PREAMBLE

1. The purpose and intent of this retrospective performance assessment document is to review the history of the developmental neurotoxicity (DNT) test method, and to demonstrate the extensive scientific efforts, including basic neurotoxicology research, inter-laboratory collaborative studies, expert workshops and validation studies, which form the foundation for this testing paradigm. The relevance, applicability and use of the DNT study in human health risk assessment is also reviewed. These considerations address the historical performance of the DNT study, in support of developing an OECD DNT Test Guideline (TG 426), that satisfies current OECD validation criteria.

2. In June 2005, the Joint Meeting declassified Guidance Document No.34 (GD34) on the “Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment” (OECD 2005a). GD34 provides guidance on issues related to the validation of new or updated test methods. It was drafted by representatives of OECD member countries, based on advice from member countries and OECD stakeholders and comprises the OECD principles for validation and for regulatory acceptance of test methods.

3. The GD34 is based on the so called “Solna Principles” for validation and regulatory acceptance. During the development of GD34 it was recognized by the Experts involved and the Working Group of National Coordinators of the Test Guidelines Programme (WNT) that the rigorous principles developed at the Solna workshop may sometimes be too stringent to meet the regulatory needs in member countries. Therefore, a section was added (paragraph 13 of the GD34) that emphasizes the importance of flexibility and adaptability in the validation process without compromising the scientific rigour:

"Given the continuing increase in the numbers and types of test methods being developed for varying purposes, the validation process should be flexible and adaptable. The extent to which these validation principles are addressed will vary with the purpose, nature, and proposed use of the test method. There are differences between in vivo assays and in vitro or ex vivo assays which should be considered in applying the validation principles. In general, the closer the linkage between the effect measured and the toxicological effect of interest, the easier it will be to establish the relevance of the assay. The more closely a test measures/observes (an) effect(s) identical to the effect(s) of concern, the greater the confidence that the test will accurately characterize or model the effect in the target species of concern.”…

4. The Developmental Neurotoxicity (DNT) study has been developed by the US EPA and has been subjected to numerous validation studies, intense peer reviews and expert judgments over the years. The US EPA has deemed the method validated for its regulatory purposes and as described by this retrospective assessment document extensive supportive material for the relevance, reliability and overall performance
of the DNT study is available. Until today, only the US EPA DNT study has been available for testing laboratories and the new TG 426 will fill a regulatory gap in OECD member countries.

5. The Expert Consultation Meeting in Tokyo, 24-26 May, 2005 (OECD 2005b) was faced with the task to make the final revisions to the draft Test Guideline 426 on Developmental Neurotoxicity, in response to the comments received after the 2nd round for comments in 2003. Based on the extensive supportive material for the performance of the method, the DNT was considered relevant and reliable for its specific regulatory purpose and use. However, the meeting emphasized that the information available on the performance of the test should be brought forward to the WNT as a supportive retrospective performance assessment document to the TG 426. It should subsequently be attached to the draft TG 426 when a 3rd revised version is circulated to the WNT for comments. This document encapsulates the enormous work that has been performed in the development of the DNT study and provides the rationale for the regulatory acceptance of the DNT as a new OECD Test Guideline.

6. The US EPA DNT guideline (OPPTS 870.6300), the prototype for TG 426, was founded upon an extensive scientific data base, including inter-laboratory validation studies, such as the Collaborative Behavioral Teratology Study, which was conducted in the mid-1980s. A separate group of experts at the Williamsburg Workshop (Kimmel et al., 1990) agreed that the methods in the DNT study are sensitive to known human developmental neurotoxicants. An Expert Consultation Meeting conducted in 2000 (OECD, 2003), discussed issues on validation, especially of individual test components versus the whole DNT test method. In doing so, they reviewed the extensive history of international validation, peer review and evaluation of DNT methods contained in the public record. Experts agreed that individual assays of the DNT test method have been shown to be relevant, reliable and sensitive, and there was agreement that there is extensive information available demonstrating the validity of individual components of the DNT test method.

INTRODUCTION

7. Developmental neurotoxicology evolved from the fields of neurotoxicology and developmental toxicology, through an extensive history of scientific research and regulatory consideration. Developmental toxicity has been defined in the OECD Guidance Document on Mammalian Reproductive Toxicity Testing and Assessment (GD 43) as

    “Any effect which interferes with normal development of the conceptus, either before or after birth, and resulting from exposure of either parent prior to conception, or exposure of the developing offspring during prenatal development, or post-natally, to the time of sexual maturation. These effects can be manifested at any point in the life-span of the organism. The major manifestations of developmental toxicity include: death of the developing organism; structural abnormality; altered growth, and; functional deficiency.”

8. The developmental neurotoxicity study is a specialized type of developmental toxicity study. Developmental neurotoxicity studies are unique among guideline toxicology studies in that they are designed to screen for adverse effects of pre- and post-natal exposure on the development and function of the nervous system and to provide dose-response characterizations of those outcomes.

9. Developmental neurotoxicity studies as described in the OECD TG 426 require administration of the test substance during gestation and lactation. Cohorts of offspring are randomly selected from control and treated litters for evaluations of gross neurological and behavioral abnormalities during postnatal development and adulthood. These include assessments of physical development, behavioral ontogeny, motor
activity, motor and sensory function, learning and memory, and post-mortem evaluation of brain weights and neuropathology.

HISTORY OF DNT TEST GUIDELINE DEVELOPMENT

10. The evolution of developmental neurotoxicity studies has its roots in scientific publications that began to appear in the early 1960s; the science has continued to develop over the past four decades. An extensive scientific literature exists of studies evaluating the potential for physical, pharmaceutical, and environmental agents to affect the development and function of the nervous system after prenatal and early postnatal exposure. This body of information, which provides a strong foundation for guideline development, implementation, and validation, is summarized in Table 1.

Table 1. Historical contributions to the Developmental Neurotoxicity Guideline

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
<th>Summary</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960's-1980's</td>
<td>Published research on developmental neurotoxicity and behavioural testing</td>
<td>Evidence that developmental exposure to chemicals and drugs can alter behavioural functioning in young and adult animals.</td>
<td>Butcher 1985</td>
</tr>
<tr>
<td>1978-84</td>
<td>Collaborative Behavioural Teratology Study (CBTS)</td>
<td>Study to examine intra- and inter-laboratory reliability and sensitivity of behavioural test methods.</td>
<td>Buelke-Sam et al., 1985 a</td>
</tr>
<tr>
<td>1989</td>
<td>Williamsburg Workshop</td>
<td>Workshop to evaluate the qualitative and quantitative comparability of animal and human data for developmental neurotoxicity.</td>
<td>Kimmel et al., 1990; Francis et al, 1990 a</td>
</tr>
<tr>
<td>1995</td>
<td>International Programme on Chemical Safety (IPCS)</td>
<td>Inter-laboratory study using neurotoxic chemicals to evaluate tests validity, reliability and measurement variability.</td>
<td>MacPhail et al, 1997; Catalano et al, 1997; Tilson et al, 1997</td>
</tr>
<tr>
<td>2000</td>
<td>International Life Sciences Institute workshop on developmental neurotoxicity testing</td>
<td>Workshop to review EPA DNT behavioural test methods, pharmacokinetics and neuropathology.</td>
<td>Milesion and Ferenc 2001; Cory-Slechta et al., 2001; Dorman et al., 2001; Garman et al., 2001</td>
</tr>
<tr>
<td>2003</td>
<td>Japanese Inter-laboratory Study</td>
<td>Inter-laboratory study using neurotoxic chemicals to determine sensitivity of behavioural measures.</td>
<td>Okazaki et al., 2003</td>
</tr>
<tr>
<td>2003 (Sept)</td>
<td>Behavioral Test Methods Workshop</td>
<td>Expert Workshop to address design, conduct and analysis of behavioural tests for neurotoxicity evaluation.</td>
<td>Slikker et al 2005</td>
</tr>
</tbody>
</table>

a Additional citations are detailed below.

11. Table 2 provides a brief historical summary of EPA and OECD DNT guideline development. While prenatal developmental toxicity test guidelines have existed for some time (e.g., OECD, 1983), the first regulatory protocol specifically designed to evaluate developmental neurotoxicity was developed and implemented by the US Environmental Protection Agency (EPA) in support of hazard evaluation for
specific solvents (US EPA, 1986). The EPA toxicology testing guidelines were developed by the Office of Toxic Substances (since renamed the Office of Pollution Prevention and Toxics [OPPT]) and the Office of Pesticide Programs (OPP), and first proposed and published for public review and comment in the US in 1986. The DNT guideline was finalized in 1991 (§83-6; US EPA, 1991). In 1998, it was revised (OPPTS 870.6300; US EPA, 1998) as part of a broader US effort to harmonize testing guidelines across OPPT and OPP, and with OECD.

12. As illustrated in Table 2, OECD initiated the development of TG 426 following the recommendations of the OECD Working Group on Reproduction and Developmental Toxicity in Copenhagen in 1995. The first draft of TG 426 was prepared following a 1996 Expert Consultation Meeting. While using the US EPA developmental neurotoxicity guideline as the basis for design of developmental neurotoxicity studies, TG 426 addressed a number of important issues and incorporated improvements recommended at the expert consultation meeting in 1996. The draft TG 426 was distributed to National Coordinators for comment in 1998, and significant technical issues that were identified by this review were further discussed at an Expert Consultation Meeting in Washington in 2000 (OECD, 2003). A revised draft was then circulated to National Experts for review and comment. Comments from member countries were addressed at a 2005 Expert Consultation meeting in Tokyo (OECD, 2005b).

**Table 2. History of the Developmental Neurotoxicity Guideline**

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>US EPA OPPT published first draft DNT protocol for peer review and public comment</td>
<td>US EPA, 1986</td>
</tr>
<tr>
<td>1995</td>
<td>OECD Working Group on Reproduction and Developmental Toxicity (Copenhagen) recommended development of OECD Developmental Neurotoxicity Test Guideline</td>
<td>OECD, 1995a</td>
</tr>
<tr>
<td>1996</td>
<td>OECD Expert Consultation Meeting (Copenhagen) provided recommendations for development of Draft OECD 426</td>
<td>OECD, 1996</td>
</tr>
<tr>
<td>2000</td>
<td>OECD Expert Consultation meeting (Washington) held to review technical issues</td>
<td>OECD, 2003</td>
</tr>
<tr>
<td>2003</td>
<td>Draft TG 426 submitted to National Coordinators for expert review and comment</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>OECD Expert Consultation Meeting (Tokyo) convened to respond to remaining comments on Draft TG 426</td>
<td>OECD 2005b</td>
</tr>
</tbody>
</table>

13. In the context of toxicological screening and testing to support human health risk assessment and chemical regulatory activities, the DNT study fills an information requirement that is not satisfied by other OECD Test Guidelines. Notably, it is the only Test Guideline that includes functional, behavioral, and anatomical evaluations of the nervous system at multiple time points, in test subjects that were exposed to
test substance during critical pre- and early postnatal periods of nervous system development. This test method has been used extensively in the past two decades on a wide variety of chemicals (Table 3).

Table 3. Number of chemicals studied using the EPA DNT Guideline or draft OECD 426 Guideline

<table>
<thead>
<tr>
<th>Chemical class</th>
<th>Number of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industrial Chemicals</td>
<td>8</td>
</tr>
<tr>
<td>Miscellaneous Agents*</td>
<td>4</td>
</tr>
<tr>
<td>Pharmaceuticals</td>
<td>3</td>
</tr>
<tr>
<td>Pesticides</td>
<td>73</td>
</tr>
<tr>
<td>Positive Control Chemicals</td>
<td>15</td>
</tr>
<tr>
<td>Solvents</td>
<td>7</td>
</tr>
</tbody>
</table>

* Food additives, cigarette smoke, dietary restriction, and maternal separation

SCIENTIFIC BASIS OF DNT GUIDELINE

14. DNT study methodology has been extensively reviewed and evaluated over the last 25 years. This has included the conduct of a number of meetings and collaborative studies involving experts from academic, industry, regulatory and public interest groups. Pivotal influences and key events in the history of the development of the DNT test guideline (Table 2) include both research on test methods development and efforts to characterize and document the sensitivity, reliability, and performance of the test methods. The development of test methods in neurotoxicology includes a long history of intra-laboratory efforts to determine the sensitivity and reliability of the test methods now used in the DNT study design. In the 1970’s Butcher and colleagues began publishing a series of papers in which rats were developmentally exposed to a variety of xenobiotics and subsequently tested during postnatal development using a battery of neurobehavioral tests (e.g., Butcher and Vorhees, 1979; Vorhees et al., 1979). At this same time Tilson and colleagues began efforts using behavioral and histological batteries, focused on sensory and motor function, in adult rodents exposed to a wide variety of neurotoxicants (e.g., Tilson et al., 1979; Pryor et al., 1983). A large body of research has provided an immense database on the ability of the functional observational battery to detect and characterize the effects of drugs and environmental chemicals (c.f., Irwin, 1968; Gad, 1982; Moser et al., 1988). This work was important in determining which methods would be most suitable for screening for neurotoxicity and developmental neurotoxicity. This early work was followed by a wide-ranging effort to understand the specificity of these test methods and the impact of both organismal and experimental factors (e.g., noise, species, strain, gender, test history) (cf., Gerber and O’Shaughnessy, 1983; Spencer et al., 1993; MacPhail et al., 1989; Levine and Butcher, 1990). Clearly, the literature is too large to properly review herein. However, the result of over 30 years of work in this area is a consensus opinion of neurotoxicologists that proper use and interpretation of the data derived from these test methods provides unique insight into the impact of xenobiotics on the developing and adult nervous system.

15. The development of test methods in neurotoxicology also includes a long history of endeavors to characterize the inter-laboratory reliability and sensitivity of the test methods now used in the DNT study design. In 1979, Butcher and colleagues published a seminal paper comparing a learning and retention method among three laboratories (Butcher et al., 1979). This was followed by the Collaborative Behavioral Teratology Study (CBTS) (Buelke-Sam et al., 1985), and the “Williamsburg Workshop” on Qualitative and Quantitative Comparability of Human and Animal Developmental Neurotoxicity (Kimmel et al., 1990). These efforts addressed various aspects of DNT study design and conduct, providing a sound scientific basis for the test method and its use in hazard evaluation. Since the 1991 publication of the US EPA DNT guideline (US EPA, 1991), there has been a continued scientific effort to review and update methodologies, for neurotoxicology in general and for developmental neurotoxicology in particular. Examples of such reviews include the IPCS collaborative study on neurobehavioral screening.
methodologies (MacPhail et al., 1997), an International Life Sciences Institute (ILSI) Risk Sciences Institute (RSI) workshop on Developmental Neurotoxicity and Risk Assessment held in 2000 (Mileson and Ferenc, 2001), a collaborative study on neurobehavioural screening in eleven Japanese laboratories (Okazaki et al., 2003), and the 2003 Behavioral Test Methods Workshop (Slikker et al., 2005). Descriptions of each of these efforts and their contributions to the scientific basis for DNT testing follow.

16. **The Collaborative Behavioral Teratology Study (CBTS)** - Several of the test procedures developed in early behavioral teratology studies underwent validation in a large inter-laboratory effort, the CBTS (Buelke-Sam et al., 1985), which compared performance of a standardized behavior test methodology in six different laboratories after *in utero* and lactational exposure to two known neurotoxicants, methylmercury and amphetamine. Conducted from 1978 to 1984, and at the time, it was the largest study ever undertaken to examine the intra- and inter-laboratory reliability and sensitivity of several behavioral test methods. The study also examined the effects of a number of other variables, including which tests had been administered to the animals, whether pups from the same litter responded more or less like their litter mates than pups from other litters, and the effects of pup gender on response to testing. The results of the CBTS were published in the peer-reviewed literature, and included descriptions of the background and overview (Kimmel and Buelke-Sam, 1985), protocol and test procedures (Adams et al., 1985c), data entry and test systems (Adams et al., 1985a), preliminary research (Adams et al., 1985b), statistical approach (Nelson et al., 1985), results (Buelke-Sam et al., 1985), and implications, current applications, and future directions (Kimmel et al., 1985). The study showed that replicability of data among laboratories using a standardized protocol was excellent, and that both positive effects (e.g., with methylmercury exposure), and the lack of effects (e.g., after low-level amphetamine exposure) were replicable. The CBTS also demonstrated that the DNT test procedures were sufficiently sensitive; no more than a 5-20% change from control values was required to detect an effect.

17. **The European Inter-Laboratory Collaborative Study** - In the 1980’s, the European Inter-laboratory Study group on Behavioural Teratology conducted an inter-laboratory study of behavioral test methods (Elsner et al., 1986). Three laboratories, one each from industry, academia and government, collaborated to examine the sensitivity and applicability of behavioral methods for routine toxicological testing. Results from animals perinatally exposed to methyl mercury indicated that behavioral tests were sensitive and that automated procedures and measures aimed at specific functional capacities were more sensitive than non-specific behavioural measures (Elsner et al., 1986, 1988).

18. **The Williamsburg Workshop** - In 1989, the US EPA held a workshop on the Qualitative and Quantitative Comparability of Human and Animal Developmental Neurotoxicity (also known as the “Williamsburg Workshop”) to provide scientific input into DNT protocol design and to evaluate its appropriateness for use in risk assessment (Kimmel et al., 1990). Expert scientists from government, industry, public interest groups, and academia reviewed a range of representative chemicals and environmental exposures including: drugs (cannabis, cocaine, methadone, and phencyclidine) (Hutchings, 1990), ethanol (Driscoll et al., 1990), the anticonvulsant phenytoin (Adams et al., 1990), as well as environmental contaminants such as methyl mercury (Burbacher et al., 1990), lead (Davis et al., 1990), polychlorinated biphenyls (Tilson et al., 1990), and ionizing radiation (Schull et al., 1990). Based on data available for these known human developmental neurotoxicants, the workshop participants concluded that DNT methodologies were adequate for detecting developmental neurotoxicity. A number of specific issues directly relevant to design and usefulness of DNT studies were extensively evaluated by participants: 1) the comparability of measures of developmental neurotoxicity in humans and laboratory animals (Stanton and Spear, 1990), 2) testing methods in developmental neurotoxicity for use in human risk assessment (Buelke-Sam and Mactutus, 1990), 3) weight of evidence and quantitative evaluation of developmental neurotoxicity data (Tyl and Sette, 1990), and 4) triggers for developmental neurotoxicity testing (Levine and Butcher, 1990). In addition, the participants were asked to address the relationship between biological endpoints specified by DNT guidelines and adverse findings observed in humans following exposure to the
developmental neurotoxic agents under consideration. A major conclusion of the workshop was that the DNT protocol would have identified each of the agents presented at the workshop as a potential developmental neurotoxicant (Francis et al., 1990), although the critical effects, and the dose at which the effects were observed, could vary across species. The predictive power of DNT guideline studies was attributed largely to the scope of neurobehavioral and neuropathological tests used that can evaluate neurological functions across multiple domains (i.e., sensory, motivational/arousal, cognitive, and motor). The laboratory animal serves as an adequate surrogate for humans because many of the biological and behavioral mechanisms underlying these neurological functions are shared between humans and laboratory animals.

19. **The International Programe on Chemical Safety (IPCS) Study** - The IPCS collaborative study, was an inter-laboratory evaluation of neurobehavioral screening methodologies used adult and developmental neurotoxicity studies (MacPhail et al., 1997). A total of 8 laboratories participated in the study, using 7 neurotoxic positive control chemicals (triethyl tin, acrylamide, parathion, p,p'-DDT, toluene, N,N'-methylene bis-acrylamide, and lead acetate) in adult male rats (Moser et al., 1997). A principal focus of the study was to examine the amount of variability that was likely to occur with the test methods, and to explore the reasons that differences occurred. The overall conclusion of this extensive study was general “agreement across laboratories in terms of their ability to detect dose-related changes in behavioral endpoints with prototypic neurotoxic agents” (Catalano et al., 1997). The study results were also reviewed at a workshop held in 1995 in Capri, Italy (Tilson et al., 1997), and were presented at the 1996 meeting of the Society of Toxicology.

20. **The ILSI workshop on DNT testing** - Enhancements to the published US EPA DNT guideline method are included in the OECD TG 426, and were implemented by the EPA Office of Pesticide Programs (OPP) when it issued Data Call In (DCI) notices for organophosphate pesticides with tolerances in 1999. Specifically, these enhancements included extension of the offspring dosing period through to the age of weaning, demonstrating that the pups are receiving the test substance when only the mother is treated (or adjusting the study protocol to ensure that this occurs), increasing the number of offspring evaluated neuropathologically, and collecting chemical class specific biomarker data. The extension of the dosing period during the lactation period raised several issues, specifically in the areas of pharmacokinetic or toxicokinetic data needs, behavioral testing, and neuropathological evaluation. To address these topics, the International Life Sciences Institute (ILSI), under a cooperative agreement with EPA, established a working group of scientists from government, industry, and academia, to assemble and evaluate the available science. The conclusions of the working group were presented in a public workshop held in Washington, DC on October 24-25, 2000, and were published in the peer reviewed literature in 2001 (Mileson and Ferenc, 2001; Cory-Slechta et al., 2001; Dorman et al., 2001; Garman et al., 2001). Overall, the working group agreed that the current DNT test protocol was based upon solid scientific principles and experience, that there were opportunities to revise and improve some aspects of the US EPA guideline study, and that further research will be valuable in providing the scientific basis for future revisions of TG 426. Further consideration of methodological issues related to the conduct of the DNT study include an ILSI workshop on the direct dosing of preweaning mammals. This workshop culminated in a monograph on direct dosing that has broad application to study design for many areas of research, e.g., pharmaceuticals, environmental pollutants, academic research, etc. (Zoetis and Walls, 2003; Moser et al., 2005).

21. **The Japanese Inter-laboratory Study** - An inter-laboratory evaluation of neurobehavioral screening methodologies (used in DNT studies as well as adult neurotoxicity studies) was carried out by laboratories in Japan (Okazaki et al., 2003). The study focused on examining technical problems in evaluating neurotoxic potential of chemicals using a common testing protocol. A total of eleven laboratories conducted a variety of neurobehavioural tests on rats after either acute or repeated (28-day) exposure to acrylamide or 3,3'-iminodipropionitrile. The overall conclusion of this study was that there
was general agreement that all laboratories detected neurotoxicity of both chemicals. The reports pointed out inter-laboratory differences and concluded that it is most important to standardize the methods and criteria, and improve observers’ skills.

22. **The Behavioral Test Methods Workshop** - In 2003, a workshop on behavioral testing was conducted in order to discuss experimental procedures and practices that could help enhance the utility of behavioral data as a reliable index of neurotoxicity and in the safety evaluation of chemical substances. Workshop participants included individuals from all sectors of the neuroscience community, include academia, government, testing laboratories, and industry. The overall conclusions from the workshop were that consensus can be reached on the fundamentals of behavioral assessment, and that aspects of behavioral assessment, including experimental design, test method selection, training of technical staff, validation, control of confounding factors, data variability, data analysis and data interpretation should be carefully considered in the planning and conduct of behavioral safety assessment (Slikker et al., 2005).

23. In summary, the scientific basis of the DNT test method has been subjected to an extensive history of international validation, peer review and evaluation which is contained in the public record. Through the various collaborative efforts and workshops that have been conducted, a number of important conclusions have been drawn. The individual test methods utilized in the DNT study have been found to be highly relevant and based upon solid scientific principles and experience. Utilizing exposures to known human developmental neurotoxicants, the DNT study has been shown to adequately identify the potential for adverse effects of chemical exposure on neurological development. The intra- and inter-laboratory reproducibility, reliability and sensitivity of the DNT test method has been demonstrated, utilizing a spectrum of test substances.

**VALUE AND USE OF THE DNT IN RISK ASSESSMENT**

24. There is a clear regulatory need for DNT testing to support risk assessments in OECD member countries. Many pesticides and other chemicals are known to affect the nervous system, and there are concerns regarding the potential for developmental neurotoxicity following early life exposures to these substances (NAS, 1993). This is particularly important since the unique behaviors and activities of children place them at greater risk for increased exposure to xenobiotics by multiple routes (Weiss et al., 2004). The call for a more rigorous assessment of the potential for developmental neurotoxicity has been issued by scientists from multiple and diverse sectors with an interest in public health protection, e.g., academia, government, non-government organizations (NGOs), and public interest groups.

25. An examination of the historical and potential uses of the DNT study in risk assessment is critical to an overall evaluation of its value in protecting human health. At this point in time, the largest collection of DNT guideline studies resides with the US EPA OPP, since this regulatory body has conducted a concerted effort to obtain information on developmental neurotoxicity for specific chemicals to satisfy the mandates of the 1996 US Food Quality Protection Act (FQPA). US EPA has furthermore engaged in a continuous, on-going scientific analysis and discourse regarding the conduct of DNT studies, the interpretation of the data from these studies, and their regulatory impact.

26. A review of twelve developmental neurotoxicity studies evaluated by the EPA Office of Pollution Pesticides and Toxic Substance (OPPTS) in support of the registration and/or use of 9 pesticides and 3 solvents was presented to a FIFRA (Federal Insecticide, Fungicide and Rodenticide Act) Scientific Advisory Panel (SAP) in 1998 (Makris et al., 1998). For the 9 pesticides examined, the EPA analyses concluded that the No-Observed-Effect-Level (NOEL) for DNT was lower than that of the fetal NOEL from the prenatal developmental toxicity study (TG 414; OECD 2001) for eight of the nine pesticides tested, and demonstrated an equivalent dose for one. The offspring NOEL for the DNT study was lower than the offspring NOEL for the reproduction study (TG 415; OECD 1983) for six of the nine pesticides,
and equivalent for one. Overall, in two of nine cases, the NOEL for DNT was lower than or equal to that for any adult or offspring endpoint from the prenatal developmental, reproduction, or neurotoxicity (TG 424; 1997) studies. Even though limited by the paucity of DNT studies available at that time, this review indicated that the DNT study includes unique endpoints which are not examined in any other Test Guideline, thereby enabling detection of neurobehavioral and neuropathological effects in offspring following exposure during sensitive periods of neurological development. Therefore, the DNT study, when present in a chemical data base, is often identified as a sensitive study and an important source of quantitative and qualitative information for risk assessment.

27. At the same SAP meeting in 1998, the Panel reviewed the use of the DNT study in risk assessment and agreed that DNT results are appropriate for use to support acute and chronic dietary risk assessments and short- and intermediate-term occupational and residential risk assessments for pesticides (FIFRA SAP, 1999). DNT endpoints have been shown to be the determining factor in the selection of endpoints and doses for risk assessment for some chemicals for which these data have been required by the US EPA. As might be expected of a study that utilizes short term exposures (approximately 25 to 40 days) during development, where a single exposure during a critical period may result in developmental insult (Rice and Barone, 2000), the predominant use of the DNT study in pesticide risk assessment has been for acute (single dose) reference doses (RfD), and for short-term (1-30 days) and intermediate-term (1-6 months) non-occupational exposures, which are especially applicable to risk assessments for children. While there is potentially a more limited use of the DNT study for chronic risk assessment (i.e., in calculating a chronic RfD for lifetime exposure to a toxicant), the DNT study has also been used for this exposure scenario when it has been shown to be the most sensitive study in the toxicology data base.

28. Since 1998, the data base of available DNT studies has increased substantially. By early 2006, approximately 114 DNT studies have been completed using either the EPA guideline or the draft OECD guidelines (Table 4). A few of these studies did not include all the endpoints recommended by EPA or OECD guideline. For example, 1,1,1,-trichloroethane was tested under a consent agreement which allowed for the addition of extensive neurophysiological testing in lieu of motor activity (USEPA, 1989). Others were conducted prior to the adoption of the early guidelines, and therefore contained limited assessments (e.g., DEET), or tested with a slightly modified protocol used for some pharmaceuticals (c.f., atorvastatin and CI-943; Henck et al., 1995; 1998). As of mid-2006, 75 DNT studies have been submitted to the EPA Office of Pesticide Programs (OPP) in support of pesticide registration. Official statistics regarding the uses of these studies in chemical risk assessments have not yet been released, pending finalization of the relevant chemical risk assessments. Nevertheless, a preliminary survey of the use of DNT studies in risk assessment in OPP was conducted in February 2004. In this analysis, the impact of the DNT study was examined by identifying its specific use in the selection of endpoints and doses for the risk assessment (as compared to the 1998 retrospective analysis which compared the NOELs of DNT studies and other studies in the chemical database).
Table 4. Examples of Chemicals Tested Using the US EPA DNT Guideline or OECD Draft 426.

<table>
<thead>
<tr>
<th>Chemicals Tested</th>
<th>Endpoints and Doses</th>
<th>Number of Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,1,1-Trichloroethane</td>
<td>Dichlorvos (DDVP) (2)</td>
<td>Methyl parathion</td>
</tr>
<tr>
<td>Abamectin</td>
<td>Dicrotrophos</td>
<td>Methylazoxymethanol (2)</td>
</tr>
<tr>
<td>Acephate</td>
<td>Dietary restriction</td>
<td>Methyl mercury</td>
</tr>
<tr>
<td>Acetamiprid</td>
<td>Dimethoate</td>
<td>Molinate</td>
</tr>
<tr>
<td>Acibenzolar-s-methyl</td>
<td>Disulfoton</td>
<td>Naled</td>
</tr>
<tr>
<td>Acrylamide</td>
<td>Emamectin</td>
<td>Nelfinavir</td>
</tr>
<tr>
<td>AE-0172747</td>
<td>Epidermal growth factor</td>
<td>Nitrous oxide</td>
</tr>
<tr>
<td>Aldicarb</td>
<td>s-Ethyl dipropylthiocarbamate</td>
<td>n-Methylneodecanamide</td>
</tr>
<tr>
<td>Alitame</td>
<td>Ethroporphos</td>
<td>Perchlorate</td>
</tr>
<tr>
<td>Amicarbazone</td>
<td>Ethylbenzene</td>
<td>Phorate (2)</td>
</tr>
<tr>
<td>Atorvastatin</td>
<td>Etofenprox</td>
<td>p-Methane-3,8-diol</td>
</tr>
<tr>
<td>Azinphos methyl</td>
<td>Fenamidone</td>
<td>Prochloraz</td>
</tr>
<tr>
<td>BAS 510F</td>
<td>Fenamiphos</td>
<td>Profenofos</td>
</tr>
<tr>
<td>BAS 670H</td>
<td>Fipronil</td>
<td>Propylthiouracil</td>
</tr>
<tr>
<td>Bifenthrin</td>
<td>Fluazinam</td>
<td>Pymetrozine</td>
</tr>
<tr>
<td>Carbaryl</td>
<td>Flubendiamide</td>
<td>Pyrasulfotole</td>
</tr>
<tr>
<td>Carbofuran</td>
<td>Flufenacet</td>
<td>Spirodiclofen</td>
</tr>
<tr>
<td>Chlorfenapyr</td>
<td>Glufosinate ammonium</td>
<td>Prothioconazole</td>
</tr>
<tr>
<td>Chlorite, sodium</td>
<td>Glyphosate trimesium</td>
<td>Styrene</td>
</tr>
<tr>
<td>Chlorpyrifos</td>
<td>GN1180 (MN rgg120/HIV-1)</td>
<td>TBBPA</td>
</tr>
<tr>
<td>CI-943</td>
<td>Hydrogen sulfide</td>
<td>Tebuconazole</td>
</tr>
<tr>
<td>Cigarette smoke</td>
<td>Imidacloprid</td>
<td>Terbufos</td>
</tr>
<tr>
<td>Clodinafop propargyl</td>
<td>Iminodipropionitrile</td>
<td>Tetrachlorvinphos</td>
</tr>
<tr>
<td>Clothianidin</td>
<td>Indoxacarb</td>
<td>Thiamethoxan</td>
</tr>
<tr>
<td>Coumaphos</td>
<td>Isopropanol</td>
<td>Thiocloprid</td>
</tr>
<tr>
<td>Cyclohexanemethanol</td>
<td>Isoxaflutole</td>
<td>Thiram</td>
</tr>
<tr>
<td>Cyfluthrin, beta-cyfluthrin</td>
<td>Lambda-cyhalothrin</td>
<td>Fentin hydroxide (TPTH)</td>
</tr>
<tr>
<td>Cymoxanil</td>
<td>Lasoxifene</td>
<td>Triallate</td>
</tr>
<tr>
<td>Octamethylcyclotetrasiloxane</td>
<td>Lead nitrate</td>
<td>Tribufos</td>
</tr>
<tr>
<td>Decamethylcyclopentasiloxane)</td>
<td>Lindane</td>
<td>Trichlorfon</td>
</tr>
<tr>
<td>DDT</td>
<td>Malathion</td>
<td>Trichloroethylene</td>
</tr>
<tr>
<td>DEET</td>
<td>Maternal separation</td>
<td>Triethylene glycol monomethyl ether</td>
</tr>
<tr>
<td>Deltamethrin</td>
<td>Methamidaphos</td>
<td>Trimethyltin</td>
</tr>
<tr>
<td>Diazepam</td>
<td>Methimazole (6)</td>
<td>Zeta-cypermethrin</td>
</tr>
<tr>
<td>Diazinon</td>
<td>Methyl bromide</td>
<td>Ziram</td>
</tr>
</tbody>
</table>

The number in parentheses represents the number of studies.

29. It was noted that for 19 pesticide chemicals where a DNT study had been considered in the weight-of-evidence review of the toxicology data base, the DNT study was utilized to select endpoints and doses for risk assessments for five of those chemicals (Makris, 2004, personal communication). Four studies were used to establish an acute RfD, one was used for a chronic RfD and three were used for short- or intermediate-term non-dietary assessment. In this analysis, a single study may have been used for multiple risk assessment scenarios. An important finding of this review was that for four of the five DNT studies the critical effects either included or were solely based upon behavioral and neuropathological parameters that are not evaluated in other guideline studies (i.e., motor activity, auditory startle habituation, learning and memory, and morphometric analysis). This is consistent with the conclusion of an earlier retrospective analysis (Makris et al., 1998) and provides further evidence of the sensitivity of the DNT study in identifying adverse effects in the young, and the importance role of DNT studies in human health risk assessments. In June 2005, a further OPP preliminary evaluation indicated that the DNT study was used for
endpoint selection in 8 out of 32 pesticide risk assessments for which the DNT study was available and included in the weight-of-evidence evaluation of the toxicology data base under consideration (US EPA, 2005), although the specific risk assessment scenarios were not identified in that review.

30. In addition to using DNT data for regulatory decisions, some regulatory agencies have also, on a case-by-case basis, incorporated an additional database uncertainty factor into their regulatory decisions because of the absence of DNT data. The use of these uncertainty factors in risk calculations reflect regulator views that DNT data are valuable in refining permissible exposure levels, and the absence of these data can increase the uncertainty about the toxicity of the chemicals (US EPA, 2002a, b).

31. Cross-laboratory comparisons of methodologies and results from DNT studies have been conducted by US EPA scientists, in an on-going evaluation of OPP study submissions (Crofton et al., 2001; Crofton et al., 2004; Raffaele et al., 2003, 2004, 2005, 2006; Sette et al., 2004; Makris et al., 2005, 2006). In an ILSI Risk Science Institute (RSI) workshop entitled "An Evaluation and Interpretation of Neurodevelopmental Endpoints for Human Health Risk Assessment" a working group consisting of scientists from government, industry, academic, and public health sectors is examining the interpretation of DNT study data and addressing a number of critical issues (i.e., public health considerations, overall data interpretation, data variability, positive control data, and statistical analysis) (Fenner-Crisp et al., 2005; Crofton et al., 2005; Tyl et al., 2005; Raffaele et al., 2005; Holson et al., 2005) with the expectation that the results of this effort will be published in the peer reviewed literature in 2007.

FUTURE ACTIVITIES

32. The US EPA DNT guideline was developed and promulgated over fifteen years ago in response to the need for regulatory-based screening methods to assess developmental neurotoxicity (US EPA, 1986, 1991). The discussion above reviewed the overall performance of the DNT test method, as well as the ability to detect effects of concern from a regulatory perspective (Buelke-Sam et al., 1985; Elsner et al., 1986; Francis et al., 1990; MacPhail et al., 1997; Makris et al., 1998; Mileson and Ferenc, 2001; Slikker et al., 2005; Tilson et al., 1997). The recent increase in the number of regulatory DNT studies being conducted (US EPA, 2002c), has refocused attention on this test method. While some have argued that some tests are insensitive (e.g., assessment of cognitive and sensory dysfunction are inadequate), others have suggested that the tests are overly sensitive and have a high rate of false positives (AIHC, 1995; Claudio et al., 1999; Claudio et al., 2000; Cory-Slechta et al., 2001; US EPA, 2006). A number of diverse groups have advocated increased testing for developmental neurotoxicity (Andersen et al., 2000; Nelson, 1986; NRC, 1992, 1993; OTA, 1990; Stein et al., 2002; Vorhees, 1986). There have also been calls to include evaluations of endpoints not currently assessed, such as social behavior (Cory-Slechta et al., 2001), pharmacokinetics and neurochemistry (Andersen et al., 2000; Dorman et al., 2001) or changes during senescence (Cory-Slechta et al., 2001). In addition, there have been extensive criticisms of the complexity of the study, accompanied by calls for deleting some test components from the protocol (Li, 2005) or utilizing screening approaches that incorporate DNT testing into other testing protocols (Ladics et al., 2005; Cooper et al., 2006). Critics also claim that variability of some endpoints (e.g., motor activity, morphometrics) is too great to be useful (CMA 1987; Nolen 1985; York et al., 2004), and that this in vivo test is not necessary to detect developmental neurotoxicity (Balls and Combes, 2005). These controversial opinions do not invalidate the DNT study as an important protocol for use in hazard identification and risk assessment, but rather they highlight the need for ongoing scientifically-based evaluation of this test method and the incorporation of appropriate revisions as scientific knowledge advances and as experience with the DNT studies warrants.

33. A number of efforts are underway, reviewing data from existing DNT studies, to identify ways to refine the DNT test and, if possible, reduce the number of animals used. It has been proposed that by
applying certain statistical approaches to the behavioral analysis, a reduction in animal use can be achieved (Chiarotti and Puopolo 2000; Puopolo and Chiarotti 2000; Puopolo et al., 2004). Reviews of historical and positive control data have pointed out the need for more standardized reporting requirements (Crofton et al., 2001), and that no one postmortem measure is more sensitive, with each providing important data (Raffaele et al., 2005). The outcome of this continuing effort will allow better data interpretation, help refine requirements for future testing, and also guide new methods development.

34. In addition to the goal of refinement of the current approach to DNT testing, there is another and more pressing driver for change in the science arena of developmental neurotoxicity. There are currently thousands of chemicals that lack even simple, basic toxicological data (e.g., High Production Volume chemicals, pesticide inert ingredients, anti-microbial pesticides), but have a high potential for human exposure (NRC, 1984). Assessing potential neurotoxicological effects for these chemicals is a challenge confronting the chemical industry, international and national regulatory agencies, and associated stakeholders. New tools and methods are required to move towards a more sustainable risk assessment paradigm for these types of chemicals. While the current DNT guideline generates useful data for risk assessment purposes, this in vivo test is costly, time consuming, and uses a relatively large number of animals when conducted as a stand-alone study (as compared to incorporating the DNT testing into other on-going studies, such as a reproductive toxicity study). A pressing goal of future research is to develop a validated true first tier screening paradigm (e.g., a high-throughput in vitro screening battery) that can rapidly screen large numbers of chemicals for their potential to cause developmental neurotoxicity (Lein et al., 2005; Cocke et al., 2006; US EPA, 2006). Coupled with development of decision frameworks (Combes et al., 2003), data from these high-throughput screens will foster prioritization of any further testing in vivo. Data generated by the current DNT test method will be vital in the validation of these high-throughput in vitro methods, providing information on the utility and limits of these methods, as well as guidance on the potential use of data from these alternative methods in a risk assessment context.

CONCLUSIONS

35. OECD Test Guidelines are periodically reviewed in the light of scientific progress, changing assessment practices and animal welfare considerations, and the TG 426 should be no exception. Currently, a number of activities are underway for development of alternative methods to TG 426, or for replacement of certain parts of TG 426. An adopted TG 426 used by OECD member countries will generate new data for risk assessment and further the development of new approaches to DNT testing. The OECD supports the 3R’s (i.e., refinement, reduction, replacement) and works on alternatives to TG 426 (or parts of it) and would welcome any revision of TG 426 that would better meet the animal welfare considerations. However, any suggestion for replacements of components of the TG 426 need to be in compliance with the OECD submission and adoption process of new or updated Test Guidelines (OECD 2006) and subjected to approval by the WNT. In addition, the performance of a revised TG will have to be demonstrated before being adopted as a new TG 426, as described by the Guidance Document No. 34 in the section on test batteries:

“...Component test methods of test batteries are treated as individual test methods for validation purposes and it is necessary to demonstrate that the combination of test methods produces reliable and relevant results and is more effective than the individual tests. In general, substitution of any component of the battery should improve its performance.”
36. The DNT TG 426 represents the best available science for assessing the potential for developmental neurotoxicity in human health risk assessment, and data generated by DNT are relevant and reliable for the assessment of these endpoints. The test methods used in the DNT have been subjected to an extensive history of international validation, peer review and evaluation which is contained in the public record. The reproducibility, reliability and sensitivity of these methods have been demonstrated, utilizing a wide variety of test substances. Multiple, independent, expert scientific peer reviews affirm these conclusions, as described in this document. The DNT TG provides an outline of behavioral domains and morphological endpoints that are relevant to human neurodevelopment that should be examined to assess potential developmental neurotoxicity of a test compound. The results from DNT studies are used for hazard/risk assessment purposes and in cases where data from a DNT study are not presented, safety factors may be employed by regulators to address the need for DNT data from a regulatory standpoint. The document shows that a variety of chemicals have been tested for DNT constituting a sampled spectrum of the chemical universe that the test is proposed to investigate. Several published reports outlined in this document show that the DNT study is robust and can be conducted in multiple laboratories with consistent performance.

37. The TG 426 is considered to meet the regulatory needs and regulatory requirements of OECD member countries as outlined in Guidance Document No. 34, and as is described by the extensive documentation of the performance of the DNT study in this document. The DNT study has received extensive validation and is considered valid for its intended purpose and regulatory use.
REFERENCES


Cooper, RL; Lamb, JC; Barlow, SM; Bentley, K; Brady, AM; Doerrrer, NG; Eisenbrandt DL; Fenner-Crisp, PA; Hines, RN; Irvine, L.; Kimmel, CA; Koeter, H; Li, AA; Makris, SL; Sheets, L; Speijers, GJA; and Whitby, K. (2006) A tiered approach to life stages testing for agricultural chemical safety assessment. Critical Reviews in Toxicology 36:69-98.


