Chapter 6
Quality Assurance and Non-Nuclear Testing

6.1 *Overview*

The goals of the U.S. nuclear weapons quality assurance (QA) programs are to validate safety, ensure required reliability, and to detect, or if possible, prevent, problems from developing for each warhead-type in the stockpile. Without nuclear testing, the current stockpile of nuclear weapons must be evaluated for quality assurance solely through the use of non-nuclear testing and surveillance. The Department of Energy (DOE) Stockpile Evaluation Program (SEP) has evolved over decades, and currently provides safety validations and reliability estimates for the stockpile. It also detects problems related to safety and reliability, permitting managers to evaluate the problems and to program required actions to resolve them. The overall quality assurance program includes: laboratory tests; flight tests; other surveillance evaluations and experiments; the reported observations from the Department of Defense (DoD) and the DOE technicians who maintain the warheads; continuous evaluation for safety validation and reliability estimates; and the replacement of defective or degrading components as required.

Because of the policy restriction on nuclear testing, no new replacement warheads have been fielded for almost two decades. During that time, sustaining the U.S. nuclear deterrent required retaining warheads well beyond their originally programmed life. As the warheads in the legacy stockpile aged, the SEP detected an increasing number of problems, primarily associated with aging non-nuclear components. This led to an expanded program of refurbishments, as required for each warhead-type, and a formal process to manage it. The SEP program has been very effective for quality assurance. Even though it has been more than 15 years since the last U.S. nuclear test, approximately one dozen different warhead-types serve as the nation's nuclear deterrent, each with annual safety validations and very high reliability estimates.

Because the warheads of the legacy stockpile continue to age, the immediate future of the stockpile will reveal age-related problems unlike any other time in the past. As a part of proactive quality assurance management, the DOE has recently established a Surveillance Transformation Project (STP). Its focus is a more knowledge-based, predictive, adaptable, and cost effective evaluation program. This chapter describes the many activities associated with the quality assurance of the U.S. nuclear weapons stockpile.

6.2 The Evolution of Quality Assurance and Sampling

The Manhattan Project, which produced one test device and two war-reserve (WR) weapons that were employed to end World War II, had no formal, structured QA program. There were no safety standards or reliability requirements to be met. QA was the sum of all precautions thought of by weapons scientists and engineers and the directives of Dr. Oppenheimer and his subordinate managers. History proves that the Manhattan Project version of QA was successful in that it accomplished an extremely difficult task without a catastrophic disaster.

The first nuclear weapons required in-flight insertion (IFI) of essential nuclear components. Once assembled, the weapons had none of the modern safety features to preclude an accidental detonation. The QA focus was on ensuring the reliability of the weapons because they would not be assembled until they were near the target. In the early-1950s, as U.S. nuclear weapons capability expanded into a wider variety of delivery systems, and because of an emphasis on more rapid response times for employment, IFI became impractical.

The development of sealed-pit weapons led to requirements for nuclear detonation safety features to be built into the warheads.[1] See Chapter 5, *Nuclear Weapons Surety*, for a detailed discussion of nuclear detonation safety and safety standards. During this time, the concern for safety and reliability caused the expansion of QA activities into a program that included random sampling of approximately 100 warheads of each type, each year. Initially, this was the *New Material and Stockpile Evaluation Program (NMSEP)*. "New Material" referred to weapons and components evaluated during a warhead's development or production phase. See Chapter 2, *Life-Cycle of Nuclear Weapons*, for a description of nuclear weapon life-cycle phases. New material tests were conducted to detect and repair problems related to design and/or production processes. The random sample warheads were used for both laboratory and flight testing, and they provided an excellent sample size to calculate reliability and to stress-test the performance of key components in various extreme environments.[2] This was unaffordable for the long term,

---

[1] Sealed-pit warheads are the opposite of IFI – they are stored and transported with the nuclear components assembled into the warhead, and they require no assembly or insertion by the military operational delivery unit.

[2] One example of a factor causing various extreme environments is temperature. Components inside a warhead, in a bunker in the summer, may have to endure relatively high temperatures, sometimes exceeding 150° F, but the same components may experience cold temperatures, below -150° F, when the warhead is carried outside an aircraft at high altitude.

and within a year or two, the program was reduced to random sampling of 44 warheads of each type. This sample size was adequate to calculate reliability for each warhead-type. Within a few more years, that number was reduced to 22 per year and remained constant for approximately a decade. Over time, the random sample number was reduced to 11 per year.

In the mid-1980s, the DOE strengthened the Significant Finding Investigation (SFI) process. Any anomalous finding or suspected defect that might negatively impact weapon safety, reliability, or control is documented as an open SFI. The QA community investigates, evaluates, and resolves SFIs.

The NMSEP is a part of today's Stockpile Evaluation Program (SEP). At the national level, random sample warheads drawn from the fielded stockpile are considered to be a part of the Quality Assurance & Reliability Testing (QART) program. Under the QART program, additional efficiencies are gained by sampling and evaluating several warhead-types as a warhead "family" if they have enough key components that are identical.[3] Normally, each warhead family has 11 random samples evaluated. The sample size of 11 per year enables the QA program to meet its current goals for each warhead-type: a) to provide an annual safety validation; b) to provide a reliability estimate semi-annually; and c) to detect any problem that affects ten percent or more of the warheads of that type, with a 90 percent assurance, within two years.

Weapons drawn for surveillance sampling are returned to the NNSA Pantex Facility near Amarillo, Texas, for disassembly. Generally, of the 11 samples selected randomly by serial number, eight are used for laboratory testing and three are used for flight testing. Surveillance testing and evaluation may be conducted at Pantex or at other NNSA facilities. Certain components are physically removed from the weapon, assembled into test configurations and subjected to electrical, explosive, or other types of performance or stress testing. The condition of the weapon and its components is carefully maintained during the evaluation process. The integrity of electrical connections remains undisturbed whenever possible. Typically, one sample per warhead family per year is subjected to non-nuclear, destructive testing of its nuclear components and cannot be rebuilt. This is called a destructive test (D-test), and the specific warhead is called a D-test unit. As long as the supply of previously produced, non-nuclear components has not been exhausted and there is a military requirement, the remaining samples are rebuilt and returned to the stockpile.

[3] For example, the B61-7 and B61-11 are sampled as one family; the B61-3, B61-4, and B61-10 are one family.

6.3 Surveillance Transformation Project (STP)

Much of the current surveillance methodology is based on the original weapon evaluation programs, relying mainly on random stockpile sampling applied to flight tests, subsystem go/no-go testing, and selected component evaluations to search for design and manufacturing "birth" or aging defects. This approach gives a current snapshot of the condition of that warhead-type but provides little ability to predict future stockpile problems. The ability to predict a problem is becoming more important as the current warheads of the legacy stockpile continue to age.

In June 2006, the Director, Office of Nuclear Weapons Stockpile, NNSA, chartered a complex-wide team to integrate efforts to develop a comprehensive plan for achieving surveillance transformation. This Surveillance Transformation Project (STP) is a plan to define a roadmap to begin transformation to more knowledge-based, predictive, adaptable, and cost effective evaluation of current and future stockpile health. It sets the nuclear weapons complex on a course to transform surveillance across four major objectives:

1. *Rigorous Requirements Basis* – create a strong technical requirements basis for stockpile evaluation;
2. *Evaluating for Knowledge* – design and execute an evaluation program that responds to changing evaluation data needs over the weapon system life-cycle;
3. *Predictive Assessment* – develop the capabilities to predict the state of health of the enduring stockpile through end-of-life projections, reliability assessments, predictive performance assessments in areas beyond reliability (i.e., safety/survivability/use control/nuclear performance), and risk-based responsiveness for replacement and refurbishment decisions; and
4. *Premier Management and Operations* – create a strong program management team to make the best decisions based on defensible cost-benefit criteria.

6.4 Stockpile Laboratory Testing (SLT)

For each warhead family, the NNSA laboratory evaluation program strives to examine: each possible operational use of the warhead; potential environmental conditions; safety and use control features; and the end-to-end process required for nuclear detonation. All aspects are verified and the data to support reliability assessments are obtained. The system-level testing program also examines safety components to determine if there is any concern for the overall safety of the weapon.

Laboratory non-destructive testing can include activities such as radiography and gas sampling. Stockpile lab testing includes: fuzing mode tests; tests of environmental sensing units; trajectory sensing device tests; functioning of firing sets tests; and for weapons so equipped, permissive action link (PAL) tests and command disable function tests. The NNSA testing program emphasizes testing at the highest possible system or subsystem levels. Diversification of tests is used as necessary to address certain aspects of weapon performance under specific use conditions and with maximum realism.

Joint Integrated Laboratory Tests (JILTs) evaluate interconnected DoD and NNSA weapon components. For example, the DoD arming, fuzing, and firing mechanism would be tested in conjunction with an NNSA de-nuclearized warhead. These system-level tests are conducted at either NNSA or DoD facilities.

Normally, the nuclear explosive package from the D-test unit is destructively tested to look for any changes in dimensions or material composition. Five key components are tested: the pit; the secondary; the detonator assembly; the high explosives; and the gas bottle system. The D-test unit is not rebuilt, and is therefore not returned to the stockpile. The remainder of the samples can be reconstructed and returned to the stockpile if replacement components are available for rebuild. If components are not available for rebuild, those warheads are eliminated from the stockpile. These reductions are called *QART consumption* in the national-level stockpile planning documents.

6.5 Stockpile Flight Testing (SFT)

Flight testing of nuclear delivery systems is accomplished using warheads with inert nuclear components known as Joint Test Assemblies (JTAs). JTAs use non-fissile nuclear components that replace the fissile components in the tested weapon. This precludes any possibility that the JTA can produce a nuclear detonation. JTA flight tests are currently conducted two to four times per year. The JTAs may be either High-Fidelity JTAs (HF-JTAs) or Instrumented JTAs (I-JTAs).

HF-JTAs replicate actual WR warheads as closely as possible, with the exception that the fissile material (plutonium and highly enriched uranium) and the tritium are removed. HF-JTAs provide some data concerning the system as a whole, while I-JTAs provide more instrumented data about individual components and sub-assemblies. HF-JTAs demonstrate the functioning of the warhead in as complete a configuration as possible without a nuclear test. I-JTAs use data-recording instruments to record the in-flight performance of certain components. Normally, I-JTAs provide much more component and sub-assembly performance data than HF-JTAs. However, in order to have

these data-recording instruments embedded in the warhead, the instruments may replace selected warhead components. Therefore, any one I-JTA will have selected warhead components replaced with data-recording instruments; while another I-JTA for the same weapon-type may have a different set of warhead components replaced with other instruments. As much as possible, the data-recording instruments are designed to have the same physical dimensions (height, width, length, weight, center-of-gravity, etc.) as the components they replace.

The Non-Nuclear Assurance Program (NNAP) ensures that actual nuclear weapons are not accidentally used in flight tests in place of the JTAs. The verification process includes inspections of tamper-evident seals and other indicators in conjunction with measurements taken by radiation detection instruments. For joint tests with the DoD, the NNSA provides the test assemblies with permanent "test" markings, tamper-evident seals, signature information, and radiation test equipment.

Flight tests are conducted at various locations in the United States including: the Tonopah Test Range in Nevada; the Utah Test and Training Range in Utah and Eastern Nevada; Vandenberg Air Force Base (AFB) in California; and Eglin AFB in Florida. Stockpile Flight Tests involve JTAs built with components from WR weapons that have already experienced stockpile handling. These tests demonstrate the continued compatibility between the warhead and the delivery vehicle and verify weapons system function throughout the Stockpile-to-Target Sequence.

6.6 Safety Validation and Reliability Estimates

Safety and reliability are evaluated based on the results of the SLT, SFT, other surveillance, computer analyses, and when required, the scientific and engineering judgment of the QA experts. The safety of each warhead-type in the stockpile is validated each year to ensure that it meets established safety standards. Safety standards and certification are discussed in detail in Chapter 5, *Nuclear Weapons Surety*. Reliability is the probability that a warhead-type will function properly if employed as intended. Reliability estimates for each warhead-type are evaluated twice per year. They are estimates, not solely statistical calculations, because the sample size is not sufficiently large to preclude the possibility that scientific and engineering judgment may be included. Reliability is estimated for each mode of operation (e.g., surface bursts, laydown).