1.0 Purpose and Background of Performance Standards

1.1 Introduction

Prior to the acceptance of a new test method for regulatory testing applications, validation studies are conducted to assess its reliability (i.e., the extent of intra- and interlaboratory reproducibility), and its relevance (i.e., the ability of the test method to correctly predict or measure the biological effect of interest) (OECD 1996, 2005; ICCVAM 1997, 2003). The purpose of performance standards is to communicate the basis by which new proprietary (i.e., copyrighted, trademarked, registered) and nonproprietary test methods have been determined to have sufficient accuracy and reliability for a specific testing purpose. These performance standards can then be used to evaluate the accuracy and reliability of other proposed test methods that are considered functionally and mechanistically similar to the accepted test method. They also allow such test methods to be evaluated with a reduced set of reference substances and tested in a minimum number of laboratories.

These test method performance standards are proposed so that modified versions of the murine local lymph node assay (LLNA) that are mechanistically and functionally similar[4] to the "traditional LLNA"[5] (ICCVAM 1999) can be effectively and efficiently evaluated for their validity by national and international validation organizations (e.g., the U.S. Interagency Coordinating Committee on the Validation of Alternative Methods [ICCVAM], the European Centre for the Validation of Alternative Methods [ECVAM], and the Japanese Center for Validation of Alternative Methods [JaCVAM]) or other organizations. The LLNA test method protocol recommended by ICCVAM (ICCVAM 1999) and accepted by U.S. regulatory agencies is consistent with procedures described in Test Guideline (TG) 429 of the Organisation for Economic Co-operation and Development (OECD; OECD 2002) and was used as the basis for development of the TG. However, the TG allows for much more procedural variation than the ICCVAM-recommended test method protocol. Therefore, the ICCVAM-recommended test method protocol is the key reference used for establishing these performance standards (**Appendix A**).[6] The test method protocol has been revised to recommend a minimum of four individual animals rather than five individual animals per group. This was based on an evaluation of data from 83 LLNA studies (275 dose groups) from six different laboratories, which indicated that a reduction in the sample size from five to four animals per group is unlikely to have a significant impact on the results of an LLNA study (**Appendix B**). This change is important since most animal-use regulations require that the minimum number of animals be used in studies. Because OECD TG 429 specifies four animals per group when pooled data are collected and five animals per group when individual animal data are collected, only pooled data have been collected in many countries.

---

[4] Components of the traditional LLNA that a mechanistically and functionally similar modified LLNA test method protocol would need to include are summarized in **Section 2.2**.

[5] The "traditional LLNA" refers to the validated ICCVAM LLNA test method protocol (ICCVAM 1999), which measures lymphocyte proliferation based on incorporation of tritiated methyl thymidine into the cells of the draining lymph nodes.

[6] **Appendix A** includes an updated version of the validated ICCVAM-recommended LLNA test method protocol (ICCVAM 1999; Dean et al. 2001), which reflects the conclusions and recommendations of an ICCVAM Independent Scientific Peer Review Panel convened in March 2008 (see http://iccvam.niehs.nih.gov/methods/immunotox/llna_PeerPanel.htm).

Modified LLNA test method protocols are expected to achieve a level of performance that is equivalent to or exceeds the accuracy and reliability of the traditional LLNA for identifying human sensitizers. All procedural modifications should be accompanied by a scientific rationale and the modified test method protocol should incorporate the essential test method components summarized in **Section 2.2** and provided in detail in **Appendix C**.

These performance standards are not proposed for evaluating other alternative test methods for measuring skin sensitization (e.g., *in vitro* methods). Additionally, these performance standards do not imply the appropriateness of performance standards for any other *in vivo* test method.

Furthermore, as more experience with the LLNA is gained and additional skin sensitization data (i.e., LLNA, guinea pig, and/or human) become available these performance standards may be updated to incorporate new information on the usefulness and limitations of the LLNA for distinguishing between sensitizers and non-sensitizers.

1.2 Elements of ICCVAM Performance Standards

Performance standards are based on an adequately validated test method and provide a basis for evaluating the comparability of a proposed test method that is mechanistically and functionally similar (ICCVAM 2003). The three elements of performance standards are:

- **Essential test method components:** These consist of essential structural, functional, and procedural elements of a validated test method that should be included in the protocol of a proposed test method that is mechanistically and functionally similar to the validated method. Essential test method components include unique characteristics of the test method, critical procedural details, and quality control measures.
- **A minimum list of reference substances:** Reference substances are used to assess the accuracy and reliability of a proposed mechanistically and functionally similar test method. These substances are a representative subset of those used to demonstrate the reliability and the accuracy of the validated test method, and are the minimum number that should be used to evaluate the performance of a proposed mechanistically and functionally similar test method.
- **Accuracy and reliability values:** These are the standards for accuracy and reliability that the proposed test method should meet or exceed when evaluated using the minimum list of reference substances.

1.3 ICCVAM Process for the Development of LLNA Performance Standards

ICCVAM established and published in 2003 the process that it follows for developing performance standards (ICCVAM 2003). ICCVAM now routinely develops draft performance standards that are proposed and considered during the ICCVAM evaluation of a new test method. However, since ICCVAM evaluated the LLNA (ICCVAM 1999) prior to establishing a process for developing performance standards, they were not developed for this test method. Accordingly, ICCVAM is now providing performance standards for the LLNA to support the development and validation of proposed modifications of the traditional LLNA.

A detailed timeline associated with the development of these performance standards is provided in **Appendix D**. ICCVAM released draft LLNA Performance Standards to the public for comment on September 12, 2007. After consideration of these comments, a revised version was made available on January 8, 2008, to the ICCVAM Independent Scientific Peer Review Panel (hereafter, "Panel") for consideration at a public meeting that was convened on March 4-6, 2008, at the Consumer Product Safety Commission Headquarters in Bethesda, MD. The revised draft LLNA Performance Standards were also made available to the public for comment before the Panel meeting, and all comments were provided to the Panel for their consideration. The Panel's conclusions and recommendations were made available to the public and to ICCVAM's Scientific Advisory Committee on Alternative Toxicological Methods (SACATM).

The Panel's report and all comments by the public and SACATM were considered by the ICCVAM Immunotoxicity Working Group and ICCVAM in preparing final LLNA performance standard recommendations for public release and submittal to U.S. Federal agencies. Performance standards adopted by U.S. Federal regulatory authorities can be provided or referenced in test guidelines. Availability of these performance standards and ICCVAM test method evaluation reports, which provide ICCVAM recommendations and a comprehensive evaluation of the usefulness and limitations of a test method, are announced in the *Federal Register*, in National Toxicology Program newsletters, and by email to ICCVAM email list members. Additional details regarding the chronology of updates to the ICCVAM LLNA performance standards and the specific changes that were made during this process can be found in **Appendix E**.

1.4 ICCVAM Development of a Performance Standard for the LLNA

1.4.1 Background on Skin Sensitization

Skin sensitization to a substance can lead to allergic contact dermatitis (ACD), a type IV hypersensitivity reaction. The development of skin sensitization occurs in two separate phases. The first phase, referred to as the induction phase, occurs when a susceptible individual is exposed topically to a sufficient quantity of a skin-sensitizing substance. Induction depends on a substance penetrating the epidermis and subsequently binding to proteins. The antigen-presenting cells in the skin (i.e., Langerhans cells) can then process the resulting hapten complex. These cells then migrate to the draining lymph nodes, where the antigen is presented to T lymphocytes, leading to their antigen-specific clonal expansion and the production of memory and effector T lymphocytes. At this point, the individual has become sensitized to the exposed substance (Basketter et al. 2003; Jowsey et al. 2006).

The second phase, referred to as the elicitation phase, occurs when the individual is exposed to the same substance at the same or different skin location. As in the induction phase, the substance penetrates the epidermis where it is processed by antigen-presenting cells. The antigen is then presented to circulating effector T lymphocytes. The T lymphocytes produce a rapid secondary immune response in the skin that can lead to ACD (Basketter et al. 2003; Jowsey et al. 2006).

ACD is a frequent occupational and environmental health problem. According to the U.S. Department of Labor Bureau of Labor Statistics, in 2005, 980 cases of ACD involved days

away from work.[7] Furthermore, ACD has also been shown to have a significant impact on quality of life in the population group affected (Hutchings et al. 2001; Skoet et al. 2003).

1.4.2 Test Methods for Assessing Skin Sensitization

There are several test methods currently recognized for evaluating skin sensitization *in vivo.* These methods are classified into two categories: adjuvant and non-adjuvant tests (see EPA 2003 for a list of acceptable test methods). Adjuvant tests use Freund's complete adjuvant (FCA) to potentiate sensitization. Examples of adjuvant tests include the guinea pig maximization test (GPMT), the Maurer optimization test, the split adjuvant test, and the FCA test. Examples of non-adjuvant tests include the Buehler test (BT), the Draize sensitization test, and the open epicutaneous test. All of these methods use the guinea pig as the test species.

For the GPMT, sensitization in guinea pigs is induced by intradermal injection of the test substance mixed with FCA at the start of the testing procedure. After 6 to 8 days, an occluded patch containing the test substance is applied to the test area and held in place with a dressing for 48 hours. After 12 to 14 days, a patch containing the test substance is applied to the test area and held in place with a dressing for 24 hours. Skin reactions (i.e., erythema, edema) are scored 24 and 48 hours after patch removal (OECD 1992).

For the BT, a test patch containing the substance is applied to the animals. Animals are exposed once a week to the test substance for 6 hours over a period of 3 weeks. Two weeks after the final treatment, a patch containing the test substance is applied for 6 hours at a location different from where the initial challenges occurred. Skin reactions (i.e., erythema and edema) are then scored 24 and 48 hours after patch removal (OECD 1992).

1.4.3 Intended Regulatory Uses for the LLNA

The LLNA can be used as a substitute for the traditional guinea pig tests (e.g., GPMT, BT),[8] where appropriate, for assessing skin sensitization. The LLNA may not be suitable for use with certain types of test materials, such as mixtures, metal compounds (particularly nickel), strong dermal irritants, and chemicals whose pharmacodynamic activity is to release dermal cytokines that cause local lymph node proliferation (e.g., certain pharmaceuticals such as imiquimod [Gaspari 2007]). Additionally, the LLNA may not be suitable for test substances that do not adhere for an acceptable length of time when applied to the dorsum of the ear during the experiment. Data to support testing of mixtures using the LLNA is currently under evaluation by ICCVAM.

1.4.4 Similarities and Differences in the Endpoints of the LLNA and Reference Skin Sensitization Test Methods

The endpoint measured in the LLNA is induction of lymphocyte proliferation, which is part of the induction phase of skin sensitization. Comparatively, the guinea pig tests described in **Section 1.4.2** involve rating skin reactions evoked by the test substance, which are part of the elicitation phase of skin sensitization. The guinea pig tests therefore allow for an assessment of the entire ACD process.

---

[7] http://www.bls.gov/IIF

[8] Of the methods listed in **Section 1.4.2**, the GPMT and BT are most widely used and are the preferred guinea pig sensitization tests as outlined in the OECD test guidelines for skin sensitization.

While the endpoints measured in the LLNA and the guinea pig test methods are different, the induction phase of skin sensitization is necessary for development of skin reactions (i.e., elicitation phase). Therefore, measurement of lymphocyte proliferation generally predicts whether the test substance will produce skin sensitization. Compared to the LLNA, which quantifies the amount of T lymphocyte proliferation, the guinea pig test methods use subjective scoring of the skin reaction (i.e., erythema, edema) observed after test substance application.