

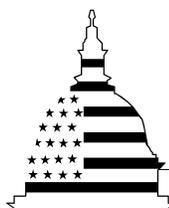
GAO

Report to the Subcommittee on Strategic
Forces, Committee on Armed Services,
House of Representatives

February 2006

NUCLEAR WEAPONS

NNSA Needs to Refine and More Effectively Manage Its New Approach for Assessing and Certifying Nuclear Weapons



G A O

Accountability * Integrity * Reliability



Highlights of [GAO-06-261](#), a report to the Subcommittee on Strategic Forces, Committee on Armed Services, House of Representatives

Why GAO Did This Study

In 1992, the United States began a unilateral moratorium on the testing of nuclear weapons. To compensate for the lack of testing, the Department of Energy's National Nuclear Security Administration (NNSA) developed the Stockpile Stewardship Program to assess and certify the safety and reliability of the nation's nuclear stockpile without nuclear testing. In 2001, NNSA's weapons laboratories began developing what is intended to be a common framework for a new methodology for assessing and certifying the safety and reliability of the nuclear stockpile without nuclear testing. GAO was asked to evaluate (1) the new methodology NNSA is developing and (2) NNSA's management of the implementation of this new methodology.

What GAO Recommends

GAO is making five recommendations to the Administrator of NNSA to (1) ensure that the three laboratories have an agreed-upon technical approach for implementing QMU and (2) improve NNSA's management of the development and implementation of QMU.

While NNSA raised concerns with some of GAO's recommendations, it agreed that it needed to better manage QMU's development and implementation. NNSA also said that GAO had not given it credit for its success in implementing QMU. GAO clarified its report to address NNSA's concerns.

www.gao.gov/cgi-bin/getrpt?GAO-06-261.

To view the full product, including the scope and methodology, click on the link above. For more information, contact Gene Aloise at (202) 512-3841 or aloisee@gao.gov.

NUCLEAR WEAPONS

NNSA Needs to Refine and More Effectively Manage Its New Approach for Assessing and Certifying Nuclear Weapons

What GAO Found

NNSA has endorsed the use of the "quantification of margins and uncertainties" (QMU) methodology as its principal method for assessing and certifying the safety and reliability of the nuclear stockpile. Starting in 2001, Los Alamos National Laboratory (LANL) and Lawrence Livermore National Laboratory (LLNL) officials began developing QMU, which focuses on creating a common "watch list" of factors that are the most critical to the operation and performance of a nuclear weapon. QMU seeks to quantify (1) how close each critical factor is to the point at which it would fail to perform as designed (i.e., the margin to failure) and (2) the uncertainty that exists in calculating the margin, in order to ensure that the margin is sufficiently larger than the uncertainty. According to NNSA and laboratory officials, they intend to use their calculations of margins and uncertainties to more effectively target their resources, as well as to certify any redesigned weapons envisioned by the Reliable Replacement Warhead program.

According to NNSA and weapons laboratory officials, they have made progress in applying the principles of QMU to the assessment and certification of nuclear warheads in the stockpile. NNSA has commissioned two technical reviews of the implementation of QMU. While strongly supporting QMU, the reviews found that the development and implementation of QMU was still in its early stages and recommended that NNSA further define the technical details supporting the implementation of QMU and integrate the activities of the three weapons laboratories in implementing QMU. GAO also found important differences in the understanding and application of QMU among the weapons laboratories. For example, while LLNL and LANL both agree on the fundamental tenets of QMU at a high level, they are pursuing different approaches to calculating and combining uncertainties.

NNSA uses a planning structure that it calls "campaigns" to organize and fund its scientific research. According to NNSA policies, campaign managers at NNSA headquarters are responsible for developing plans and high-level milestones, overseeing the execution of these plans, and providing input to the evaluation of the performance of the weapons laboratories. However, NNSA's management of these processes is deficient in four key areas. First, NNSA's existing plans do not adequately integrate the scientific research currently conducted across the weapon complex to support the development and implementation of QMU. Second, NNSA has not developed a clear, consistent set of milestones to guide the development and implementation of QMU. Third, NNSA has not established formal requirements for conducting annual, technical reviews of the implementation of QMU at the three laboratories or for certifying the completion of QMU-related milestones. Finally, NNSA has not established adequate performance measures to determine the progress of the three laboratories in developing and implementing QMU.

Contents

Letter		1
	Results in Brief	4
	Background	7
	The QMU Methodology Is Highly Promising but Still in the Early Stages of Development	12
	NNSA's Management of the Development and Implementation of QMU Is Deficient in Four Key Areas	22
	Conclusions	33
	Recommendations for Executive Action	34
	Agency Comments and Our Evaluation	35

Appendixes

Appendix I: Comments from the National Nuclear Security Administration	39
Appendix II: GAO Contact and Staff Acknowledgments	51

Tables

Table 1: Nuclear Weapons in the Enduring Stockpile	8
Table 2: NNSA Funding for the Scientific Campaigns, Fiscal Years 2001–2005	11
Table 3: NNSA Funding Requests and Projections for the Scientific Campaigns, Fiscal Years 2006–2010	12
Table 4: NNSA Level-1 Milestones Related to the Development and Implementation of QMU	25
Table 5: Primary Campaign Level-2 Milestones Related to the Development and Implementation of QMU	26

Abbreviations

ASC	Advanced Simulation and Computing
DOE	Department of Energy
ICF	Inertial Confinement Fusion
LANL	Los Alamos National Laboratory
LLNL	Lawrence Livermore National Laboratory
NNSA	National Nuclear Security Administration
Primary	Primary Assessment Technologies
QMU	quantification of margins and uncertainties
RRW	Reliable Replacement Warhead
Science Council	NNSA's Office of Defense Programs Science Council
Secondary	Secondary Assessment Technologies
SNL	Sandia National Laboratory

This is a work of the U.S. government and is not subject to copyright protection in the United States. It may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.



United States Government Accountability Office
Washington, D.C. 20548

February 3, 2006

The Honorable Terry Everett
Chairman
The Honorable Silvestre Reyes
Ranking Minority Member
Subcommittee on Strategic Forces
Committee on Armed Services
House of Representatives

In 1992, the United States began a unilateral moratorium on the testing of nuclear weapons. Prior to the moratorium, underground nuclear testing was a critical component of the evaluation and certification of the performance of a nuclear weapon. Confidence in the continued performance of stockpiled weapons relied heavily on the expert judgment of weapon designers who had significant experience with successful nuclear tests. In addition, the training of new weapon designers depended on continued nuclear testing. In 1993, the Department of Energy (DOE), at the direction of the President and the Congress, established the Stockpile Stewardship Program to ensure the preservation of the United States' core intellectual and technical competencies in nuclear weapons without testing.¹ The National Nuclear Security Administration (NNSA), a separately organized agency within DOE, is now responsible for carrying out the Stockpile Stewardship Program, which includes activities associated with the research, design, development, simulation, modeling, and nonnuclear testing of nuclear weapons. The three nuclear weapons design laboratories—Lawrence Livermore National Laboratory (LLNL) in California, Los Alamos National Laboratory (LANL) in New Mexico, and Sandia National Laboratories (SNL) in California and New Mexico—use the results of these activities to annually assess the safety and reliability of the nation's nuclear weapons stockpile and to certify to the President that the resumption of underground nuclear weapons testing is not needed.

When the moratorium began in 1992, DOE (and subsequently NNSA) faced several challenges in fulfilling its new mission of stockpile stewardship. For example, since both expected and unexpected changes occur as the nuclear stockpile ages, NNSA has become more concerned with gaining a detailed understanding of how such changes might affect the safety and

¹The National Defense Authorization Act for Fiscal Year 1994, Pub. L. No. 103-160, § 3135 (1993), directed DOE to establish the Stockpile Stewardship Program.

reliability of stockpiled weapons. However, unlike the rest of a nuclear weapon, the nuclear explosive package—which contains the primary and the secondary²—cannot be tested simply by evaluating individual components. Specifically, because the operation of the nuclear explosive package is highly integrated, nonlinear, occurs during a very short period of time, and reaches extreme temperatures and pressures, there are portions of the nuclear explosive package that cannot be tested outside of a nuclear explosion. In addition, although the United States conducted about 1,000 nuclear weapons tests prior to the moratorium, only a few tests were designed to collect data on uncertainties associated with a particular part of the nuclear explosive package. As a result, much of the scientific basis for the examination of an exploding nuclear weapon must be extrapolated from other phenomena. Finally, since nuclear testing is no longer available to train new weapons designers, NNSA and the weapons laboratories are faced with the need to develop a rigorous, transparent, and explainable approach to all aspects of the weapon design process, including the assessment and certification of the performance of nuclear weapons.

To address these challenges, in 1999, DOE established 18 programs—which it referred to as “campaigns”—six of which were intended to develop the scientific knowledge, tools, and methods required to provide confidence in the assessment and certification of the safety and reliability of the nuclear stockpile in the absence of nuclear testing. These scientific campaigns include the (1) Primary Assessment Technologies (Primary), (2) Secondary Assessment Technologies (Secondary), (3) Advanced Simulation and Computing (ASC), (4) Advanced Radiography, (5) Dynamic Materials Properties, and (6) Inertial Confinement Fusion and High Yield (ICF) campaigns. In particular, the Primary and Secondary campaigns are designed to analyze and understand the different scientific phenomena that occur in the primary and secondary stages of a nuclear weapon during detonation. As such, the Primary and Secondary campaigns are intended to set the requirements for the computer models and experimental data provided by the other campaigns that are needed to assess and certify the safety and reliability of nuclear weapons.

While the campaign structure brought increased organization to the scientific research conducted across the weapons complex, NNSA still lacked a coherent strategy for relating the scientific research conducted by

²Modern nuclear weapons have two stages: the primary, which is the initial source of energy, and the secondary, which is driven by the primary and provides additional explosive energy.

the weapons laboratories to the needs of the nuclear stockpile and the Stockpile Stewardship Program. Consequently, in 2001, LLNL and LANL began developing what is intended to be a common framework for a new methodology for assessing and certifying the safety and reliability of warheads in the nuclear stockpile in the absence of nuclear testing.

The Stockpile Stewardship Program is now over 10 years old, NNSA's campaign structure is in its sixth year, and 4 years have passed since LLNL and LANL began their effort to develop a new assessment and certification methodology. As the weapons in the nuclear stockpile continue to age, and as more experienced weapon designers and other scientists and technicians retire, NNSA is faced with increased urgency in meeting the goals of the Stockpile Stewardship Program. Furthermore, NNSA has recently created an effort, known as the Reliable Replacement Warhead (RRW) program, to study a new approach to maintaining nuclear warheads over the long term. The RRW program would redesign weapon components to be easier to manufacture, maintain, dismantle, and certify without nuclear testing, potentially allowing NNSA to transition to a smaller, more efficient weapons complex. NNSA's ability to successfully manage these efforts will have a dramatic impact on the future of the U.S. nuclear stockpile and, ultimately, will affect the President's decision of whether a return to nuclear testing is required to maintain confidence in the safety and reliability of the stockpile.

In this context, you asked us to evaluate (1) the new methodology NNSA is developing for assessing and certifying the safety and reliability of the nuclear stockpile in the absence of nuclear testing and (2) NNSA's management of the implementation of this methodology.

To evaluate the new methodology NNSA is developing for assessing and certifying the safety and reliability of the nuclear stockpile in the absence of nuclear testing, we reviewed relevant policy and planning documents from NNSA and the three weapons laboratories, including implementation plans and program plans for the six scientific campaigns. We focused our work principally on the Primary and Secondary campaigns because the primary and secondary are the key components of the nuclear explosive package and because the Primary and Secondary campaigns are intended to set the requirements for the experimental data and computer models needed to assess and certify the performance of nuclear weapons. We also reviewed relevant reports, including those from NNSA's Office of Defense

Programs Science Council, the MITRE Corporation's JASON panel,³ University of California review committees for LANL and LLNL, and the Strategic Advisory Group Stockpile Assessment Team for U.S. Strategic Command. In addition, we interviewed officials from NNSA headquarters and site offices, as well as contractors who operate NNSA sites. Our primary source of information was NNSA's Office of Defense Programs. We also met with officials at LANL, LLNL, and SNL. Finally, we interviewed nuclear weapons experts, senior scientists, and other relevant officials outside of NNSA and the laboratories, including members of NNSA's Office of Defense Programs Science Council, the JASON panel, University of California review committees for LANL and LLNL, the Strategic Advisory Group Stockpile Assessment Team for U.S. Strategic Command, and the Deputy Assistant to the Secretary of Defense (Nuclear Matters) for the Department of Defense.

To evaluate NNSA's management of the implementation of its new methodology to assess and certify the safety and reliability of nuclear weapons in the absence of nuclear testing, we reviewed relevant NNSA policy, planning, and evaluation documents, including the Office of Defense Program's *Program Management Manual*, campaign program and implementation plans, contractor performance evaluation plans and reports, and internal reviews of NNSA management. We also reviewed contractor planning and evaluation documents, including LANL, LLNL, and SNL performance evaluation plans and reports. Finally, we met with campaign managers and other officials at NNSA headquarters and site offices, LANL, LLNL, and SNL. We performed our work between August 2004 and December 2005 in accordance with generally accepted government auditing standards.

Results in Brief

NNSA has endorsed the use of the "quantification of margins and uncertainties" (QMU) methodology as its principal method for assessing and certifying the safety and reliability of the existing nuclear stockpile in the absence of nuclear testing. The QMU methodology focuses on creating a "watch list" of factors that, in the judgment of nuclear weapon experts, are the most critical to the operation and performance of a nuclear weapon. Starting in 2001, LANL and LLNL officials began developing QMU, which they described as a common methodology for quantifying how close

³JASON is a group of nationally known scientists who advise government agencies on defense, energy, and other technical issues.

each critical factor is to the point at which it would fail to perform as designed (i.e., the margin to failure), as well as quantifying the uncertainty that exists in calculating the margin, in order to ensure that the margin is sufficiently greater than the uncertainty. According to NNSA and laboratory officials, the weapons laboratories intend to use their calculations of margins and uncertainties to more effectively target their resources to either increasing the margin in a nuclear weapon or reducing the uncertainties associated with calculating the margin. In addition, they said that QMU will be vital to certifying any redesigned weapons, such as those envisioned by the RRW program.

NNSA and laboratory officials told us that they have made progress in applying the principles of QMU to the certification and assessment of nuclear warheads in the stockpile. However, QMU is still in its early stages of development, and important differences exist among the three laboratories in their application of QMU. To date, NNSA has commissioned two technical reviews of the implementation of QMU at the weapons laboratories. While strongly supporting QMU, the reviews found that the development and implementation of QMU was still in its early stages. For example, one review stated that, in the course of its work, it became evident that there were a variety of differing and sometimes diverging views of what QMU really was and how it was working in practice. The reviews recommended that NNSA take steps to further define the technical details supporting the implementation of QMU and integrate the activities of the three weapons laboratories in implementing QMU. However, NNSA and the weapons laboratories have not fully implemented these recommendations. Beyond the issues raised in the two reports, we also found differences in the understanding and application of QMU among the three laboratories. For example, LLNL and LANL officials told us that the QMU methodology only applies to the nuclear explosive package and not to the nonnuclear components that control the use, arming, and firing of the nuclear warhead. However, SNL officials told us that they have been applying their own version of QMU to nonnuclear components for a long time. In addition, we found that while LLNL and LANL both agree on the fundamental tenets of QMU at a high level, their application of the QMU methodology differs in some important respects. Specifically, LLNL and LANL are pursuing different approaches to calculating and combining uncertainties. While there will be methodological differences among the laboratories in the detailed application of QMU to specific weapon systems, it is fundamentally important that these differences be understood and, if need be, reconciled, to ensure that QMU achieves the goal of the common methodology NNSA has stated it needs to support the continued

assessment of the existing stockpile or the certification of redesigned nuclear components under the RRW program.

NNSA relies on its Primary and Secondary campaigns to manage the development and implementation of QMU. According to NNSA policies, campaign managers at NNSA headquarters are responsible for developing campaign plans and high-level milestones, overseeing the execution of these plans, and providing input to the evaluation of the performance of the weapons laboratories. However, NNSA's management of these processes is deficient in four key areas. First, the planning documents that NNSA has established for the Primary and Secondary campaigns do not adequately integrate the scientific research currently conducted that supports the development and implementation of QMU. Specifically, a significant portion of the scientific research that is relevant to the Primary and Secondary campaigns, and the implementation of QMU, is funded and carried out by a variety of campaigns and other programs within the Stockpile Stewardship Program. Second, NNSA has not developed a clear, consistent set of milestones to guide the development and implementation of QMU. For example, while one key campaign plan envisions a two-stage path to identify and reduce key uncertainties in nuclear weapon performance using QMU by 2014, the performance measures in NNSA's fiscal year 2006 budget request call for the completion of QMU by 2010. Third, NNSA has not established formal requirements for conducting annual, technical reviews of the implementation of QMU at the three weapons laboratories or for certifying the completion of QMU-related milestones. Finally, NNSA has not established adequate performance measures to determine the progress of the laboratories in developing and implementing QMU. Specifically, NNSA officials were not able to show how they are able to measure progress toward current performance targets related to the development and implementation of QMU (e.g., NNSA's statement that the development and implementation of QMU was 10 percent complete at the end of fiscal year 2004). As a result of these deficiencies, NNSA cannot fully ensure that it will be able to meet key deadlines for implementing QMU.

GAO is making five recommendations to the Administrator of NNSA to (1) ensure that the three weapons laboratories have an agreed upon technical approach for implementing QMU and (2) improve NNSA's management of the development and implementation of QMU.

We provided NNSA with a draft of this report for their review and comment. Overall, NNSA generally agreed that there was a need for an

agreed-upon technical approach for implementing QMU and that NNSA needed to improve the management of QMU through clearer, long-term milestones and better integration across the program. However, NNSA stated that QMU had already been effectively implemented and that we had not given NNSA sufficient credit for its success. In addition, NNSA raised several issues about our conclusions and recommendations regarding their management of the QMU effort. We have modified our report to more fully recognize that QMU is being used by the laboratories to address stockpile issues and to more completely characterize its current state of development. NNSA also made technical clarifications, which we incorporated in this report as appropriate.

Background

Most modern nuclear warheads contain a nuclear explosive package, which contains the primary and the secondary, and a set of nonnuclear components.⁴ The nuclear detonation of the primary produces energy that drives the secondary, which produces further nuclear energy of a militarily significant yield. The nonnuclear components control the use, arming, and firing of the warhead. All nuclear weapons developed to date rely on nuclear fission to initiate their explosive release of energy. Most also rely on nuclear fusion to increase their total energy yield. Nuclear *fission* occurs when the nucleus of a heavy, unstable atom (such as uranium-235) is split into two lighter parts, which releases neutrons and produces large amounts of energy. Nuclear *fusion* occurs when the nuclei of two light atoms (such as deuterium and tritium) are joined, or fused, to form a heavier atom, with an accompanying release of neutrons and larger amounts of energy.

The U.S. nuclear stockpile consists of nine weapon types. (See table 1.) The lifetimes of the weapons currently in the stockpile have been extended well beyond the minimum life for which they were originally designed—generally about 20 years—increasing the average age of the stockpile and, for the first time, leaving NNSA with large numbers of weapons that are close to 30 years old.

⁴The terms “nuclear warhead” and “nuclear weapon” have different technical meanings. For example, a nuclear weapon, in the case of a reentry vehicle, includes the warhead and certain Department of Defense components, such as fuses and batteries. However, for purposes of this report, we often use the terms “warhead” and “weapon” interchangeably.

Table 1: Nuclear Weapons in the Enduring Stockpile

Warhead or bomb mark	Description	Date of entry into stockpile	Laboratory	Military service
B61 3/4/10	Tactical bomb	1979/1979/1990	LANL, SNL	Air Force
B61 7/11	Strategic bomb	1985/1996	LANL, SNL	Air Force
W62	ICBM warhead ^a	1970	LLNL, SNL	Air Force
W76	SLBM warhead ^b	1978	LANL, SNL	Navy
W78	ICBM warhead ^a	1979	LANL, SNL	Air Force
W80 0/1	Cruise missile warhead	1984/1982	LLNL, SNL	Air Force / Navy
B83 0/1	Strategic bomb	1983/1993	LLNL, SNL	Air Force
W87	ICBM warhead ^a	1986	LLNL, SNL	Air Force
W88	SLBM warhead ^b	1989	LANL, SNL	Navy

Source: NNSA.

Note: The dates of entry into the enduring nuclear stockpile are based on when the weapon reached phase 6 of the weapons development and production cycle. As of 2005, responsibility for the W80 0/1 was transferred from LANL to LLNL.

^aICBM = intercontinental ballistic missile.

^bSLBM = submarine launched ballistic missile.

Established in 1993, the Stockpile Stewardship Program faces two main technical challenges: provide (1) a better scientific understanding of the basic phenomena associated with nuclear weapons and (2) an improved capability to predict the impact of aging and remanufactured components on the safety and reliability of nuclear weapons. Specifically,

- An exploding nuclear weapon creates the highest pressures, greatest temperatures, and most extreme densities ever made by man on earth, within some of the shortest times ever measured. When combined, these variables exist nowhere else in nature. While the United States conducted about 1,000 nuclear weapons tests prior to the moratorium, these tests were conducted mainly to look at broad indicators of weapon performance (such as the yield of a weapon) and were often not designed to collect data on specific properties of nuclear weapons physics. After more than 60 years of developing nuclear weapons, while many of the physical processes are well understood and accurately modeled, the United States still does not possess a set of completely known and expressed laws and equations of nuclear weapons physics that link the physical event to first principles.

-
- As nuclear weapons age, a number of physical changes can take place. The effects of aging are not always gradual, and the potential for unexpected changes in materials causes significant concerns as to whether weapons will continue to function properly. Replacing aging components is, therefore, essential to ensure that the weapon will function as designed. However, it may be difficult or impossible to ensure that all specifications for the manufacturing of new components are precisely met, especially since each weapon was essentially handmade. In addition, some of the manufacturing process lines used for the original production have been disassembled.

In 1995, the President established an annual assessment and reporting requirement designed to help ensure that nuclear weapons remain safe and reliable without underground testing.⁵ As part of this requirement, the three weapons laboratories are required to issue a series of reports and letters that address the safety, reliability, performance, and military effectiveness of each weapon type in the stockpile. The letters, submitted to the Secretary of Energy individually by the laboratory directors, summarize the results of the assessment reports and, among other things, express the directors' conclusions regarding whether an underground nuclear test is needed and the adequacy of various tools and methods currently in use to evaluate the stockpile.

To address these challenges, in 1999 DOE developed a new three-part program structure for the Stockpile Stewardship Program that included a series of campaigns, which DOE defined as technically challenging, multiyear, multifunctional efforts to develop and maintain the critical capabilities needed to continue assessing the safety and reliability of the nuclear stockpile into the foreseeable future without underground testing. DOE originally created 18 campaigns that were designed to focus its efforts in science and computing, applied science and engineering, and production readiness. Six of these campaigns currently focus on the development and improvement of the scientific knowledge, tools, and methods required to provide confidence in the assessment and certification of the safety and reliability of the nuclear stockpile in the absence of nuclear testing. These six campaigns are as follows:

⁵The Defense Authorization Act for Fiscal Year 2003, Pub. L. No. 107-314, § 3141 (2002), established a statutory requirement for annual stockpile assessments.

-
- The Primary and Secondary campaigns were established to analyze and understand the different scientific phenomena that occur in the primary and secondary stages of a nuclear weapon during detonation. As such, the Primary and Secondary campaigns are intended to support the development and implementation of the QMU methodology and to set the requirements for the computers, computer models, and experimental data needed to assess and certify the performance of nuclear weapons.
 - The ASC campaign provides the leading-edge supercomputers and models that are used to simulate the detonation and performance of nuclear weapons.
 - Two campaigns—Advanced Radiography and Dynamic Materials Properties—provide data from laboratory experiments to support nuclear weapons theory and computational modeling. For example, the Advanced Radiography campaign conducts experiments that measure how stockpile materials behave when exposed to explosively driven shocks. One of the major facilities being built to support this campaign is the Dual Axis Radiographic Hydrodynamic Test Facility at LANL.
 - The ICF campaign develops experimental capabilities and conducts experiments to examine phenomena at high temperature and pressure regimes that approach but do not equal those occurring in a nuclear weapon. As a result, scientists currently have to extrapolate from the results of these experiments to understand similar phenomena in a nuclear weapon. One of the major facilities being built as part of this campaign is the National Ignition Facility at LLNL.

The other two program activities associated with the Stockpile Stewardship Program are “Directed Stockpile Work” and “Readiness in Technical Base and Facilities.” Directed Stockpile Work includes the activities that directly support specific weapons in the stockpile, such as the Stockpile Life Extension Program, which employs a standardized approach for planning and carrying out nuclear weapons refurbishment activities to extend the operational lives of the weapons in the stockpile well beyond their original design lives. The life extension for the W87 was completed in 2004, and three other weapon systems—the B61, W76, and W80—are currently undergoing life extensions. Each life extension program is specific to that weapon type, with different parts being replaced or refurbished for each weapon type. Readiness in Technical Base and Facilities includes the physical infrastructure and operational readiness

required to conduct campaign and Directed Stockpile Work activities across the nuclear weapons complex. The complex includes the three nuclear weapons design laboratories (LANL, LLNL, and SNL), the Nevada Test Site, and four production plants—the Pantex Plant in Texas, the Y-12 Plant in Tennessee, a portion of the Savannah River Site in South Carolina, and the Kansas City Plant in Missouri.

From fiscal year 2001 through fiscal year 2005, NNSA spent over \$7 billion on the six scientific campaigns (in inflation-adjusted dollars). (See table 2.) NNSA has requested almost \$7 billion in funding for these campaigns over the next 5 years. (See table 3.)

Table 2: NNSA Funding for the Scientific Campaigns, Fiscal Years 2001–2005

Dollars in millions						
	FY 2001	FY 2002	FY 2003	FY 2004	FY 2005	Total
Primary	\$49.8	\$52.4	\$48.7	\$41.2	\$73.4	\$265.5
Secondary	43.7	42.1	49.2	54.6	57.2	246.8
ASC	770.9	692.2	799.3	738.9	685.9	3,687.2
Advanced Radiography	85.7	100.3	74.2	53.5	52.7	366.4
Dynamic Materials Properties	79.4	80.7	85.2	87.8	74.2	407.3
ICF	515.7	593.3	518.9	480.1	492.1	2,600.1
Total	\$1,545.2	\$1,561.0	\$1,575.5	\$1,456.1	\$1,435.5	\$7,573.3

Source: NNSA.

Note: In constant dollars, base year 2005.

Table 3: NNSA Funding Requests and Projections for the Scientific Campaigns, Fiscal Years 2006–2010

Dollars in millions

	FY 2006	FY 2007	FY 2008	FY 2009	FY 2010	Total
Primary	\$45.2	\$47.5	\$48.9	\$48.7	\$45.6	\$235.9
Secondary	61.3	63.9	65.0	65.0	65.0	320.2
ASC	660.8	666.0	666.0	666.0	666.0	3,324.8
Advanced Radiography	49.5	42.7	39.5	38.7	41.9	212.3
Dynamic Materials Properties	80.9	85.1	86.5	87.4	87.4	427.3
ICF	460.4	461.6	461.6	461.6	461.6	2,306.8
Total	\$1,358.1	\$1,366.8	\$1,367.5	\$1,367.4	\$1,367.5	\$6,827.3

Source: DOE, FY 2006 Congressional Budget Request, February 2005.

Within NNSA, the Office of Defense Programs is responsible for managing the campaigns and the Stockpile Stewardship Program in general. Within this office, two organizations share responsibility for overall management of the scientific campaigns: the Office of the Assistant Deputy Administrator for Research, Development, and Simulation and the Office of the Assistant Deputy Administrator for Inertial Confinement Fusion and the National Ignition Facility Project. The first office oversees campaign activities associated with the Primary and Secondary campaigns—as well as the ASC, Advanced Radiography, and Dynamic Materials Properties campaigns—with a staff of about 13 people. The second office oversees activities associated with the ICF campaign with a single staff person. Actual campaign activities are conducted by scientists and other staff at the three weapons laboratories. LANL and LLNL conduct activities associated with the nuclear explosive package, while SNL performs activities associated with the nonnuclear components that control the use, arming, and firing of the nuclear warhead.

The QMU Methodology Is Highly Promising but Still in the Early Stages of Development

NNSA has endorsed the use of a new common methodology, known as the quantification of margins and uncertainties, or QMU, for assessing and certifying the safety and reliability of the nuclear stockpile. NNSA and laboratory officials told us that they have made progress in applying the principles of QMU to the certification and assessment of nuclear warheads in the stockpile. However, QMU is still in its early stages of development, and important differences exist among the three laboratories in their application of QMU. To date, NNSA has commissioned two technical reviews of the implementation of QMU at the weapons laboratories. While

strongly supporting QMU, the reviews found that the development and implementation of QMU was still in its early stages. The reviews recommended that NNSA take steps to further define the technical details supporting the implementation of QMU and integrate the activities of the three weapons laboratories in implementing QMU. However, NNSA and the weapons laboratories have not fully implemented these recommendations. Beyond the issues raised in the two reports, we also found differences in the understanding and application of QMU among the three laboratories.

NNSA Has Endorsed QMU as a New, Common Methodology for Assessing and Certifying Stockpile Safety and Reliability

When the Primary and Secondary campaigns were established in 1999, they brought some organization and overall goals to the scientific research conducted across the weapons complex. For example, as we noted in April 2005, the Primary campaign set an initial goal in the 2005 to 2010 time frame for certifying the performance of the primary of a nuclear weapon to within a stated yield level.⁶ However, according to senior NNSA officials, NNSA still lacked a coherent strategy for relating the scientific work conducted by the weapons laboratories under the campaigns to the needs of the nuclear stockpile and the overall Stockpile Stewardship Program. This view was echoed by a NNSA advisory committee report, which stated in 2002 that the process used by the weapons laboratories to certify the safety and reliability of nuclear weapons was ill defined and unevenly applied, leading to major delays and inefficiencies in programs.⁷

Starting in 2001, LLNL and LANL began developing what is intended to be a common methodology for assessing and certifying the performance and safety of nuclear weapons in the absence of nuclear testing. In 2003, the associate directors for nuclear weapons at LLNL and LANL published a white paper—entitled “National Certification Methodology for the Nuclear Weapon Stockpile”—that described this new methodology, which they referred to as the quantification of margins and uncertainties or QMU. According to the white paper, QMU is based on an adaptation of standard engineering practices and lends itself to the development of “rigorous, quantitative, and explicit criteria for judging the robustness of weapon system and component performance at a detailed level.” Moreover, the

⁶GAO, *Nuclear Weapons: Preliminary Results of Review of Campaigns to Provide Scientific Support for the Stockpile Stewardship Program*, GAO-05-636R (Washington, D.C.: Apr. 29, 2005).

⁷National Nuclear Security Administration Advisory Committee, “Science and Technology in the Stockpile Stewardship Program,” Mar. 1, 2002.

quantitative results of this process would enable NNSA and the weapons laboratories to set priorities for their activities and thereby make rational decisions about allocating program resources to the nuclear stockpile.

The process envisaged in the white paper focuses on creating a “watch list” of factors that, in the judgment of nuclear weapons experts, are the most critical to the operation and performance of a nuclear weapon. These factors include key operating characteristics and components of the nuclear weapon. For each identified, critical factor leading to a nuclear explosion, nuclear weapons experts would define performance metrics. These performance metrics would represent the experts’ best judgment of what constitutes acceptable behavior—i.e., the range of acceptable values for a critical function to successfully occur or for a critical component to function properly—as well as what constitutes unacceptable behavior or failure. To use an analogy, consider the operation of a gasoline engine. Some of the events critical to the operation of the engine would include the opening and closing of valves, the firing of the spark plugs, and the ignition of the fuel in each cylinder. Relevant performance metrics for the ignition of fuel in a cylinder would include information on the condition of the spark plugs (e.g., whether they are corroded) and the fuel/air mixture in the cylinder.

Once nuclear experts have identified the relevant performance metrics for each critical factor, according to the 2003 white paper, the goal of QMU is to quantify these metrics. Specifically, the QMU methodology seeks to quantify (1) how close each critical factor is to the point at which it would fail to perform as designed (i.e., the performance margin or the margin to failure) and (2) the uncertainty in calculating the margin. According to the white paper, the weapons laboratories would be able to use their calculated values of margins and uncertainties as a way to assess their confidence in the performance of a nuclear weapon. That is, the laboratories would establish a “confidence ratio” for each critical factor—they would divide their calculated value for the margin (“M”) by their calculations of the associated uncertainty (“U”) and arrive at a single number (“M/U”). According to the white paper, the weapons laboratories would only have confidence in the performance of a nuclear weapon if the margin “significantly” exceeds uncertainty for all critical issues. However, the white paper did not define what the term “significantly” meant.

In a broad range of key planning and management documents that have followed the issuance of the white paper, NNSA and the weapons laboratories have endorsed the use of the QMU methodology as the

principal tool for assessing and certifying the safety and reliability of the nuclear stockpile in the absence of nuclear testing. For example, in its fiscal year 2006 implementation plan for the Primary campaign, NNSA stated as a strategic objective that it needs to develop the capabilities and understanding necessary to apply QMU as the assessment and certification methodology for the nuclear explosive package. In addition, in its fiscal year 2006 budget request, NNSA selected its progress toward the development and implementation of QMU as one of its major performance indicators. Finally, in the plans that NNSA uses to evaluate the performance of LANL and LLNL, NNSA has established an overall objective for LANL and LLNL to assess and certify the safety and reliability of nuclear weapons using a common QMU methodology.

Officials at NNSA and the weapons laboratories have also stated that QMU will be vital to certifying any weapon redesigns, such as are envisioned by the RRW program. For example, senior NNSA officials told us that the Stockpile Stewardship Program will not be sustainable if it only involves the continued refurbishment in perpetuity of existing weapons in the current nuclear stockpile. They stated that the accumulation of small changes over the extended lifetime of the current nuclear stockpile will result in increasing levels of uncertainty about its performance. If NNSA moves forward with the RRW program, according to NNSA documents and officials, the future goal of the weapons program will be to use QMU to replace existing stockpile weapons with an RRW whose safety and reliability could be assured with the highest confidence, without nuclear testing, for as long as the United States requires nuclear forces.

The Development and Implementation of QMU Is at an Early Stage and Important Differences Exist Among the Weapons Laboratories in their Application of QMU

According to NNSA and laboratory officials, the weapons laboratories have made progress in applying the principles of QMU to the certification of life extension programs and to the annual stockpile assessment process. For example, LLNL officials told us that they are applying QMU to the assessment of the W80, which is currently undergoing a life extension.⁸ They said that, in applying the QMU methodology, they tend to focus their efforts on identifying credible “failure modes,” which are based on observable problems, such as might be caused by the redesign of components in a nuclear weapon, changes to the manufacturing process for components, or the performance of a nuclear weapon under aged

⁸LLNL first applied QMU in its certification of the life extension of the W87, which was completed in November 2004.

conditions. They said that, for the W80 life extension program, they have developed a list of failure modes and quantified the margins and uncertainties associated with these failure modes. Based on their calculations, they said that they have increased their confidence in the performance of the W80.

Similarly, LANL officials told us that they are applying QMU to the W76, which is also currently undergoing a life extension and is scheduled to finish its first production unit in 2007. They said that, in applying the QMU methodology, they tend to focus their efforts on defining “performance gates,” which are based on a number of critical points during the explosion of a nuclear weapon that separate the nuclear explosion into natural stages of operation. The performance gates identify the characteristics that a nuclear weapon must have at a particular time during its operation to meet its performance requirements (e.g., to reach its expected yield). LANL officials told us that they have developed a list of performance gates for the W76 life extension program and are beginning to quantify the margins and uncertainties associated with these performance gates.

Despite this progress, we found that QMU is still in its early stages of development and that important differences exist among the weapons laboratories in their application of QMU. To date, NNSA has commissioned two technical reviews of the implementation of QMU at the weapons laboratories. The first review was conducted by NNSA’s Office of Defense Programs Science Council (Science Council)—which advises NNSA on scientific matters across a range of activities, including those associated with the scientific campaigns—and resulted in a March 2004 report.⁹ The second review was conducted by the MITRE Corporation’s JASON panel and resulted in a February 2005 report.¹⁰ Both reports endorsed the use of QMU by the weapons laboratories and listed several potential benefits that QMU could bring to the nuclear weapons program. For example, according to the Science Council report, QMU will serve an important role in training the next generation of nuclear weapon designers and will quantify and increase NNSA’s confidence in the assessment and certification of the nuclear stockpile. According to the JASON report, QMU could become a useful management tool for directing investments in a given weapon

⁹NNSA Defense Programs Science Council, “Report on the Friendly Reviews of QMU at the NNSA Laboratories,” March 2004.

¹⁰JASON, The MITRE Corporation, *Quantification of Margins and Uncertainties (QMU)*, JSR-04-330, Feb. 17, 2005.

system where they would be most effective in increasing confidence, as required by the life extension programs. In addition, the JASON report described how LANL and LLNL officials had identified potential failure modes in several weapon systems and calculated the associated margins and uncertainties. The report noted that, for most of these failure modes, the margin for success was large compared with the uncertainty in the performance.

However, according to both the Science Council and the JASON reports, the development and implementation of QMU is still in its early stages. For example, the JASON report described QMU as highly promising but unfinished, incomplete and evolving, and in the early stages of development. Moreover, the chair of the JASON panel on QMU told us in June 2005 that, during the course of his review, members of the JASON panel found that QMU was not mature enough to assess its reliability or usefulness. The reports also stated that the weapons laboratories have not fully developed or agreed upon the technical details supporting the implementation and application of QMU. For example, the JASON report stated that, in the course of its review, it became evident that there were a variety of differing and sometimes diverging reviews of what QMU really was and how it was working in practice. As an example, the report stated that some of the scientists, designers, and engineers at LANL and LLNL saw the role of expert judgment as an integral part of the QMU process, while others did not. In discussions with the weapons laboratories about the two reports, LANL officials told us that they believed that the details of QMU as a formal methodology are still evolving, while LLNL officials stated that QMU was “embryonic” and not fully developed.

While supporting QMU, the two reports noted that the weapons laboratories face challenges in successfully implementing a coherent and credible analytical method based on the QMU methodology. For example, in its 2004 report, the Science Council stated that, in its view, the QMU methodology is based on the following core assumptions:

- Computer simulations can accurately predict the behavior of a complex nuclear explosive system as a function of time.
- It is sufficient for the assessment of the performance of a nuclear weapon to examine the simulation of the time evolution of a nuclear explosive system at a number of discrete time intervals and to determine whether the behavior of the system at each interval is within acceptable bounds.

-
- The laboratories' determinations of acceptable behavior can be made quantitatively—that is, they will make a quantitative estimate of a system's margins and uncertainties.
 - Given these quantitative measures of the margins and uncertainties, it is possible to calculate the probability (or confidence level) that the nuclear explosive system will perform as desired.

However, the Science Council's report noted that extraordinary degrees of complexity are involved in a rational implementation of QMU that are only beginning to be understood. For example, in order for the QMU methodology to have validity, it must sufficiently identify all critical failure modes, critical events, and associated performance metrics. However, as described earlier, the operation of an exploding nuclear weapon is highly integrated and nonlinear, occurs during a very short period of time, and reaches extreme temperatures and pressures. In addition, the United States does not possess a set of completely known and expressed laws and equations of nuclear weapons physics. Given these complexities, it will be difficult to demonstrate the successful implementation of QMU, according to the report. In addition, the Science Council stated that it was not presented with any evidence that there exists a method—even in principle—for calculating an overall probability that a nuclear explosive package will perform as designed from the set of quantitative margins and uncertainties at each time interval.

To address these and other issues, the two reports recommended that NNSA take steps to further define the technical details supporting the implementation of QMU and to integrate the activities of the three weapons laboratories in implementing QMU. For example, the 2004 Science Council report recommended that NNSA direct the associate directors for nuclear weapons at LANL and LLNL to undertake a major effort to define the details of QMU. In particular, the report recommended that a trilaboratory team be charged with defining a common language for QMU and identifying the important performance gates, failure modes, and other criteria in the QMU approach. The report stated that this agreed-upon "reference" set could then be used to support all analyses of stockpile issues. In addition, the report recommended that NNSA consider establishing annual or semiannual workshops for the three weapons laboratories to improve the identification, study, and prioritization of potential failure modes and other factors that are critical to the operation and performance of nuclear weapons.

Similarly, the 2005 JASON panel report noted that the meaning and implications of QMU are currently unclear. To rectify this problem, the report recommended that the associate directors for nuclear weapons at LANL and LLNL write a new, and authoritative, paper defining QMU and submit it to NNSA. Furthermore, the report recommended that the laboratories establish a formal process to (1) identify all failure modes and performance gates associated with QMU, using the same methodology for all weapon systems, and (2) establish better relationships between the concepts of failure modes and performance gates for all weapon systems in the stockpile.

However, NNSA and laboratory officials have not fully implemented these recommendations, particularly the recommendations of the Science Council. For example, while LLNL and LANL officials are drafting a new “white paper” on QMU that attempts to clarify some fundamental tenets of the methodology, officials from SNL are not involved in the drafting of this paper. In addition, NNSA has not required the three weapons laboratories to hold regular meetings or workshops to improve the identification, prioritization, and integration of failure modes, performance gates, and other critical factors.

According to NNSA’s Assistant Deputy Administrator for Research, Development, and Simulation, NNSA has not fully implemented the recommendations of the Science Council’s report partly because the report was intended more to give NNSA a sense of the status of the implementation of QMU than it was to provide recommendations. For example, the 2004 report states that the “friendly review,” as the report is referred to by NNSA, would not have budget implications and that the report’s findings and recommendations would be reported only to the senior management of the weapons laboratories. As a result, the Assistant Deputy Administrator told us that he had referred the recommendations to the directors of the weapons laboratories and told them to implement the recommendations as they saw fit.

Furthermore, LLNL and LANL officials disagreed with some of the statements in the Science Council report and stressed that, in using QMU, they do not attempt to assign an overall probability that the nuclear explosive package will perform as desired. That is, they do not attempt to add up calculations of margins and uncertainties for all the critical factors to arrive at a single estimate of margin and uncertainty, or a single confidence ratio, for the entire nuclear explosive package. Instead, they said that they focus on ensuring that the margin for each identified critical

factor in the explosion of a nuclear weapon is greater than the uncertainty. However, they said that, for a given critical factor, they do combine various calculations of individual uncertainties that contribute to the total amount of uncertainty for that factor.

In addition, in addressing comments in the JASON report, LLNL and LANL officials stressed that QMU has always relied, and will continue to rely heavily, on the judgment of nuclear weapons experts. For example, LLNL officials told us that since there is no single definition of what constitutes a threshold for failure, they use expert judgment to decide what to put on their list of failure modes. They also said that the QMU methodology provides a way to make the entire annual assessment and certification process more transparent to peer review. Similarly, LANL officials said that they use expert judgment extensively in establishing performance metrics and threshold values for their performance gates. They said that expert judgment will always be a part of the scientific process and a part of QMU.

Beyond the issues raised in the two reports, we found that there are differences in the understanding and application of QMU among the three laboratories. For example, the three laboratories do not agree about the application of QMU to areas outside of the nuclear explosive package. Specifically, LLNL officials told us that the QMU methodology, as currently developed, only applies to the nuclear explosive package and not to the nonnuclear components that control the use, arming, and firing of the nuclear warhead. According to LLNL and LANL officials, SNL scientists can run hundreds of experiments to test their components and, therefore, can use normal statistical analysis in certifying the performance of nonnuclear components. As a result, according to LLNL and LANL officials, SNL does not have to cope with real uncertainty and does not “do” QMU. Furthermore, according to LLNL officials, SNL has chosen not to participate in the development of QMU with LLNL and LANL.

However, SNL officials told us that while some of the nonnuclear components are testable to a degree, SNL is as challenged as the other two weapons laboratories in certifying the performance of their systems without actual testing. For example, SNL officials said that they simply do not have enough money to perform enough tests on all of their nonnuclear components to be able to rely completely on statistical analysis to meet their safety performance levels. In addition, SNL scientists are not able to test their components under the conditions of a nuclear explosion but are still required to certify the performance of the components under these

conditions. Thus, SNL officials told us that they had been using their own version of QMU for a long time.

SNL officials told us that they define QMU as a way to make risk-informed decisions about the effect of variabilities and uncertainties on the performance of a nuclear weapon, including the nonnuclear components that control the use, arming, and firing of the nuclear warhead. Moreover, they said that this kind of risk-informed approach is not unique to the nuclear weapons laboratories and is used extensively in areas such as nuclear reactor safety. However, they told us that they have been left out in the development of QMU by the two other weapons laboratories. Specifically, they said that while SNL scientists have worked with other scientists at LANL and LLNL at a “grass roots” level, there has only been limited cooperation and dialogue between upper-level management at the three laboratories concerning the development and implementation of QMU.

In addition, we found that while LLNL and LANL both agree on the fundamental tenets of QMU at a high level, their application of the QMU methodology differs in some important respects. For example, LLNL and LANL officials told us that, at a detailed level, the two laboratories are pursuing different approaches to calculating and combining uncertainties. For the W80 life extension program, LLNL officials showed us how they combined calculations of individual uncertainties that contributed to the total uncertainty for a key failure mode of the primary—the amount of primary yield necessary to drive the secondary. However, they said that the scientific support for their method for combining individual calculations of uncertainty was limited, and they stated that they are pursuing a variety of more sophisticated analyses to improve their current approach.

Moreover, the two laboratories are taking a different approach to generating a confidence ratio for each critical factor, as described in the 2003 white paper on QMU. For example, for the W80 life extension program, LLNL officials showed us how they calculated a single confidence ratio for a key failure mode of the primary, based on their calculations of margin and uncertainty. They said that the weapon systems for which they are responsible have a lot of margin built into them, and they feel comfortable generating this number. In contrast, in discussions with LANL officials about the W76 life extension program, LANL officials told us that they prefer not to calculate a single confidence ratio for a performance gate, partly because they are concerned that their customers (e.g., the

Department of Defense) might think that the QMU methodology is more formal than it is currently.

In commenting on the differences between the two laboratories, NNSA officials stated that the two laboratories are pursuing complementary approaches, and that these differences are part of the rationale for a national policy decision to maintain two nuclear design laboratories. In addition, they stated that the confidence in the correctness of scientific research is improved by achieving the same answer through multiple approaches. LLNL officials also made similar comments, stating that the nation will benefit from some amount of independence between the laboratories to assure that the best methodology for assessing the stockpile in the absence of nuclear testing is achieved.

NNSA's Management of the Development and Implementation of QMU Is Deficient in Four Key Areas

NNSA relies on its Primary and Secondary campaigns to manage the development and implementation of QMU. According to NNSA policies, campaign managers at NNSA headquarters are responsible for developing campaign plans and high-level milestones, overseeing the execution of these plans, and providing input to the evaluation of the performance of the weapons laboratories. However, NNSA's management of these processes is deficient in four key areas. First, the planning documents that NNSA has established for the Primary and Secondary campaigns do not adequately integrate the scientific research currently conducted that supports the development and implementation of QMU. Second, NNSA has not developed a clear, consistent set of milestones to guide the development and implementation of QMU. Third, NNSA has not established formal requirements for conducting annual, technical reviews of the implementation of QMU or for certifying the completion of QMU-related milestones. Finally, NNSA has not established adequate performance measures to determine the progress of the laboratories in developing and implementing QMU.

Campaign Planning Documents Do Not Adequately Integrate the Scientific Activities Supporting QMU

As part of its planning structure, NNSA requires the use of program and implementation plans to set requirements and manage resources for the campaigns and other programs associated with the Stockpile Stewardship Program. Program plans are strategic in nature and identify the long-term goals, high-level milestones, and resources needed to support a particular program over a 7-year period, while implementation plans establish performance expectations for the program and each participating site for

the current year of execution. According to NNSA policies, program and implementation plans should flow from and interact with each other using a set of cascading goals and requirements.

NNSA has established a single program plan, which it calls the “Science campaign program plan,” that encompasses the Primary and the Secondary campaigns, as well as two other campaigns—Advanced Radiography and Dynamic Materials Properties. NNSA has also established separate implementation plans for each of these campaigns, including the Primary and Secondary campaigns. According to NNSA, it relies on these plans—and in particular the plans related to the Primary and Secondary campaigns—to manage the development and implementation of QMU, as well as to determine the requirements for the experimental data and computer modeling needed to analyze and understand the different scientific phenomena that occur in a nuclear weapon during detonation.

However, the current Primary and Secondary campaign plans do not contain a comprehensive, integrated list of the relevant scientific research being conducted across the weapons complex to support the development and implementation of QMU. For example, according to the NNSA campaign manager for the Primary campaign, he had to hold a workshop in 2005 with officials from the weapons laboratories in order to catalogue all of the scientific activities that are currently performed under the heading of “primary assessment” regardless of the NNSA funding source. According to this official, the existing Primary campaign implementation plan does not provide the integration across NNSA programs that is needed to achieve the goals of the Primary campaign and to develop and implement QMU.

According to NNSA officials, the lack of integration has occurred in large part because a significant portion of the scientific research that is relevant to the Primary and Secondary campaigns is funded and carried out by different campaigns and other programs. Specifically, different NNSA campaign managers use different campaign planning documents to plan and oversee research and funding for activities that are directly relevant to the Primary and Secondary campaigns and the development and implementation of QMU. For example, the ASC campaign provides the supercomputing capability that the weapons laboratories use to simulate and predict the behavior of an exploding nuclear weapon. Moreover, the weapons laboratories rely on ASC supercomputers to quantify their uncertainties with respect to the accuracy of these computer simulations—a key component in the implementation of QMU. As a result,

the ASC campaign plans and funds activities that are critical to the development and implementation of QMU.

To address this problem, according to NNSA officials, NNSA is taking steps to establish better relationships among the campaign plans. For example, NNSA is currently drafting a new plan—which it calls the Primary Assessment Plan—in an attempt to better coordinate the activities covered under the separate program and implementation plans. The draft plan outlines high-level research priorities, time lines, and proposed milestones necessary to support (1) NNSA’s responsibilities for the current stockpile, (2) primary physics design for the development of an RRW, and (3) certification of an RRW in the 2012 time frame and a second RRW in the 2018 time frame. According to NNSA officials, they expect to finalize this plan by the third quarter of fiscal year 2006. In addition, they expect to have a similar plan for the Secondary campaign finalized by December 2006 and are considering combining both plans into a full-system assessment plan. According to one NNSA official responsible for the Primary and Secondary campaigns, NNSA will revise the existing campaign program and implementation plans to be consistent with the Primary Assessment Plan.

More fundamentally, some nuclear weapons experts have suggested that NNSA’s planning structure should be reorganized to better reflect the use of QMU as NNSA’s main strategy for assessing and certifying the performance of nuclear weapons. For example, the chair of the LLNL Defense and Nuclear Technologies Director’s Review Committee—which conducts technical reviews of LLNL’s nuclear weapons activities for the University of California—told us that the current campaign structure has become a series of “stovepipes” that NNSA uses to manage stockpile stewardship. He said that in order for NNSA to realize its long-term goals for implementing QMU, NNSA is going to have to reorganize itself around something that he called an “uncertainty spreadsheet” for each element of a weapon’s performance (e.g., implosion of the primary, transfer of energy to the secondary, etc.), leading to the weapon’s yield. He said that the laboratories should develop a spreadsheet for each weapon in the stockpile that (1) identifies the major sources of uncertainty at each critical event in their assessment of the weapon’s performance and (2) relates the laboratory’s scientific activities and milestones to these identified sources of uncertainty. He said that the development and use of these spreadsheets would essentially capture the intent of the scientific campaigns and make them unnecessary.

NNSA Does Not Have a Clear, Consistent Set of QMU-Related Milestones

NNSA has established a number of milestones that relate to the development and implementation of QMU. Within the Science campaign program plan, NNSA has established a series of high-level milestones, which it calls “level-1” milestones. According to NNSA policies, level-1 milestones should be sufficient enough to allow strategic integration between sites involved in the campaigns and between programs in NNSA. Within the implementation plans for the Primary and Secondary campaigns, NNSA has established a number of lower-level milestones, which it calls “level-2” milestones, which NNSA campaign managers use to track major activities for the current year of execution. The level-1 milestones related to QMU are shown in table 4, and the level-2 milestones related to QMU for the Primary campaign are shown in table 5.

Table 4: NNSA Level-1 Milestones Related to the Development and Implementation of QMU

Due date	Milestone number	Milestone description
FY2007	M46	Publish documented plan to reduce major sources of uncertainty. (Cycle I)
FY2010	M47	Accounting for simulation and experimental uncertainties, assess ability to reproduce the full underground test data sets for a representative group of nuclear tests with a consistent set of models.
FY2011	M48	Publish documented plan to reduce the major sources of uncertainty assessed in fiscal year 2010. (Cycle II)
FY2014	M20	Accounting for simulation and experimental uncertainties, reassess ability to reproduce the full underground test data sets for a representative group of nuclear tests with a consistent set of models.

Source: NNSA, FY2006 Science campaign program plan.

Table 5: Primary Campaign Level-2 Milestones Related to the Development and Implementation of QMU

Due date	Milestone description
FY2004	Analyze specific underground test events in the support of QMU.
FY2004	Develop QMU certification logic to support the W76.
FY2004	Develop QMU certification logic to support the W88.
FY2005	Analyze specific underground test events in the support of QMU.
FY2005	Predict primary performance and identify major sources of uncertainty for the W-76 LEP. Quantify these sources where possible or develop requirements of a plan to do so.
FY2005	Develop probabilistic tools and methods to combine various sources of uncertainty for primary performance.

Source: NNSA Primary campaign implementation plans, fiscal years 2004 and 2005.

According to NNSA officials, the level-1 milestones in table 4 represent a two-stage path to systematically identify uncertainties and reduce them through analyzing past underground test results, developing new experimental capabilities, and performing new experiments to understand the relevant physical processes. According to these level-1 milestones, NNSA expects to complete the second stage or “cycle” of this process by fiscal year 2014 (i.e., milestone M20), at which time NNSA will have sufficiently reduced major sources of uncertainties and will have confidence in its ability to predict the performance of nuclear weapons in the absence of nuclear testing.

However, we identified several problems with the NNSA milestones related to the development and implementation of QMU. Specifically, the level-1 milestones in the Science campaign program plan have the following problems:

- The milestones are not well-defined and never explicitly mention QMU. According to NNSA officials responsible for overseeing the Primary campaign, these milestones are too qualitative and too far in the future to enable NNSA to effectively plan for and oversee the implementation of QMU. They described these milestones as “fuzzy” and said that they need to be better defined. However, NNSA officials also stated that these milestones are not just for QMU but for the entire Science campaign, of which QMU is only a part.

-
- The milestones conflict with the performance measures shown in other important NNSA management documents. Specifically, while the Science campaign program plan envisions a two-stage path to identify and reduce key uncertainties related to nuclear weapon operations using QMU by 2014, the performance measures in NNSA's fiscal year 2006 budget request and in Appendix A of the Science campaign program plan call for the completion of QMU by 2010.
 - The milestones have not been integrated with other QMU-related level-1 milestones in other planning documents. For example, the current ASC campaign program plan contains a series of level-1 milestones for completing the certification of several weapon systems—including the B61, W80, W76, and W88—with quantified margins and uncertainties by the end of fiscal year 2007. However, these milestones do not appear in and are not referenced by the Science campaign program plan. Moreover, the ASC campaign manager told us that, until recently, he was not aware of the existence of the level-1 milestones for implementing QMU that are contained in the Science campaign program plan.

In addition, we found that neither the Science campaign program plan nor the Primary campaign implementation plan describe how the level-2 milestones on QMU in the Primary campaign implementation plan are related to the level-1 milestones on QMU in the Science campaign program plan. Consequently, it is unclear how the achievement of specific level-2 milestones—such as the development of probabilistic tools and methods to combine various sources of uncertainty for primary performance—will result in the achievement of level-1 milestones for the implementation of QMU or how NNSA expects to certify several major nuclear weapon systems using QMU before the QMU methodology is fully developed and implemented.

NNSA, as well as laboratory officials, agreed that there are weaknesses with the current QMU milestones. According to NNSA officials, when NNSA established the current tiered structure for campaign milestones in 2003, the different tiers of milestones served different purposes and, therefore, were never well-integrated. For example, NNSA officials said that the level-1 milestones were originally created to reflect measures that were deemed to be important to senior NNSA officials, while level-2 milestones were created to be used by NNSA campaign managers to perform more technical oversight of the weapons laboratories. Furthermore, according to NNSA officials, the current level-2 milestones are only representative of campaign activities conducted by the weapons

laboratories. That is, the level-2 milestones were never designed to cover the entire scope of work being conducted by the weapons laboratories and are, therefore, not comprehensive in scope.

To address these problems, according to NNSA officials, NNSA is taking steps to develop better milestones to track the implementation of the QMU methodology. For example, in the draft Primary Assessment Plan, NNSA has established 19 “high-level” milestones that cover the time period from fiscal year 2006 to fiscal year 2018. According to these draft milestones, by fiscal year 2010, NNSA expects to “complete the experimental work and methodology development needed to demonstrate the ability of primary certification tools to support certification of existing stockpile system and RRW.” In addition, NNSA expects to certify a RRW in fiscal year 2012 and a second RRW in fiscal year 2018.

NNSA Has Not Established Formal Requirements for Conducting Technical Reviews or Certifying the Completion of QMU-Related Milestones

According to NNSA policies, campaign managers are required to track the status of level-1 and level-2 milestones and provide routine, formal reports on the status of their programs. For example, campaign managers are required to track, modify, and score the status of level-1 and level-2 milestones through the use of an Internet-based application called the Milestone Reporting Tool. On a quarterly basis, campaign managers assign one of four possible scores for each milestone listed in the application: (1) “blue” for completed milestones, (2) “green” for milestones that are on track to be finished by the end of the fiscal year, (3) “yellow” for milestones that may not be completed by the end of the fiscal year, and (4) “red” for milestones that will not be completed by the end of the fiscal year. At quarterly program review meetings, campaign managers brief senior-level NNSA officials on the status of major milestones, along with cost and expenditure data for their programs. In addition, campaign managers are responsible for conducting technical reviews of the campaigns for which they are responsible, at least annually, to ensure that campaign activities are being executed properly and that campaign milestones are being completed.

However, NNSA campaign managers have not met all of the NNSA requirements needed to effectively oversee the Primary and Secondary campaigns. For example, we found that the campaign managers for the Primary and Secondary campaigns have not established formal requirements for conducting annual, technical reviews of the implementation of QMU at the three weapons laboratories. Moreover, these officials have not established requirements for certifying the completion of

level-2 milestones that relate to QMU. They could not provide us with documentation showing the specific activities or outcomes that they expected from the weapons laboratories in order to certify that the laboratories had completed the level-2 milestones for QMU. Instead, they relied more on ad hoc reviews of campaign activities and level-2 milestones as part of their oversight activities for their campaigns. According to the Primary campaign manager, the officials at the weapons laboratories are the principal managers of campaign activities. As a result, he views his role as more of a “sponsor” for his program and, therefore, does not require any written reports or evidence from the laboratories to certify that they have completed specific milestones.

In contrast, we found that the ASC campaign manager has established formal requirements for a variety of reoccurring technical reviews of activities associated with the ASC campaign. Specifically, the ASC campaign relies on semiannual reviews conducted by the ASC Predictive Science Committee—which provides an independent, technical review of the status of level-2 milestones—as well as on annual “principal investigators” meetings that provide a technical review of every program element within the ASC campaign. The ASC campaign manager told us that he relies on these technical reviews to oversee program activities because the quarterly program review meetings are not meant to help him manage his program but are really a way for senior-level NNSA officials to stay informed.

In addition, the ASC campaign manager has established detailed, formal requirements for certifying the completion of level-2 milestones for the ASC campaign. Specifically, the fiscal year 2006 implementation plan for the ASC campaign contains a detailed description of what NNSA expects from the completion of each level-2 milestone, including a description of completion criteria, the method by which NNSA will certify the completion of the milestone, and an assessment of the risk level associated with the completion of the milestone. The ASC campaign manager told us that, when NNSA officials created the level-2 milestones for the campaigns in 2003, the milestones were really just “sentences” and lacked the detailed criteria that would enable NNSA managers to adequately track and document the completion of major milestones. As a result, the ASC campaign has made a major effort in recent years to develop detailed, formal requirements to support the completion of ASC level-2 milestones.

NNSA Has Not Established Adequate Measures to Determine the Laboratories' Performance in Developing and Implementing QMU

NNSA uses performance measurement data to inform resource decisions, improve the management and delivery of products and services, and justify budget requests. According to NNSA requirements, performance measurement data should explain in clear, concise, meaningful, and measurable terms what program officials expect to accomplish for a specific funding level over a fixed period of time. In addition, performance measurement data should include annual targets that describe specific outputs that can be measured, audited, and substantiated by the detailed technical milestones contained in documentation such as campaign implementation plans.

With respect to QMU, NNSA has established an overall annual performance target to measure the cumulative percentage of progress toward the development and implementation of the QMU methodology. Specifically, in its fiscal year 2006 budget request to the Congress, NNSA stated that it expects to complete the development and implementation of QMU by 2010 as follows:

- 25 percent complete by the end of fiscal year 2005,
- 40 percent complete by the end of fiscal year 2006,
- 55 percent complete by the end of fiscal year 2007,
- 70 percent complete by the end of fiscal year 2008,
- 85 percent complete by the end of fiscal year 2009, and
- 100 percent complete by the end of fiscal year 2010.

According to NNSA, it had progressed 10 percent toward its target of completing QMU by the end of fiscal year 2004. However, NNSA officials could not document how they can measure progress toward the performance target for developing and implementing QMU. Moreover, NNSA officials could not explain how the 2010 overall performance target for the completion and implementation of QMU is related to the level-1 milestones for QMU in the Science campaign program plan, which describes a two-stage process to identify and reduce key uncertainties in nuclear weapon performance using QMU by 2014. According to one NNSA official responsible for overseeing the Primary campaign, NNSA created this annual performance target because the Office of Management and Budget requires agencies to express some of their annual performance

targets in percentage terms. However, this official said the actual percentages are not very meaningful, and he does not have any specific criteria for how to measure progress to justify the use of the percentages in the budget request.

NNSA has also established broad performance measures to evaluate the performance of LANL and LLNL. Specifically, in its performance evaluation plans for LANL and LLNL for fiscal year 2006, NNSA has established the following three performance measures:

- Use progress toward quantifying margins and uncertainty, and experience in application, to further refine and document the QMU methodology.
- Demonstrate application of a common assessment methodology (i.e., QMU) in major warhead assessments and the certification of Life Extension Program warheads.
- Complete the annual assessment of the safety, reliability, and performance of all warhead types in the stockpile, including reaching conclusions on whether nuclear testing is required to resolve any issues.

However, the plan that NNSA uses to evaluate the performance of SNL does not contain any performance measures or targets specifically related to QMU, and the performance evaluation plans for LANL and LLNL do not contain any annual targets that can be measured and linked to the specific performance measures related to QMU. Instead, the plans state that NNSA will rely on LLNL and LANL officials to develop the relevant targets and related dates for each performance measure, as well as to correlate the level-1 and level-2 milestones with these measures. When asked why these plans do not meet NNSA's own requirements, NNSA officials said that they have not included specific annual performance targets in the plans because to do so would make it harder for them to finalize the plans and adjust to changes in NNSA's budget. However, they said that NNSA is planning on implementing more stringent plans that will include annual performance targets when the next contract for LANL and LLNL is developed. In addition, NNSA officials told us that they recognize the need to develop performance measures related to QMU for SNL and anticipate implementing these changes in the fiscal year 2007 performance evaluation plan.

NNSA officials told us that they have used more specific measures, such as the completion of level-2 milestones, in their assessment of the weapons laboratories' performance since fiscal year 2004. However, we also found problems with the way NNSA has assessed the performance of the weapons laboratories in implementing QMU. For example, in NNSA's annual performance appraisal of LANL for fiscal year 2004, NNSA states that LANL had completed 75 percent of the work required to develop "QMU logic" for the W76 life extension by the end of fiscal year 2004. However, NNSA officials could not document how they are able to measure progress toward the development and implementation of QMU logic for the W76 life extension. Again, an NNSA official responsible for overseeing the Primary campaign told us that the actual percentages are not very meaningful, and that he did not have any specific criteria for how to measure progress to justify the use of the percentage in the appraisal.

In a recent report, we recognized the difficulties of developing useful results-oriented performance measures for programs such as those geared toward research and development programs.¹¹ For programs that can take years to observe program results, it can be difficult to identify performance measures that will provide information on the annual progress they are making toward achieving program results. However, we also recognize that such efforts have the potential to provide important information to decision makers.

NNSA officials told us that they recognize the need for developing appropriate measures to ensure that adequate progress is being maintained toward achieving the goals and milestones of the campaigns. However, according to NNSA, very few products of the scientific campaigns involve the repetition of specific operations whose costs can be monitored effectively as a measure of performance. As a result, the best measure of progress for the scientific campaigns is through scientific review by qualified technical peers at appropriate points in the program. However, NNSA has not established any performance measures or targets for implementing QMU that require periodic scientific peer reviews or define what is meant by "appropriate" points in the program.

¹¹GAO, *Performance Budgeting: PART Focuses Attention on Program Performance, but More Can Be Done to Engage Congress*, GAO-06-28 (Washington, D.C.: Oct. 28, 2005).

Conclusions

Faced with an aging nuclear stockpile, as well as an aging workforce, NNSA needs a methodologically rigorous, transparent, and explainable approach for how it will continue to assess and certify the safety and reliability of the nation's nuclear weapons stockpile, now and into the foreseeable future, without underground testing. After over a decade of conducting stockpile stewardship, NNSA's selection of QMU as its methodology for assessment and certification represents a positive step toward a methodologically rigorous, transparent, and explainable approach that can be carried out by a new cadre of weapons designers. However, important technical and management details must be resolved before NNSA can say with certainty that it has a sound and agreed upon approach.

First, NNSA must take steps to ensure that all three nuclear weapons laboratories—not just LANL and LLNL—are in agreement about how QMU is to be defined and applied. While we recognize that there will be methodological differences between LANL and LLNL in the detailed application of QMU to specific weapon systems, we believe that it is fundamentally important that these differences be understood and, if need be, reconciled, to ensure that QMU achieves the goal of a common methodology with rigorous, quantitative, and explicit criteria, as envisioned by the original 2003 white paper on QMU. More importantly, we believe that SNL has an important role in the development and application of QMU to the entire warhead, and we find the continuing disagreement over the application of QMU to areas outside of the nuclear explosive package to be disconcerting. There have been several recommendations calling for a new, technical paper defining QMU, as well as the establishment of regular forums to further develop the QMU methodology and reconcile any differences in approach. We believe the NNSA needs to fully implement these recommendations.

Second, NNSA has not made effective use of its current planning and program management structure to ensure that all of the research needed to support QMU is integrated and that scarce scientific resources are being used efficiently. We believe that NNSA must establish an integrated management approach involving planning, oversight, and evaluation methods that are all clearly linked to the overall goal of the development and application of QMU. In particular, we believe that NNSA needs clear, consistent, and realistic milestones and regular, technical reviews of the development of QMU in order to ensure sound progress. Finally, while we support the development of QMU and believe it must be effectively managed, we also believe it is important to recognize and acknowledge that

the development and application of QMU, especially the complexities involved in analyzing and combining uncertainties related to potential failure modes and performance margins, represents a daunting research challenge that may not be achievable in the time constraints created by an aging nuclear stockpile.

Recommendations for Executive Action

To ensure that the weapons laboratories will have the proper tools in place to support the continued assessment of the existing stockpile or the certification of redesigned nuclear components under the RRW program, we recommend that the Administrator of NNSA take the following two actions:

- Require the three weapons laboratories to formally document an agreed upon, technical description of the QMU methodology that clearly recognizes and reconciles any methodological differences.
- Establish a formal requirement for periodic collaboration between the three weapons laboratories to increase their mutual understanding of the development and implementation of QMU.

To ensure that NNSA can more effectively manage the development and implementation of QMU, we recommend that the Administrator of NNSA take the following three actions:

- Develop an integrated plan for implementing QMU that contains (1) clear, consistent, and realistic milestones for the development and implementation of QMU across the weapons complex and (2) formal requirements for certifying the completion of these milestones.
- Establish a formal requirement for conducting annual, technical reviews of the scientific research conducted by the weapons laboratories that supports the development and implementation of QMU.
- Revise the performance evaluation plans for the three weapons laboratories so that they contain annual performance targets that can be measured and linked to specific milestones related to QMU.

Agency Comments and Our Evaluation

We provided NNSA with a draft of this report for their review and comment. Overall, NNSA agreed that there was a need for an agreed-upon technical approach for implementing QMU and that NNSA needed to improve the management of QMU through clearer, long-term milestones and better integration across the program. However, NNSA stated that QMU had already been effectively implemented and that we had not given NNSA sufficient credit for its success. In addition, NNSA raised several issues about our conclusions and recommendations regarding their management of the QMU effort. The complete text of NNSA's comments on our draft report is presented in appendix I. NNSA also made technical clarifications, which we incorporated in this report as appropriate.

With respect to whether QMU has already been effectively implemented, during the course of our work, LANL and LLNL officials showed us examples of where they used the QMU methodology to examine specific issues associated with the stockpile. At the same time, during our discussions with laboratory officials, as well as with the Chairs of the JASON panel on QMU, the Office of Defense Programs Science Counsel, and the Strategic Advisory Group Stockpile Assessment Team of the U.S. Strategic Command, there was general agreement that the application of the QMU methodology was still in the early stages of development. As NNSA pointed out in its letter commenting on our report, to implement QMU, the weapons laboratories need to make a number of improvements, including techniques for combining different kinds of uncertainties, as well as developing better models for a variety of complex processes that occur during a nuclear weapon explosion. In addition, the successful implementation of QMU will continue to rely on the expert judgment and the successful completion of major scientific facilities such as the National Ignition Facility. We have modified our report to more fully recognize that QMU is being used by the laboratories to address stockpile issues and to more completely characterize its current state of development. At the same time, however, because QMU is still under development, we continue to believe that NNSA needs to make more effective use of its current planning and program management structure.

NNSA raised several specific concerns about our conclusions and recommendations. First, NNSA disagreed with our conclusion and associated recommendations that NNSA take steps to ensure that all three nuclear weapons laboratories are in agreement about how QMU is to be defined and applied. NNSA stated that we overemphasized the differences between LANL and LLNL in implementing QMU and that, according to

NNSA, LANL and LLNL have a “common enough” agreement on QMU to go forward with its implementation. Moreover, NNSA stated that our recommendations blur very clear distinctions between SNL and the two nuclear design labs. According to NNSA, QMU is applied to issues regarding the nuclear explosive package, which is the mission of LANL and LLNL.

While we believe that some of the technical differences between the laboratories remain significant, we have revised our report to more accurately reflect the nature of the differences between LANL and LLNL. With respect to SNL, we would again point out that SNL officials are still required to certify the performance of nuclear weapon components under the conditions of a nuclear explosion and, thus, use similar elements of the QMU methodology. Therefore, we continue to believe that all three laboratories, as well as NNSA, would benefit from efforts to more formally document the QMU methodology and regularly meet to increase their mutual understanding. As evidence of the benefits of this approach, we would note that LLNL and LANL are currently developing a revised “white paper” on QMU, and that in discussions with one of the two authors, he agreed that inclusion of SNL in the development of the draft white paper could be beneficial.

Second, NNSA made several comments with respect to our recommendation that NNSA develop an integrated plan for implementing QMU that contains clear, consistent, and realistic milestones. For example, NNSA stated that they expect to demonstrate the success of the implementation of QMU and the scientific campaigns by the performance of a scientifically defensible QMU analysis for each required certification problem. In addition, NNSA stated that the 2010 budget target and the 2014 milestone were developed for different purposes and measure progress at different times. According to NNSA, the 2010 target describes developing QMU to the point that it can be applied to certification of a system (e.g., the W88) without underground testing, while the 2014 milestone is intended to be for the entire Science campaign effort.

However, as we state in our report, and as acknowledged by NNSA officials responsible for the Primary and Secondary campaigns, there continue to be problems with the milestones that NNSA has established for implementing QMU. Among these problems is the fact that these milestones are not well-defined and conflict with other performance measures that NNSA has established for QMU. Moreover, in its comments on our report, NNSA agreed that better integration and connectivity of milestones between

various program elements would improve the communications of the importance of program goals and improve the formality of coordination of program activities, “which is currently accomplished in an informal and less visible manner.” Given this acknowledgment by NNSA, we continue to believe that an integrated plan for implementing QMU, rather than NNSA’s current ad hoc approach, is warranted.

Third, NNSA made several comments regarding our recommendation that NNSA establish a formal requirement for conducting annual, technical reviews of the scientific research conducted by the weapons laboratories that supports the development and implementation of QMU. NNSA stated that it believes the ad hoc reviews it conducts, such as the JASON review, provide sufficient information on scientific achievements, difficulties, and required redirection to manage these programs effectively. As a result, NNSA stated that it has not selected a single review process to look at overall success in the implementation of QMU but expects to continue to rely on ad hoc reviews.

We agree that reviews, such as the JASON review, are helpful, and we relied heavily on the JASON review, as well as other reviews as part of our analysis. However, as we point out in the report, the issue is that the campaign managers for the Primary and Secondary campaigns do not meet all of NNSA’s own requirements for providing effective oversight, which include the establishment of formal requirements for conducting technical reviews of campaign activities. Therefore, we believe that NNSA needs to take steps to implement its own policies. In addition, we believe that the ASC campaign provides a good role model for how the Primary and Secondary campaigns should be managed.

Finally, NNSA made several comments with respect to our recommendation for NNSA to revise the performance evaluation plans for the laboratories so that they contain annual performance targets that can be measured and linked to specific milestones related to QMU. Specifically, NNSA stated that the implementation of QMU is an area where it is difficult to establish a meaningful metric. According to NNSA, since QMU is implicitly evaluated in every review of the components of the science campaign, NNSA does not believe it is necessary to formally state an annual QMU requirement. However, as we point out in the report, the current performance evaluation plans for LANL and LLNL do not meet NNSA’s own requirements for the inclusion of annual performance targets that can be measured and linked to the specific performance measures related to QMU. More fundamentally, since NNSA has placed such

emphasis on the development and implementation of QMU in the years ahead, we continue to believe that NNSA needs to develop more meaningful criteria for assessing the laboratories' progress in developing and implementing QMU.

We are sending copies of this report to the Administrator, NNSA; the Director of the Office of Management and Budget; and appropriate congressional committees. We also will make copies available to others upon request. In addition, the report will be available at no charge on the GAO Web site at <http://www.gao.gov>.

If you or your staff have any questions about this report or need additional information, please contact me at (202) 512-3841 or aloisee@gao.gov. Contact points for our Offices of Congressional Relations or Public Affairs may be found on the last page of this report. GAO staff who made major contributions to this report are listed in appendix II.

A handwritten signature in black ink that reads "Gene Aloise". The signature is written in a cursive style with a large, looped initial "G".

Gene Aloise
Director, Natural Resources
and Environment

Comments from the National Nuclear Security Administration



Department of Energy
National Nuclear Security Administration
Washington, DC 20585

JAN 10 2006

Mr. Gene Aloise
Director
Natural Resources and Environment
U.S. Government Accountability Office
Washington, D.C. 20548

Dear Mr. Aloise:

The National Nuclear Security Administration (NNSA) appreciates the opportunity to review the Government Accountability Office's (GAO) draft report, GAO-06-261, "NUCLEAR WEAPONS: NNSA Needs to Refine and More Effectively Manage Its New Approach for Assessing and Certifying Nuclear Weapons." NNSA understands that the House Strategic Forces Subcommittee, Committee on Armed Services, originally requested GAO to determine how NNSA currently defines the scientific research portion of its campaign that is intended to provide a safe and reliable stockpile. During the course of this audit, the scope of the audit evolved into a review of the Quantification of Margins and Uncertainties (QMU) methodology for assessing and certifying the stockpile.

While NNSA agrees that there must be an agreed-upon technical approach to QMU implementation and that NNSA should always strive to improve the management of QMU implementation, we believe that QMU has been implemented as an effective approach to stockpile certification. The present implementation of QMU is highly effective in bringing science to stockpile issues and used for weapons certification issues across the stockpile, as well as being the basis for the Laboratory Directors' recommendations in the annual assessment reports on the stockpile.

Ad hoc scientific reviews conducted by panels such as JASONS, the University of California Science and Technology Panel, and the Strategic Commands' Strategic Advisory Group Stockpile Assessment Team (SAGSAT) are appropriate *fora* for assessing scientific programs in areas depending on the implementation of QMU. Those reviews have demonstrated the steady and rapid progress in the application of QMU to weapons certification since the initial 2003 white-paper on QMU implementation. The success in the development of QMU is an accomplishment resulting from a decade of scientific progress since the establishment of Stockpile Stewardship in 1995. Continued progress in key science areas in primary and



Printed with soy ink on recycled paper

Appendix I
Comments from the National Nuclear
Security Administration

2

secondary physics, materials science, and high energy density physics, including the National Ignition Campaign, and computational advances are required to sustain future certification requirements.

We have enclosed two documents for GAO's consideration prior to the publication of the final report. The first document addresses the background for QMU and what the Program believes to be the maturity level of the QMU process. The second document is detailed technical comments for your consideration.

Should you have any questions related to this response, please contact Richard Speidel, Director, Policy and Internal Controls Management.

Sincerely,



Michael C. Kane
Associate Administrator
for Management and Administration

Enclosures

cc: Deputy Administrator for Defense Programs
Senior Procurement Executive
Director, Service Center

NNSA Response to the GAO report, GAO-06-261, "NUCLEAR WEAPONS: NNSA Needs to Refine and More Effectively Manage Its New Approach for Assessing and Certifying Nuclear Weapons."

Executive Summary

Because of a successful record of progress in the development of the Quantification of Margins and Uncertainties (QMU) approach for certifying nuclear warheads, the National Nuclear Security Administration (NNSA) is not seeking further refinements beyond the currently envisioned program of work. NNSA will, however, seek management improvements in implementing this approach.

Despite the conclusions of the GAO audit of NNSA's QMU program, the NNSA has already implemented QMU as an effective approach to stockpile certification.

- The present implementation of QMU is highly effective in bringing science to stockpile issues.
- QMU is now used for weapons certification issues across the stockpile and is the basis for the Laboratory Directors' recommendations in the annual assessment reports on the stockpile.
- The *ad hoc* scientific reviews conducted by panels such as JASONs, University of California Science and Technology Panel, and the Strategic Commands' Strategic Advisory Group Stockpile Assessment Team (SAGSAT) are appropriate *forums* for assessing scientific programs in areas depending on the implementation of QMU.
- These reviews have demonstrated steady and rapid progress in the application of QMU to weapons certification since the initial 2003 white paper on QMU implementation.
- The success in developing QMU is a key accomplishment resulting from a decade of outstanding scientific progress since the establishment of Stockpile Stewardship in 1995.
- Continued progress in key science areas in primary and secondary physics, materials science, and high energy density physics, including the National Ignition Campaign, and computational advances will be required to sustain future certification requirements.

NNSA recognizes that design and certification of a Reliable Replacement Warhead (RRW) as well as transformation of the nuclear weapons complex to meet newly identified responsive infrastructure goals pose new challenges. These will require a careful review of science campaign priorities and will require better integration across NNSA activities. The recent completion of the revised Work Breakdown Structure for Advanced Simulation and Computing (ASC) and the completion of the Primary Assessment Plan are initial steps in that process. The QMU approach can be managed as an integrating influence across program components.

**THE NATIONAL NUCLEAR SECURITY ADMINISTRATION (NNSA) HAS
IMPLEMENTED QMU**

Despite the contention of the GAO that because of management shortcomings NNSA is likely to have difficulties in implementing QMU, the Department of Energy maintains that it has successfully implemented QMU. This methodology constitutes the framework by which the Directors of the Nuclear Weapons National Laboratories, through the Secretaries of Energy and Defense, execute their statutory responsibility to assure the President of the United States of the safety, security and reliability of the U.S. nuclear deterrent. It has visibility, oversight and management from the highest levels of the government, the national laboratories, and the august scientific bodies that provide advice to the Administration, to its agencies, and to the Congress. Not appreciating the demonstrated success in the implementation of QMU has led to unfounded conclusions that because of management failings QMU is likely to fail in the future and that important efforts to transform the stockpile may be at risk.

QMU is a framework for connecting the scientific method to a variety of questions regarding assessment of the stockpile and for presenting the results. Review of QMU requires a scientific evaluation of progress in providing objective, technically based answers to complex questions that arise in the prediction of the performance, safety and reliability of the stockpile. Its utility is best judged not in the abstract but in the context of the ability to solve specific problems, in this case, the set of specific issues that must be settled in order to certify specific devices.

The report is critical of *ad hoc* reviews to measure the progress in QMU. Because the value of QMU is most meaningfully weighed by evaluating technical progress in specific applications, however, NNSA relies on *ad hoc* expert reviews of the application of QMU to specific problems as the best review mechanism. The JASON review of QMU and an ongoing JASON review on pit lifetimes, which is an application of QMU in a vital area, are both examples of such reviews. Several Strategic Advisory Group Stockpile Assessment Team (SAGSAT) reviews of specific stockpile certification issues are additional examples of useful reviews.

Each year the NNSA and other organizations conduct numerous reviews that cover the broad gamut of efforts within the science campaign and in particular on subjects where QMU plays a vital role. While there are a number of drivers to conduct reviews, NNSA is cognizant of the high programmatic costs on the laboratories to support these and is hesitant to add to their number unless given good justification. NNSA believes that the reviews it conducts and those of which it has cognizance provide sufficient information on scientific achievements, difficulties and required redirection to manage these programs effectively. This GAO audit itself relies in part on the results of those same *ad hoc* reviews, but appears to undervalue them.

Despite the characterization by the GAO report, the development of QMU and its present application to the broad range of certification issues facing the national laboratories is a significant and vital accomplishment. It represents progress brought about through a

Appendix I
Comments from the National Nuclear
Security Administration

sustained decade long effort in implementing the charge of the FY 1994 National Defense Authorization Act which directed the Secretary of Energy to "establish a stewardship program to ensure the preservation of the core intellectual and technical competencies of the United States in nuclear weapons, including weapons design, system integration, manufacturing, security, use control, reliability assessment, and certification."

In response, DOE developed the 1995 Science Based Stockpile Stewardship program, which set out the vision that DOE has subsequently followed, with few modifications. Important efforts included the establishment of Accelerated Strategic Computing Initiative (ASCI, now ASC), revitalization of the Inertial Confinement Fusion (ICF) program including high energy density physics, efforts in hydrodynamic experiments and facilities, and a variety of experimental efforts to improve understanding of materials properties crucial to prediction of weapons performance.

In order to better organize the program, establish more specific goals, track progress, and provide a level of transparency to its sponsors, DOE created the campaign program management structure in 1999, creating the six science campaign efforts that are the subject of this GAO report. Efforts begun in response to the 1995 program had by 2002 achieved substantial improvements in capabilities. These included; the development of primary and secondary burn codes and the improved computational capability provided by Advanced Simulation and Computing (ASC); improved understanding of underlying phenomenology through experimental successes in hydrotesting and the subcritical experiments program; successes in the area of high energy density physics; improved understanding of the properties and aging of nuclear weapon materials and components; and, improved analysis of historical underground nuclear tests.

The progress in these underlying capabilities enabled the development of QMU requested by NNSA and described in the seminal 2003 QMU paper by Dr. Bruce Goodwin and Dr. Ray Juzaitis referred to in the report. Subsequent progress has been rapid, from the partial level of application of QMU shown in the NNSA Science Council Review (the 2004 Friendly Review), the increased progress shown in the 2004 JASON review of QMU, finally, to the application in 2005 of QMU to all annual weapons systems assessment and certifications.

In the implementation of the underlying 1994 Congressional charge and 1995 program, QMU represents a transformation from certification based on the individual judgment of designers grounded in the success of the underground test program, to more quantitative and objective results. As the report notes, "QMU seeks to quantify (1) how close each critical factor is to the point at which it would fail to perform as designed (i.e., the margin to failure) and (2) the uncertainty that exists in calculating the margin, in order to ensure that the margin is sufficiently larger than the uncertainty." It is in the formal development and presentation of quantitative analyses that QMU enables the articulation of institutional conclusions where results can be analyzed, repeated, and any differences can be reconciled through inter-laboratory peer review processes.

Appendix I
Comments from the National Nuclear
Security Administration

While this statement of the basics of QMU is relatively simple and easily understood, the complexity occurs in the detailed application to specific weapons systems and performance issues because, for each weapons system, the potential failure modes for the system must be identified, margins for each established, and a thorough analysis of uncertainties in establishing those margins must be applied. The uncertainty analysis is complex because it must convolve information about manufacturing variability, uncertainties of physical understanding of complex physical phenomenon, and the limitations of a sparse set of underground and aboveground test data.

GAO claims that "absent the prompt resolution of remaining disagreements over the definition and implementation of QMU, it is unclear whether the weapons laboratories will have the common framework they say they need to support the continued assessment of the existing stockpile or the certification of redesigned nuclear components under the RRW program." NNSA believes this statement is incorrect. The laboratories have a common enough agreement on QMU, and the definitional differences are well enough appreciated that any difficulties in certifying an RRW, or any other system will be the result of a currently unforeseen fundamental technical issue rather than a definitional dispute. The outcome of a scientific debate does not depend upon the definition of words, but evidence developed to answer a question. NNSA's observation at working sessions with the laboratories is that scientists now quickly get down to the hard work of understanding fundamental differences in the outcome of experiments or predictions of computer models rather than argue over approaches.

To be precise, the Directors of the Nuclear Weapons National Laboratories certify that a nuclear warhead will meet the "Military Characteristics" under a specified "Stockpile to Target Sequence." While it is nuclear weapons that are certified and not individual components, QMU within the context of the six science campaign efforts that were the subject of the GAO report, is applied to issues regarding the nuclear explosives package, which is the mission of the nuclear design laboratories, Lawrence Livermore National Laboratory and Los Alamos National Laboratory. While Sandia National Laboratories has its own applications of the QMU methodology, and communications and sharing of techniques occurs among all three laboratories, the GAO recommendations blur very clear distinctions between Sandia and the two nuclear design laboratories.

Nevertheless, the Sandia approach is reconcilable with the approach used by the nuclear design laboratories, as is to be expected since the Sandia manager who was responsible for developing their certification methodologies is now at LANL responsible for the development and implementation of LANL's certification methodology. Nevertheless, the specific problems to be solved are different, using different codes, models and experimental tools in very different physical regimes. In those areas where interchange can usefully occur, it happens and will continue, such as the development of statistical techniques.

In overemphasizing the level of difference, the GAO also underplays the role of complementary approaches, the importance of which is the rationale for a national policy decision to maintain two nuclear design laboratories. First, as noted, weapons certification involves scientific research where the outcome is not a *priori* known and

Appendix I
Comments from the National Nuclear
Security Administration

confidence in the correctness of the result is improved by achieving the same answer through multiple approaches. The implication that lack of an identical approach motivates a need to "refine" the approach, suggests a misunderstanding of the nature of scientific approaches to complex problems. Generally, a scientific result is accepted not because of uniformity of methods, but because multiple researchers can reproduce the same result using different techniques and approaches.

Likewise, GAO misunderstands LLNL's approach in finding that the LLNL combines QMU ratios for different failure modes into a single ratio for the entire warhead. LLNL does not do this. (LLNL combines uncertainties within a given failure mode.) In fact, all labs treat each failure mode separately. Inadequate margin for any failure mode represents a risk that the weapon will fail to meet requirements. Excess margin in one area, e.g., primary yield cannot generally compensate for inadequate margin in another area, e.g., one-point safety.

To be sure, NNSA continues to perform research because there are vital questions that are unanswered and vital capabilities that must be improved to ensure the long-range health of the deterrent as the stockpile ages or as replacement systems are introduced. Nevertheless, the GAO statement that "the weapons laboratories face extraordinary technical challenges in successfully implementing a coherent and credible analytical method based on QMU" is without context. Without noting the significant achievements to date, the statement leaves the reader with the conclusion that success in this vital area is unlikely and efforts to certify an RRW will likely fail despite the noted successes to date. NNSA disagrees.

There is no question that improving the ability to meet future certification requirements will require further improvements in a number of areas. The development of QMU is only one aspect of the science campaign effort that includes, importantly, the underlying scientific effort to improve physical understanding and reduce uncertainties resulting from data and models. Methods for combining differing kinds of uncertainties constrained by the sparse data set from underground testing is groundbreaking research at the forefront of statistical science. Models for boost, mix, the high-pressure equation of state of plutonium, the behavior of dense plasmas, and a range of other physical phenomena require refinement. Numerical methods of computation require improvement, the 2nd axis of DARHT must be commissioned and NIF ignition must be achieved. Certification requirements are a significant consideration in decisions regarding the future of LANSCE. To state that current certification methods require refinement is to state the obvious. But it is wrong to portray the program as lacking clear goals, or technically defensible standards for success.

Success is demonstrated by the performance of a scientifically defensible QMU analysis for each required certification problem. While there can be other measures of program efficiency, there is no other comparable measure to determine whether or not the program is achieving its scientific goals.

Although the GAO report is focused on QMU, this grew out of an audit of six science efforts, including the science campaign, ASC, and the ICF program. These programs

Appendix I
Comments from the National Nuclear
Security Administration

have multiple goals and achievements beyond the specific focus of QMU although they are supportive of that goal. One of the functions of QMU, as it is further applied will be to identify research priorities within the science campaign. As the JASON QMU report cautions, however, "prioritization of efforts has to be modulated by the need to maintain expertise across the entire weapon system, and its processes. That is, a baseline of effort needs to be maintained across all activities, including those judged to be of lower priority." Of course, less effort should be put into lower-priority activities (i.e., those bearing on processes with higher margins relative to uncertainties), but there needs to be enough ongoing activity even regarding "reliable" (high-margin) processes in order to maintain expertise and to allow for the possibility of revising previous estimates of reliability (and responding to those revisions) or to address unforeseen conditions (e.g., significant findings in surveillance).

The United States has, since the inception of the Manhattan Project, relied upon world-class science to support confidence in the nation's nuclear deterrent, and is likely to continue to do so. QMU is the current framework for the application of that science base to establishing confidence. Questions and issues regarding safety, performance and reliability of the stockpile will, so far as one can foresee, continue to occur, and therefore the continued development and refinement of QMU or some follow-on certification methodology will continue to be required. Therefore, the impression that one can define certain milestones, the achievement of which will indicate that the development of QMU is finished, is misleading.

The NNSA fundamentally disagrees with the methodology of trying to measure scientific progress through an audit largely reliant on review of administrative documents and disputes some of the conclusions derived therefrom. The QMU methodology has already proven successful and is unlikely to be the source of future failings in the program. However, more detailed responses to a few of the specific findings and management recommendations, not already covered, are provided.

RESPONSE TO THE GAO FINDINGS ON MANAGEMENT OF SCIENTIFIC EFFORTS

The GAO report criticizes the management tools and methods used to administer the Science Campaign, leading one to conclude that the Science Campaign lacks clear goals and has lacked substantial achievements. In fact, the Science Campaign Program Plan has had clear statements of long-term goals that have remained largely unchanged since the inception of the Science Campaign, and important progress has been made in key areas. For instance, the campaign has achieved a vastly improved understanding of plutonium properties under extreme conditions resulting from the subcritical experiments program, and increased accuracy of plutonium equation of state data obtained from the recently commissioned JASPER experiment. Significant new insight has been gained on an important problem in understanding the energy balance in nuclear weapons. Understanding of mix sensitivities has been vastly improved, and these insights will provide direction for future experimental and modeling efforts. New materials damage

Appendix I
Comments from the National Nuclear
Security Administration

models have been developed and implemented into ASC codes and experimental data is being acquired to establish important parameters in that model. Kinetics models for high explosives performance have been developed and implemented into weapons codes.

The underlying assumption that the science campaigns should respond to and be measured by a directed set of milestones provides an incomplete picture. By way of contrast, however, the GAO correctly states, "The Primary and Secondary campaigns were established to analyze and understand the different scientific phenomena that occur in the primary and secondary stages of a nuclear weapon during detonation. As such, the Primary and Secondary campaigns are intended to support the development and implementation of the QMU methodology and to set the requirements for the computers, computer models, and experimental data needed to assess and certify the performance of nuclear weapons." While these campaigns have long-term goals towards which they are making progress they also perform required research to determine the comprehensive requirements for other elements of the program.

Before this audit had begun, the NNSA had identified that in view of the recent progress in areas such as QMU, long-term goals need to be refined and restated, and better integration across the program is required. Therefore the NNSA had begun the development of the Primary Assessment Plan referred to in the GAO report. This plan identifies key level 1 milestones that must be supported by primary certification capabilities, and indicates priorities for achieving improvements in those science areas that will be required to support those goals. A key focus is on Reliable Replacement Warhead certification. The next step will be to identify those level 2 milestones that are needed to support the long-term goals, though in the immediate future, these are unlikely to change to a significant degree from present goals.

NNSA has the following responses to some of the specific report findings:

Finding: First, the planning documents that NNSA has established for the Primary and Secondary campaigns do not adequately integrate the scientific research currently conducted that supports the development and implementation of OMU.

Response: The NNSA agrees with this statement; this is the motivation for the development of the Primary Assessment Plan and the subsequently planned Secondary Assessment Plan. In addition, NNSA will develop further guidance to the program on science integration associated with QMU.

Finding: Second, NNSA has not developed a clear, consistent set of milestones to guide the development and implementation of OMU. For example, while one key campaign plan envisions a two-stage path to complete the development of OMU by 2014, the performance measures in NNSA's fiscal year 2006 budget request call for the completion of OMU by 2010.

Appendix I
Comments from the National Nuclear
Security Administration

Response:

NNSA agrees that better integration and connectivity of milestones between various program elements would improve the communications of the importance of program goals and improve the formality of coordination of program activities, which is currently accomplished in an informal and less visible manner. This will be done in part through more careful coordination of level one and level two milestones. An NNSA Headquarters team will provide additional program guidance on science integration supporting QMU and will seek to clarify PART measures.

At the same time, the GAO analysis of the milestones shown in Table 4 of the report is not entirely accurate. The table shows the level-one milestones for Science Campaign for the period from 2007 to 2014. These are milestones are not just for QMU but for the entire science campaign, of which QMU is only a part. For instance GAO cites the FY 2014 milestone "accounting for simulation and experimental uncertainties, reassess the ability to reproduce the full underground test data sets for a representative group of nuclear tests with a consistent set of models." To meet this milestone, NNSA must have completed the development of a full set of improved physics models, including improved mix and boost models, improved plutonium damage and equation of state models, and improved models for secondary performance. These models must have been validated and incorporated into ASC codes. This also requires developing techniques, under QMU, to perform the required uncertainty analysis. The milestone anticipates success and integration of all of these factors.

The GAO also claim that NNSA lacks milestones for the development and application of QMU, but the report itself lists level-two milestones for the development of certification plans for the W76 and W88 based on QMU, milestones of national significance, which have recently been completed.

The 2010 milestone and 2014 milestone were developed for different purposes and measure progress at different times. The 2010 milestone was developed to respond to a requirement of the Office of Management and Budget (OMB) under the government-wide Performance and Rating Tool (PART) system to establish and report on a few programmatically significant long-term milestones. A list of accomplishment was developed with annual progress goals and a completion date of 2010 as directed by the OMB. The PART target describes developing QMU to the point that it can be applied to certification of a system without underground testing (e.g. LANL manufactured W88 pit). The 2014 milestone refers, however, to a more complete development and more complex application of this approach for a series of weapons tests. Therefore, saying that the OMB PART target would be completed in 2010 is a target distinct from the statement that the broader Science Campaign milestone would be completed in 2014.

Appendix I
Comments from the National Nuclear
Security Administration

Finding: Third. NNSA has not established formal requirements for conducting annual, technical reviews of the implementation of QMU at the three weapons laboratories or for certifying the completion of QMU-related milestones.

The issue of *ad hoc* reviews has been addressed in the overview. The programs at the national laboratories are reviewed on a frequent basis established to meet a wide variety of customer requirements, and QMU is integral to most of those reviews. Relevant periodic reviews include the University of California Division Review Committees, the Strategic Command Strategic Advisory Committee's Stockpile Assessment Team (SAGSAT), periodic reviews of the W76 LEP, and W88 pit certification. A recent review of the Subcritical Experiments Advisory Committee to ensure the subcriticality of the proposed Unicorn experiment was a review of the QMU methodology applied to this important safety question and noted excellent progress in the application of the QMU methodology.

The GAO cites with approval the "Predictive Science Panel" chartered under the ASC program, which is a panel of outside experts, not NNSA staff. The purview of this panel encompasses exactly those parts of both ASC and the science campaigns that are relevant to the development of tools, models and methods that support the development of predictive capabilities, and therefore QMU.

NNSA has not selected a single review process to look at overall success in the implementation of QMU but expects to continue to rely on *ad hoc* reviews.

Finding: Finally, NNSA has not established adequate performance measures to determine the progress of the laboratories in developing and implementing QMU.

Response: The NNSA has established level 1 milestones in the Primary Assessment Plan which incorporate, implicitly QMU goals. The extensive set of external reviews discussed on page 2 of this response provide ample opportunity to determine the progress in implementing QMU.

Finding: According to NNSA, very few products of the scientific campaigns involve the repetition of specific operations whose costs can be monitored effectively as a measure of performance. As a result, the best measure of progress for the scientific campaigns is through scientific review by qualified technical peers at appropriate points in the program. However, NNSA has not established any performance measures or targets for implementing QMU that require periodic scientific peer reviews or define what is meant by "appropriate" points in the program.

Response: Scientific peer reviews will be continued to evaluate progress in addressing scientific issues. One weighs the scientific information that has been developed against the problem to be solved. As stated, NNSA does have targets for accomplishing certain specific tasks, such as writing certification plans. But to have a metric or quantifiable

Appendix I
Comments from the National Nuclear
Security Administration

targets suggests that one already has an answer or enough of one that one can define a meaningful measurable outcome.

For those things that can be scheduled and usefully counted, the Science Campaign already does so. For instance the NNSA has established a detailed plan for completing the DARHT 2nd axis with well-defined milestones. NNSA tracks the operating days at LANSCE, again, because this is an important indicator of facility operating efficiency. NNSA tracks the number of experiments performed on JASPER and the costs thereof because this bears on the productivity of the facility and also is a surrogate for the rate of progress in accumulating important plutonium equation of state data. In none of these cases, however, does the metric substitute for an actual evaluation of scientific knowledge gained.

The implementation of QMU is one of those examples where it is difficult to establish a meaningful metric. NNSA chartered a review, in this case the JASON QMU review, to examine the application of QMU in specific instances, evaluate its adequacy, look at weakness and suggest future directions. A future additional review by JASON will be considered. Since QMU is implicitly evaluated in every review of the components of the science campaign, NNSA does not view it as necessary to formally state an annual QMU requirement.

In summary, NNSA believes that it has achieved substantial progress to date in developing both QMU and meeting other goals of the science campaign, through appropriate management focus and oversight. At the same time, NNSA agrees and has recognized that the growing immediacy of meeting new requirements for both the Reliable Replacement Warhead and responsive infrastructure require a reevaluation of the level of coordination and integration of goals and milestones across all NNSA programs. The completion of the Primary Assessment Plan was one step in a number of envisioned efforts to reassess priorities and improve the level of coordination.

GAO Contact and Staff Acknowledgments

GAO Contact

Gene Aloise (202) 512-3841

**Staff
Acknowledgments**

In addition to the individual named above, James Noel, Assistant Director; Jason Holliday; Keith Rhodes; Peter Ruedel; and Carol Herrnsstadt Shulman made key contributions to this report.

GAO's Mission

The Government Accountability Office, the audit, evaluation and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's Web site (www.gao.gov). Each weekday, GAO posts newly released reports, testimony, and correspondence on its Web site. To have GAO e-mail you a list of newly posted products every afternoon, go to www.gao.gov and select "Subscribe to Updates."

Order by Mail or Phone

The first copy of each printed report is free. Additional copies are \$2 each. A check or money order should be made out to the Superintendent of Documents. GAO also accepts VISA and Mastercard. Orders for 100 or more copies mailed to a single address are discounted 25 percent. Orders should be sent to:

U.S. Government Accountability Office
441 G Street NW, Room LM
Washington, D.C. 20548

To order by Phone: Voice: (202) 512-6000
TDD: (202) 512-2537
Fax: (202) 512-6061

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Web site: www.gao.gov/fraudnet/fraudnet.htm

E-mail: fraudnet@gao.gov

Automated answering system: (800) 424-5454 or (202) 512-7470

Congressional Relations

Gloria Jarmon, Managing Director, JarmonG@gao.gov (202) 512-4400
U.S. Government Accountability Office, 441 G Street NW, Room 7125
Washington, D.C. 20548

Public Affairs

Paul Anderson, Managing Director, AndersonP1@gao.gov (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, D.C. 20548