Up-and-Down Procedure (UDP)

Peer Panel Report

August 21, 2001 Meeting

1.0 INTRODUCTION

This report provides the conclusions and recommendations of an independent scientific peer review panel (Panel) evaluation of a revised version of the Up-and-Down Procedure (UDP) (July 2001). The Panel convened in a public teleconference meeting on August 21, 2001, at the National Institute of Environmental Health Science (NIEHS), Research Triangle Park, North Carolina, U.S. The Panel reviewed the following:

- The revised draft UDP, modified in response to recommendations from the July 2000 Panel meeting;
- A proposed procedure for calculating the confidence interval (CI) for the estimated LD50; and
- A software program to aid in establishing test doses, determining when to stop the test, estimating the LD50, and providing a CI for the LD50. (see **Appendix C**).

The meeting was organized by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). *Federal Register* notices relevant to the meeting include a Notice of Availability and Request for Comments (NIEHS, 2001a) and Notice and Agenda of Public Teleconference (NIEHS, 2001b).

The UDP was proposed by the U.S. Environmental Protection Agency (U.S. EPA) to ICCVAM in April 2000 as an alternate for the existing conventional LD50 test (EPA 870.1100, 1998; OECD TG 401, 1987) used to evaluate the acute oral toxicity of chemicals. A previous version of the draft UDP test guideline was reviewed by the UDP Peer Review Panel at a meeting on July 25, 2000 organized by the NICEATM and ICCVAM (Final Report Section I). The revised draft UDP reviewed on August 21, 2001 incorporated modifications made in response to the conclusions and recommendations of the Panel at the July 2000 meeting.

1.1 Objectives of the Peer Panel Evaluation

The Panel was charged with evaluating the following:

- the extent to which the revised draft UDP test guideline (July 12, 2001) addressed the Panel's recommendations at the July 25, 2000 Peer Review Panel meeting
- the appropriateness and adequacy of the proposed procedure for calculating a CI for the LD50; and
- the adequacy and consistency of the software program for use in the revised draft UDP test guideline.

1.2 Conduct of the Meeting and Reports

The UDP Peer Panel Review Meeting, which was open to the public, was conducted via teleconference on August 21, 2001 (**Appendix E-2**). The meeting began with an introduction including an overview of the ICCVAM Test Method Review Process. The Panel convened and evaluated the appropriateness and suitability of the further revised draft UDP test guideline, the approach for obtaining the CI, and the suitability of the software program. Following an opportunity for public comment, the Panel provided conclusions and adjourned. A written report, summarizing the discussions, recommendations, and conclusions from the teleconference, was provided to ICCVAM/NICEATM and is included in this final report (Final Report Section II).

2.0 REVISED DRAFT UP-AND-DOWN PROCUDURE TEST GUIDELINE

Based on the conclusions and recommendations of the Panel from their meeting in July 2000, the UDP Technical Task Force revised the test method guideline for the proposed UDP Primary and Limit Tests, deleted the UDP Supplemental Test, and included a procedure for calculating the CI for the estimated LD50. This revised draft UDP test guideline (GUIDELINE FOR THE TESTING OF CHEMICALS: Acute Oral Toxicity: Revised Up-and-Down Procedure. Draft, July 12, 2001; **Appendix C-1**) was developed by UDP Technical Task Force and submitted to ICCVAM on July 12, 2001. (Note: The slope of the dose-response curve was not addressed by the revised draft UDP test guideline.)

2.1 Panel Agreement on Guideline Revisions

The Panel concluded many of the recommended and requested changes had been appropriately considered and all members concurred with the current modifications. However, several previous recommendations appeared to have not been adequately addressed in the revised draft UDP test guideline, including the following:

- To increase flexibility and adaptability in animal use, the use of either sex or the more sensitive sex (if information is available indicating that one sex is more sensitive) should be permitted. The Panel unanimously re-affirmed this previous recommendation.

- The body weight of an animal on day 1 of dosing should be within 20% of the mean body weight of all previously dosed animals. The Panel chose to withdraw this recommendation based on the revised language included in paragraph 14 of the revised draft UDP test guideline as follows, "At the commencement of its dosing, each animal should be between 8 and 12 weeks old and its weight should fall in an interval ±20 % of the mean initial weight of all previously dosed animals" (**Appendix C-1)**.

- Additional guidance for use of pre-start data (data available before the acute toxicity test is conducted) to aid in determining the starting dose level should be included. The revised draft UDP test guideline addresses this recommendation in paragraph 4 as follows: "All available information on the test substance should be considered by the testing laboratory prior to conducting the study. Such information will include the identity and chemical structure of the substance; its physical chemical properties; the results of any other *in vitro* or *in vivo* toxicity tests on the substance or mixtures; toxicological data on structurally related substances or similar mixtures; and the anticipated use(s) of the substance. This information is useful to determine the relevance of the test for the protection of human health and the environment, and will help in the selection of an appropriate starting dose" (**Appendix C-1)**.

- Several Panel members stated this type of information was more appropriate for inclusion in a training session or guidance document, rather than a test guideline. The rationale for this recommendation was to help provide a better idea of the types of information or data to consider when selecting a starting dose level and to provide an alternative for the default starting dose level. The Panel unanimously recommended the following modification to the revised draft UDP test guideline, paragraph 4: All available information on the test substance should be considered by the testing laboratory prior to conducting the study. Such information <u>may</u> include the identity and chemical structure of the substance; its physical chemical properties; the results of any other *in vitro* or *in vivo* toxicity tests on the substance or mixtures; toxicological data on structurally related substances or similar mixtures; and the anticipated use(s) of the substance. This information is useful to determine the relevance of the test for the protection of human health and the environment. <u>This information may be valuable in selecting a dose other than the default starting dose</u>.

- The Panel unanimously re-affirmed their previous recommendation for a practicability evaluation of the revised UDP test guideline.

- A separate section in the revised UDP test guideline describing how the revised UDP Primary Test addresses reduction, refinement, and replacement of animals compared to the previous tests should be provided. The UDP Technical Task force formed the following response to this recommendation: *The Guideline significantly reduces the number of animals used in comparison to OECD Test Guideline 401, which often required at least 20 animals in a test: 1) the stopping rule limits the number of animals in a test; 2) sequential dosing introduces further efficiencies in animal use; 3) initial dosing is now set to be below the LD50, increasing the percentage of animals in which dosing levels will be sub lethal and thereby providing some reduction in pain and distress; and 4) the use of a single sex reduces the number of animals needed and minimizes the variability in the test population. Theoretically using females only could lead to an oversupply of males. However, the use of male rats in animal research greatly exceeds that of females and, thus, the preference for females in acute toxicity testing may well result in a better overall balance of the use of both genders. Importantly, the guideline contains a requirement to follow the OECD Guidance Document on the Recognition, Assessment, and Use of Clinical Signs as Humane Endpoints for Experimental Animals used in Safety Evaluation (2000) that should reduce the overall suffering of animals used in this type of toxicity test.*

- The removal of gender specific references or the addition of the acceptability to use either gender (as per the preceding recommendation) was suggested and unanimously agreed upon by the Panel (see the underlined sentences in the above paragraph). This information should be included in the revised UDP test guideline.

- In paragraph 17a of the revised draft UDP test guideline, constant concentration should be used unless there is scientific or regulatory need for using constant volume. If constant volume is used in the performance of the UDP, concentrations used should also be provided. The Panel unanimously recommended that this statement be added to the revised UDP test guideline.

2.2 Recommendations

- The use of either sex of animals or the more sensitive sex (if information is available indicating one sex is more sensitive) should be permitted.
- Additional guidance pertaining to the use of pre-start data (data available before the acute toxicity test is conducted), which may be helpful in determining the starting dose level, should be provided.
- A practicability evaluation of the revised UDP test guideline should be conducted.
- A separate section detailing how the revised UDP Primary Test addresses reduction, refinement, and replacement of animals compared to the previous tests should be included.
- The Panel continues to express concerns that sufficient explanation is not included in the revised draft UDP test guideline describing the need and use of slope and CI for risk assessment and extrapolation to low doses for any purpose.

In addition to the above recommendations, the Panel identified the following editorial recommendations for the revised draft UDP test guideline:

- Check the text for the use of both “half-log unit” and “dose progression factor of 3.2” in the same sentence.
- Check whether the sentence in paragraph 10 should read “A test dose of 2000” rather than “A test dose of up to 2000”.
- Check for inconsistency in the number of stopping criteria. Annex 3 indicates four stopping criteria, but only three are described in the text.

- Check page 12 for the requirement of supplying a slope.
- Check to ascertain whether differences truly exist in the manner in which the 2000 mg/kg limit test is conducted compared to the 5000 mg/kg limit test. One test indicates dosing one animal at a time and the other indicates dosing in pairs. If the guideline is correct as written, a sentence concerning the rationale for the difference should be included.
- Check paragraph 27 and Annex 2 for consistency. Paragraph 27 suggests increasing the progression factor if the slope is <2.5. No recommendations are made for circumstances in which the slope >2.5, although Annex 2 details such cases. If smaller dose progression factors are recommended for steep slopes, a statement of this information should be included; otherwise, Annex 2 should be amended to accommodate only shallow slopes.
- Check paragraph 36 for clarity. Parts of paragraph 36 are unclear and the reference to paragraph 39 is not helpful. Perhaps a better explanation would be "An estimate of the log of the true LD50 is given by the value of mu (μ) to maximize the likelihood L."
- Clarify statements which include "OECD" (paragraphs 8, 38 and 40 for example). There is confusion about what the documents are called and how many exist.
- Include optional clinical chemistry in paragraph 34.
- Include an explanation for the use of 5 animals in the limit test.
- Check page 16, Stopping Rule. Consider including reference to both paragraphs 5 and 28.

3.0 CONFIDENCE INTERVAL PROCEDURE

Calculation of confidence intervals (CI) provides a basis for evaluating how to incorporate test results into regulatory applications. Therefore, a CI calculation was included in previous versions of the UDP guideline (OECD 1998a and ASTM 1998). Following deletion of the proposed supplemental procedure from the previous draft Revised UDP as per recommendation by the July 2000 Panel review, another method was needed to assist the investigator using the UDP to calculate a CI for the LD50. Based on this need, the U.S. EPA developed a proposed procedure for obtaining the CI; this procedure is a statistical calculation that does not require the use of test animals beyond what is needed to estimate the LD50 (**Appendix C-2)**. Further, the procedure helps to place the estimated LD50 in a statistical context for hazard and risk assessment purposes.

The UDP Panel charged Drs. Condon, Flournoy, and Stallard (the Panel's biostatisticians) with developing the Panel's position for this section by determining the appropriateness and adequacy of the procedure for calculating a CI for use with the revised draft UDP test guideline. It was recommended that language be added to the revised UDP test guideline to specifically indicate the shortcomings, uncertainties, and limitations of the CI procedure. Further, the procedure should be modified accordingly as more is learned about the use of these types of statistical methods.

3.1 Recommendations

1. Circumstances in which the proposed method does not perform well should be stated. The addition of non-statistical language and the outlining of specific situations in which the procedure does not perform well (e.g., shallow slopes) should be included in the revised UDP test guideline and the software program documentation. To aid in this task, appropriate references as suggested by the Panel included Jennison and Turnbull, 2000; Woodroofe, 1982; Liu, 1997; and Shiryaev and Spokoiny, 2000.

2. A very strong cautionary statement concerning the use of results for extrapolation to responses at lower dose levels is needed.

3. The fact that infinite confidence bounds can be obtained by this method should be stated.

4. A stronger cautionary statement pertaining to the utilization of a starting dose at the LD50 should be provided. If the LD50 is used as the starting dose level, a much wider confidence interval is obtained than if a higher or lower starting dose were used.

5. The revised UDP test guideline should state that evaluation of this method and examination of alternative approaches, such as nonparametric methods, should be encouraged.

4.0 SOFTWARE PROGRAM

A software program was designed and made publicly available to aid in the UDP test guideline procedures, to facilitate performance of the UDP, and to mitigate its complexity for the user (**Appendix C-3**). The U.S. EPA developed the "Acute Oral Toxicity (Guideline 425) Statistical Program" (AOT425StatPgm) to perform the statistical calculations associated with the OECD GUIDELINE FOR THE TESTING OF CHEMICALS, Section 4: Health Effects Test No. 425, Acute Oral Toxicity: Up-and-Down Procedure (OECD TG 425). The program may also be used with the revised draft UDP test guideline. The AOT425StatPgm program performs the calculations required to complete the test procedure by calculating 1) the doses for the test animals, 2) when to stop dosing animals, and 3) the specified LD50 and a confidence interval for the LD50. Additionally, the U.S. EPA conducted quality assurance testing and simulation testing to assess the performance of the software program and to determine the statistical performance of the OECD TG 425 procedure under various conditions.

With the charge of determining the sufficiency of the software, the Panel unanimously agreed that the software program to accompany the UDP is adequate and consistent with the procedures in the revised draft UDP test guideline. In the future, the program may need minor revisions as related to the evaluation of this method and examination of alternative approaches, such as nonparametric methods, as recommended in Section 3.1.

5.0 References

American Society for Testing and Materials (ASTM). 1998. Standard Test Method for Estimating Acute Oral Toxicity in Rats. ASTM E1163-98. In: Annual Book of ASTM Standards, Philadelphia.

Jennison, C. and B.W. Turnbull. 2000. Group Sequential Methods with Applications to Clinical Trials. Chapman & Hall/CRC: Boca Raton, FL.

Liu, A. 1997. On the MLE for the drift of Brownian motion following a symmetric SPRT. Comm. Statist.-Theory Meth. 26: 977-989.

National Institute of Environmental Health Sciences (NIEHS). 2001a. National Toxicology Program: The Revised Draft Up-and-Down Procedure for Assessing Acute Oral Toxicity. Notice of Availability and Request for Public Comments. 66 FR 33550. June 22, 2001.

NIEHS. 2001b. National Toxicology Program: The Revised Draft Up-and-Down Procedure for Assessing Acute Oral Toxicity. Notice of Peer Review Meeting. 66 FR 36294. July 11, 2001.

Organisation for Economic Co-operation and Development (OECD). 2000. Guidance Document on the Recognition, Assessment, and Use of Clinical Signs as Humane Endpoints for Experimental Animals used in Safety Evaluation, OECD Environmental Health and Safety Publications, Series on Testing and Assessment, No. 19. OECD, Paris. 44 pp. Available: http://www.oecd.org//ehs/test/monos.htm.

OECD. 1998a. OECD Guideline for Testing Chemicals 425: Acute Oral Toxicity – Up-and-Down Procedure. OECD, Paris. [updated in 2000]. Available: http://www.oecd.org/ehs/test/health.htm.

OECD. 1987. OECD Guideline for Testing Chemicals 401: Acute Oral Toxicity. OECD, Paris.

Shiryaev, A.N and V.G. Spokoiny. 2000. Statistical Experiments and Decisions. Statistical inference for autoregressive models of the first order asymptotic theory. vol 8, chapter 5. World Scientific Publ., London, Singapore.

U.S. EPA. Health Effects Test Guidelines, OPPTS 870.1100, Acute Oral Toxicity. Washington, DC: U.S. Environmental Protection Agency, 1998. Available: http://www.epa.gov/docs/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Series/ [cited September 12, 2001].

Woodroofe, M. 1982. Nonlinear Renewal Theory in Sequential Analysis. BMS-NSF Regional Conference Series in Applied Mathematics 39, Section 1.3. SIAM.