Regulatory Toxicology and Pharmacology 68 (2014) 96–107

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

Workshop Report

Approaches and considerations for the assessment of immunotoxicity for environmental chemicals: A workshop summary

Darrell R. Boverhof[a,*], Greg Ladics[b], Bob Luebke[c], Jane Botham[d], Emanuela Corsini[e], Ellen Evans[f], Dori Germolec[g], Michael Holsapple[h], Scott E Loveless[i], Haitian Lu[j], Jan Willem van der Laan[k], Kimber L. White Jr.[l], Yung Yang[m]

[a] The Dow Chemical Company, Midland, MI, United States
[b] DuPont Pioneer, Wilmington, DE, United States
[c] Cardiopulmonary and Immunotoxicology Branch, US Environmental Protection Agency, Research Triangle Park, NC, United States
[d] Syngenta Ltd., Macclesfield, UK
[e] Laboratory of Toxicology, Dipartimento di Scienze Farmacologiche e Biomolecolari, Università degli Studi di Milano, Milan, Italy
[f] Pfizer, Inc., Groton, CT, United States
[g] National Toxicology Program, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC, United States
[h] Battelle Memorial Institute, Columbus, OH, United States
[i] DuPont Haskell Global Centers for Health and Environmental Sciences, Newark, DE, United States
[j] Dow AgroSciences, Indianapolis, IN, United States
[k] Medicines Evaluation Board, Utrecht, The Netherlands
[l] ImmunoTox® Inc., Richmond, VA, United States
[m] Office of Pesticides Programs, US Environmental Protection Agency, Washington, DC, United States

ARTICLE INFO

Article history:
Received 28 October 2013
Available online 23 November 2013

Keywords:
Immunotoxicity
Environmental chemicals
Testing

ABSTRACT

As experience is gained with toxicology testing and as new assays and technologies are developed, it is critical for stakeholders to discuss opportunities to advance our overall testing strategies. To facilitate these discussions, a workshop on practices for assessing immunotoxicity for environmental chemicals was held with the goal of sharing perspectives on immunotoxicity testing strategies and experiences, developmental immunotoxicity (DIT), and integrated and alternative approaches to immunotoxicity testing. Experiences across the chemical and pharmaceutical industries suggested that standard toxicity studies, combined with triggered-based testing approaches, represent an effective and efficient approach to evaluate immunotoxic potential. Additionally, discussions on study design, critical windows, and new guideline approaches and experiences identified important factors to consider before initiating DIT evaluations including assay choice and timing and the impact of existing adult data. Participants agreed that integrating endpoints into standard repeat-dose studies should be considered for fulfilling any immunotoxicity testing requirements, while also maximizing information and reducing animal use. Participants also acknowledged that *in vitro* evaluation of immunosuppression is complex and may require the use of multiple assays that are still being developed. These workshop discussions should contribute to developing an effective but more resource and animal efficient approach for evaluating chemical immunotoxicity.

© 2013 Elsevier Inc. All rights reserved.

Abbreviations: 2,4-D, 2,4-dichlorophenoxyacetic acid; ACD, allergic contact dermatitis; ACSA, agricultural chemical safety assessment; AFC, antibody forming cell; CP, cyclophosphamide; CT, carbon tetrachloride; DART, developmental and reproductive toxicology; DIT, developmental immunotoxicity; ELISA, enzyme linked immunosorbent assay; EOGRTS, extended one-generation reproductive toxicity study; EPA, environmental protection agency; EU, European union; FDA, food and drug administration; HESI, health and environmental sciences institute; ICH, international conference on harmonization; ILSI, international life sciences institute; KLH, keyhole limpet hemocyanin; LLNA, local lymph node assay; NK, natural killer; NIEHS, national institute of environmental health sciences; NIOSH, national institute for occupational safety and health; NOAEL, no observed adverse effect level; NRC, national research council; NTP, national toxicology program; OECD, organisation for economic cooperation and development; PND, postnatal day; RIVM, dutch national institute of public health and the environment; SRBC, sheep red blood cells; STS, standard toxicology studies; TDAR, T-cell dependent antibody reaction; US, United States; WoE, weight-of-evidence.

* Corresponding author. Address: The Dow Chemical Company, 1803 Building, Midland, MI 48674, United States. Fax: +1 989 638 9305.
E-mail address: rboverhof@dow.com (D.R. Boverhof).

0273-2300/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.yrtph.2013.11.012

1. Introduction

Evaluation of immunotoxicity is an important component of the hazard evaluation and safety assessment process for both pharmaceuticals and environmental chemicals, including pesticides. Although standard toxicity studies provide valuable data for evaluating immunotoxicity, endpoints such as organ weights, histopathology, hematology, and additional endpoints that involve characterization of the cellular and functional status of the immune system can provide additional information for the assessment of immunotoxic potential. Such studies may include assessment of the ability to respond to immunization (e.g., the T-cell dependent antibody response (TDAR)), the capacity to destroy neoplastic cells (e.g., the natural killer (NK) cell assay), the relative abundance of lymphocyte subpopulations, or a variety of other functional and observational assays. Due to the additional information these studies can provide, immunotoxicity testing guidance has been developed for pharmaceuticals and environmental chemicals, although with different requirements for incorporation into their respective testing paradigms (ICH, 2005; US EPA, 2007). Under chemical regulations, pesticide registrations require the completion of a substantial number of toxicity studies with a recent additional requirement for the conduct of specific immunotoxicity assays (US EPA, 2007). In contrast, guidance for pharmaceuticals uses a weight of evidence (WoE) approach that only requires specific immunotoxicity assays if there is cause for concern identified in standard toxicity studies.

While animal toxicity testing is a critical component for assessing the hazard potential of both environmental and pharmaceutical chemicals, it is recognized that the approach is time-consuming, expensive, requires extensive use of animals, and may not take full advantage of emerging technologies and biological knowledge. Such sentiments are conveyed in the National Research Council (NRC) report entitled: "Toxicity Testing in the 21st Century: A Vision and a Strategy" (NRC, 2007). This report has highlighted a potential strategy to move away from animal-focused testing to an approach that uses high-throughput methods with *in vitro* human model systems. While it will take years for such a vision to be completely developed, evaluated and effectively implemented, it has provided increased focus and discussion on the need for improving both the efficiency and effectiveness of our current testing approaches (Andersen and Krewski, 2010). In the shorter term, there are opportunities to refine our approaches to toxicity testing through progressive tiered-evaluation approaches and integrated testing approaches that can be used to refine and optimize animal use and data generation. Similarly, retrospective analyses can be valuable approaches to evaluate the impact of current testing strategies as a means to prioritize areas for improvement, guide and support changes in current data requirements and study designs, and to enhance our approaches to data interpretation and utilization. Recent examples that are relevant for pesticides include analyses and discussions on the one-year dog study requirement (Dellarco et al., 2010), the two-generation rat reproductive study (Piersma et al., 2011; Rorije et al., 2011), and the mouse carcinogenicity study (Billington et al., 2010).

As additional experience is gained with current testing requirements and as new assays and technologies are developed, it is critical for all stakeholders to engage in active dialog about potential opportunities to advance our current testing approaches. To facilitate these discussions in the area of immunotoxicology, a workshop hosted by the International Life Sciences Institute-Health and Environmental Sciences Institute (ILSI–HESI) was held on the evaluation of current practices for the assessment of immunotoxicity for environmental chemicals. The goal of this workshop was to share current perspectives from experts in the field on approaches for the assessment of immunotoxicity, with a focus on immunosuppression. Diverse perspectives were captured from various sectors to ensure a broad consideration of different approaches and experiences with the use of standard endpoints and functional assays, as well as tiered-based testing strategies and developing assays and study designs. Major themes of the workshop included discussions on current immunotoxicity assessment strategies and experiences, developmental immunotoxicity assessment, and integrated and alternative approaches to immunotoxicity testing. What follows is a summary of the key messages and discussions in these areas that took place during the workshop.

2. Current immunotoxicology assessment approaches and experiences

Immunotoxicity is a term used to describe the alteration of the normal structure and/or function of the immune system as determined by established immunological and toxicological approaches. Studies in laboratory animals have provided information on the types of immunotoxic effects that chemicals may induce, and that information has been used to assess the sensitivity and predictability of toxicological testing approaches for the identification of immunotoxicity (Luster et al., 1988; Luster et al., 1992a; Luster et al., 1993; Vos and Van Loveren, 1987). In 1979, under the auspices of the United States National Toxicology Program (US NTP), a panel of experts gathered to prioritize a list of immunological assays that would be suitable for use in rodent toxicology studies. Four laboratories participated in the ensuing validation effort to determine whether the tests selected by the panel had the required sensitivity and reproducibility to successfully detect subtle alterations in immune function and host resistance in mice (Luster et al., 1988). Subsequent studies used this testing panel to evaluate approximately 50 chemicals and established correlations between specific immune function and host resistance tests (Luster et al., 1993; Luster et al., 1992b). In addition to these comprehensive examinations with mouse models, the rat has also been a focus in immunotoxicity testing, primarily because of its standardized use in preclinical toxicity studies. In the late 1970s, a testing panel using the rat based on the Organisation for Economic Cooperation and Development (OECD) 407 guideline was developed at the Dutch National Institute of Public Health and the Environment (RIVM) (van Loveren and Vos, 1989; Vos, 1977; Vos, 1980). The utility of the rat model for immunotoxicity testing was further evaluated and validated through a number of inter-laboratory studies with known immunosuppressive agents (ICICIS Group Investigators, 1998; Richter-Reichhelm et al., 1995; White, 1992). Over time, the screening paradigms from both the NTP and RIVM have been updated to include additional endpoints, such as "enhanced histopathology" and routine enumeration of lymphocyte subsets, and new techniques (particularly *in vitro* methods) are continuously being considered and evaluated for their utility as predictors of potential toxicity to the immune system. Importantly, the early work from these groups, in terms of immunotoxicology assay development, evaluation and implementation, played a critical role in shaping the development of immunotoxicology guidelines for both pharmaceuticals and environmental chemicals.

2.1. Pharmaceutical industry guidelines for immunotoxicity and experience

Guidelines for the assessment of the immunotoxicity of pharmaceuticals emerged independently and differently within the European Union (EU) and the United States. The EU Committee

for Proprietary Medicinal Products requested routine evaluation of all new chemical entities for their potential to cause adverse effects on the immune system and immune function in preclinical species (Gore, 2006). These studies were required in addition to the standard toxicology studies (STS) and included immunological assays such as lymphocyte phenotyping, NK cell assays and the TDAR assay. In contrast, the US Food and Drug Administration (FDA) used a "cause for concern" approach, where the results from STS were used to provide evidence of the potential for immunotoxicity which, if identified, would be followed up with additional studies, including functional immunotoxicity assays.

While both agencies supported a tiered approach in which any initial evidence of immunotoxicity would be followed up with additional immunological studies, the differences in the mandatory requirement for specific immunological studies between the agencies prompted action by the International Conference on Harmonization (ICH) to discuss potential harmonization of immunotoxicity testing requirements. This decision led to the formation of an Immunotoxicology Expert Working Group whose objective was to develop best practices for assessment of unintended immunosuppression (Weaver et al., 2005). A research project was initiated to review the correlation between STS and functional immune assays. The consensus output from this analysis was that routine testing, with specific studies that evaluate immune function, for all new drugs did not offer significant additional information beyond the STS and should not be a mandatory testing requirement. This opinion and associated recommendations were subsequently implemented into the ICH S8 guidance on immunotoxicity studies for human pharmaceuticals (ICH, 2005).

The ICH S8 guidance states that "all new human pharmaceuticals should be evaluated for the potential to produce immunotoxicity". The foundations of this evaluation are the test results from the STS, which contain numerous immune system endpoints (e.g., clinical signs of infection, hematologic evidence of inflammation or cytopenias, globulin levels, changes in bone marrow or lymphoid tissue, inflammatory infiltrates or pathogenic/opportunistic organisms seen in tissues). No single parameter is enough to rule immunotoxicity "in" or "out"; all parameters should be considered together in a WoE approach to determine if there may be an effect on the immune system. The review of the STS data, in conjunction with other factors, may prompt additional testing. The guidance lists six factors that should be considered in the evaluation of potential immunotoxicity (Table 1). These factors are to be assessed in a WoE approach to determine if a cause for concern exists. A finding of sufficient magnitude in a single area should trigger functional immunotoxicity studies. As part of the guideline, the evaluation of the immunotoxic potential of a pharmaceutical has been structured using a decision tree approach. The established approach would not lead to a blockade for development of the pharmaceutical compound if the potential for immunotoxicity was identified; however, it provides a signal to be included in the risk assessment/management approach in the clinical situation. This approach was perceived as an important step forward for putting the immunotoxicological findings into a translational perspective.

In practice, the ICH S8 WoE approach to immunotoxicity assessment has proven to be a practical yet science-based approach to pharmaceutical risk assessment. STS, which are the foundation of the guidance, have many endpoints which assess the immune system from different perspectives and follow-up studies should be selected based on specific concerns determined in a WoE review. In practice, one potential controversial area for the assessment of immunotoxicity is the contribution of stress responses to lymphoid changes. The ICH S8 appropriately cautions against dismissing such effects without the appropriate scientific rationale. The release of endogenous corticosteroids under stressful conditions can influence lymphocyte trafficking, production, and survival. However, without corroborating evidence of "stress," a direct effect on the immune system should be considered. Stress-related changes to immunological endpoints in STS can usually be differentiated from direct immunotoxicity through various indicators, including signs of increased endogenous corticosteroid release (e.g., stress leukogram of neutrophilia and lymphopenia, hyperglycemia, and adrenal cortex hyperplasia) and findings suggesting overt toxicity (e.g., anorexia, weight loss, decreased body weight gain, hunched appearance, organ failure). If the WoE review suggests a need for immunotoxicity assessment beyond STS, studies are conducted to determine the mechanism, the no observed adverse effect level (NOAEL), potential reversibility, translation to human risk, and overall product marketability. In some instances this review may also be a trigger to screen back-up or alternative candidate molecules. Experience has shown that few molecules demonstrate sufficient cause for concern to trigger the second tier of testing (i.e., beyond STS). Of these, very few have been found to be immunotoxic in humans, or are discontinued from development due to immunotoxicity concerns (Bugelski et al., 2010; Weaver et al., 2005).

2.2. Chemical industry guidelines for immunotoxicity and experience

Exposure scenarios for drugs and chemicals are quite different. In the case of pharmaceuticals, therapeutic exposure is typically well-understood, a risk/benefit assessment for the individual patient can be conducted, and individuals are monitored, with any effects documented, in human clinical trials. In contrast, chemical regulations mandate minimizing or preventing human exposure to ensure the chemicals do not pose an unreasonable risk to health and the environment. For the chemical industry, only pesticides have specific regulations and there are extensive data requirements that are used in pesticide hazard evaluation and risk assessment. In 1989, the US Environmental Protection Agency (EPA) Scientific Advisory Panel recommended evaluation of immunotoxicity as a part of the hazard assessment process for pesticide registrations. Subsequently, an immunotoxicity test guideline (OPPTS 870.7800) was developed and published (US EPA, 1998). Currently, the US EPA is the only chemical regulatory agency to have specific immunotoxicity testing requirements. The assessment of immunotoxicity became a part of the required studies in the revised 40 CFR Part 158 Toxicology Data Requirements for conventional pesticide

Table 1
Summary of the ICHS8 guideline factors to consider in the evaluation of potential Immunotoxicity.

Factor	Context
Standard toxicity studies	Are there signs of immunotoxic potential based on the result of previous toxicology studies?
Pharmacological properties	Do the pharmacological properties of a test compound indicate it has the potential to affect immune function (e.g., anti-inflammatory drugs)?
Intended patient population	Is the intended patient population already in an immunocompromised state?
Structural similarity	Is there structural similarity to other known immunotoxicants?
Disposition of the drug	Is the drug retained at high concentrations in cells of the immune system?
Signs observed in clinical trials or clinical use	Did clinical findings or signs suggest potential immunotoxicity?

registration for food and non-food uses (US EPA, 2007) and more recently for antimicrobial pesticide registration (US EPA, 2013). The immunotoxicity test guideline is intended to provide information on suppression of the immune system which might occur as a result of repeated exposure to a test chemical. These data are to be used along with results from STS to provide more accurate information on the risk to the immune system. The test methods are designed to assess the TDAR to evaluate the functional responsiveness of major components of the immune system to a T-cell dependent antigen (e.g., sheep red blood cells (SRBC), with either an antibody forming cell (AFC) assay or an enzyme linked immunosorbent assay (ELISA)). In the event that the test substance produces significant suppression of the TDAR, phenotypic analysis for lymphocyte subpopulations may be required to characterize which cell types are affected by the test substance. If the test substance has no significant effect on the TDAR, a functional test for NK cell activity may be performed to evaluate the chemical's effect on non-specific (innate) immunity to ensure the integrity of the immune system.

In 2009, EPA's Office of Pesticide Programs modified the above immunotoxicity testing strategy in agreement with the registrants and the pesticide stakeholders to simplify immunotoxicity testing. The new strategy uses a WoE approach to determine the testing needs beyond the TDAR study. For example, if the test substance produces dose-related suppression of TDAR, then the test substance is considered immunotoxic and no further testing is required. On the other hand, if the TDAR is negative and there is suggestive evidence of adverse effect on the immune system endpoints from the STS (e.g., changes in lymphoid organ weights, hematology, or histopathology), an NK cell assay is required to evaluate any effect on innate immune function. However, if there is no evidence to suggest that the immune system is a primary target organ of the test substance, a functional NK cell assay may not be necessary (Fig. 1).

Since this requirement came into effect, pesticide registrants have completed and submitted more than 80 guideline-compliant TDAR studies which provided an opportunity to evaluate the potential implications of these required immunotoxicity studies on hazard identification and risk assessment. To assess the regulatory impact, Crop Life America member companies participated in a retrospective analysis that included a dataset of 82 immunotoxicity studies conducted on a diverse set of 78 unique pesticide chemicals (Gehen et al., submitted for publication; US EPA, 2013). This dataset represented a wide distribution of product types and chemical classes, and also included a balanced diversity in terms of species, strains, gender and methods of TDAR quantitation to help ensure that the results of the analysis did not arise from a narrow set of TDAR study conditions. The objective of the Crop Life America analysis was to determine the frequency of positive assay responses and the potential for the TDAR to impact hazard assessment and human health risk assessment. The analysis revealed that for 77 of 82 studies the TDAR was unaltered at all dose levels, while in the remaining five studies, a reduction was observed but only at the high-dose level (Gehen et al., submitted for publication; US EPA, 2013). In all of the cases, a follow up NK assay was deemed unnecessary by the chemical registrant based on the results from the TDAR as well as the STS. To assess the impact of these data, TDAR NOAELs were compared to the toxicity endpoints that were previously selected for use in the specific pesticide risk assessments. For all 78 chemicals, the TDAR NOAELs were greater than the critical values selected for the existing risk assessments. Based on this Crop Life America analysis, routine conduct of the TDAR assay appears to have had a limited impact on chemical pesticide hazard identification and human health risk assessment. These data suggest that alternative approaches, that include an initial WoE analysis for immunotoxic potential, should be considered prior to conducting animal-based functional immunotoxicity testing for chemical pesticides, similar to what is done for the hazard assessment of pharmaceuticals.

2.3. Opinions from workshop participants

Given the differences in the approaches and requirements for chemicals and pharmaceuticals, discussions during the workshop session on current immunotoxicology testing approaches and experiences focused on consideration of whether characterization of the immunotoxic potential of a chemical should be routinely examined, or if there is sufficient scientific knowledge to use a triggered approach. All participants agreed that a trigger-based approach for environmental chemicals, as used for pharmaceuticals under ICH S8 guidance, should be considered; however, participants also realized that defined criteria need to be available to ensure the consistent and transparent interpretation of the potential need for additional immunotoxicity testing. This assessment should consist of a WoE assessment of the available data, including that from STS, along with consideration of toxicokinetics, use patterns, and the actual potential for exposure. Such an approach is not unlike that provided in the ICH S8 guidance for pharmaceuticals. Additional retrospective analyses may help to solidify the appropriateness of this approach while also defining the most informative triggers that can be used as part of the risk assessment.

The participants also agreed that if additional immunotoxicity testing is triggered, based on the initial WoE assessment, then the assays selected for use should be those that have shown good performance in terms of both sensitivity and specificity based on previous validation and retrospective analyses. From a practical standpoint, the assays should not be overly complex, to ensure broad and reliable applicability across laboratories. Participants also acknowledged that current data indicate that the TDAR is a good starting point for assay consideration given that it is one of the few immunotoxicological endpoints that requires many of the cellular components of an immune response (e.g., B-cells, T-cells, macrophages) and thus, is a sensitive indicator of a chemical's immunotoxic potential (Luster et al., 1992b). However, it was also realized that assay selection for a specific chemical should also consider findings from the existing data such that the studies can be selected to address the identified cause for concern, as applicable. Overall, participants agreed that these consensus points deserve further consideration and discussion through focused assessments and expert consultations for implementation into an approach for the assessment of immunotoxicity for chemicals management.

3. Assessment of developmental immunotoxicity

Over the last two decades there has been a growing interest in establishing an increased understanding of age-related susceptibilities to toxicity with considerable efforts devoted to understanding potential differences between infants and children when compared to adults. During this time, attention to DIT has grown significantly in scope. This interest is based on the concept that the developing immune system may be qualitatively and/or quantitatively more susceptible to xenobiotic perturbation when compared to the adult. There have been numerous workshops, roundtables, symposia, and sponsored research to address questions pertaining to DIT (i.e., triggers, protocols, endpoints and assays to measure DIT, hazard identification, and risk assessment) (Burns-Naas et al., 2008; DeWitt et al., 2012; Dietert et al., 2000; Holsapple, 2002; Holsapple et al., 2005; Holsapple et al., 2007; Ladics et al., 2005; Ladics et al., 2000; Luster et al., 2003). One of the goals of this workshop was to

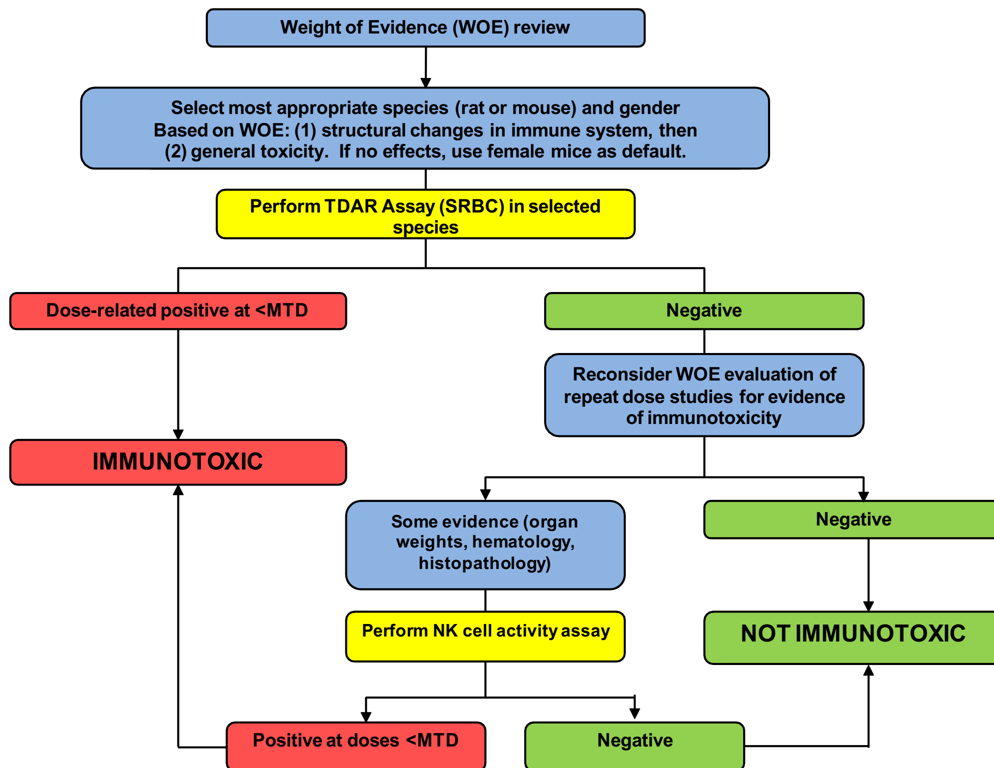

Fig. 1. EPA guidance for a weight of evidence (WoE) decision tree approach for immunotoxicity testing. The existing standard toxicity studies (STS) are first reviewed for any evidence of immunotoxicity that would help to identify the most sensitive species (rat/mouse) and gender for subsequent immunotoxicity testing. If no immunotoxic effects are identified, general toxicity parameters are used to define the most sensitive species and gender, with the female mouse as the default in the absence of appropriate toxicity discriminators. The TDAR is subsequently conducted in the most sensitive species and gender. If the test material produces a dose-related suppression of the TDAR, it is considered 'immunotoxic' and no further testing is required. If the available STS and WoE suggest no potential for immunotoxicity, then no further testing is required and the test material is considered 'not immunotoxic'. However, if the data suggest some potential for effects on the immune system an NK cell assay may be required to further characterize any immunotoxic potential. (MTD = maximum tolerated dose).

share the most current perspectives from experts in the field on approaches and considerations for the assessment of DIT.

3.1. Background on developmental immunotoxicity testing

The basic tenet behind the interest in assessing the risk to human development is that children differ significantly from adults in their biological and/or physiological responses to environmental exposures (Holsapple et al., 2004). In the context of the development of the human immune system, common infectious diseases can occur more often, and are usually more severe in the very young when compared to adolescents and adults, and infants are more susceptible to immune toxicities and immune manipulations. Yet, it is important to note that infants can mount a vigorous immune response to tissue and organ allografts, and to vaccines (i.e., there is immunocompetence in the very young).

In the context of DIT, the specific hypothesis to be considered is that the developing immune system demonstrates greater susceptibility to chemical perturbation than the adult immune system. From a regulatory perspective, there is concern that a well-designed guideline immunotoxicity study conducted in adult animals could miss a potentially greater susceptibility associated with the critical stages of development of the immune system. As noted by Burns-Naas et al. (2008), this susceptibility can be manifest as a qualitative difference (e.g., developing immune system affected, but not the adult immune system), a quantitative difference (e.g., developing immune system affected at lower doses than the adult immune system), or a temporal difference (e.g., effects on the developing immune system are more persistent than those seen with the adult immune system). Luebke et al. (2006) presented evidence that these types of differences between the developing and adult immune system have been identified through the specific discussion of five example compounds.

From a scientific perspective, there has been a steadily increasing interest in this topic with over 40 workshops, symposia, and scientific sessions focused in-whole or in-large part on DIT since 2000. In particular, two workshops were organized in 2001 that both took place in Washington, DC. The outcome from a workshop organized by ILSI-HESI was summarized by Holsapple (2002), and the outcome from a workshop organized by the NIEHS and the National Institute for Occupational Safety and Health (NIOSH) was summarized by Luster et al. (2003). While these two workshops used different approaches, they addressed similar objectives, and came forward with remarkably consistent conclusions and recommendations, which included the following:

- Both permanent and transient effects should be considered.
- All critical stages of immune system development should be assessed as part of a standard study design and specific stages should be further studied, if an effect is seen.
- A DIT protocol could be extended to include juveniles and young adults, as it is critical to continue exposure until possible effects on the immune system can be assessed.
- Although the mouse is the primary animal model for immunological studies, and was the species used in the studies by Luster et al. (1988, 1992b, 1993), the rat is considered to be the species of choice for DIT studies (Holsapple, 2003) because of historical experience and preference for rats in developmental and reproductive toxicology (DART) studies (OECD, 1995; OECD, 1996). However, it is important to consider the species differences in

developmental timelines as part of the evaluations (Holsapple, 2003). This point is illustrated in Table 2 for the development of the mouse and human immune systems, which clearly shows that the development of the mouse immune system lags behind that of the human immune system in the context of the full gestation period for each species. Importantly, the developmental timelines for the immune systems of the mouse and the rat are similar.

The HESI workshop recognized that most 'classical' protocols to assess the potential for DIT (e.g., those conducted prior to 2000) used experimental designs in which pregnant dams were exposed under a variety of experimental conditions (usually during some period(s) of gestation) and the possible effects on the immune system were assessed in the F1 generation when they became young adults (Fig. 2A). This assessment approach was based on the fact that immune tests had been optimized in young adult rodents. Ladics et al. (2000) investigated an 'alternative' protocol for an assessment of DIT (Fig. 2B), in which 10-day old and 21-day old rats pups were evaluated using standard immune tests, such as phenotyping and the TDAR. Their results served as a confirmation of the profound immaturity of the rodent immune system (Table 2), when it was assessed shortly after parturition and/or weaning. In light of these results, and the recommendations from both the HESI workshop and the NIEHS/NIOSH workshop, it was proposed that that all stages of development should be assessed under a single study design and that it is critical to continue exposure until possible effects on the immune system can be assessed. From this, a new DIT protocol was 'proposed' (Fig. 2C) in which pregnant dams were exposed, and the exposure was continued until the F1 generation became young adults, at which time the immune tests were conducted. The HESI workshop was the first to recommend that DIT study designs should be integrated into already required DART protocols, to the extent possible. A subsequent paper by Ladics et al. (2005) began to address the issues associated with including neurotoxicology and immunotoxicology assessments in standard DART studies. In spite of the recognized challenges associated with this approach, it was concluded that a proposed framework for DIT should be based on the integration of DIT study designs into standard DART protocols, to the extent possible (Holsapple et al., 2005). This paper also emphasized that a proposed framework for DIT should address all critical windows of development (e.g., in the rat, evaluate a time point that allows for the assessment of any effects induced from gestation to postnatal day (PND) 45) and specific windows should be further studied if an effect is seen (Fig. 2C). It was also emphasized that animals should be exposed throughout the treatment protocol and that the triggers for assessing DIT may include information from structure-activity relationships, results from other toxicity studies, and a consideration of the intended use of the drug or chemical.

The integration of immunotoxicity endpoints and functional assays into these studies has been accomplished in the so-called "juvenile pesticide studies," as reported by Chapin et al. (1997)

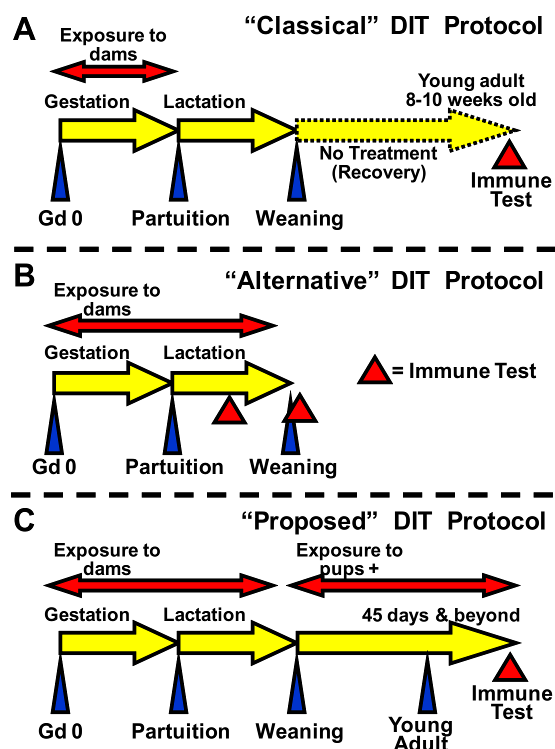

Fig. 2. Comparison of Multiple Protocols to Assess the Potential for DIT. The 'Classical' DIT Protocol (2A) depicts the state-of-the-science before and around 2000, in which pregnant dams were exposed under a variety of experimental conditions – usually various periods of gestation – and the exposed F1 generation was assessed for effects on the immune system when they became young adults, the age at which these parameters had been optimized. An 'Alternative' DIT Protocol (2B) was evaluated because of the possibility that the 'Classical' DIT Protocol was primarily an assessment of the recovery of effects on the immune system. The application of this protocol has been limited because of the relative immaturity of the rodent immune system, when compared to the human immune system (See Table 2). Based on the recommendations from two 'benchmark' workshops in 2001, a new DIT Protocol (2C) was 'Proposed' such that the exposure extends to at least Day 42 to cover the juvenile stage.

and by Smialowicz et al. (2001), and was a critical element of the HESI Agricultural Chemical Safety Assessment (ACSA) framework, as described by Cooper et al. (2006). Importantly, the ACSA framework became the basis for the OECD Test Guideline 443 study design (i.e., the so-called extended one-generation reproductive toxicity study (EOGRTS)).

3.1.1. OECD extended one-generation reproductive toxicity study (EOGRTS)

A guideline approach for the evaluation of developmental immunotoxicity is available through the conduct of the EOGRTS (OECD, 2011c). The OECD Test Guidance 443 is based on initial

Table 2
Timing of immune system development in mice and humans.

Event	Mouse (days) (% of term)	Human (weeks) (% of term)
Appearance of T cells in fetal liver	14 (67%)	6–8 (15–20%)
Organogenesis of thymus begins	11 (52%)	6 (15%)
Lymph nodes evident	10.5 (50%)	8–12 (20–30%)
Spleen develops	13 (62%)	10–14 (25–35%)
B cell lymphopoiesis begins in bone marrow	17 (81%)	12 (30%)
B lymphocytes detectable in blood	13 (62%)	12 (30%)
CD4+ and CD8+ T cells detectable in spleen	19 (91%)	14 (35%)
Thymus development completed	13 (62%)	15–16 (37–40%)
Bone marrow becomes the major site of hematopoiesis	17.5 (83%)	22 (55%)
T cell receptor expression in periphery	Early post-natal	23 (58%)

study design described by Cooper et al. (2006) and proposal by the HESI ACSA Technical Committee. The study assesses parental fertility and reproductive function and the development of offspring to sexual maturity with an assessment of sexual landmarks. An evaluation of the developing nervous and/or immune system is also included and a second generation can be triggered if any effects requiring further evaluation are identified in the first generation (OECD, 2011b). The design provides the opportunity to evaluate life stages not covered by other study types while minimizing the use of experimental animals. Overall, the study represents a highly integrated, and potentially logistically and technically complicated, study design that includes an assessment of DIT.

Prior to the formal adoption of the guideline, studies were undertaken by members of the European Crop Protection Association to evaluate the draft study design for technical feasibility. In addition, a study was conducted in response to a request by the US EPA for reproduction and developmental neurotoxicity data. The outcome of these studies has been summarized (Fegert et al., 2012). An evaluation of immunotoxicity was included in some of the studies reported and results for the studies on lead acetate and 2,4-dichlorophenoxyacetic acid (2,4-D) are discussed below.

OECD 443 specifies that determination of the primary IgM response to a TDAR is required on PND 56 (±3 days) if the DIT option is exercised. The antigens suggested are SRBC or keyhole limpet hemocyanin (KLH), although the latter is not yet recognized by US EPA as an acceptable antigen, and the evaluation methods specified are AFC or serum antibody titers as measured by ELISA. The draft OECD Guidance Document 151 (OECD, 2011a) discusses the choice of appropriate antigen in more detail and highlights the potential of investigating antibody class switching, a biological mechanism that changes a B cell's antibody production from one class to another (e.g., IgM to IgG), if deemed necessary.

The study design for OECD 443 includes a pre-mating treatment period of 2–4 weeks followed by a mating period of 2 weeks. The majority of mating would be expected to occur during the first estrus cycle, but some pairs may mate during the second cycle. As a result, littering (PND 0) would be expected to occur over at least 4, and possibly 8, days assuming no effect of test substance on the duration of the cycle or gestation length. The resulting variability in the date at which each litter reaches PND 56 can complicate TDAR assay conduct and may influence the choice of TDAR assay undertaken (i.e., AFC assay or ELISA based approaches). The overall integrated study design and associated complexity are depicted in Fig. 3.

Two EOGRTSs that included immunotoxicity evaluations were presented as part of the workshop, one with lead acetate and the other with 2,4-D. The evaluation of lead acetate was conducted in Wistar rats and included an evaluation of the T-cell dependent IgM antibody response to SRBC using an ELISA approach. One animal/sex/litter was used (a total of 20/sex/group), with blood taken on PND 70. The group mean values for serum anti-SRBC IgM indicated a reduction in circulating anti-SRBC IgM in top dose males (27% decrease relative to controls) while the positive control, cyclophosphamide, demonstrated a greater than 95% decrease in the response (Fegert et al., 2012). Previous developmental toxicity studies with lead acetate (Bunn et al., 2001; IARC, 2006; Miller et al., 1998) have demonstrated a modulation of select immune responses in both Fischer 344 and Sprague-Dawley rats; however, an effect on the antibody response to KLH was not identified in these studies, consistent with the lack of a significant effect in the EOGRTS evaluation. Importantly, this study demonstrated the initial feasibility for the conduct of DIT as a part of the EOGRTS.

The evaluation of 2,4-D was conducted with Sprague-Dawley rats and included an evaluation of the TDAR to SRBC on PND 70–74 using the AFC assay (10 animals/sex/group). In addition, an evaluation of NK cell activity was conducted in 10 animals/sex/group on PND 87–93., All positive controls responded appropriately and there was no evidence of 2,4-D related immunotoxicity in the TDAR assay and the NK cell assay, consistent with previous evaluations (Fegert et al., 2012; Marty et al., 2013). However, the TDAR assay data was confounded by added temporal variability across the dose groups due to the conduct of the assay on different days which was not balanced by dose group due to litter assignments and the defined age windows for the study and this particular endpoint (PND 70–74). This confounding factor is a disadvantage of using the AFC assay for evaluation of the TDAR response, as the assay needs to be conducted immediately following animal sacrifice (Ladics, 2007). Therefore, due to the study design age constraints and the assay conduct constrains, a clear "day" effect was observed that was unbalanced across groups. While this result did not impact the overall interpretation for the endpoint, it highlighted the need for planning and flexibility in the age window to ensure a more balanced assessment if using the AFC assay.

Overall, using the draft OECD 443 guideline, the EOGRTS has been successfully conducted by several laboratories in the US and Europe (Fegert et al., 2012). As part of the study designs, an evaluation of DIT was included and examples of the TDAR assay with SRBC, as measured by the AFC assay and ELISA approaches, have been reviewed. The data suggest that for this study, the ELISA assay offers the advantage that blood samples can be taken from animals on the specified sampling dates and stored frozen allowing samples to be analyzed concurrently for antibody levels. This approach reduces the logistical difficulties of this highly complex study and eliminates temporal variability and the possibility of an imbalance across groups introduced by the need to use fresh tissue as required by the AFC assay. Bleeding and retaining animals for further investigation of antibody switching is also possible using the ELISA assay as interim blood samples can be taken without sacrificing the animal. The successful completion of these studies clearly highlights the feasibility of the conduct of the EOGRTS; however, the logistical and technical complexity of the study design, along with the potential impact of the data for advancing the risk assessment for the molecule, should be carefully considered prior to the request and/or conduct of the study.

3.2. Opinions from workshop participants

One of the objectives of this workshop was to provide a forum in which participants could discuss the current state-of-the-science of DIT. Specific questions on DIT were posed during several breakout sessions to help guide the discussion. The first question considered the value or advantages and disadvantages of DIT vs. adult testing. DIT was considered valuable by workshop members in that it specifically covers multiple critical windows of immune system development, and therefore, DIT was able to maximize sensitivity to detect any potential immunotoxic effects. DIT also afforded the opportunity to look at different developmental periods, if necessary, as well as the prospect for integrated studies (e.g., OECD 443 EOGRTS) and the associated advantages of integrated assessment of endpoints (i.e., concurrent evaluation of other endpoints and the sensitivity of other developing systems such as the endocrine or nervous systems). The overall consensus of participants was that if a DIT was conducted, there was no need to conduct a separate immunotoxicology study with adult animals. Likewise, if an adult immunotoxicology study was available, it was suggested that additional information be considered to justify the need for a DIT study (e.g., additional specific cause-for-concern or endocrine activities).

A number of potential disadvantages or challenges associated with conducting a DIT were noted, particularly the current limited understanding of the relative sensitivity of the developing immune system in rodents and humans. Other identified issues included

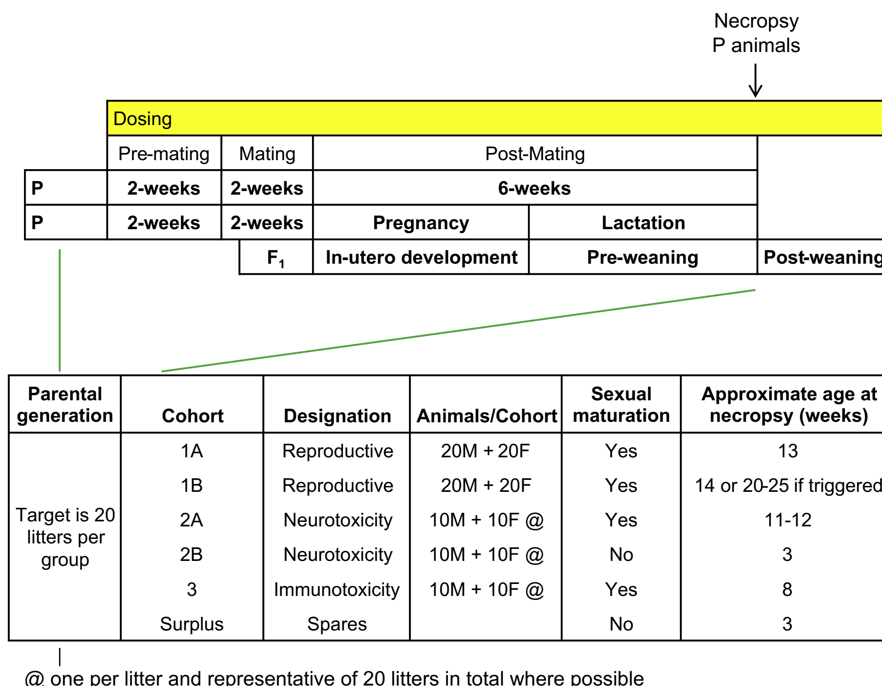

Fig. 3. Study Schematic for the OECD 443 Extended One-Generation Reproductive Toxicity Study. This study design is intended to provide an evaluation of reproductive and developmental effects that may occur as a result of pre- and postnatal chemical exposure as well as an evaluation of systemic toxicity in pregnant and lactating females and young and adult offspring. In the assay, sexually-mature male and female rodents (parental (P) generation) are exposed to doses of the test substance starting 2 weeks before mating and continuously through mating, gestation and pup weaning (F_1 generation). At weaning, pups are selected and assigned to cohorts of animals for reproductive/developmental toxicity testing (cohort 1), developmental neurotoxicity testing (cohort 2) and developmental immunotoxicity testing (cohort 3). The F_1 offspring receive further treatment with the test substance from weaning to adulthood. Clinical observations and pathology examinations are performed on all animals for signs of toxicity, with special emphasis on the integrity and performance of the male and female reproductive systems and the health, growth, development and function of the offspring. Part of cohort 1 (cohort 1B) may be extended to include an F_2 generation; in this case, procedures for F_1 animals will be similar to those for the P animals.

cost and complexity of study design; dosing issues (e.g., placental and/or lactational transfer); a lack of global alignment of DIT testing approaches and the need for conducting such studies; the lack of a specific testing guideline; and a loss of context of the WoE data associated with adult exposures from other toxicology studies.

The endpoints and assays to be conducted in a DIT study have been discussed extensively (Burns-Naas et al., 2008; Holsapple et al., 2005). Participants agreed that the age at the time of evaluation is critical in determining which types of endpoints are to be assessed. For example, prior to PND21, assessment of the TDAR is not feasible, as the immune system is not fully developed. The TDAR can be assessed, albeit not at optimized levels, at and after PND21 (Ladics et al., 2000). Workshop members indicated that a TDAR can be conducted successfully when animals are six to eight weeks of age (i.e., PND 42–56) (Ladics et al., 2000). Some participants suggested that cell-mediated immunity (e.g., the cytotoxic T-lymphocyte assay) and NK cells should also be evaluated in young adult animals following exposure throughout all life stages. A number of gaps or problems were also raised in regard to the implementation and/or interpretation of DIT. Specifically, for DIT purposes, an EOGRTS should not be viewed as a routine assay, as it is already quite complex with multiple dose groups and endpoints. Therefore, if a DIT is required, a standalone DIT protocol may be preferred. In addition, some participants questioned whether there were any compounds known to be positive in a DIT study but negative in an adult immunotoxicology study, as the extent to which this was known to occur would help to further define the value in the conduct of DIT studies. To date, few such compounds have been identified (e.g., 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (Gehrs and Smialowicz, 1999; Smialowicz et al., 1994)).

Further discussion addressed factors and data that may provide relevant triggers for conducting a DIT study. Participants agreed that both adult and DIT studies should be triggered by an initial WoE that would determine and consider the need and impact of the study data (i.e., a cause for concern approach), unless conducting an EOGRTS that includes the DIT option. The WoE considerations discussed for DIT included: the xenobiotic and its intended use (i.e., is there potential for exposure at earlier life stages?); absorption, distribution, metabolism, and excretion data; effects observed in developmental toxicology and/or multigenerational reproduction studies; and adult immunotoxicity findings, if available. Some participants suggested that the same information (i.e., triggers) evaluated for adults should also be considered as triggers for a DIT study. Overall participants agreed that DIT studies can play an important role in characterization of immunotoxicity; however, the requests and designs for this study should be made after careful consideration of the study objectives and needs.

4. Integrated and alternative approaches to immunotoxicity testing

In addition to the presentation and discussion of regulatory testing requirements and DIT, the workshop also addressed current and emerging approaches that could be implemented to help further inform immunotoxicity potential as part of chemical safety assessments. To accomplish this goal, a focus was placed on the application of new approaches and technologies that would provide additional data while reducing or replacing animal use. Although many approaches and technologies may be available, the workshop focused on the potential for the use of integrated study designs and *in vitro* assays for the assessment of immunotoxic potential.

4.1. Concurrent assessment of TDAR in repeat dose toxicology studies

Integrated testing strategies involve the assessment of multiple endpoints within a single toxicity study and represent an important approach for reducing animal use and streamlining testing approaches. In the field of immunotoxicology, research has been conducted to explore the ability to integrate the TDAR assay with SRBCs into a standard 90-day repeat dose toxicity study. To understand the feasibility and scientific validity of such an approach, a series of studies were conducted to address whether injection with SRBC would alter clinical pathology values or prevent the detection of standard toxicological responses, such as alterations in organ weights and histopathology. Two chemicals were used in the evaluation: cyclophosphamide (CP), a known immunosuppressive compound, and carbon tetrachloride (CT), a hepatotoxicant whose primary target is not the immune system. Male CD rats ($n = 10–16$/group) were dosed either with vehicle or CP by intraperitoneal injection (study 1) or with vehicle or CT by oral gavage (study 2) for 30 or 90 days. Six days prior to study termination, half of the animals from each treatment group were injected with SRBCs for evaluation of the TDAR response. The remaining animals from each dose group served as the non-SRBC injected controls to allow for evaluation of the impact of SRBCs on the standard toxicological endpoints. At study termination, endpoints evaluated included hematology/clinical chemistry, serum anti-SRBC antibody (ELISA), organ weights and histopathology (brain, heart, liver, thymus, kidneys, testes, spleen) and spleen cell number. Splenocyte subpopulations were also determined by flow cytometry following CT exposure. The results indicated that injection of SRBC in rats administered CP over 30 or 90 days did not alter hematology or clinical chemistry parameters compared to rats not receiving SRBC. With the expected exception of the spleen, SRBC administration did not alter the weights or morphology of tissues routinely collected or spleen cell numbers. Furthermore, the immunosuppressive effects of CP were not masked by the administration of SRBC at either time point (Ladics et al., 1995).

Similar to the results with CP, the injection of SRBCs in rats administered CT over 30 or 90 days did not alter hematology, clinical chemistry, organ weights or histopathology, with the exception of the spleen, when compared to rats not receiving SRBC. Injection with SRBC also had no effect on relative or absolute numbers of splenic lymphocyte populations. Furthermore, injection of SRBC in either 30- or 90-day study animals did not mask or alter the ability to characterize the mild hepatotoxic effects of CT on endpoints, such as increased sorbitol dehydrogenase activity, liver hypertrophy, or centrilobular fatty changes (Ladics et al., 1998). The results of these studies support the conclusion that assessment of the functional responsiveness of humoral immunity can be conducted in animals from standard toxicity studies without altering standard toxicological endpoints.

Complementary studies supporting this conclusion were completed in an independent laboratory using female Fischer 344 rats that were treated with CP or vehicle and with or without SRBC injections. In this examination, the impact of SRBC immunization on spleen and thymus weights and histopathology was examined, along with the ability to evaluate the spleen AFC assay response and spleen histopathology in the same animal via spleen sectioning approaches. Consistent with the previous evaluations, immunization with SRBCs did not impact organ weights. Histopathology of the spleen and thymus were also unaffected, with the exception of minor increases in the number of germinal centers in the spleen. The AFC response in vehicle and CP treated rats was similar between whole spleen samples and sectioned spleen, which support the ability to evaluate both spleen histopathology and the AFC response in a single animal (Woolhiser et al., 2007).

4.1.1. Opinions from workshop participants

In aggregate, these data clearly demonstrated that assessment of immunosuppression, as measured by a variety of endpoints (e.g., spleen histopathology; SRBC antibody response) can be successfully performed simultaneously in repeated dose studies as part of an integrated study design, without the need for a separate evaluation using additional animals. Workshop participants agreed that the application of integrated study designs should be considered in designing studies to fulfill requirements for the assessment of immunotoxicity. Integration of assays to detect immunosuppression into existing testing schemes (e.g., 90-day standard toxicity tests for pesticides in adult animals and the use of excess offspring generated in a reproductive toxicity study for lymphocyte phenotyping and TDAR, as described in the OECD Test Guidance 443) appear to be feasible, and will, if adopted, produce results that can be used for immunotoxicity risk assessment. Several scenarios on how to achieve the integrated testing approach were discussed: 1) conduct the TDAR assay at the conclusion of a 28 day or 90-day study (AFC or ELISA approach); 2) collect serum on day 28 and freeze and then complete standard 90-day study on the same animals. If 90-day study shows immune organ effects, then conduct the IgM ELISA; or 3) alternatively, one could analyze 28-day serum and, based on results, also evaluate secondary IgG anti-SRBC response at the end of the 90-day study in the same animals. Successful implementation of these integrated testing strategies will reduce animal use and overall testing costs, and likely decrease the time required to assess potential immunotoxicity.

4.2. Current considerations for application of In vitro approaches for the assessment of immunotoxicity

Growing political and practical resistance to toxicity testing in animals has driven the development of alternative methods for the screening and prioritization of toxicants, including those causing immunosuppression and allergic hypersensitivity. Significant progress has been made in developing *in vitro* assays that will reduce animal use and testing costs while increasing throughput in screening and prioritization efforts (dos Santos et al., 2009; Galbiati et al., 2010; Martin et al., 2010). Development efforts have focused on assays to detect hypersensitivity and immunosuppression to meet existing regulatory demands, although assessment of DIT, immune stimulation and autoimmunity are also needed (Gennari et al., 2005). Existing *in vitro* methods have some limitations, including the fact that they may not capture the complex cellular interactions that take place within the *in vivo* setting. In addition, the use of primary cells for *in vitro* research can entail challenges and inconsistencies due to constraints around securing human samples or the need to differentiate progenitor cells. Although continuous cell lines are not the physiological equivalents of primary cells, many have proven to be valid surrogates but require appropriate characterization to ensure they are accurately recapitulating the normal immunological responses and functions.

As part of the evaluation of the use of *in vitro* models for assessment of potential immunotoxicity, there are several overarching factors that always need to be considered. Efforts to develop *in vitro* screening assays for potential allergic contact sensitizers were aided by the relatively limited and well-understood immunological events that lead to contact sensitization and as a result, a number of assays are undergoing validation efforts. These include assays to assess the ability to interact with host proteins or cells in ways that initiate and propagate contact hypersensitivity. However, as the underlying mechanisms of chemical immunomodulation are often complex and impact multiple interdependent cell types and processes, it is likely that a battery of tests, that reflect various modes of action will be required to detect and characterize

Table 3
Key events in chemical-induced skin sensitization and *in vitro* opportunities.

Key event	*In vitro* Opportunities	References
1. Skin penetration	Human skin biopsis, pig skin; reconstituted human epidermis	(Basketter et al., 2007; OECD, 2004)
2. Binding to macro-molecules (i.e., proteins)	QSAR/Expert systems; peptide binding assay	Reviewed by (Gerberick et al., 2008; Patlewicz et al., 2007)
3. Local trauma and generation of danger signals	KeratinoSens^{TM}; KC activation; NCTC2544 IL-18 assay; KC gene expression profile	Reviewed by (Aeby et al., 2010; Corsini and Roggen, 2009; Galbiati et al., 2010)
4. Langerhans cells maturation and migration	DC-like up-regulation of class II antigens and costimulatory molecules, i.e., CD54, CD86; Cytokine release, i.e., IL-8; LC-like MUTZ-3 cells migration assay; DC-like gene expression profile	Reviewed by (Casati et al., 2005; dos Santos et al., 2009; Galbiati et al., 2010)
5. Antigen presentation to T_H cells and memory T-cell generation	*In vitro* T-cell activation	Reviewed by (Martin et al., 2010)

chemical-induced imunosuppression or immunostimulation. Before starting with *in vitro* tests, bioavailability should be considered. If the compound does not have appreciable bioavailability, immunotoxicity is unlikely to occur. Furthermore, it should be recognized that *in vitro* systems may lack the appropriate metabolic components that could enhance or reduce the immunotoxic potential in an *in vivo* system. As a general strategy, *in vitro* testing for direct immunotoxicity should be done in a tiered approach. A potential **first tier** in this process could include measuring myelotoxicity (bone marrow toxicity) in response to the test material of interest. Compounds that are capable of damaging or destroying the bone marrow will often have a profound immunotoxic effect by limiting or eliminating immune effector cells. Therefore, if a compound is myelotoxic the material will be a *de facto* immunotoxicant. A methodology for evaluating myelotoxicity *in vitro* using bone marrow culture systems is well characterized (Haglund et al., 2010; Pessina et al., 2001).

A negative finding in the first tier would not necessarily be conclusive, because some chemicals that spare bone marrow may damage or destroy lymphocytes, which are the primary effectors and regulators of acquired immunity. Compounds could therefore be tested for lymphotoxicity as part of the **second tier**. Toxicity may result from the destruction of rapidly dividing cells by necrosis or apoptosis; alternatively, chemicals may interfere with cell activation affecting signal transduction pathways. A variety of methods are available for assessing cell viability (e.g., colorimetric, flow cytometric assays, etc.). After myelotoxicity and overt cytotoxicity are excluded as endpoints, basic immune cells functionality should then be assessed by performing specific functional assays to characterize the nature of the immunotoxic effects as part of the **third tier**. Using non-cytotoxic concentrations of the tested chemicals is an important component of these functional evaluations (viability >80%). Table 3 presents some key targets in chemical-induced immunosuppression and *in vitro* opportunities to assess relevant immune functional parameters, including proliferative responses, cytokine production, NK cell activity, and the T-dependent antibody response or TDAR (Gennari et al., 2005; House, 2000; Koeper and Vohr, 2009). Consistent with the important role that the TDAR plays in most regulatory guidelines described elsewhere in this paper, studies have shown that the *in vitro* TDAR has a high sensitivity and specificity, and is a promising assay for the prediction of immunosuppressive properties of chemicals and drugs (Koeper and Vohr, 2009). Moreover, Fischer et al. (2011) concluded that the investigation of *in vitro* antibody responses is a sensitive and reliable approach for detection of compound - specific effects on the immune system, and that the implementation of this endpoint in routine toxicology also enables refinement of existing *in vivo* studies by reducing the numbers of animals.

While the primary focus of this workshop was on immunosuppression, it is important to note that chemically-induced skin sensitization potential is an endpoint that needs to be assessed within the framework of existing and forthcoming legislation (e.g., the EU REACH regulations, the 7th Amendment to the EU Cosmetics Directive) (Corsini and Roggen, 2009), and, as a result, there has been a lot of progress towards integrating *in vitro* and alternative approaches. Allergic contact dermatitis (ACD) is to a considerable extent a preventable disease. A reduction of ACD can be achieved by: correct detection of skin sensitizers; characterization of potency; understanding of human skin exposure; and, the application of adequate risk assessment and management strategies. A range of *in vivo* methods have been proven to be very accurate in terms of identifying chemicals that possess skin sensitizing properties, including the murine local lymph node assay (LLNA) (Kimber and Weisenberger, 1989). However, the current challenge is to obtain the same quality of information on the potency of skin sensitizing chemicals using *in silico* and *in vitro* methods. In the last decade important progress has resulted in the development of alternative test methods that could make a valuable contribution to the replacement of the existing animal models (dos Santos et al., 2009; Galbiati et al., 2010; Luebke, 2012; Martin et al., 2010).

At present, several non-animal test methods, namely the Direct Peptide Reactivity Assay, the KeratinoSens assay, the Myeloid U937 Skin Sensitization Test and the human Cell Line Activation Test are under formal validation at European Centre for the Validation of Alternative Methods (EURL-ECVAM), for their potential to predict skin sensitization potential. The Myeloid U937 Skin Sensitization Test has been, however, recently stopped due to transferability problems. Results are anticipated by the end of 2013, although it will likely take at least another 7–9 years for the full replacement of the *in vivo* animal models currently used to assess sensitization (Adler et al., 2011; Aeby et al., 2010). Another method, the THP-1 IL-8 Luc assay (Takahashi et al., 2011), is currently under validation by the Japanese Center for the Validation of Alternative Methods, and results are anticipated in 2013.

4.2.1. Opinions from workshop participants

Meeting participants agreed that pursuing *in vitro* approaches to immunotoxicity hazard is warranted, and that published proof-of-concept studies suggest that a number of assay methods hold particular promise. The relatively straightforward primary events leading to induction of allergic contact hypersensitivity have helped to simplify the technical and theoretical aspects of assay design and as a result assays for this endpoint have seen good progression. However, the complicated nature of other immunotoxicity responses has limited standardization of general and functional assays. Overall, participants agreed that at present, *in vitro* data are not suitable for risk assessment. Additional development and evaluation of *in vitro* methods was deemed necessary, coincident with development and publication of immunotoxicity testing databases that would allow for a more standardized evaluation of the predictive value of the developing *in vitro* assays. When feasible, concurrent testing using *in vitro* and traditional *in vivo* methods were recommended to determine the accuracy and precision of new methods. In addition, groups concluded that progress in

developing useful *in vitro* approaches would be aided by experts providing recommendations for the application and execution of *in vitro* assays.

5. Conclusions

Overall, the workshop provided an invaluable opportunity for key stakeholders from government, academia and industry (both chemical and pharmaceutical) to discuss current practices for the assessment of immunotoxicity for environmental chemicals. Diverse perspectives were shared from these sectors on experiences with the use of standard toxicology testing endpoints and functional immune assays, as well as tiered-based testing strategies and developing assays and study designs.

With respect to the concept of routine functional immunotoxicity testing for environmental chemicals, participants agreed that emerging evidence may indicate that a trigger-based approach is appropriate, which is consistent with the experience and approach used by the pharmaceutical industry under the ICH S8 guidance. However, it was also recognized that defined criteria need to be established to ensure the consistent and transparent interpretation of the potential need for additional immunotoxicity testing. Overall, these perspectives are supported both by a recent retrospective analysis of 171 immunotoxicity studies (155 chemicals) that was completed by the EPA (US EPA, 2013) and the retrospective analysis from Crop Life America described above (Gehen et al., submitted for publication).

With regard to DIT, participants agreed that assessment of DIT represents a valuable means to evaluate multiple critical windows during immune system development. Extensive discussion and evolution in proposed DIT test approaches has occurred over the last decade, including discussions on study design and evaluation age. The EOGRTS represents one approach that allows for the evaluation of DIT as part of an integrated study design; however, the complexity of the EOGRTS is such that requests for this assay should only be made after careful consideration of the study objectives and needs.

With respect to integrating functional immunotoxicity endpoints into standard adult rodent repeat dose studies (Ladics et al., 1998; Ladics et al., 1995), participants agreed that this approach should be considered as appropriate for fulfilling any immunotoxicity testing requirements, while also maximizing information collection and reducing overall animal use.

Participants also recognized the significant efforts in the area of *in vitro* approaches for evaluating toxicity, along with the numerous factors that continue to drive research in this area. In the case of immunotoxicity, considerable advancements have been made with *in vitro* testing approaches for contact sensitization, as compared to *in vitro* assays for immunosuppressant hazard identification. Efforts to develop reliable alternative methods to identify contact sensitizers have benefitted from the knowledge and compartmentalization of key steps in the development of skin sensitization, whereas assays for immunosuppression will need to address multiple complex pathways. This realization is consistent with the opinions of an OECD immunotoxicity workgroup report on alternative testing methods for immunotoxicity assessment that is nearing finalization.

An active dialog among toxicologists continues regarding potential opportunities to change our approaches to chemical hazard characterization and risk assessments (Andersen and Krewski, 2010; Patlewicz and Lander, 2013). In the case of pesticides, current registration requirements involve the completion of a substantial number of toxicity studies. As additional experience is gained with current testing requirements and as new assays and technologies are developed that can advance our testing approaches, it is critical for all stakeholders to engage in active dialog about potential opportunities to advance our current testing approaches to increase efficiency while maintaining the utility of the data for hazard identification and risk assessment. The information shared and the discussions that took place during this workshop have provided an important exchange that will help to shape the future of immunotoxicity testing for environmental chemicals.

6. Disclaimers

This article may be the work product of an employee or group of employees of the National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH), however, the statements, opinions or conclusions contained therein do not necessarily represent the statements, opinions or conclusions of NIEHS, NIH or the United States government.

This report has been reviewed by the Environmental Protection Agency's Office of Research and Development and Office of Pesticides Programs, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Agency nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

References

Adler, S. et al., 2011. Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. Arch. Toxicol. 85, 367–485.
Aeby, P. et al., 2010. Identifying and characterizing chemical skin sensitizers without animal testing: Colipa's research and method development program. Toxicol. In Vitro 24, 1465–1473.
Andersen, M.E., Krewski, D., 2010. The vision of toxicity testing in the 21st century: moving from discussion to action. Toxicol. Sci. 117, 17–24.
Basketter, D. et al., 2007. Skin sensitisation and epidermal disposition: the relevance of epidermal disposition for sensitisation hazard identification and risk assessment. The report and recommendations of ECVAM workshop 59. Altern. Lab. Anim. 35, 137–154.
Billington, R. et al., 2010. The mouse carcinogenicity study is no longer a scientifically justifiable core data requirement for the safety assessment of pesticides. Crit. Rev. Toxicol. 40, 35–49.
Bugelski, P.J. et al., 2010. Critical review of preclinical approaches to evaluate the potential of immunosuppressive drugs to influence human neoplasia. Int. J. Toxicol. 29, 435–466.
Bunn, T.L. et al., 2001. Exposure to lead during critical windows of embryonic development: differential immunotoxic outcome based on stage of exposure and gender. Toxicol. Sci. 64, 57–66.
Burns-Naas, L.A. et al., 2008. What's so special about the developing immune system? Int. J. Toxicol. 27, 223–254.
Casati, S. et al., 2005. Dendritic cells as a tool for the predictive identification of skin sensitisation hazard. Altern. Lab. Anim. 33, 47–62.
Chapin, R.E. et al., 1997. The effects of perinatal/juvenile methoxychlor exposure on adult rat nervous, immune, and reproductive system function. Fundam. Appl. Toxicol. 40, 138–157.
Cooper, R.L. et al., 2006. A tiered approach to life stages testing for agricultural chemical safety assessment. Crit. Rev. Toxicol. 36, 69–98.
Corsini, E., Roggen, E.L., 2009. Immunotoxicology: opportunities for non-animal test development. Altern. Lab. Anim. 37, 387–397.
Dellarco, V.L. et al., 2010. A retrospective analysis of toxicity studies in dogs and impact on the chronic reference dose for conventional pesticide chemicals. Crit. Rev. Toxicol. 40, 16–23.
DeWitt, J.C. et al., 2012. Current status of developmental immunotoxicity: early-life patterns and testing. Toxicol. Pathol. 40, 230–236.
Dietert, R.R. et al., 2000. Workshop to identify critical windows of exposure for children's health: immune and respiratory systems work group summary. Environ. Health Perspect. 108 (Suppl 3), 483–490.
dos Santos, G.G. et al., 2009. Progress on the development of human in vitro dendritic cell based assays for assessment of the sensitizing potential of a compound. Toxicol. Appl. Pharmacol. 236, 372–382.
Fegert, I. et al., 2012. Feasibility of the extended one-generation reproductive toxicity study (OECD 443). Reprod. Toxicol. 34, 331–339.
Fischer, A. et al., 2011. Specific antibody responses of primary cells from different cell sources are able to predict immunotoxicity in vitro. Toxicol. In Vitro 25, 1966–1973.
Galbiati, V. et al., 2010. Present and future of in vitro immunotoxicology in drug development. J. Immunotoxicol. 7, 255–267.
Gehen, S. C., et al., in submission. Retrospective evaluation of the impact of functional immunotoxicity testing on pesticide hazard identification and risk assessment.

Gehrs, B.C., Smialowicz, R.J., 1999. Persistent suppression of delayed-type hypersensitivity in adult F344 rats after perinatal exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Toxicology 134, 79–88.
Gennari, A. et al., 2005. The use of in vitro systems for evaluating immunotoxicity: the report and recommendations of an ECVAM workshop. J. Immunotoxicol. 2, 61–83.
Gerberick, F. et al., 2008. Chemical reactivity measurement and the predicitve identification of skin sensitisers. The report and recommendations of ECVAM workshop 64. Altern. Lab. Anim. 36, 215–242.
Gore, E.R., 2006. Immune function tests for hazard identification: a paradigm shift in drug development. Basic Clin. Pharmacol. Toxicol. 98, 331–335.
Haglund, C. et al., 2010. The FMCA-GM assays, high throughput non-clonogenic alternatives to CFU-GM in preclinical hematotoxicity testing. Toxicol. Lett. 194, 102–107.
Holsapple, M.P., 2002. Developmental immunotoxicology and risk assessment: a workshop summary. Hum. Exp. Toxicol. 21, 473–478.
Holsapple, M.P., 2003. Developmental immunotoxicity testing: a review. Toxicology 185, 193–203.
Holsapple, M.P. et al., 2005. A proposed testing framework for developmental immunotoxicology (DIT). Toxicol. Sci. 83, 18–24.
Holsapple, M.P. et al., 2004. Symposium summary: children's health risk–what's so special about the developing immune system? Toxicol. Appl. Pharmacol. 199, 61–70.
Holsapple, M.P. et al., 2007. Development of a framework for developmental immunotoxicity (DIT) testing. In: Luebke, R. et al. (Eds.), Immunotoxicology and Immunopharmacology. CRC Press, Boca Raton, pp. 347–361.
House, R.V., 2000. An overview of in vitro/ex vivo assays for preclinical evaluation of immunomodulation. Hum. Exp. Toxicol. 19, 246–250.
IARC, 2006. Inorganic and Organic Lead Compounds. vol. 87. World Health Organization International Agency for Research on Cancer, Lyon, France
ICH, 2005. Harmonised tripartite guideline: immunotoxicity studies for human pharmaceuticals, S8.
ICICIS Group Investigators, 1998. Report of validation study of assessment of direct immunotoxicity in the rat. Toxicology 125, 183–201.
Kimber, I., Weisenberger, C., 1989. A murine local lymph node assay for the identification of contact allergens. Assay development and results of an initial validation study. Arch. Toxicol. 63, 274–282.
Koeper, L.M., Vohr, H.W., 2009. Functional assays are mandatory for a correct prediction of immunotoxic properties of compounds in vitro. Food Chem. Toxicol. 47, 110–118.
Ladics, G.S., 2007. Primary immune response to sheep red blood cells (SRBC) as the conventional T-cell dependent antibody response (TDAR) test. J. Immunotoxicol. 4, 149–152.
Ladics, G.S. et al., 2005. Developmental toxicology evaluations–issues with including neurotoxicology and immunotoxicology assessments in reproductive toxicology studies. Toxicol. Sci. 88, 24–29.
Ladics, G.S. et al., 2000. Characterization of an approach to develoopmental immunotoxicology assessment in the rat using SRBC as the antigen. Toxicol. Mech. Methods 10, 283–311.
Ladics, G.S. et al., 1998. Further evaluation of the incorporation of an immunotoxicological functional assay for assessing humoral immunity for hazard identification purposes in rats in a standard toxicology study. Toxicology 126, 137–152.
Ladics, G.S. et al., 1995. Possible incorporation of an immunotoxicological functional assay for assessing humoral immunity for hazard identification purposes in rats on standard toxicology study. Toxicology 96, 225–238.
Luebke, R., 2012. Immunotoxicant screening and prioritization in the twenty-first century. Toxicol. Pathol. 40, 294–299.
Luebke, R.W. et al., 2006. The comparative immunotoxicity of five selected compounds following developmental or adult exposure. J. Toxicol. Environ. Health B Crit. Rev. 9, 1–26.
Luster, M.I. et al., 2003. Consensus workshop on methods to evaluate developmental immunotoxicity. Environ. Health Perspect. 111, 579–583.
Luster, M.I. et al., 1988. Development of a testing battery to assess chemical-induced immunotoxicity: national toxicology program's guidelines for immunotoxicity evaluation in mice. Fundam. Appl. Toxicol. 10, 2–19.
Luster, M.I. et al., 1992a. Qualitative and quantitative experimental models to aid in risk assessment for immunotoxicology. Toxicol. Lett. 64–65, 71–78, Spec No.
Luster, M.I. et al., 1993. Risk assessment in immunotoxicology. II. Relationships between immune and host resistance tests. Fundam. Appl. Toxicol. 21, 71–82.
Luster, M.I. et al., 1992b. Risk assessment in immunotoxicology. I. Sensitivity and predictability of immune tests. Fundam. Appl. Toxicol. 18, 200–210.
Martin, S.F. et al., 2010. T-cell recognition of chemicals, protein allergens and drugs: towards the development of in vitro assays. Cell. Mol. Life Sci. 67, 4171–4184.
Marty, M. S., Neal, B.H., Zablotny, C.L., Yano, B.L., Andrus, A.K., Woolhiser, M.R., Boverhof, D.R., Saghir, S.A., Perala, A.W., Passage, J.K., Lawson, M.A., Bus, J.S., Lamb, J.C., and L. Hammond, 2013. An F1-extended one generation reproductive toxicity study (EOGRTS) in CRL:CD(SD) rats with 2,4-dichlorophenoxyacetic acid (2,4-D). in submission.
Miller, T.E. et al., 1998. Developmental exposure to lead causes persistent immunotoxicity in Fischer 344 rats. Toxicol. Sci. 42, 129–135.
NRC, 2007. Toxicity Testing in the 21st Century: A Vision and a Strategy. National Academies Press, Washington, DC.
OECD, 1995. Guideline for the Testing of Chemicals 421: Reproduction/Developmental Toxicity Screening Test. Organisation for Economic Cooperation and Development, Paris, France.
OECD, 1996. Guideline for the Testing of Chemicals 422: Combined Repeated Dose Toxicity Study with the Reproduction/Developmental Toxicity Screening Test. Organisation for Economic Cooperation and Development, Paris, France.
OECD, 2004. Guideline for the Testing of Chemicals 428: Skin Absorpotion. In Vitro Method. Organisation for Economic Cooperation and Development, Paris, France.
OECD, 2011a. Draft Guidance Document 151 in Support of OECD Test Guideline 443 on an Extended one Generation Reproductive Toxicity Study. Organisation for Economic Cooperation and Development, Paris, France.
OECD, 2011b. Guidance Document 117 on the Current Implementation of Internal Triggers in Test Guideline 443 for an Extended one Generation Reproductive Toxicity Study in the United States and Canada., vol. 117.
OECD, 2011c. Guideline for the Testing of Chemicals 443: Extended One-Generation Reproductive Toxicity Study. Organisation for Economic Cooperation and Development, Paris, France.
Patlewicz, G. et al., 2007. An evaluation of selected global (Q)SARs/expert systems for the prediction of skin sensitisation potential. SAR QSAR Environ. Res. 18, 515–541.
Patlewicz, G.Y., Lander, D.R., 2013. A step change towards risk assessment in the 21st century. Front Biosci. (Elite Ed). 5, 418–434.
Pessina, A. et al., 2001. Prevalidation of a model for predicting acute neutropenia by colony forming unit granulocyte/macrophage (CFU-GM) assay. Toxicol. In Vitro 15, 729–740.
Piersma, A.H. et al., 2011. Combined retrospective analysis of 498 rat multi-generation reproductive toxicity studies: on the impact of parameters related to F1 mating and F2 offspring. Reprod. Toxicol. 31, 392–401.
Richter-Reichhelm, H.B. et al., 1995. Validation of a modified 28-day rat study to evidence effects of test compounds on the immune system. Regul. Toxicol. Pharmacol. 22, 54–56.
Rorije, E. et al., 2011. On the impact of second generation mating and offspring in multi-generation reproductive toxicity studies on classification and labelling of substances in Europe. Regul. Toxicol. Pharmacol. 61, 251–260.
Smialowicz, R.J. et al., 1994. Effects of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) on humoral immunity and lymphocyte subpopulations: differences between mice and rats. Toxicol. Appl. Pharmacol. 124, 248–256.
Smialowicz, R.J. et al., 2001. The effects of perinatal/juvenile heptachlor exposure on adult immune and reproductive system function in rats. Toxicol. Sci. 61, 164–175.
Takahashi, T. et al., 2011. An in vitro test to screen skin sensitizers using a stable THP-1-derived IL-8 reporter cell line, THP-G8. Toxicol. Sci. 124, 359–369.
US EPA, OPPTS Harmonized Test Guideline: OPPTS 870.7800. Immunotoxicity. EPA/712/C- 98/351. United States Environmental Protection Agency, Office of Chemical Safety and Pollution Prevention, Washington, DC, 1998.
US EPA, 2007. Environmental Protection Agency, 40 CFR Parts 9 and 158. Pesticides; Data Requirements for Conventional Chemicals. Federal Register, Vol. 72, No 207 / Friday, October 26.
US EPA, 2013. A Retrospective Analysis of the Immunotoxicity Studies (OCSPP Test Guildeline No. 870.7800). United States Environmental Protection Agency, Health Effects Division, Office of Pesticide Programs
van Loveren, H., Vos, J.G., 1989. Immunotoxicological considerations: a practical approach to immunotoxicity testing in the rat. In: Dayan, A.D., Paine, A.J. (Eds.), Advances in Applied Toxicology. Taylor & Francis, London, pp. 143–165.
Vos, J.G., 1977. Immune suppression as related to toxicology. CRC Crit. Rev. Toxicol. 5, 67–101.
Vos, J.G., 1980. Immunotoxicity assessment: screening and function studies. Arch. Toxicol. Suppl. 4, 95–108.
Vos, J.G., Van Loveren, H., 1987. Immunotoxicity Testing. In: Burger, E.J., et al., (Eds.), The Rat. Advances In Modern Environmental Toxicology. vol. 13. Princeton Scientific Publishing, Princeton, N.J., pp. 167–180.
Weaver, J.L. et al., 2005. Development of the ICH guidelines for immunotoxicology evaluation of pharmaceuticals using a survey of industry practices. J. Immunotoxicol. 2, 171–180.
White, K.L., 1992. Validation: state of the art United States national toxicology program & fischer 344 rat cyclosporine–a study. J. Toxicol. Clin. Exp. 12, 429–447.
Woolhiser, M.R., et al., 2007. Feasibility study to conduct the antibody plaque-forming cell assay and spleen histopathology simultaneously using Fischer 344 rats. 46th Annual Meeting of the Society of Toxicology, Charlotte, NC.