Sandra Mounier-Jack, Bounlay Phommasack, Weerasak Putthasri, Chin-Shui Shih, Sok Touch, and Richard Coker, "Pandemic Influenza Preparedness and Health Systems Challenges in Asia: Results from Rapid Analyses in 6 Asian Countries," *BMC Public Health*, Vol. 10, 2010, p. 322.

Hatry, Harry P., *Performance Measurement: Getting Results*, 2nd ed., Washington, D.C.: Urban Institute Press, 2007.

Hecht, Robert, and Raj Shah, "Recent Trends and Innovations in Development Assistance for Health," in Dean T. Jamison, Joel G. Breman, Anthony R. Measham, George Alleyne, Mariam Claeson, David B. Evans, Prabhat Jha, Anne Mills, and Philip Musgrove, eds., *Disease Control Priorities in Developing Countries*, 2nd ed., Washington, D.C.: World Bank, 2006, pp. 243–257.

Hess, Glenn, "Biosecurity Efforts Expands to Africa," *Chemical & Engineering News*, Vol. 89, No. 15, April 2011.

Ijaz, Kashef, Eric Kasowski, Ray R. Arthur, Frederick J. Angulo, and Scott F. Dowell, "International Health Regulations—What Gets Measured Gets Done," *Emerging Infectious Diseases*, July 2012.

Institute of Medicine, *The Future of the Public's Health in the 21st Century*, Washington, D.C.: National Academies Press, 2002.

Johns, Matthew C., and David L. Blazes, "International Health Regulations (2005) and the U.S. Department of Defense: Building Core Capacities on a Foundation of Partnership and Trust," *BMC Public Health*, Vol. 10, No. S4, 2010.

Katz, Rebecca L., Jose A. Fernandez, and Scott J. N. McNabb, "Disease Surveillance, Capacity Building and Implementation of the International Health Regulations," *BMC Public Health*, Vol. 10, No. S4, 2010.

Keeney, Ralph L., *Value-Focused Thinking: A Path to Creative Decisionmaking*, Cambridge, Mass.: Harvard University Press, 1992.

Kimball, Ann Marie, Melinda Moore, Howard Matthew French, Yuzo Arima, Kumnuan Ungchusak, Suwit Wibulpolprasert, Terence Taylor, Sok Touch, Alex Leventhal, "Regional Infectious Disease Surveillance Networks and Their Potential to Facilitate the Implementation of the International Health Regulations," *Medical Clinics of North America*, Vol. 92, No. 6, 2008, pp. 1459–1471.

Lenne, B., and H. Cleland, "Describing Program Logic," *Program Evaluation Bulletin*, Vol. 2, 1987.

McNabb, Scott J. N., Stella Chungong, Mike Ryan, Tadesse Wuhib, Peter Nsubuga, Wondi Alemu, Vilma Carande-Kulis, and Guenael Rodier, "Conceptual Framework of Public Health Surveillance and Action and Its Application in Health Sector Reform," *BMC Public Health*, Vol. 2, January 2002.

Michaud, Josh, Kellie Moss, and Jen Kates, *The U.S. Department of Defense and Global Health*, Menlo Park, Calif.: Henry J. Kaiser Family Foundation, September, 2012.

Miller, Rebecca, and Scott F. Dowell, *Investing in a Safer United States: What Is Global Health Security and Why Does It Matter?* Washington, D.C.: Center for Strategic and International Studies, August 2012.

Moore, Melinda, Edward Chan, Nicole Lurie, Agnes Gereben Schaefer, Danielle M. Varda, and John A. Zambrano, "Strategies to Improve Global Influenza Surveillance: A Decision Tool for Policymakers," *BMC Public Health*, Vol. 8, 2008.

Moore, Melinda, David J. Dausey, Bounlay Phommasack, Sok Touch, Lu Guoping, Soe Lwin Nyein, Kumnuan Ungchusak, Nguyen Dang Vung, and Moe Ko Oo, "Sustainability of Sub-Regional Disease Surveillance Networks," *Global Health Governance*, Vol. 5, No. 2, 2012.

Moroney, Jennifer D. P., and Joe Hogler, *Building Partner Capacity to Combat Weapons of Mass Destruction*, Santa Monica, Calif.: RAND Corporation, MG-783-DTRA, 2009. As of May 21, 2014: http://www.rand.org/pubs/monographs/MG783.html

Nacht, Michael, "Biological Threat Reduction Program Strategic Planning Guidance and Program Objectives," Memorandum to Programs, Assistant to the Secretary of Defense for Nuclear and Chemical and Biological Defense, August 18, 2009.

NAS—*See* National Academy of Sciences.

National Academy of Sciences, *Global Security Engagement: A New Model for Cooperative Threat Reduction*, Washington, D.C.: National Academies Press, 2009.

———, *Improving Metrics for the Department of Defense Cooperative Threat Reduction Program*, Washington, D.C.: National Academies Press, 2012.

National Academy of Sciences, Institute of Medicine, and National Research Council, *Controlling Dangerous Pathogens: A Blueprint for U.S.-Russian Cooperation, A Report to the Cooperative Threat Reduction Program of the U.S. Department of Defense*, Washington, D.C.: National Academies Press, 1997.

National Intelligence Council, *Strategic Implications of Global Health*, Washington, D.C., ICA 2008-10D, December 2008.

National Research Council, *An Assessment of the International Science and Technology Center: Redirecting Expertise in Weapons of Mass Destruction in the Former Soviet Office*, Washington, D.C.: National Academies Press, 1996.

———, "Letter Report on the Threat Agent Detection Response System Database," Washington, D.C.: National Academies Press, 2006.

———, *The Biological Threat Reduction Program of the Department of Defense: From Foreign Assistance to Sustainable Partnerships*, Washington, D.C.: National Academies Press, 2007.

———, *Countering Biological Threats: Challenges for the Department of Defense's Nonproliferation Program Beyond the Former Soviet Union*, Washington, D.C.: National Academies Press, 2009.

National Security Council, *National Strategy for Countering Biological Threats*, Washington, D.C., November 2009.

National Security Staff, *Promoting Global Health Security: Guidance and Principles for U.S. Government Departments and Agencies to Strengthen IHR Core Capacities Internationally*, June 2011.

Nikitin, Mary Beth D., and Amy F. Woolf, *The Evolution of Cooperative Threat Reduction: Issues for Congress*, Washington, D.C.: Congressional Research Service, July 8, 2013.

Nikitin, Mary Beth, Paul K. Kerr, and Steven A. Hildreth, *Proliferation Control Regimes: Background and Status*, Washington, D.C.: Congressional Research Service, October 25, 2012.

NRC—*See* National Research Council.

Nsubuga, Peter, Mark E. White, Stephen B. Thacker, Mark A. Anderson, Stephen B. Blount, Claire V. Broome, Tom M. Chiller, Victoria Espitia, Rubina Imtiaz, Dan Sosin, Donna F. Stroup, Robert V. Tauxe, Maya Vijayaraghavan, and Murray Trostle, "Public Health Surveillance: A Tool for Targeting and Monitoring Interventions," in Dean T. Jamison, Joel G. Breman, Anthony R. Measham, George Alleyne, Mariam Claeson, David B. Evans, Prabhat Jha, Anne Mills, and Philip Musgrove, eds., *Disease Control Priorities in Developing Countries*, 2nd ed., Washington, D.C.: World Bank, 2006, pp. 997–1015.

Organisation for Economic Co-operation and Development, *Paris Declaration on Aid Effectiveness*, 2005. As of April 21, 2011:
http://www.oecd.org/dataoecd/11/41/34428351.pdf

———, *Accra Agenda for Action*, 2008. As of April 21, 2011:
http://www.oecd.org/dataoecd/11/41/34428351.pdf

Oshitani, Hitoshi Taro Kamigaki, and Akira Suzuki, "Major Issues and Challenges of Influenza Pandemic Preparedness in Developing Countries," *Emerging Infectious Diseases*, Vol. 14, No. 6, June 2008, pp. 875-880.

Pappaioanou, M., M. Malison, K. Wilkins, B. Otto, R. Goodman, R.E. Churchill, et al., "Strengthening Capacity to Use Data for Public Health Decision Making: The Data for Decision Making Project," *Social Science and Medicine*, Vol. 57, No. 10, November 2003, pp. 1925–1937.

Peake, James B., J. Stephen Morrison, Michèle M. Ledgerwood, and Seth E. Gannon, *The Defense Department's Enduring Contributions to Global Health: The Future of the U.S. Army and Navy Overseas Medical Research Laboratories*, Washington, D.C.: Center for Strategic and International Studies, 2011.

Pesenti, Pete, "Science, Research and Operational Biosurveillance," unpublished CBEP briefing, June 17, 2013.

Public Law 107-107, National Defense Authorization Act for Fiscal Year 2002, December 28, 2001.

Public Law 110-181, National Defense Authorization Act for Fiscal Year 2008, January 28, 2008.

Public Law 111-84, National Defense Authorization Act for Fiscal Year 2010, October 28, 2009.

Reich, Michael R., and Keizo Kakemi, "G8 and Strengthening of Health Systems: Follow-Up to the Tokayo Summit," *Lancet*, Vol. 373, February 7, 2009, pp. 508–515.

Rockefeller Foundation, *Disease Surveillance Networks Initiative*, 2011. As of April 21, 2011: http://www.rockefellerfoundation.org/uploads/files/d3588732-0562-4dad-9aa0-a080862bec05-dsn.pdf

Rogers, P. J., A. Petrosino, T. A. Huebner, and T. A. Hacsi, "Program Theory Evaluation: Practice, Promise, and Problems," *New Directions for Evaluation*, Vol. 87, No. Fall, 2000, pp. 5–13.

Salaam-Blyther, Tiaji, *U.S. and International Responses to the Global Spread of Avian Flu: Issues for Congress*, Washington, D.C.: Congressional Research Service, 2006.

Schön, D. A., "Theory-of-Action Evaluation," paper presented to the Harvard Evaluation Task Force, April 1997.

Shelton, Shoshana R., Christopher D. Nelson, Anita W. McLees, Karen Mumford, and Craig Thomas, "Building Performance-Based Accountability with Limited Empirical Evidence: Performance Measurement for Public Health Preparedness," *Disaster Medicine and Public Health Prepredness*, 2013.

Smithson, Amy E., *Toxic Archipelago: Preventing Proliferation from the Former Soviet Chemical and Biological Weapons Complexes*, Washington, D.C.: Henry L. Stimson Center, December 1999.

Stroot, Philippe, and Ursula Jenal, "A New Approach: Contributing to BWC Compliance via Biosafety, Biosecurity, and Biorisk Management," *Nonproliferation Review*, Vol. 18, No. 3, 2011a.

———, "A New Approach: Contributing to BWC Compliance via Biosafety, Biosecurity, and Biorisk Management," *Nonproliferation Review*, Vol. 18, No. 3, 2011b.

Suchman, Edward Allen, *Evaluative Research: Principals and Practice in Public Service and Social Action Programs*, New York: Russell Sage Foundation, 1967.

Taboy, Celine H., Will Chapman, Adilya Albetkova, Sarah Kennedy, and Mark A. Rayfield, "Integrated Disease Investigations and Surveillance Planning: A Systems Approach to Strengthening National Surveillance and Detection of Events of Public Health Importance in Support of the International Health Regulations," *BMC Public Health*, Vol. 10, 2010.

Thacker, S. B., "Historical Development," in S. T. Teutsch, and R. E. Churchill, eds., *Principles and Practice of Public Health Surveillance*, New York: Oxford University Press, 2000.

Thompson, Donald F., "The Role of Medical Diplomacy in Stabilizing Afghanistan," *Defense Horizons*, Vol. 63, 2008.

United Nations, Revised Forms for the Submission of the Confidence-Building Measures, 2011.

United Nations Security Council 1540 Committee, *The 1540 Matrix*, 2005.

U.S. Department of Defense, *Cooperative Threat Reduction Annual Report to Congress Fiscal Year 2012*, Washington, D.C., December 31, 2010a.

———, *Quadrennial Defense Review*, Washington, D.C., February 2010b.

———, *Fiscal Year 2012 Budget Estimate Cooperative Threat Reduction Program*, Washington, D.C.: February 2011.

U.S. Department of Health and Human Services, *National Health Security Stategy of the United States of America*, 2009.

U.S. Department of Homeland Security, *Target Capabilities List: A Companion to the National Preparedness Guidelines*, September 2007.

———, *National Preparedness Goal*, September, 2011.

U.S. Global Health Initiative, *Guidance for Global Health Initiaitve Country Strategies: GHI Guidance 2.0*, Washington, D.C., May 2011.

U.S. Senate, Hearing Before the Committee on Armed Services United States Senate, One Hundred Twelfth Congress, First Session on S. 1253, to Authorize Appropriations for Fiscal Year 2012 for Military Activities of the Department of Defense and for Military Construction, to Prescribe Military Personnel Strengths for Fiscal Year 2012, and for Other Purposes: Part 5 Emerging Threats and Capabilities, May 10, 2011.

Voronova-Abrams, Marina, "Biosecurity 2.0: Enduring Threats in the Former Soviet Union," *Bulletin of the Atomic Scientists*, Vol. 67, No. 4, 2011.

Weaver, l. Michael, "Biosafety and Biosecurity Activities of the International Science and Technology Center in the Republics of the Former Soviet Union: Accomplishments, Challenges, and Prospects," *Applied Biosafey*, Vol. 15, No. 2, 2010.

Weber, Andrew C., *Strategic Implementation Guidance for the Cooperative Biological Engagement Program*, August 21, 2013.

Weiss, C. H., "How Can Theory-Based Evaluation Make Greater Headway?" *Evaluation Review*, Vol. 21, 1997, pp. 501–524.

White House, *The United States Global Health Initiative: Strategy Document*, May 2009.

———, *National Security Strategy*, May 2010.

———, S*ustaining U.S. Global Leadership: Priorities for 21st Century Defense*, January 3, 2012a.

———, *National Strategy for Biosurveillance*, July 2012b.

WHO—*See* World Health Organization.

Wickel, Kevin, "Cooperative Threat Reduction: Reducing Biological Risks in East Africa," Stimson Center, March 9, 2011.

Wilkins, K., P. Nsubuga, J. Mendlein, D. Mercer, and M. Pappaioanou, "The Data for Decision Making Project: Assessment of Surveillance Systems in Developing Countries to Improve Access to Public Health Information," *Public Health*, Vol. 122, 2008, pp. 914–922.

Woolf, Amy F., *Nonproliferation and Threat Reduction Assistance: U.S. Programs in the Former Soviet Union*, Congressional Research Service, March 6, 2012.

World Health Assembly, *Global Public Health Response to Natural Occurrence, Accidental Release or Deliberate Use of Biological and Chemical Agents or Radionuclear Material That Affect Health*, WHA Resolution 55.16, May 18, 2002.

———, *Revision of the International Health Regulations*, WHA Resolution 58.3, 2005.

World Health Organization, *Laboratory Biosafety Manual*, 3rd ed., 2004.

———, *International Health Regulations*, 2005.

———, *Biorisk Management: Laboratory Biosecurity Guidance*, WHO/CDS/EPR/2006.6, September 2006a.

———, *Assessing Content, Process Output and Outcomes of Rapid Response Team Training Course for Rapid Response Avian and Pandemic Influenza: Assessment Tools*, December 2006b. As of July 15, 2014:
http://209.61.208.138/LinkFiles/Rapid_Response_RRT_Trng_Assmt_Tool.pdf

———, *A Safer Future: Global Public Health Security in the 21st Century*, 2007.

———, *Health Metrics Network: Framework and Standards for Country Health Information Systems*, 2nd ed., Geneva, 2008.

———, "Maximizing Positive Synergies Collaborative Group: An Assessment of Interactions Between Global Health Initiatives and Country Health Systems," *Lancet*, Vol. 373, June 20, 2009, pp. 2137–2169.

———, *Protocol for Assessing National Surveillance and Response Capacities for the International Health Regulations (2005): A Guide for Assessment Teams*, December 2010.

———, *IHR Core Capacity Monitoring Framework: Checklist and Indicators for Monitoring Progress in the Development of IHR Core Capacities in States Parties*, February 2011.

The Cooperative Biological Engagement Program (CBEP) is the biological threat component of the Cooperative Threat Reduction program. It grew out of efforts to address risks associated with legacy biological agents, related materials, and technical expertise developed as part of the biological weapon program in the former Soviet Union. CBEP now partners with about 20 countries in different regions around the world and works with them to address diverse threats to international security, including terrorist organizations seeking to acquire pathogens of security concern; human, animal, and agricultural facilities operating with inadequate safety and security safeguards; and the spread of diseases with potential security or economic consequences. As the program has evolved since its inception two decades ago, so too have its content and approaches to performance measurement. The objective of the research reported here was to build on existing work to develop a comprehensive evaluation framework and recommend metrics for assessing and communicating progress toward CBEP's goals. The report ultimately recommends a number of qualitative and quantitative indicators of CBEP performance, some that can be implemented immediately, some to be implemented later.

NATIONAL DEFENSE RESEARCH INSTITUTE

www.rand.org

$37.50

ISBN-10 0-8330-8693-6
ISBN-13 978-0-8330-8693-8

RR-660-OSD