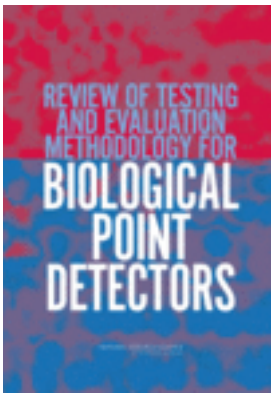

Review of Testing and Evaluation Methodology for Biological Point Detectors: Abbreviated Summary

Committee on the Review of Testing and Evaluation Methodology for Biological Point Detectors, National Research Council

ISBN: 0-309-53119-5, 38 pages, 6 x 9, (2005)

REVIEW OF TESTING AND EVALUATION METHODOLOGY FOR BIOLOGICAL POINT DETECTORS

ABBREVIATED SUMMARY

Committee on the Review of Testing and Evaluation Methodology for Biological Point Detectors

Board on Chemical Sciences and Technology

Division on Earth and Life Studies

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, N.W. Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

Support for this study was provided by the U.S. Department of Defense under contract number MD972-03-1-0009.

All opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for this project.

International Standard Book Number 0-309-09179-9

Additional copies of this report are available from:

The National Academies Press
500 Fifth Street, N.W.
Box 285
Washington, DC 20055
800-624-6242
202-334-3313 (in the Washington metropolitan area)
http://www.nap.edu

Printed in the United States of America

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce M. Alberts is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Wm. A. Wulf is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce M. Alberts and Dr. Wm. A. Wulf are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

COMMITTEE ON THE REVIEW OF TESTING AND EVALUATION METHODOLOGY FOR BIOLOGICAL POINT DETECTORS

J. Patrick Fitch, Lawrence Livermore National Laboratory, Chair
Mark P. Buttner, University of Nevada, Las Vegas
Wyett Colclasure II, Booz Allen Hamilton
Margaret E. Coleman, Syracuse Research Corporation*
Philip E. Coyle III, Science Strategies
Herbert H. Hill, Jr., Washington State University
Nancy B. Jackson, Sandia National Laboratories
Darrell L. Jan, Jet Propulsion Laboratory, California Institute of Technology
Bernadette Johnson, Massachusetts Institute of Technology Lincoln Laboratory
Timothy Moshier, Massachusetts Institute of Technology Lincoln Laboratory
Stephen N. Rudnick, Harvard University
Basil I. Swanson, Los Alamos National Laboratory
David R. Walt, Tufts University

Staff

Christopher K. Murphy, Program Officer
Tina Masciangioli, Program Officer
David C. Rasmussen, Program Assistant
Dorothy Zolandz, Director, Board on Chemical Sciences and Technology

*Formerly U.S. Department of Agriculture.

BOARD ON CHEMICAL SCIENCES AND TECHNOLOGY

Co-chairs

ARNOLD F. STANCELL (NAE), Georgia Institute of Technology, Atlanta, GA (emeritus)
A. WELFORD CASTLEMAN, JR. (NAS), Pennsylvania State University, University Park, PA

Members

PAUL T. ANASTAS, Green Chemistry Institute, Washington, DC
DENISE M. BARNES, Independent Consultant, Snellville, GA
ANDREA W. CHOW, Caliper Life Sciences, Mountain View, CA
MARK E. DAVIS (NAE), California Institute of Technology, Pasadena, CA
JEAN DE GRAEVE, Université de Liège, Liège, Belgium
MILES P. DRAKE, Air Products & Chemical Company, Allentown, PA
CATHERINE C. FENSELAU, University of Maryland, College Park, MD
GEORGE W. FLYNN (NAS), Columbia University, New York, NY
MAURICIO FUTRAN (NAE), Bristol-Myers Squibb Company, New Brunswick, NJ
LOU ANN HEIMBROOK, Merck & Company, Inc., Rahway, NJ
ROBERT HWANG, Brookhaven National Laboratory, Upton, NY
JAY V. IHLENFELD, 3M Research & Development, St. Paul, MN
JAMES L. KINSLEY (NAS), Rice University, Houston, TX
MARTHA A. KREBS, Science Strategies, Los Angeles, CA
WILLIAM A. LESTER, JR., University of California, Berkeley, CA
GREGORY O. NELSON, Eastman Chemical Company, Kingsport, TN
DONALD PROSNITZ, Lawrence Livermore National Laboratory, Livermore, CA
ELSA REICHMANIS (NAE), Lucent Technologies, Murray Hill, NJ
MATTHEW V. TIRRELL (NAE), University of California, Santa Barbara, CA

National Research Council Staff

KAREN LAI, Research Associate, Christine Mirzayan Graduate Fellow (through December 2004)
TINA M. MASCIANGIOLI, Program Officer
ERICKA M. MCGOWAN, Research Associate
CHRISTOPHER K. MURPHY, Program Officer
SYBIL A. PAIGE, Administrative Associate
DAVID C. RASMUSSEN, Project Assistant
DOROTHY ZOLANDZ, Director

Acknowledgment of Reviewers

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making the published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their review of this report:

Robert J. Cotter, Johns Hopkins School of Medicine
Stephen W. Drew, Merck & Company, Inc. (retired)
Emil Gotschlich, The Rockefeller University
Michael Ladisch, Purdue University
Janet M. Macher, California Department of Health Services
Edwin P. Przybylowicz, Eastman Kodak Company (retired)
Dean Wilkening, Stanford University

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations nor did they see the final draft of the report before its release. The review of this report was overseen by Royce W. Murray, University of North Carolina, and Warner D. North, NorthWorks, Inc. Appointed by the National

Research Council, they were responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

Preface

The Department of Defense Joint Program Executive Office for Chemical and Biological Defense (DoD JPEO-CBD) requested that a committee of the National Research Council review current biological point-detection system testing protocols and integrated evaluation methodologies. Specific consideration of the Joint Biological Point Detection System (JBPDS) and a whole system live agent test and evaluation (T&E) strategy were requested.

The focus of the final report is a specific proposal from the West Desert Test Center, Dugway Proving Ground (DPG) for a whole system live agent testing facility (WSLAT). The committee considered scientific, technological and regulatory aspects of the WSLAT proposal in its critique and offers suggestions for improvement if WSLAT were selected as part of the near term test and evaluation strategy. Because of identified schedule and scientific risks, we suggest an alternate approach that focuses test and evaluation efforts on leveraging existing data, improving agent simulants, representing appropriate inhibitors and backgrounds, and modeling for performance prediction to extrapolate from relatively controlled environments with simulant testing to complex environments with live agent operations. The committee believes that the DOD, in fact, needs significant resources to expand the T&E strategy to include elements of both approaches.

The committee's report was originally transmitted to the sponsor on July 23, 2004, requesting confirmation that the report is in fact unclassified and can be made available to the public without restriction, to include the possibility of posting it on the National Academy of Sciences' (NAS) world wide web site. On September 24, 2004, BG Stephen V. Reeves, U.S. Army, Joint Program Executive Officer for Chemical and Biological Defense, informed the NAS that JPEO-CBD

"recommends that the distribution of the subject report be restricted to U.S. Government Agencies and their contractors only."[1] Further, the letter states "Though the report is unclassified, per Department of Defense Directive 5230.24, the distribution must be restricted because it contains information concerning keystone test equipment, characterized in the regulation as "critical technology." Additionally, potential military applications, system test and evaluation limitations, and biological defense vulnerabilities are discussed throughout the document. The document should also be marked export-controlled per Department of Defense Directive 5230.25."

Section 15 of the Federal Advisory Committee Act (FACA) provides that the Academy shall make its final report available to the public unless the Academy determines that the report would disclose matters described in one or more of the exemption provisions under the Freedom of Information Act (FOIA). If the Academy determines that the report would disclose matters described in one or more of the FOIA exemptions, "the Academy shall make public an abbreviated version of the report that does not disclose such matters."

Paragraph 4.5 of DoD Directive 5230.25 states "The authority provided herein may not be used to withhold from public disclosure unclassified information regarding DoD operations, policies, activities, or programs, including the costs and evaluations of performance and reliability of military and space equipment. When such information does contain technical data subject to this Directive, the technical data shall be excised from that which is disclosed publicly." Accordingly, all technical data has been excised from the attached Summary.

The Academy has determined that this Summary does not disclose matters described in any of the FOIA exemptions.

All of the committee's recommendations remain identical to the recommendations in the version of the report delivered to JPEO-CBD.

[1]Letter, BG Stephen V. Reeves, U.S. Army, Joint Program Executive Officer for Chemical and Biological Defense to Mr. Kevin Hale, Director, Program Security, The National Academies, September 24, 2004.

Contents

Review of Testing and Evaluation Methodology for Biological Point Detectors

Final Report

Executive Summary

Biodetector validation is based on a range of metrics that should increase confidence in the system and its ability to meet specific performance goals. In the end the warfighter wants confidence in a system's performance and the sponsor wants assurance that the funds are well spent. Both of these goals require managing risks.

In its Statement of Work (see Appendix) the committee was asked to "review current biological point-detection system testing protocols and integrated evaluation methodologies", and "assess the feasibility and benefits of the DOT&E [Director, Operational Test and Evaluation] requirement for a systems-level, active agent testing capability." Despite the importance of the Joint Biological Point Detection System (JBPDS) production decision, the committee is not convinced that the planned Whole System Live Agent Testing (WSLAT) facility alone will provide the information required for that decision. Without an integrated test and evaluation methodology and a suite of optimized and dependable simulants, the agent/simulant correlation measurements intended to be conducted in WSLAT and elsewhere will not be meaningful. Also, changes in threat requirements and unexpected results with live agents in WSLAT or other facilities will require dependable simulants and a consistent test methodology for resolution. The WSLAT performance requirements document equivocates on critical threat details, and was changed during the committee's work.

A preliminary review and assessment of these issues was made by the committee in its interim report. In Chapter 1 of this report, the committee provides a general introduction and a summary of the Interim Report that addressed the first item of the committee's Statement of Work.

With additional data available to the committee, these issues are examined in further detail in Chapter 2. Additionally, the committee was asked in its Statement of Work to review the Whole System Live Agent Testing proposal provided under a study from the West Desert Test Center, Dugway Proving Ground. This review is provided in Chapter 3 of this report.

In its overall evaluation the committee agrees with the concept of WSLAT, or a WSLAT-like facility, in which biological detectors will be evaluated against threat representative challenges. However, the committee concludes that the purposes and need for WSLAT must be clarified and defined more rigorously, since the proposed WSLAT program carries high schedule risk, high cost and questionable feasibility of adequate system performance. The WSLAT program does not present a clear plan for the testing to be done, data to be produced, or how the data will be analyzed and applied to the decision. These issues have not yet been addressed and included in schedule and budget plans for WSLAT.

The committee concludes that in parallel with the planned WSLAT program, an essential matching effort is needed by the DOD to develop a well-defined testing and evaluation methodology.

In Chapter 2, the committee provides the following three high-level recommendations (recommendations 2-1, 2-2, and 2-3), and justification relative to the Interim Report and information provided to the committee subsequent to the release of the Interim Report.

Recommendation 2-1: A scientifically credible integrated test and evaluation methodology should be established and implemented to guide sensor validation and subsequent production decisions.

This test and evaluation methodology is detailed further in recommendations 2-4, 2-5, and 2-6.

Recommendation 2-2: Facilities are needed for validation of JBPDS and other biodetection systems, and should be developed in a phased manner where each phase adds to the ability to predict performance capability of the system under testing.

This facility could be a WSLAT-like facility that evolves from the original concept to support changes in the test and evaluation methodology and to address shortcomings identified by the committee in its review. The recommended strategy is to work backwards from the integrated T&E plan to the data and analysis needed and then to specification of appropriate facilities and infrastructure. A mobile whole system T&E capability is recommended to address T&E in representative environments and backgrounds.

Recommendation 2-3: An independent expert advisory committee should be created to provide guidance on JBPDS testing and subsequent detection system development.

Based on the collective expertise and experience of the committee, the technical challenges inherent in the development and implementation of evaluation methodology for the JBPDS warrant outside scientific advice through an independent team. Because this team could be composed of individuals both inside and outside DOD, more wide ranging scientific expertise can assist project managers in moving forward with additional confidence in the scientific rationale for the decisions that DOD needs to make.

The integrated T&E plan should be externally peer reviewed as should any anticipated changes to the WSLAT proposal.

Each of the following recommendations related to testing and evaluation methodology is developed in detail in Chapter 2.

Recommendation 2-4: A suite of optimized and dependable simulants should be developed.

Recommendation 2-5: Modeling and correlation analyses of components and whole systems should be developed.

Recommendation 2-6: Testing and modeling should be constructed in complex, operationally realistic environments.

With the exception of component testing of simulants and live agents, it is clear that Dugway Proving Ground does not currently have adequate test facilities. It is for this reason that the committee recommends that a Whole System Killed Related Agent Testing (WSKRAT) facility be constructed and used to begin implementation of the test methodology described above.

Recommendation 2-7: A WSKRAT facility should be constructed that would permit the development of predictive models and the testing of the adequacy of killed related-strain agents as simulants for live agents.

The use of killed related strain agents provide the best opportunity to both mimic live agents and to pass regulatory hurdles that may limit testing. Another advantage of such a facility is that WSKRAT would be a fully contained system with no open-air release, implying that some testing can begin immediately. In addition, by avoiding the use of live agents, it should be possible to construct a mobile WSKRAT facility that can test in realistic operational environments.

Recommendation 2-8: A WSKRAT mobile test facility should be constructed to assess the ability of system models to predict whole system behavior in operational environments.

While the committee is convinced that the development of WSKRAT is essential as part of a complete set of test and evaluation facilities, a WSLAT-like facility also will be necessary. However, without the needed data provided by a

graded approach that includes WSKRAT, a WSLAT itself will not represent an adequate test strategy.

Recommendation 2-9: A test facility should be constructed for final validation of JBPDS or any future point detection system.

In Chapter 3, the committee provides its evaluation of the WSLAT feasibility study, the second item in the committee's Statement of Work, including the following recommendation and a summary of the justification developed more fully in Chapter 3.

Recommendation 3-1: An analysis of alternative WSLAT design approaches that would include comparisons of the various key performance requirements in each design approach should be undertaken. The analysis should include a thorough discussion and analysis of risk, mitigation methods, and development of a sound risk mitigation plan.

The committee is not optimistic that the proposed WSLAT approach will succeed. The design has major unresolved technical questions. The schedule is highly compressed and is already behind. The committee also concluded that risk identification, analysis, and proposed management were particularly weak. The committee's observations on the feasibility study are presented in three groups. The first addresses the requirements analysis, the second addresses the scientific and engineering approach, and the third group of comments addresses regulatory and policy concerns.

In addition to these comments specifically about the WSLAT feasibility study, the committee notes that a credible testing program requires more than testing facilities. Supporting methodology, materials (e.g., agents and simulants), and validated processes are also needed. To be useful, WSLAT should be operated according to an overall, integrated test and evaluation methodology. The live agents it will use must be validated and correlated with a suite of optimized and dependable simulants. Neither the methodology nor the suite of simulants is adequate, and WSLAT alone will not suffice.

In Chapter 4, the committee suggests an alternative and complementary process for the testing and evaluation of biopoint detection systems in light of the recommendations and evaluations provided by the committee in this report.

Assuming that simulants can be selected that mimic the biological agents[1] of concern, the predictive capability of the system performance testing protocol increases in the following order (this list is in order of increasing confidence levels):

[1]Biological agents are live microorganisms or toxins that can incapacitate or kill humans and animals, and damage crops.

Level 1	Component testing with simulants
Level 2	Component testing with live agents
Level 3	Whole system testing with simulants alone
Level 4	Whole system testing with live agents
Level 5	Whole system testing with simulants under environmentally representative conditions
Level 6	Whole system testing with live agents under environmentally representative conditions

Recommendation 4-1: Assuming good simulants exist, the committee recommends testing to proceed from component testing with simulants (Level 1), and live agents (Level 2), through whole system testing with simulants (Level 3), and then to whole system testing with simulants under environmentally representative conditions (Level 5).

With an ideal simulant, there will be sufficient confidence in the simulant's ability to faithfully mimic the agent(s) of concern that whole system live agent testing will not be required. If such simulants do not exist then Level 4 (WSLAT) is necessary before going to Level 5. The committee's position is that testing at Level 5 should be the minimum for system certification. The committee assumes that Level 6 will not be permitted as part of any evaluation procedure.

A WSLAT-like capability likely may be necessary in the future because adequate simulants for all agents of concern will not be found, or because adequate simulants may prove to be too time consuming or costly to develop for all agents of concern, or because new agents of concern may arise for which a WSLAT capability is needed to support the warfighter in the field on short notice, or because the modeling may not develop dependable relationships in all cases.

A WSLAT facility alone will not obviate the needs for an integrated T&E plan, method development including better simulants, and modeling and analysis for predicting operational performance against live agents. The commitment to develop a feasible WSLAT plan, therefore, necessitates parallel investments in these other areas.

1

Introduction

INTERIM REPORT SUMMARY

Prior to the present Final Report, the committee issued an Interim Report that addressed several of the committee's tasks for this study (the committee's statement of task can be found in the Appendix). A brief summary of the Interim Report is provided below.

The committee developed a technical framework that considers the decision process for test and evaluation, including the WSLAT option. The risk mitigating value of live agent simulants and component-level testing were evaluated in the context of biological point detection systems in general and JBPDS in particular. Three specific questions are integral to the decision framework.

1. Are simulants good surrogates for live agents?
2. Can whole system performance against live agents be predicted from whole system and component testing with simulants?
3. Are testing methodology and facilities adequate and complete?

Are Simulants Good Surrogates for Live Agents?

The four most commonly used biothreat agent simulants under consideration, *Bacillus subtilis* var *niger* (BG), *Erwinia herbicola* (EH), *Male Specific Coliphage* (MS2), and Ovalbumin (OV) were evaluated for effectiveness as live agent surrogates. The consensus of the committee is that BG is the only reasonable surrogate for a full range of biothreat agent properties (e.g. physical, antigenic).

A single simulant is unlikely to meet the requirements for all live agents of concern; therefore, the committee believes that representative simulants may be selected for classes of similar or related live agents.

When available, killed related strains provide effective surrogates to biothreat agents, with the benefit of less stringent handling. The component tests can be conducted at lower biosafety levels than the corresponding live-agent tests, thereby greatly increasing the flexibility and variability of component testing (although the live-agent component testing is still required).

The committee supports the use of the current simulants in open-air whole-system tests as part of operational checkout, i.e., environmental tolerances, maintenance and logistical support.

The use of killed related strains may not obviate the need for a WSLAT facility. It remains to be determined whether whole system performance can be extrapolated from component testing.

Can Whole System Performance Against Live Agents Be Predicted from Whole System and Component Testing with Simulants?

Determination of the relationships between component and whole system measurements requires a high degree of systematic testing and analysis. Multiple performance models relating system response to simulants and live agents should be considered to address the different metrics, including sensitivity and false alarm rates, in component and whole system testing.

To address component-whole system correlation at an appropriate level of detail, there should be an integrated experimental plan, coordination of data analysis reports across component and system T&E, system modeling, and creation of an integrated evaluation team. Data will be needed for each simulant category from components and whole system testing, as well as data from component testing for selected agents of concern.

The committee considered the risks and benefits of several whole system testing approaches. The committee is of the opinion that whole system killed related agent testing (WSKRAT) may provide nearly the same level of confidence as WSLAT, while conferring low risk of failure because much of the testing can be done in existing facilities. This opinion assumes that killed related agents can be identified and demonstrated for all live biological agents of concern.

The committee has also examined the options for whole-system testing, including simulant only, killed agent, killed related strain, and live agent. Assuming that comparative tests are conducted and validate the use of killed related strain pathogens, they should be used when possible so that the opportunities for aerosol testing can be increased to include tunnel tests of the whole system.

Are Testing Methodology and Facilities Adequate and Complete?

Both the test methodology and the facilities appear to need additional resources to pursue WSLAT or any of the other options the committee investigated. A WSLAT facility would provide value to both military and counterterrorism programs. However, the committee has concerns that regulatory requirements and meeting all legal stipulations will delay construction and use of a WSLAT facility in time to impact JBPDS decisions. Even if the schedule were guaranteed, extensive testing will be required. Even more important from the point of view of facilities is the fact that there are pathogens that simply cannot be done with a WSLAT strategy.

To date there has been insufficient work to define "representative" backgrounds that could then be reproduced in controlled testing facilities.

A possible means for achieving this in the near term is to inject "qualitative" background aerosols into low containment whole system testing. In parallel, the DOD should undertake efforts to develop quantitatively defined, representative background aerosols that can be reproducibly injected into whole system tests.

To mitigate the risk of WSLAT the committee recommends an investigation of the predictive value of using killed related agents as simulants. A Whole System Killed Related Agent Test (WSKRAT) strategy significantly reduces schedule risks and possibly OT&E costs compared to WSLAT. In the committee's opinion, an alternative approach to WSKRAT, using a Whole System Killed Agent (WSKAT) strategy, would not significantly reduce the costs or schedule risks compared with WSLAT. In addition to the challenges associated with biological agents, the T&E strategy should formally incorporate system performance variation associated with changing backgrounds and inhibitors that are representative of deployment environments, and should decrease technical risks.

THE FINAL REPORT

The Whole System Live Agent Testing (WSLAT) proposal was not specifically evaluated in the Interim Report, but was generally considered as one option to assist in the JBPDS full-rate-production (FRP) decision process.[1] At the time of drafting its Interim Report the committee had little documentation to support its answers to the questions addressed in that report; subsequently the committee received several briefings, reports, data summaries, and communications with representatives of the Joint Program Executive Office for Chemical and Biological Defense (JPEO-CBD). Although it has not seen complete information, the committee believes that it is sufficiently aware of the types of data that have been generated to elaborate on and complete its recommendations. Thus, the purpose of this final report is to expand upon the recommendations provided in the interim report and to evaluate the specific WSLAT proposal under consideration.

[1]"Full-rate production" is defined as contracting for economic production quantities following stabilization of the system design and validation of the production process.

The remainder of the report is structured into three chapters. Chapter 2 summarizes the committee's high-level recommendations that result from its evaluation of data, briefings, and reports that became available after the release of the Interim Report and that address the need for development of a scientifically credible evaluation plan. Note that no information obtained subsequent to the release of the Interim Report has altered the primary recommendations that we made; rather, any new information has served primarily to strengthen our confidence in the credibility of these recommendations. Chapter 3 examines the WSLAT proposal under current consideration, addressing the questions;

1. **Will construction of the WSLAT facility, as defined in the current proposal, provide a capability for JBPDS testing that currently does not exist?** Yes, it would be a unique facility that will require an associated enhanced infrastructure.

2. **Is this capability technically necessary and sufficient?** The capability is necessary to provide a basis for a JBPDS FRP decision on a reasonable time scale, but not sufficient. Additional capabilities will be needed in new methods, modeling and analysis to extrapolate WSLAT data to detector performance in operational environments.

3. **Can this facility be constructed within the schedule constraints imposed by the full-rate production decision?** Uncertain, because the committee has identified areas for technical improvement in the specific proposal as well as in the overall T&E plan that increase schedule risks.

4. **What is the technical risk to the JBPDS program of not fabricating the WSLAT facility as proposed?** The technical risk to the JBPDS program is the risk of not having the methodology to properly use the results from WSLAT. Without good methodology there is also the risk of accepting (or rejecting) the JBPDS based on incomplete or faulty test data. If the technical evaluation cannot be completed, then the JBPDS FRP slips day for day until they can develop another strategy. Technical and schedule risk for WSLAT can be reduced by addressing the areas of concern the committee identifies in Chapter 3 and holding an external design review for the updated plan.

Chapter 4 provides an outline of recommended evaluation strategies that the committee believes are both technically defensible and that will more adequately support the Army's needs for biological agent[2] detection system testing and evaluation than would the current WSLAT proposal alone.

[2]Biological agents are live microorganisms or toxins that can incapacitate or kill humans and animals, and damage crops.

2

Overview and High-Level Recommendations

There are three high-level recommendations that derive from the committee's assessment of the current testing and evaluation methodologies.

Recommendation 2-1: A scientifically credible integrated test and evaluation methodology should be established and implemented to guide sensor validation and subsequent production decisions.

Recommendation 2-2: Facilities needed for validation of JBPDS and other biodetection systems should be developed in a phased manner where each phase adds to the ability to predict performance capability of the system under testing.

Recommendation 2-3: An independent expert advisory committee should be created to provide guidance on JBPDS testing and subsequent detection system development.

Each of these recommendations will be discussed in detail in the following sections. Additional recommendations related to these high-level recommendations are also in these sections.

TEST AND EVALUATION METHODOLOGY

As the committee received information on the process by which the JBPDS program had identified performance requirements, presented testing criteria, and evaluated early sensor constructs, it became clear that a systematic and scientifically credible test methodology would still be needed to guide production deci-

sions for JBPDS. As detailed in the interim report, this test methodology should comprise

- component testing with optimized simulants and live agents;
- whole system testing with simulants and, where necessary, live agents;
- predictive correlation analyses of component to whole system performance; and
- testing and correlation analyses in realistic operational environments.

Although these testing elements appear straightforward, and much has been done consistent with the committee's recommendations, it appears that these testing elements have not been applied in a systematic and rigorous manner to the evaluation process as a whole. Therefore, the committee examined each of these testing elements in order to provide further guidance and elaboration on their justification and implementation.

The committee was not provided protocols and corresponding data to assess the performance of JBPDS components or whole system upon which to base a decision for operational deployment of these systems. The first fundamental need for a credible integrated test methodology for JBPDS is the development of a suite of well-correlated simulants for each live agent of concern or for appropriate classes of live agents. The second need is the development of predictive models and correlation analyses between component response and whole system performance.

Recommendation 2-4: A suite of optimized and dependable simulants should be developed.

It is important that the term "simulant" be clearly defined for the purposes of this report. Simulants are nonpathogenic organisms that are employed to minimize the exposure risks to testing personnel. In fact, simulants are anything that provides useful, evaluative information on the performance of the system under test, short of the actual biologically active, weaponized warfare agents themselves. Other terms used frequently in this report are defined as follows:

Live Agent—One or more of the viable pathogenic organisms or active toxic compounds as cited in the International Task Force-6 (ITF-6) report of 9 February 1990.

Killed Agent—The nonviable or inactive forms of the aforementioned live agents.

Live Related-Strain Agent—One or more viable organisms or active compounds that closely resembles, genetically or chemically, a live agent as described above EXCEPT that it is of low to zero pathogenicity, virulence or toxicity.

Killed Related-Strain Agents—Nonviable or inactivated form of related-strain agent.

The better the simulant (or surrogate) mimics the biological warfare agents of concern (i.e., creates the same response in the system under test) the more confidence evaluators and warfighters will have in the system's ability to fulfill its mission. Equally important to the biological verisimilitude of the simulant is the ability for testers to employ the simulant in a threat representative manner and in operationally relevant environments. In addition, simulants should be easily quantifiable so that evaluators can in turn provide a quantitative measure of the tested system's performance. In summary, the ideal simulant will have the following properties:

- It will interact with the system under test in a manner that can be directly correlated with the analogous warfare agent.
- It can be presented to the system under test in a threat representative manner, and under operationally relevant environments.
- It supports quantitative measure of simulant challenges to the system under test.
- It should provide greater flexibility in whole system testing and evaluation compared with actual agents.

The committee's recommendations for better simulants derives entirely from the need to exercise the system in as realistic a fashion as possible so that results derived from simulant testing can be used to predict live-agent performance. Optimization of the actual JBPDS performance itself should be an ongoing but separate effort.

The committee suggests that the best simulants to mimic the behavior of live agents would be killed agents or killed related strains of live agents.

Provided that the biodetection component response to these killed related strain agents can be correlated with that of their pathogenic live counterparts, their relative ease of growth, preparation, and handling will provide for more realistic and frequent whole-system testing.

Recommendation 2-5: Modeling and correlation analyses of components and whole systems should be developed.

Substantial emphasis was placed in the Interim Report on the importance of developing both component and whole system performance prediction models. The assumptions are twofold: (1) The component response to simulants can be statistically correlated to those of live agents, and (2) whole system models can be developed with good predictive value even in realistic environments. If these assumptions are true, then component-to-whole-system modeling *may* enable live agent performance prediction from simulant whole system testing, and therefore obviate the need for WSLAT under some conditions.

The committee notes that additional expertise in system modeling may be required to supplement the current team that is slated to execute the tests; options for providing this essential expertise should be explored.

Recommendation 2-6: Testing and modeling should be conducted in complex, operationally realistic environments.

Degradation of performance when complex systems are taken to the field is the rule rather than the exception, and this is the reason for the committee's consensus that testing in realistic environments is as important, *if not more important*, than whole system testing against live agents. The integrated test methodology should include a path forward for eventual tests in more challenging operational environments.

While the committee recognizes the complexity of such modeling and analyses outlined above, we believe they are within the state of practice in the field. The critical question then becomes, "Which facilities and expertise are necessary and sufficient to support the testing and model validation?"

TESTING AND VALIDATION FACILITY REQUIREMENTS

In reviewing the available test facilities at DPG and in assessing what will be needed for system validation it is clear that with the exception of component testing of simulants and live agents, DPG does not currently have adequate test facilities. It is for this reason that the committee recommends that a Whole System Killed Related Agent Testing (WSKRAT) facility be constructed and used to begin implementation of the test methodology presented in the previous section. The WSKRAT would be a fully contained system with no open-air release. Techniques for introduction of transient environmental interferents should be developed to provide more realistic conditions for whole system testing.

Recommendation 2-7: A WSKRAT facility should be constructed that would permit the development of predictive models and the testing of the adequacy of killed related-strain agents as simulants for live agents.

As noted previously, it is essential to ultimately test in realistic operational environments. While much can be learned from the introduction of collected background air into the WSKRAT facility, it is difficult to reproduce the natural variation found in the broad range of likely JBPDS deployment environments. Therefore a mobile test facility should also be constructed that can accommodate simulants that mimic live agents. The mobile WSKRAT would enable optimized simulant challenges (to mimic live agents) in varying complex operational environments. This system would offer capability in assessing operational performance as well as in validating the whole system correlation models.

Recommendation 2-8: A WSKRAT mobile test facility should be constructed to assess the ability of system models to predict whole system behavior in operational environments.

The cornerstones of the committee's recommended approach are high correlation between component results for killed related strain and live agents and the ability of predictive models to reproduce whole system behavior in operational environments. A phased approach to test facility development has the advantages of immediately proceeding with implementation of the systematic test and evaluation methodology proposed while also providing valuable data to guide the development of an end-point facility.

While the committee is convinced that the development and construction of both stationary and mobile WSKRAT facilities are essential, we suspect that a WSLAT facility also will be necessary as part of a complete set of test and evaluation facilities. However, while WSLAT may become necessary, by itself it will not represent an adequate test strategy. The proposed graded approach—including WSKRAT—will provide the needed data.

Recommendation 2-9: A test facility should be constructed for final validation of JBPDS or any future point detection system.

The arguments for a WSLAT-like facility are to

1. complete the validation of system models.
2. test systems or agents not amenable to modeling.
3. test agents not amenable to the killed related strain strategy.
4. establish the facilities and expertise.

The committee stresses that design and construction of an appropriate test facility requires generation of an extensive scientific body of evidence, as well as substantial analysis and modeling, before an effective WSLAT facility can be developed; reliance on the current WSLAT proposal alone will not adequately address the needs for test validation to support JBPDS testing, evaluation, and production. In summary, the committee proposes that the order of priorities in test and evaluation should be

1. implementation of a scientifically credible, integrated test and evaluation methodology;
2. component testing of simulants and live agents (and the development of models to correlate the two);
3. component and whole-system testing in WSKRAT (and the development of predictive system performance models);
4. testing and evaluation in complex operational environments (with model refinement as required); and

5. whole system live-agent testing and evaluation in a WSLAT-like facility (with model validation).

ESTABLISHMENT OF AN INDEPENDENT ADVISORY COMMITTEE

The need for an independent expert committee to advise the JPEO on the development of credible integrated test and evaluation methodology was apparent by the lack of submission of any protocols and systematic methodologies for the committee to review. The scientific, management, and regulatory demands of such methods development are daunting, and JPEO would be well served by a formal advisory committee to accelerate the progress and enhance the utility of future products of the program.

The overall evaluation methodology for the JBPDS (and biological point detectors in general) should be documented and submitted for peer review.

Recommendation 2-3: An independent expert advisory committee should be created to provide guidance on JBPDS testing and subsequent detection system development.

Based on the collective expertise and experience of the committee, the technical challenges inherent in the development and implementation of evaluation methodology for the JBPDS warrant outside scientific advice through an independent team. Because this team could be composed of individuals both inside and outside DOD, more wide ranging scientific expertise can assist project managers in moving forward with additional confidence in the scientific rationale for the decisions that DOD needs to make.

This team should include a broad spectrum of expertise spanning the technical breadth required for the development, testing, and evaluation of systems such as the JBPDS. Although the details of the charter for this team should be determined by the JPEO-CBD and DOT&E, the committee recommends that this advisory body be used to assess future proposals for test facilities and test and evaluation protocols for JBPDS and next generation systems. The committee expects that the proposed independent team will complement and greatly assist the expertise resident in the existing programs.

3

Evaluation of the WSLAT Feasibility Study

This chapter presents the results of the committee's review of the WSLAT feasibility study[1] that was specifically requested in the Statement of Work. The WSLAT feasibility study responds to the directive dated July 9, 2002, from the Director, Operation Test and Evaluation (DOT&E) to incorporate whole system testing with live biological warfare agents, both before and after the full-rate production decision for JBPDS.[2] The committee's consideration of the overall need for WSLAT, and a recommended alternative approach, are presented elsewhere in this report. This section is directed at the feasibility study of February 2004 and the requirement on which it is based, and is independent of the committee's other recommendations.

The WSLAT feasibility study proposes a single design approach to satisfy the requirements of the DOT&E directive. It concludes with a plan and construction schedule to meet the technical and operational deadlines, with the clear proviso that certain decisions and funding should be made in a timely manner. The committee notes that the time for decision and funding is now several months overdue. As a result, the risks of not establishing WSLAT successfully and not

[1]Feasibility Study for the Whole System Live Agent Testing Program, Battelle Memorial Institute, 505 King Avenue, Columbus, OH 43201. WDTC Document No. WDTC-FS-03-125, West Desert Test Center, U.S. Army Dugway Proving Ground, UT 84022-5000. February 2004. This document is referred to as the WSLAT feasibility study in this report.

[2]Memo, July 9, 2002, Office of the Secretary of Defense, Operational Test and Evaluation, Deputy Director, SUBJECT: Joint Biological Point Detection System (JBPDS) Test and Evaluation (T&E) Strategy.

obtaining data and results to support the JBPDS decision schedule are already very high and becoming higher as the schedule compresses.

The committee is not optimistic that the proposed WSLAT approach will succeed. The design has major unresolved technical questions. The schedule is highly compressed and already behind. The committee also concluded that risk identification, analysis, and proposed management were particularly weak. The committee's observations on the feasibility study will be presented in three primary groups. The first addresses the requirements analysis, the second addresses the scientific and engineering approach, and the third group of comments addresses regulatory and policy concerns.

In addition to these comments specifically about the WSLAT feasibility study, the committee has noted in the report that a credible testing program requires more than testing facilities. It also needs supporting methodology, materials (e.g., agents and simulants), and validated procedures. To be useful, the WSLAT should be operated according to an overall, integrated test and evaluation methodology. The live agents it will use must be validated and correlated with a suite of optimized and dependable simulants. Neither the methodology nor the suite of simulants is adequate, and WSLAT alone will not suffice.

SCIENTIFIC AND ENGINEERING CONCERNS ON THE PROPOSED WSLAT

The scientific and engineering concerns arise from the apparent lack of engineering design and analysis, even at a late stage of the schedule. The schedule itself includes many elements of high risk, and is already behind. The study has insufficient engineering detail, so the comments are more general in nature than they would have been had detailed design information been available for the committee. If a detailed engineering design for WSLAT is prepared, then the independent advisory panel recommended in this report could serve as a review body for the design.

IMPACT OF POLICIES AND REGULATIONS

In its evaluation of the regulatory aspects of WSLAT, the committee has no comment regarding DOD regulations and procedures, nor on DPG's obligations to the State of Utah. The committee's comments address only implications of compliance with the National Environmental Policy Act requirements.

The committee concludes that to be complete, the WSLAT feasibility study should address environmental requirements, and the potential for controversy and subsequent delay, in a rigorous fashion.

SUMMARY COMMENTS AND RECOMMENDATION

The committee is concerned that the WSLAT feasibility study is not mature and that the methodology, technology, time, and management attention available are not adequate to provide a reasonable expectation of success. The study writers themselves refer frequently to data that are not available (but essential), and the need to develop additional technology and methodology. The project is very high risk, and is already behind schedule. The difficulty the committee has experienced in obtaining information and data critical to its own study, cited in the Interim Report, is not a good indicator that the funds, attention, and management needed to execute this program will be made available.

Recommendation 3-1: An analysis of alternative WSLAT design approaches that would include comparisons of the various key performance requirements in each design approach should be undertaken. The analysis should include a thorough discussion and analysis of risk, mitigation methods, and development of a sound risk mitigation plan.

The committee considers these to be essential features before realistic funding and schedule decisions can be made.

4

Recommended Evaluation Strategies

As discussed in Chapter 2, the committee's two primary concerns are that the current simulant suite does not adequately mimic live agents and that whole system testing in pristine environments will not predict successful operation in the field. Many chemical and biological detectors that perform well under laboratory conditions fail when faced with real-world challenges. Thus, in addition to the evaluation of the current WSLAT proposal as described in Chapter 3, the committee has suggested alternative and complementary protocols for the evaluation of biopoint detection systems.

The Statement of Work for the committee directs it to assess the WSLAT proposal's ability "to support test and evaluation of current and near-future biological point detection systems."[1] The question that derives from this directive is this: "*How does the WSLAT fit into the Government's total evaluation plan?*" The committee's alternative protocols were built around this central question.

DEVELOPMENT OF OVERALL SYSTEM ASSESSMENT

It is imperative then that the top-level question of how the WSLAT contributes to the overall system assessment be addressed by DOD before committing to an expensive and high-risk course of action. It is the committee's belief that the OTA's concerns of evaluating the JBPDS in an operationally relevant environment, and with threat representative challenges, can best be met in the near future

[1]Statement of Work from the JPEO-CBD to the National Research Council. (see Appendix A).

with a WSKRAT-type facility, a facility that will also serve as a mechanism for developing critical enabling methodologies for a WSLAT-like test facility.

The committee notes that killed[2] related agents, such as those used for vaccines, are readily available and provide agent simulants (or surrogates) that most nearly mimic agent activity without pathogenecity. In addition, these killed related agents can be evaluated in BSL-2 facilities where the detectors can be challenged with representative environments.

Assuming that simulants can be selected that mimic the agents of concern, the predictive capability of the system performance testing protocol increases in the following order (this list is in order of increasing confidence levels):

Level 1	Component testing with simulants
Level 2	Component testing with live agents
Level 3	Whole system testing with simulants alone
Level 4	Whole system testing with live agents (WSLAT)
Level 5	Whole system testing with simulants under environmentally representative conditions
Level 6	Whole system testing with live agents under environmentally representative conditions

Recommendation 4-1: Assuming good simulants exist, the committee recommends testing to proceed from component testing with simulants (Level 1), and live agents (Level 2), through whole system testing with simulants (Level 3), and then to whole system testing with simulants under environmentally representative conditions (Level 5).

With an ideal simulant, there will be sufficient confidence in the simulant's ability to faithfully mimic the agent(s) of concern that whole system live agent testing will not be required. If such simulants do not exist then Level 4 (WSLAT) is necessary before going to Level 5. The committee's position is that testing at Level 5 should be the minimum for system certification. The committee assumes that Level 6 will not be permitted as part of any evaluation procedure.

A WSLAT-like capability likely *may* be necessary in the future because adequate simulants for all agents of concern will not be found, or because adequate simulants may prove to be too time consuming or costly to develop for all agents of concern, or because new agents of concern may arise for which a WSLAT capability is needed to support the warfighter in the field on short notice, or because the modeling may not develop dependable relationships in all cases.

[2]"Killed" agents are microbes or toxins that are rendered biologically inactive, and hence non-pathogenic and non-reproducing. Typical "kill" or inactivation protocols use gamma-irradiation or chemical inactivation. The goal is to render the biologics inactive and non-pathogenic/non-toxic, but still identifiable by the sensor under test. "Live" agents retain their pathogenic/toxic and/or reproductive capacities.

An additional advantage of the correlated simulant approach is that transportable test facilities (referred to as "mobile WSKRAT" in this report) can be taken to a specific location and the detector tested directly under real environmental conditions.

The committee understands that creating a representative environment inside a BSL-3 or higher facility will be expensive and difficult to accomplish. Nevertheless, a WSLAT facility in which detectors can be challenged with representative environments would provide a high level of confidence in the detector. The committee repeats its earlier recommendation, however, that a WSLAT-like facility should be considered an end stage in a progressively more realistic test facility development and should not be viewed as sufficient in and of itself to provide detection system evaluation.

SUMMARY: THE ARGUMENT FOR AN IMMEDIATE START ON WSKRAT

The committee's proposed strategy offers significant advantages and risk reductions over a strategy based on the WSLAT program alone. The committee stresses that with its strategy constructive work could begin immediately on the development of an evaluation plan, challenge definition and methodology development. At a future time work on testing could take place to support the full rate production decision.

Appendix

Statement of Work

The DOD Joint Program Executive Office for Chemical and Biological Defense (JPEO-CBD) commissions the National Research Council to evaluate the requirement for whole-system live agent testing (WSLAT) for biological point-detection systems in general, and the efficacy of its WSLAT proposal. Resolution of this issue impacts fielding to support battlefield missions as well as homeland defense missions. The National Research Council will:

- Review current biological point-detection system testing protocols and integrated evaluation methodologies, as well as potential near-future development and operational testing protocols, and assess the feasibility and benefits of the DOT&E requirement for a system-level, active agent testing capability. In particular, the review will consider use of (1) active agent testing up through the system component level, (2) inactivated agents (including nonpathogenic and gamma-irradiated), and (3) simulants. In evaluating the use of simulants, it will consider the nature and robustness of agent-simulant correlations (including correlation between various challenge scenarios, including both ambient breeze tunnel and field trials.) The review should consider the additional knowledge and confidence gained at each level of testing. The study will also identify risks associated with executing whole-system live agent testing for biological point detectors.

- Review the WSLAT proposal provided under a study from the West Desert Test Center, Dugway Proving Ground. This review will consider the scientific, technological, and regulatory aspects of the WSLAT pro-

posal to support test and evaluation of current and near-future biological point-detection systems. In particular, the review shall independently ascertain the ability of the WSLAT proposal to support an evaluation of biological point-detection system technical and operational requirements, including risk associated with design, development, and verification/validation of the WSLAT proposal. This review will include consideration of the procedures needed to operate a large bio-level 3 facility, and the disposal and decontamination of tested items.