Appendix D
Nuclear and Non-Nuclear Testing

D.1 Overview

From 1945 to 1992, the United States conducted both nuclear and non-nuclear testing. After 1992, the United States developed a robust program to certify the continued safety, security, and effectiveness of nuclear weapons without the use of nuclear explosive testing.

D.2 U.S. Nuclear Testing Program

The U.S. nuclear testing program began with the *Trinity* test on July 16, 1945, at a location approximately 55 miles northwest of Alamogordo, New Mexico, now called the Trinity Site. The test confirmed the *Fat Man* implosion design weapon would function to produce a nuclear detonation and also gave the Manhattan Project scientists their first look at the effects of a nuclear detonation.

The United States conducted five additional nuclear tests between 1946 and 1948. By 1951, the United States had increased the ability to produce nuclear devices for

testing and conducted 16 nuclear tests that year. Between 1951 and 1958, the United States conducted 188 nuclear tests. Increasing the knowledge and data associated with nuclear physics and weapon design was the main purpose of most of these tests. Some tests were designed to develop nuclear weapons effects data while a few were safety experiments. These tests were a mixture of underground, aboveground, high-altitude, underwater, and above-water detonations.

In 1958, the United States instituted a self-imposed moratorium on nuclear tests. Nuclear testing resumed in 1961 and the United States conducted an average of approximately 27 tests per year over the next three decades. These included 24 joint tests with the United Kingdom;[1] 35 tests for peaceful purposes as part of the Plowshare program;[2] seven to increase the capability to detect, identify, and locate nuclear tests as part of the Vela Uniform program; four to study nuclear material dispersal in possible accident scenarios; and post-fielding tests of specific weapons. By 1992, the United States had conducted a total of 1,054 nuclear tests. In 1992, Congress passed legislation that prohibited the U.S. from conducting an underground nuclear test and led to the current policy restriction on nuclear explosive testing.

D.2.1 Early Years of the U.S. Nuclear Testing Program

The first six nuclear tests represented the infancy stage of the U.S. nuclear testing program. The first test at the Trinity Site in New Mexico provided the confidence required for an identical weapon to be employed at Nagasaki. The second and third tests, both in 1946, used identical *Fat Man* design devices to evaluate the effects of airdrop and underwater detonations in the vicinity of Bikini Island, located in the Pacific. The next three tests were conducted in 1948 on towers on the Enewetak Atoll in the Pacific, testing three different weapon designs. These first six tests began with no previous data and, by today's standards, very crude test measurement equipment and computational

[1] The United States and the United Kingdom were preparing to conduct a 25th test when President George H. W. Bush announced a moratorium on underground nuclear testing in 1992. Until that point, the nuclear relationship between the United States and the United Kingdom, as defined by the *1958 Mutual Defense Agreement*, allowed for the conduct of joint tests between the two nations. This was a great benefit to the United Kingdom—especially following the atmospheric testing moratorium of 1958—because the nation did not have the same access to land that could be used for underground nuclear testing as the United States and the Soviet Union. Following the 1992 testing moratorium, the United Kingdom formally undertook to end nuclear testing in 1995 and they ratified the *Comprehensive Nuclear-Test-Ban Treaty* in April 1998. See *Chapter 9: International Nuclear Cooperation*, for a more detailed discussion of the nuclear relationship between the United States and the United Kingdom.

[2] The Plowshare program was primarily intended to evaluate the use of nuclear detonations for constructive purposes (e.g., to produce craters for the rapid and effective creation of canals).

capabilities. Because of this, only limited amounts of scientific data were gained in each of these events.

The 188 nuclear tests conducted between 1951 and 1958 included 20 detonations above one megaton (MT), one detonation between 500 kilotons (kt) and one MT, 13 detonations between 150 and 500 kt, and 17 tests that produced zero or near-zero-yields, primarily as safety experiments. Many of these tests produced aboveground detonations, which were routine at the time. The locations for these tests included the Nevada Test Site (NTS), Enewetak Atoll, Bikini Island, the Pacific Ocean, and Nellis Air Force Range in Nevada. Some of the highest yield detonations were produced by test devices far too large to be used as deliverable weapons. For example, the *Mike* device, which produced a 10.4 MT detonation on November 1, 1952, at Enewetak, was almost seven feet in diameter, 20 feet long, and weighed 82 tons. On February 28, 1954, the *Bravo* test on Bikini Island produced a surface burst detonation of approximately 15 MT, the highest yield ever produced by the United States. The *Bravo* device was a two-stage design in a weapon-size device, using enriched lithium as fusion fuel in the secondary stage. **Figure D.1** shows the *Bravo* fireball shortly after detonation.

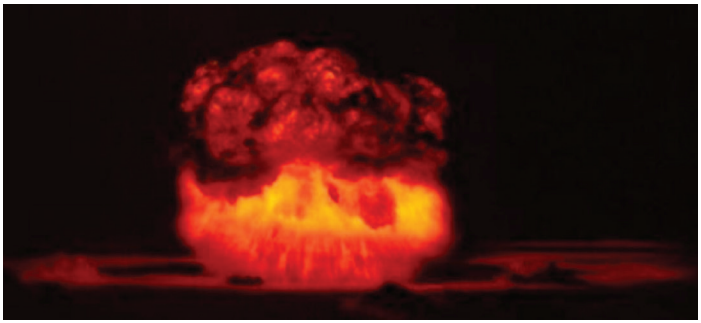

Figure D.1 Bravo Nuclear Test

During this period, as the base of scientific data grew and as sensor technology, test measurement, and diagnostic equipment became more sophisticated and more capable, the amount of data and scientific information gained from each test increased. The initial computer codes, used to model fissile material compression, fission events, and the like, were based on two-dimensional models. These computer models became more capable as the scientific data base expanded and computing technology evolved.

D.2.2 Transition to Underground Nuclear Testing

Between October 31, 1958, and September 14, 1961, the United States conducted no nuclear tests because of a self-imposed testing moratorium. The United States resumed nuclear testing on September 15, 1961 and conducted 100 tests over the next 14 months

to include underground, underwater, and aboveground detonations. These tests included nine detonations above one MT, eight detonations between 500 kt and one MT, and four detonations between 150 and 500 kt. The locations for these tests included the NTS, the vicinity of Christmas Island in the East Indian Ocean, the Pacific Ocean, Johnston Island in the Pacific, and Carlsbad, New Mexico. The last four tests of this group were conducted during a nine-day period between October 27 and November 4, 1962. These were the last U.S. nuclear tests that produced aboveground or surface burst detonations.

In compliance with the 1968 *Limited Test Ban Treaty* (LTBT), all subsequent U.S. nuclear test detonations were conducted deep underground. Initially, some thought this restriction would have a negative impact on the program to develop accurate data on the effects of nuclear weapons. The Atomic Energy Commission (AEC) and the Defense Atomic Support Agency (DASA)[3] responded with innovative ways to minimize the impact of this restriction. Through the use of long and deep horizontal tunnels, and with the development of specialized sensors and diagnostic equipment to meet the need, the effects testing program continued successfully.

In the 30 years between November 9, 1962, and September 23, 1992, the United States conducted 760 deep underground nuclear tests (UGT).[4] The locations for these tests included the NTS, Nellis Air Force Range in Nevada, and the vicinities of Fallon, Nevada; Hattiesburg, Mississippi; Amchitka, Alaska; Farmington, New Mexico; Grand Valley, Colorado; and Rifle, Colorado.[5] The tests during the period between November 1962 and April 1976 included four detonations above one MT, 14 detonations between 500 kt and one MT, and 88 detonations between 150 and 500 kt.[6] Of the 1,054 total U.S. nuclear tests, 63 had simultaneous detonations of two or more devices while 23 others had zero or near-zero yield.

Generally, a device for a weapons-related UGT (for physics research, to refine a warhead design in engineering development, or for a post-fielding test) was positioned down a deep vertical shaft in one of the NTS test areas. Informally, this type of test was called a "vertical test." Typically, a large instrumentation package would be lowered into the shaft and positioned relatively close to the device with electrical wires running back

[3] While the AEC was a forerunner organization to the current NNSA the DASA served as a precursor to the current Defense Threat Reduction Agency (DTRA).
[4] Four of these were surface experiments, without a nuclear detonation, to study plutonium scattering.
[5] After May 17, 1973, all U.S. nuclear tests were conducted at the NTS.
[6] 81 of the 90 tests are listed in the unclassified record with a yield between 20 and 200 kt.

to aboveground recording instruments. The vertical shaft was covered with earth and structural support was added to prevent the weight of the earth from crushing the instrumentation package or the device. This closed the direct opening to the surface and precluded the fireball from pushing hot radioactive gases up the shaft into the atmosphere. When the detonation occurred, the hundreds or thousands of down-hole instruments momentarily transmitted data but were almost immediately consumed in the fireball. The preparation for a vertical UGT took months and included drilling the vertical shaft and preparation of the instrumentation package, which was constructed vertically, usually within 100 meters of the shaft. The instrumentation package was typically 40 to 80 feet high, several feet in diameter, and surrounded by a temporary wooden structure. The structure would have levels, approximately seven to eight feet apart, and a temporary elevator to take technicians to the various floors to place and prepare the instruments. The test device would be lowered into the shaft, followed by the cylindrical instrument package. After the test, the ground above the detonation would often collapse into the cavity left by the cooling fireball, forming a subsidence crater on the surface directly over the test location.[7] See **Figure D.2** for a photograph of a preparation site for an underground nuclear test.

Figure D.2 Underground Nuclear Test Preparation

Generally, a UGT device for an effects test was positioned in a long, horizontal tunnel deep in the side of one of the mountains in the Yucca Mountain Range, located at the north end of the NTS. Informally, this type of test was called a "horizontal test." The tunnels were relatively large, usually more than 30 to 40 feet across, and ran several miles into the side of the mountain. Typically, the

[7] The collapse that caused the subsidence crater could occur at any time, from minutes to months, after the detonation, making the time of the collapse unpredictable.

tunnel had a small-scale railroad track running from the entrance to the deepest part of the main tunnel, which included a train to support the logistics movement of workers and equipment. The main tunnel would have many long branches, called "side-drifts," each of which could support a UGT. Instruments were positioned at various distances from the device and a huge blast door was constructed to permit the instantaneous effects of nuclear and thermal radiation, X-rays, and electromagnetic pulse to travel to instruments at greater distances but to close prior to the arrival of the blast wave. After the detonation, instruments outside the blast door would be recovered and the side-drift would be closed and sealed with a large volume of earth.

For both vertical and horizontal UGTs, the device would be prepared in a laboratory environment and transported to the test site, usually only a few days prior to the test date. On the test date, the NTS operations center would continuously monitor wind direction and speed to determine where any airborne radioactive particles would travel in the unlikely event of a "venting" incident.[8] If the wind conditions could blow venting gases to a populated area, the test was delayed until the wind conditions changed. Frequently, UGTs were delayed hours or days.

In 1974, the *Threshold Test Ban Treaty* (TTBT) was signed by the United States. The treaty would not be ratified until 1990 but, in 1976, the United States announced it would observe the treaty pending ratification. The treaty limited all future tests to a maximum yield of 150 kt. This presented a unique problem because, at the time, each of the three legs of the nuclear triad required new warheads with yields exceeding 150 kt and this compelled the weapons design community to make two major changes to nuclear weapons development.

First, new warhead designs were limited to using tested and proven secondary stage components, which provide most of the yield in high-yield weapons. The rationale for this change was that if previous testing had already determined the X-ray output required from the primary stage to ignite or drive the secondary and if testing had also determined the output of the secondary, then all that would be needed was a test to determine if the new primary would produce a yield large enough to drive the secondary. Of the

[8] Venting incidents occurred very few times during the history of U.S. underground nuclear testing. Venting occurs when a vertical UGT shaft is close enough to an unknown deep underground cave system that leads to the surface and permits the expanding fireball to push hot radioactive gases through the underground cave system to the surface and into the air. Instruments to determine geology thousands of feet underground were not precise enough to detect all possible underground caves or cavities. Venting can also occur if the blast door for a horizontal UGT is not strong enough to contain the blast wave.

1,054 U.S. nuclear tests, at least 82 had yields that exceeded 150 kt. Another 79 may have had yields exceeding 150 kt but are listed in unclassified source documents only as being between 20 to 200 kt. Many of these tests provided the data for scientists to determine the required information (e.g., ignition threshold, yield output) to certify several different secondary stage designs, which would produce yields greater than 150 kt. See **Figure D.3** for a summary of U.S. nuclear tests by yield.

Time Period	Yield					
	Zero or Near-Zero	> 0 to 150 kt	Possible > 150 kt	> 150 to 500 kt	> 500 kt to 1 MT	> 1 MT
1945–1948	0	6	0	0	0	0
1951–1958	17	137	0	13	1	20
1961–11/04/62 *	0	79	0	4	8	9
11/9/62–03/17/76 **	5	391	79	9	14	4
5/76–1992	1	257	0	0	0	0
Total:	**23**	**870**	**79**	**26**	**23**	**33**

* Last U.S. aboveground or surface detonation.
** Last U.S. detonation above 150 kt.

Grand Total: 1,054 Nuclear Tests

Figure D.3 U.S. Nuclear Tests by Yield

The second change was that, in order to test any new warhead with a yield greater than 150 kt, the warhead would have to be reconfigured to ensure it would not produce a yield in excess of 150 kt. Thus, the newest strategic warheads would not have a nuclear test, in its new configuration, for any yields above 150 kt.

By the 1980s, the U.S. nuclear testing program had evolved into a structure that categorized tests as physics research, effects, warhead development engineering, and post-fielding tests. Physics research tests contributed to the scientific knowledge and technical data associated with general weapons design principles. The effects tests contributed to the base of nuclear effects data and to testing the vulnerability of key weapons and systems to the effects of nuclear detonations. Development tests were used to test or refine key aspects of specific designs to increase yield output or to improve certain nuclear detonation safety features. Post-fielding tests were conducted

to provide stockpile confidence and ensure safety. For each warhead-type, a stockpile confidence test (SCT) was conducted between six and 12 months after fielding. This was intended to check the yield to ensure any final refinements in the design added after the last development test and any imperfections that may have resulted from the mass-production process did not corrupt the designed yield. Post-fielding tests were also used to confirm or repair safety or yield problems when non-nuclear testing, other surveillance, or computer simulation detected possible problems, especially unique abnormalities with the fissile component. If a problem was confirmed and a significant modification applied, a series of nuclear tests could be used to validate the modification to ensure that fixing one problem did not create a new issue.

D.2.3 Transition to 3-D Codes

By the early 1980s, the United States had conducted more than 970 nuclear tests, most of which had the basic purpose of increasing the scientific data associated with weapon design or refining specific designs. The national laboratories had acquired the most capable computers of the time and were expanding the computer codes to analyze, for example, fissile material compression and fission events in a three-dimensional (3-D) model. By the mid-1980s, use of 3-D codes had become routine. The 3-D codes provided more accurate estimates of what would be achieved with new designs or what might happen, for nuclear detonation safety considerations, in an abnormal environment.

With the 3-D codes, the national laboratories evaluated a broader range of abnormal environments for fielded warhead-types (e.g., the simultaneous impact of two high-velocity fragmentation pieces). This led to safety experiments and improvements that might not have otherwise occurred.[9] The increased computational modeling capability with the 3-D codes also helped scientists to refine the near-term nuclear testing program to include tests that would enhance the base of scientific knowledge and data. Each year, the results of the nuclear testing program increased the computational modeling capabilities.

D.2.4 End of Underground Nuclear Testing

In 1992, in anticipation of a potential comprehensive test ban treaty, the United States voluntarily suspended underground nuclear testing. Public Law (Pub. L.) 102-377, the legislation prohibiting U.S. underground nuclear testing, had several key elements.

[9] For example, an interim fix for one of the Army warheads was fielding a "horse-blanket" to be draped over the container to provide fragmentation/projectile shielding for transportation and storage; the ultimate fix put the shielding inside the container.

Preparation for Divider, the Last U.S. Underground Nuclear Test

These included a provision for 15 additional nuclear tests to be conducted by the end of September 1996 for the primary purpose of applying three modern safety features (enhanced nuclear detonation safety (ENDS), insensitive high explosive (IHE), and fire-resistant pit (FRP)) to those warheads planned for retention in the reduced stockpile under the proposed *Strategic Arms Reduction Treaty (START) II*.

With a limit of 15 tests within less than four years, there was no technically credible way, at the time, to certify design modifications that would incorporate any of the desired safety features into existing warhead-types. Therefore, the legislation was deemed too restrictive to achieve the objective of improving the safety of those warhead-types lacking all of the available safety enhancements and it was decided the United States would not conduct any further tests. The last U.S. underground nuclear test, *Divider*, was conducted on September 23, 1992.

The *National Defense Authorization Act* (*NDAA*) *for Fiscal Year (FY) 1994* (Pub. L. 103-160) called on 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." The Stockpile Stewardship Program, a science-based approach to ensure the preservation of competencies as mandated by the FY 1994 NDAA, has served as a substitute for underground nuclear testing since 1992. For more information on the Stockpile Stewardship Program, see *Chapter 4: U.S. Nuclear Weapons Infrastructure*.

D.3 Quality Assurance and Non-Nuclear Testing

The goals of the U.S. nuclear weapons quality assurance (QA) programs are to validate safety, ensure required reliability, and 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 QA only through the use of non-nuclear testing, surveillance, and, to the extent applicable, modeling. The DOE/NNSA Stockpile Evaluation Program (SEP) has evolved over decades and currently provides the information to support stockpile decisions and assessments of the safety, reliability, and performance of the stockpile. This program is designed to detect stockpile defects, understand margins at a component level, understand and evaluate changes (e.g., aging), and, over time, predictably assess the stockpile. The overall QA program includes laboratory tests, flight tests, component and material evaluations, other surveillance evaluations and experiments, the reported observations from DoD and DOE/NNSA

technicians who maintain the warheads, continuous evaluation for safety validation and reliability estimates, and the replacement of defective or degrading components as required.

No new replacement warheads have been fielded by the United States for over two decades. During that time, sustaining the nuclear deterrent has required the United States to retain warheads well beyond their originally designed life. As warheads in the stockpile age, the stockpile evaluation has detected an increasing number of problems, primarily ones associated with non-nuclear components. This led to an expanded program of refurbishments, as required for each warhead-type.

Because the warheads of the stockpile continue to age beyond any previous experience, it is anticipated the stockpile will reveal age-related problems unlike any other time in the past. As part of proactive QA management, the DOE/NNSA maintains a surveillance program to ensure effectiveness of the U.S. stockpile. These surveillance activities take place in multiple DOE/NNSA locations, including the Pantex Plant in Amarillo, Texas (**Figure D.4**).

Figure D.4 Pantex Plant

D.3.1 Evolution of Quality Assurance and Sampling

The Manhattan Project, which produced one test device and two war reserve (WR) weapons, *Little Boy* and *Fat Man*, employed to end World War II, had no formal, structured QA program and no safety standards or reliability requirements to be met. Rather, QA resulted from all precautions thought of by weapons scientists and engineers and the

directives of Dr. J. Robert Oppenheimer and his subordinate managers. History proves the Manhattan Project approach to quality 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, until which time the weapons were unusable. Once assembled in flight, the weapons had none of the modern safety features to preclude an accidental detonation. The early 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 the 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 to replace IFI weapons led to requirements for nuclear detonation safety features to be built into the warheads.[10] See *Chapter 7: Nuclear Surety*, for a detailed discussion of nuclear detonation safety and surety 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 called 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 *Appendix B: U.S. Nuclear Weapons Life-Cycle*, 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 provided a sample size to calculate reliability and stress-test the performance of key components in various extreme environments. This sample size was unsustainable for the long term, 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, the number was reduced to 22 per year and remained constant for approximately a decade. Over time, the random sample number was once again reduced to 11 per year to reflect fiscal and logistical realities. Each weapon system was re-evaluated with respect to the approach to sampling, accounting for the specific technical needs of each system, and new approaches to evaluation tests being implemented. As a result, some system samples were reduced from 11 per year to lower numbers.

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

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 or reliability is documented as an SFI. Weapon system engineers and surveillance engineers investigate, evaluate, and resolve SFIs.

At the national level, random sample warheads drawn from the fielded stockpile are considered part of the Surveillance Program. Under the this program, additional efficiencies are gained by sampling and evaluating several warhead-types as a warhead "family" if there are enough identical key components. Until 2006, each warhead family had 11 random samples evaluated each year under what was called the Quality Assurance and Reliability Testing (QART) program. The sample size enabled the QA program to provide an annual safety validation, supply a reliability estimate semi-annually, and sample any randomly occurring problem that was present in 10 percent or more of that warhead-type (with a 90 percent assurance, within two years).

Weapons drawn for surveillance sampling are returned to the DOE/NNSA Pantex Facility for disassembly. Generally, of the samples selected randomly by serial number, two to three are used for flight testing and the remainder are used for laboratory testing and/or component and material evaluation (CME). Surveillance testing and evaluation may be conducted at Pantex or at other DOE/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. Depending on the availability of non-nuclear components and the military requirement to maintain stockpile quantities, the remaining samples may be rebuilt and returned to the stockpile.

D.3.2 Stockpile Surveillance

The Surveillance Program is composed of the Stockpile Evaluation Program and the Enhanced Surveillance Subprogram. The SEP conducts evaluations of both the existing stockpile (stockpile returns) and new production (i.e., Retrofit Evaluation System Test Units). The Enhanced Surveillance Subprogram provides diagnostics, processes, and other tools to the SEP to enable prediction and detection of initial or age-related defects,

reliability assessments, and component and system lifetime estimates. These two program elements work closely together to execute the current Surveillance Program and develop new surveillance capabilities at the system, component, and material levels.

The evaluations conducted as part of the SEP are either system-level tests or laboratory tests. System-level testing can be high-fidelity Joint Test Assemblies (JTAs), instrumented JTAs, Weapons Evaluation Test Laboratory (WETL) testbeds, or Joint Integrated Laboratory Test (JILT) units. System-level tests may occur jointly with the Air Force or the Navy and use combinations of existing weapons and/or new production units, which are modified into JTAs. Some JTAs contain extensive telemetry instrumentation, while others contain high-fidelity mock nuclear assemblies to recreate, as closely as possible, the mass properties of WR. These JTAs are flown on the respective DoD delivery platform to gather the requisite information to assess the effectiveness and reliability of both the weapon and the launch or delivery platform and the associated crews and procedures. Stockpile laboratory tests conducted at the component level assess major assemblies and components and, ultimately, the materials that compose the components (e.g., metals, plastics, ceramics, foams, and explosives). This surveillance process enables detection and evaluation of aging trends and anomalous changes at the component or material level. The SEP consists of four elements:

- **Disassembly and Inspection**—Weapons sampled from the production lines or returned from the DoD are inspected during disassembly. Weapon disassembly is conducted in a controlled manner to identify any abnormal conditions and preserve the components for subsequent evaluations. Visual inspections during dismantlement can also provide "state-of-health" information.
- **Flight Testing**—After disassembly and inspection, selected weapons are reconfigured into JTAs and rebuilt to represent the original build to the extent possible. However, all special nuclear material (SNM) components are replaced with either surrogate materials or instrumentation. The JTA units are flown by the DoD operational command responsible for the system. JTA configurations vary from high-fidelity units, which essentially have no onboard diagnostics, to fully instrumented units, which provide detailed information on component and subsystem performance.
- **Stockpile Laboratory Testing**—Test bed configurations are built to enable prescribed function testing of single parts or subsystems using parent unit hardware from stockpile weapon returns. The majority of this testing occurs at the WETL, which is operated by Sandia National Laboratories at Pantex and involves electrical and

mechanical testing of the systems. The Air Force JILT facility, located at Hill Air Force Base in Utah, also conducts evaluations of joint test beds to obtain information regarding delivery platform-weapon interfaces.

- **Component Testing and Material Evaluation**—Components and materials from the disassembly and inspection process undergo further evaluations to assess component functionality, performance margins and trends, material behavior, and aging characteristics. The testing can involve both non-destructive evaluation techniques (e.g., radiography, ultrasonic testing, and dimensional measurements) and destructive evaluation techniques (e.g., tests of material strength and explosive performance, as well as chemical assessments).

Surveillance requirements, as determined by the national laboratories for the weapon systems, in conjunction with the Air Force and the Navy for joint testing, result in defined experiments to acquire the data that support the Surveillance Program. The national laboratories, in conjunction with the DOE/NNSA and the nuclear weapons production facilities, continually refine these requirements, based on new surveillance information, annual assessment findings, and analysis of historical information using modern assessment methodologies and computational tools.

The Enhanced Surveillance Subprogram assesses the impact of material behavior changes on weapon performance and safety. This joint science and engineering effort provides material, component, and subsystem lifetime assessments and develops predictive capabilities for early identification and assessment of stockpile aging issues. The Subprogram identifies aging issues with sufficient lead time to ensure the DOE/NNSA has the refurbishment capability and capacity in place when required. Typically, the lifetime assessments include efforts to understand basic aging mechanisms and interactions of materials in components, assemblies, and subassemblies. Accelerated aging experiments are used to obtain data beyond that available from traditional stockpile surveillance. Experiments are also used to validate broader, more age-aware models developed to support lifetime assessments and predictions pertinent to life extension programs. In addition, the subprogram provides new or improved diagnostic techniques and technologies to detect and quantify aging degradation mechanisms in the stockpile. The capabilities and knowledge gained are applied to assess and develop candidate replacement materials, through separate technology and component maturation program efforts, for future stockpile insertion.