

## The Accuracy of Extended Histopathology to Detect Immunotoxic Chemicals

D. R. Germolec,<sup>\*,1</sup> M. Kashon,<sup>†</sup> A. Nyska,<sup>‡</sup> C. F. Kuper,<sup>§</sup> C. Portier,<sup>¶</sup> C. Kommineni,<sup>||</sup> K. A. Johnson<sup>|||</sup> and M. I. Luster<sup>||||</sup>

<sup>\*</sup>Laboratory of Molecular Toxicology/National Toxicology Program, National Institute of Environmental Health Sciences, RTP, North Carolina; <sup>†</sup>Biostatistics Branch, National Institute for Occupational Safety and Health, Morgantown, West Virginia; <sup>‡</sup>Laboratory of Experimental Pathology, National Institute of Environmental Health Sciences, RTP, North Carolina; <sup>§</sup>TNO Nutrition and Food Research, Zeist, The Netherlands; <sup>¶</sup>Laboratory of Computational Biology and Risk Assessment, National Institute of Environmental Health Sciences, RTP, North Carolina; <sup>||</sup>Pathology and Physiology Research Branch, National Institute for Occupational Safety and Health, Morgantown, West Virginia; <sup>|||</sup>Toxicology & Environmental Research and Consulting, The Dow Chemical Company, Midland, Michigan; and <sup>||||</sup>Toxicology and Molecular Biology Branch, National Institute for Occupational Safety and Health, Morgantown, West Virginia

Received May 17, 2004; accepted August 24, 2004

The accuracy of extended histopathology to detect immunotoxic chemicals in female B6C3F1 mice was evaluated under the auspices of the National Toxicology Program (NTP). A workgroup was formed consisting of four pathologists who conducted extended histopathological evaluation of lymphoid tissues obtained from a subset of NTP toxicology studies, in which previously detailed immunotoxicity assessment was performed. In addition, a positive control data set of three known immunosuppressive agents, one negative control data set, and an additional negative control group composed of the vehicle only treated groups were included. Data obtained from extended histopathology evaluations were compared to more traditional immune test results (both functional and nonfunctional) from previously conducted immunotoxicity assessments. Analyses of the data indicated that the ability to identify immunotoxic chemicals using histological endpoints decreased linearly as the level of stringency used to determine significant histopathological changes increased. A relatively high (80%) accuracy level was achieved when histological changes were considered *in toto* (i.e., any histological abnormality in the three tissues examined), using minimal or mild criteria for scoring. When minimal or mild histological changes were considered significant for a specific tissue, a 60% level of accuracy in identifying immunotoxic chemicals was obtained as compared to a 90% accuracy level that was achieved with this data set using the antibody plaque forming cell response, considered to represent the most predictive functional test. A minimal classification was obtained in the analyses of the negative control groups, suggesting that use of the minimal classification for hazard identification is inappropriate as it will likely result in a high incidence of false positives. This was not the case when mild classifications were used as an indicator of significance, which in most instances allowed the successful identification of negatives. When moderate to marked histopathological changes were used to identify immunotoxic chemicals, the level of accuracy that could be achieved was poor. A considerably higher level of accuracy

was obtained for the positive control data set than the test chemical data set suggesting that the ability to detect an immunotoxic agent histologically is proportional to the potency of the immunotoxic agent. Comparison of immune function test results and histopathological results obtained from the high-dose treatment groups and the lower-dose treatment group did not reveal any significant differences between the two endpoints to predict immunotoxicity as a function of dose. Of the three lymphoid organs examined, (i.e., lymph node, thymus, and spleen), the most consistent and discernible histological lesions were observed in the thymus cortical region. These lesions correlated with thymus: body weight ratios and to a slightly lesser extent, the antibody plaque forming cell response. Addition of general toxicological endpoints such as body weight and leukocyte counts did not significantly improve the sensitivity of extended histopathology for this data set. Taken together, these data suggest that, while not as sensitive as functional analyses, extended histopathology may provide a reasonable level of accuracy as a screening test to identify immunotoxic chemicals, provided the level of stringency used to score histological lesions is carefully considered to allow for detection of immunotoxic agents while limiting false positives.

**Key Words:** extended histopathology; immunotoxicity; screening tests; B6C3F1 mice; safety assessment.

Just as a well-functioning immune system is central to good health, maladaptive immunological alterations may influence the etiology, progression, and/or severity of a broad range of disorders that include infectious diseases, certain cancers, autoimmune disease, and chronic inflammatory disorders. As such, development of sensitive measures of the immune response in humans and experimental animal models is of continuing interest not only to those in the fields of immunotoxicology and safety assessment, but also to investigators in the areas of AIDS, geriatric immunology, allergy, neuroimmunology, and rheumatology. Attempts to establish immune measures has led some investigators to propose the use of extended histopathology assessment, which involves a detailed examination of key primary and secondary lymphoid organs in experimental animals

<sup>1</sup> To whom correspondence should be addressed at Laboratory of Molecular Toxicology, National Institute of Environmental Health Sciences, 111 Alexander Drive, PO Box 12233, Research Triangle Park, NC 27709. Fax: (919) 541-0870. E-mail: germolec@niehs.nih.gov.

for chemical-induced lesions. Extended histopathology, alone or in combination with other standard subchronic toxicity tests, could provide an opportunity to readily assess immunotoxicity as part of routine toxicology studies without the need to conduct special tests or incorporate additional animals as is commonly suggested (Hastings, 2002; ICICIS, 1998; OECD, 1995, USEPA, 1998). However, since extended histopathology does not directly measure immune function, i.e., response to antigenic stimulation, there is considerable debate regarding its predictive value.

To help determine the utility of the extended histopathology approach a validation effort was initiated under the auspices of the National Toxicology Program (NTP), in which a workgroup consisting of four pathologists was formed to conduct extended histopathology using a previously recommended protocol (Kuper *et al.*, 2000). This evaluation was conducted on 10 different test chemicals, previously evaluated in depth by the NTP for functional immunotoxicity and published elsewhere (Burns *et al.*, 1994; Cao *et al.*, 1990; Karrow *et al.*, 2000a,b; NTP, 1988a,b, 1989; Phillips *et al.*, 1997; Sikorski *et al.*, 1989). In an earlier publication, which focused on the utility of specific histologic parameters and the agreement found between pathologists, it was observed that the ability to identify histopathological changes in lymphoid organs was dependent on the experience/training of the individual pathologist, the severity of the lesion, and the particular lymphoid organ involved (Germolec *et al.*, 2004). Agreement between pathologists was highest when evaluating the thymus, in particular thymus cortical cellularity, and lower within all of the compartments examined in the spleen and lymph nodes. In the present studies, the ability of enhanced pathology to predict immunotoxicity is addressed by comparing the results obtained from extended histopathology evaluation with those previously obtained from in-depth immunotoxicity testing from the experimental (test) and positive control data sets.

## MATERIALS AND METHODS

**Experimental, vehicle control, and positive control data sets.** Lymphoid tissues were examined from four to eight animals in each treatment group from a subset of chemical test studies assessed previously in the NTP immunotoxicology program. Briefly, the experimental data set consisted of 10 test chemicals that were selected from a larger data set based upon completeness of the study. The subset of selected chemicals included compounds that were known to inhibit humoral, cell-mediated, or innate immune function; known to have no effects on immune function or that had been shown to be immunostimulatory (Table 1). All studies were conducted in female B6C3F1 mice at three dose levels, with the highest dose selected to be slightly below that which caused evidence of overt toxicity (e.g., decreased body weight gain, liver enzyme changes) in separately conducted dose-range studies (unpublished data). Although routine toxicological assessment in rodent models often tests compounds at the maximum tolerated dose (MTD), there is evidence that high doses may produce a neuroendocrine stress response that does not occur at lower doses (Brown *et al.*, 1988; Clement, 1985; Kunimatsu *et al.*, 1996). These nonspecific stress responses can lead to erroneous identification of the test compound as immunotoxic (Pruett *et al.*, 1993, 1999, 2000). Within the NTP, preliminary dose-range studies are routinely

conducted prior to immunotoxicity studies and, for the reasons stated above, the highest dose is set slightly below the MTD and at doses where body weight changes would not be  $\geq 10\%$ . Nine of the 10 chemicals examined were classified as immunotoxic, based upon whether the test material produced a significant dose-response effect ( $p \leq 0.05$ ) in any one or more of a number of previously defined immune parameters or significantly ( $p \leq 0.05$ ) altered two or more immune function tests at the highest dose of the chemical evaluated (Luster *et al.*, 1992). Eight of these caused suppression in functional assays, and one, thalidomide, demonstrated immunostimulatory activities (Table 1). A negative chemical, aldicarb oxime, was also included in the experiment. To help determine the potential likelihood of falsely identifying chemicals as immunotoxic and to provide accurate estimates for background levels of response, an additional negative control data set was established which was derived from tissues of animals treated with the vehicles for the 10 chemicals studied. Histopathological examination and analysis also included tissue from animals administered the positive control chemicals, cyclophosphamide, methotrexate, and sodium arsenite, which were used in the initial screening studies. The positive controls were examined at only one dose level and were evaluated along with six of the 10 chemicals in the experimental data set, with cyclophosphamide as the positive control for four of the test chemicals. Each of these was treated as an independent study for the analysis of the positive control data set. Detailed information on each of the chemicals and doses used in these studies can be found in Table 1 and in the Supplementary Appendix of the publication by Germolec *et al.* (2004) available online at [www.toxsci.oupjournals.org](http://www.toxsci.oupjournals.org). More extensive information on the duration, route of exposure, and other study parameters can be found in the following references: Burns *et al.*, 1994; Cao *et al.*, 1990; Karrow *et al.*, 2000a,b; NTP, 1988a,b, 1989; Phillips *et al.*, 1997; Sikorski *et al.*, 1989.

**Immune and host resistance tests.** The immune assays used to identify immunotoxic chemicals included both functional (i.e., IgM antibody plaque forming cell [PFC] response, natural killer [NK] cell activity, cytotoxic T lymphocyte [CTL] cytotoxicity and delayed hypersensitivity responses [DHR]), and quantitative immune measures (i.e., lymphoid organ weights, lymphocyte phenotype analysis, white blood cell counts [WBC]), and have been described in detail elsewhere (Luster *et al.*, 1988). Host resistance tests data were available for eight of the chemicals tested and were included in the analysis independent of the specific host resistance model employed. Details of the host resistance models employed in evaluation of immunotoxicity have recently been reviewed (Germolec, 2004). For the test chemicals in this data set host resistance assays included infection with either *Listeria monocytogenes*, *Streptococcus pneumoniae*, or *Plasmodium yoelii*, or challenge with either PYB6 fibrosarcoma or B16F10 melanoma tumor cells, with the selection being based upon the immune test results.

**Tissue preparation.** Tissues represented archived samples from NTP immunotoxicity studies collected at the termination of each study under GLP and AALAC guidelines according to Standard Operating Procedures developed under an NTP contract. Thymus, spleen, and the complete chain of the superior mesenteric lymph nodes were collected and fixed in 10% neutral-buffered formalin. One middle cross section from the spleen, both lobes of the thymus, and the mesenteric lymph nodes were embedded in paraffin, and five 5–6 micron sections were prepared and stained with hematoxylin and eosin (H&E) for histopathological evaluation.

**Histological analysis.** Details of the histological analyses and workgroup's strategy have been previously published (Germolec *et al.*, 2004). Briefly, for each pathologist, a slide set was generated for each of the chemicals. Although no information was provided concerning the chemical identity of the test compounds before commencing their evaluation of the tissues, each pathologist received data identifying positive control, negative control, and test groups, including dose levels as well as organ weights. A semiquantitative assessment was used to estimate the histopathological changes within different anatomical compartments of the lymphoid tissues. The grading scheme consisted of ordinal categories ranging from 0 (no effect) to 4 (severe) and an indicator as to whether the effect was increased or decreased relative to normal tissue. Histopathological evaluations took into consideration changes in cell density or changes in the anatomical compartment size. The pathologists were also instructed to add comments that were not quantifiable, but considered important for proper

TABLE 1  
Functional Test Calls

Endpoint measured	1	7.5	15	50	150	300	P_Nitrotoluene (mg/kg/day)	2',3'_Dideoxy- inosine (mg/kg/day)	2',4'_Diamino- toluene (mg/kg/day)	Dideoxy- adenosine (mg/kg/day)	Gallium arsenide (mg/kg/day)	Ribavirin (mg/kg/day)	TCDD (µg/kg/day)	Thalidomide (mg/kg/day)
PFC/spleen	0	0	0	0	0	0	0	1	1	1	1	1	1	1
PFC/10 <sup>6</sup> cells	0	1	0	0	0	0	1	1	1	1	1	1	1	1
Host resistance	0	0	0	0	0	0	0	*	0	0	1	*	0	0
CTL	0	0	0	1	1	1	1	1	1	0	0	0	*	1
MLR	0	0	0	0	0	0	0	1	1	0	1	0	*	1
NK	1	0	0	0	0	0	0	0	1	1	1	0	0	0
Flow_B	1	1	1	0	0	0	0	0	1	1	1	1	*	0
Flow_T	0	0	0	0	0	0	1	0	0	0	0	0	*	0
Spleen/BW ratio	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Spleen cells	0	0	0	1	0	0	0	0	0	0	1	0	0	0
Thymus/BW ratio	0	0	0	0	0	0	0	0	0	0	1	0	1	0
WBC	0	1	0	0	0	0	0	0	0	0	0	0	*	0
Body weight	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Note. Table 1 is a summary showing the immunotoxic call or biological effect for each chemical and dose level for each of the endpoints examined. 0 = No significant effects at the dose indicated. 1 = Statistically significant alteration at  $p < 0.05$  by Student's  $t$ -test. \* = No data for this endpoint or dose.

histopathological assessment, such as “focal increased cellularity of outer thymic cortex” or “increased tingible body macrophages in the thymic cortex.” These comments were not included in the analysis and none were remarkable in nature. Four endpoints were evaluated in the lymph node (L): grade of cellularity in the follicles (FGCD), paracortical areas (PAC), medullary cords (MCC), and sinuses. Five endpoints were evaluated in the spleen (S): cellularity of periarteriolar lymphoid sheaths (PALS), lymphoid follicles (FC), marginal zone (MZ), red pulp (RP), and the number of germinal centers (GC). Three endpoints were evaluated in the thymus (T): cortex cellularity (CC), medullary cellularity (MC), and the cortico-medullary (CM) ratio. Following the microscopic examination, the coded data were transferred to an electronic format, and a formal quality control was conducted on the data entry of the entire set of findings. Lymph node tissue was not available for thalidomide or aldricarb oxime.

**Data analysis.** Histology scores for each chemical/dose combination were generated for each tissue examined and were determined by multiplying the histological grade assigned by the pathologist for the lesion by a weighting factor (derived from the frequency of animals within a dose group with a specific grade) by the histological grade assigned by the pathologist for the lesion. Thus, if two of four animals (50%), in the high dose group were given a grade of 2, and the remaining 50% had a histological grade of 3, the score for that dose group was equal to  $10[(2 \times 2) + (2 \times 3)]$ . The score for each dose group was then corrected for background (subtracting histological scores obtained in animals treated with vehicle only). The final histological score was reduced to five categories as follows: 0 (no effect) histological score equal to 0; 1 (minimal) histological score from 1 to 3; 2 (mild) histological score from 4 to 6; 3 (moderate) histological score from 7 to 9; and 4 (marked) histological score greater than 9 (see Table 2).

Accuracy, defined as the ratio of the number of correctly identified compounds using extended histopathology relative to the total number of compounds evaluated, was determined for both individual tissue-specific endpoints and composite (*in toto*) histopathological scores by evaluating agreement with individual immune tests and the overall “immunotoxic call” which was described previously (Luster *et al.*, 1992, 1993). The Phi correlation coefficient for binary variables was used to calculate the correlation coefficient between the functional test calls and the histopathology calls (Liebetrau, 1983).

## RESULTS

The accuracy of each of the immune tests to identify immunotoxic chemicals, which was based upon previously defined criteria (Luster *et al.*, 1992; Table 1), is shown in Figure 1A for the test chemical data set. There was a large variability in the ability of the individual immune endpoints to accurately identify immunotoxic chemicals, with WBCs flagging only 1 of 10 chemicals and the antibody PFC assay identifying 9 of 10 chemicals in this data set. However, the level of accuracy achieved for each of the tests to identify immunotoxic chemicals was comparable to that reported earlier using a larger ( $n = 51$ ) chemical data set (Luster *et al.*, 1992). An accuracy plot showing the ability of each of the immune tests to predict immunotoxicity derived from the positive control data set is depicted in Figure 1B. In this case, a relatively high degree of accuracy was obtained for most of the immune tests with the exception of NK cell activity. The differences in the level of accuracy obtained between these two data sets most likely reflects the fact that the positive controls included compounds known to significantly suppress immune responses, and represent a very small data set ( $n = 6$ ).

There is no general agreement on the level of histopathological change (number of endpoints altered or severity of lesion) that would constitute a biologically significant immune effect.

Therefore, a continuous scale was developed to allow for determining the accuracy of histopathology to predict immunotoxic chemicals at various levels of severity. Figure 2 provides accuracy plots (i.e., ability to predict a chemical classified as immunotoxic) derived from the average score provided by the four pathologists for each histopathological parameter. Taking the least stringent approach, and accepting a minimal classification as immunotoxic based on any histopathological observation at any grade (i.e., a score of  $>0$ ) within a specific lymphoid organ/endpoint, the accuracy level for the large majority of histopathological endpoints ranged between 40 and 60%. The highest level of accuracy for any specific histological endpoint was observed in the measurement of the thymus cortico-medullary ratio (62%). Taking more stringent criteria to classify immunotoxic chemicals by considering histopathological scores  $\geq 3$  (mild level), the accuracy for the thymus cortico-medullary ratio, while still the highest, were reduced to 47%. At the moderate to marked levels ( $\geq 7$ ), the accuracy ranged from 20 to 40% for thymus endpoints and 5 to 25% for histological measures in the spleen or lymph node. Of the three lymphoid organs evaluated, histopathological measures in the thymus allowed for the highest degree of accuracy, while those in the lymph node the least. It should be noted that the lymph node parameters were not evaluated for thalidomide or aldricarb oxime. Addition of body weight as a general indicator of toxicity to the analyses did not significantly improve the accuracy ratings (data not shown).

Figure 3 provides accuracy plots derived from average scores provided by the four pathologists for each histopathological parameter using the results from the positive control data set. The accuracy with which histopathological measures predict immunotoxicity is significantly improved in this data set compared to the data set derived from the test chemicals. For the positive controls the most accurate parameter was again thymus cortico-medullary ratio, with an accuracy of 86% using a minimal histological score, 69% using mild, and 60% using a moderate score (Fig. 3C). The least accurate predictor for immunotoxicity was the lymph sinus at 33% using a minimal category, 12.5% for mild, and 0% for moderate (Fig. 3A). For all lymphoid organs examined, there were only two specific endpoints (thymus cortico-medullary ratio and thymus cortical cellularity) that achieved accuracy levels greater than 50% using the most stringent criteria (i.e., score of  $>9$ ; marked).

Figure 4 provides accuracy plots derived from the histopathological scores for each of the pathologists obtained from combining all endpoints (*in toto*). Using the least stringent criteria (minimal), which would constitute a lesion in any tissue, at any dose and at any level of severity, the level of accuracy ranged from 70–90%. Considering compounds that would be classified as causing mild histological changes (i.e., score of  $\geq 3$ ), the level of accuracy ranged from 70–80%. For histological scores considered moderate and marked, the level of accuracy among the pathologists was variable, ranging from 20 to 80% and 10 to 50%, respectively.

TABLE 2  
Immunotoxic Calls for each Compound by Dose and Pathologist

Chemical		Aldicarb Oxime mg/kg/day			Oxymetholone mg/kg/day			P_Nitrotoluene mg/kg/day			2',3'_ Dideoxyinosine mg/kg/day			2',4'_ Diaminotoluene mg/kg/day			Dideoxyadenosine mg/kg/day			Gallium Arsenide mg/kg/day			Ribavirin mg/kg/day			TCDD ug/kg/day			Thalidomide mg/kg/day						
Pathologist	Parameter	1	7.5	15	50	150	300	200	400	600	250	500	1000	25	50	100	87.5	175	350	50	100	200	75	150	300	0.1	1	10	50	100	150				
Rater 1	L_FGCD				0	0	2			0	0	1	0	3	0	0	0	1	2	2	1	0	4	0	0	0	0	0	0	0	0				
Rater 2	L_FGCD				0		2					0	0	3	0	0	0		0	2	0	0	0	0	0	0	0	0	0	0	0				
Rater 3	L_FGCD				0	1	1			1	1	1	1	1	0	0	0		0	1	0	1	1	0	0	0	0	0	0	2					
Rater 4	L_FGCD				0	1	1			1	1	0	1	1	0	0	0		0	0	1	0	0	0	0	0	1	3	3						
Rater 1	L_MCC				0	0	1			0	0	0	0	4	0	0	0		0	1	2	0	0	0	0	0	0	0	0	0	0				
Rater 2	L_MCC				0		0			0	1	2	0	0	0	0			1	1	0	0	0	0	0	0	0	0	0	0	0				
Rater 3	L_MCC				0	0	0			0	0	1	0	0	0	0	0		0	0	0	1	0	0	0	0	0	0	0	0	0				
Rater 4	L_MCC				0	2	1			1	1	0	1	1	0	0	0		0	1	0	1	0	0	0	0	0	0	0	0	0				
Rater 1	L_PAC				0	0	2			0	0	0	1	4	0	0	0		0	2	2	0	0	0	0	0	0	0	0	0	0				
Rater 2	L_PAC				0		0			0	0	0	1	0	0	0	0		0	2	0	0	0	0	0	1	0	0	0	0					
Rater 3	L_PAC				0	0	0			1	1	1	1	2	0	0	0		0	1	0	1	0	0	0	0	0	0	0	0	0				
Rater 4	L_PAC				0	1	1			0	0	0	1	1	0	0	0		0	0	1	0	0	0	0	0	0	2	2						
Rater 1	L_SINUS				0	0	0			0	0	0	0	0	0	0	0		0	2	0	0	0	0	0	0	0	0	0	0	0				
Rater 2	L_SINUS				0		0			0	0	1	0	0	0	0	0		1	2	0	0	0	0	0	0	0	0	0	0	0				
Rater 3	L_SINUS				0	0	0			0	0	1	0	1	0	0	0		0	0	0	0	0	0	0	1	0	0	0	0	0				
Rater 4	L_SINUS				0	2	1			1	1	0	1	0	0	0	0		0	1	0	1	0	0	0	0	0	0	2						
Rater 1	S_FC	0	0	0	0	0	0	0	0	2	3	0	2	4	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0			
Rater 2	S_FC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	1	0	1	0	0	0			
Rater 3	S_FC	0	1	0	0	0	1			0	0	0	1	2	0	1	0		0	1	0	1	1	0	0	0	1	1	0	0	1				
Rater 4	S_FC	0	0	1	0	2	2			1	1	0	0	0	0	0	0	1	1	1	2	2	1	0	0	0	0	0	3	0	0	0			
Rater 1	S_GC	0	0	0	0	0	0	0	0	2	4	0	2	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3				
Rater 2	S_GC	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	2	0	0	0	0	0	0	1	1	2				
Rater 3	S_GC	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	1		1	0	0	0	0	0	1	1	0	0	0	0	1	1			
Rater 4	S_GC	0	0	0	0	2	1			1	1	0	0	0	0	0	2	1	1	1	2	2	1	0	0	0	0	2	0	0	0	0			
Rater 1	S_MZ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
Rater 2	S_MZ	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0			
Rater 3	S_MZ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	1	0	0	0	0	0			
Rater 4	S_MZ	0	0	0	0	2	2			0	0	0	2	0	0	0	0	1	1	2	2	0	0	0	0	0	0	3	0	0	0	0			
Rater 1	S_PALS	0	0	0	0	0	0	0	0	2	0	0	2	4	0	0	0	0	0	2	0	0	0	0	0	0	0	1	3	0	0	0			
Rater 2	S_PALS	0	0	1	0	0	0	1	1	2	0	1	2	0	0	0	0	1	1	1	1	3	1	0	0	0	1	1	2	0	1	0			
Rater 3	S_PALS	0	0	0	0	0	1			0	0	0	0	1	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0			
Rater 4	S_PALS	0	0	1	0	2	2			2	2	0	0	2	0	0	0	0	0	0	1	2	0	0	0	0	0	3	0	0	0	0			
Rater 1	S_RP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Rater 2	S_RP	0	0	2	0	0	1	1	2	4	0	0	0	0	0	1	0	0	1	2	2	2	0	1	1	0	1	2	0	1	1	1			
Rater 3	S_RP	0	1	2	0	0	0			1	3	0	1	0	0	0	0		2	1	0	2	2	0	2	2	0	2	0	0	0	0	0		
Rater 4	S_RP	0	0	0	0	2	2			0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0		
Rater 1	T_CC	0	0	0	0	0	2	0	0	0	1	2	4	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Rater 2	T_CC	0	0	0	0	0	0	0	0	0	1	3	3	0	0	0	0	0	0	4	0	0	0	0	0	2	0	1	2	0	0	0	0		
Rater 3	T_CC	0	0	0	0	0	1	0	0	1	0	1	2	0	0	1	0	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Rater 4	T_CC	0	0	0	0	1	0	0	2	2	0	2	2	0	0	0	0	1	1	2	2	2	0	0	0	0	0	0	1	0	0	0	0		
Rater 1	T_CM_RATIO	0	0	0	0	0	3	0	0	0	0	1	4	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	1	0	0	0	0		
Rater 2	T_CM_RATIO	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	1	4	1	3	2	0	0	2	0	0	3	0	0	0	0	0		
Rater 3	T_CM_RATIO	0	0	0	0	0	2	0	0	0	0	1	4	0	1	2	0	0	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
Rater 4	T_CM_RATIO	0	0	0	0	1	0	0	2	2	0	2	2	0	0	0	0	1	1	2	2	2	0	0	0	0	0	0	1	0	0	0	0	0	
Rater 1	T_MC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Rater 2	T_MC	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	
Rater 3	T_MC	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rater 4	T_MC	0	0	0	0	1	0	0	2	2	0	1	2	0	0	0	0	1	1	2	2	2	0	0	0	0	0	0	1	0	0	0	0	0	0

This table summarizes histopathology scores from each pathologist for each chemical and does level and provides an overview of rater consistency. As described in the methods, the severity of each lesion was adjusted to remove background alterations observed in control tissues. For most parameters the effects were scored as either not evident or minimal to mild. Four chemicals showed little, if any histological changes. These included aldicarb oxime, 2,4-DAT, Ribavirin and thalidomide. Chemicals that showed moderate immunohistological changes included TCDD, p-nitrotoluene, oxymetholone and gallium arsenide. The chemicals that showed the most extensive changes were DDI and DDA.

0 = No immunotoxic call – When the histopathology score is equal to 0.

1 = Minimal immunotoxic call – When the histopathology score does not exceed 3.

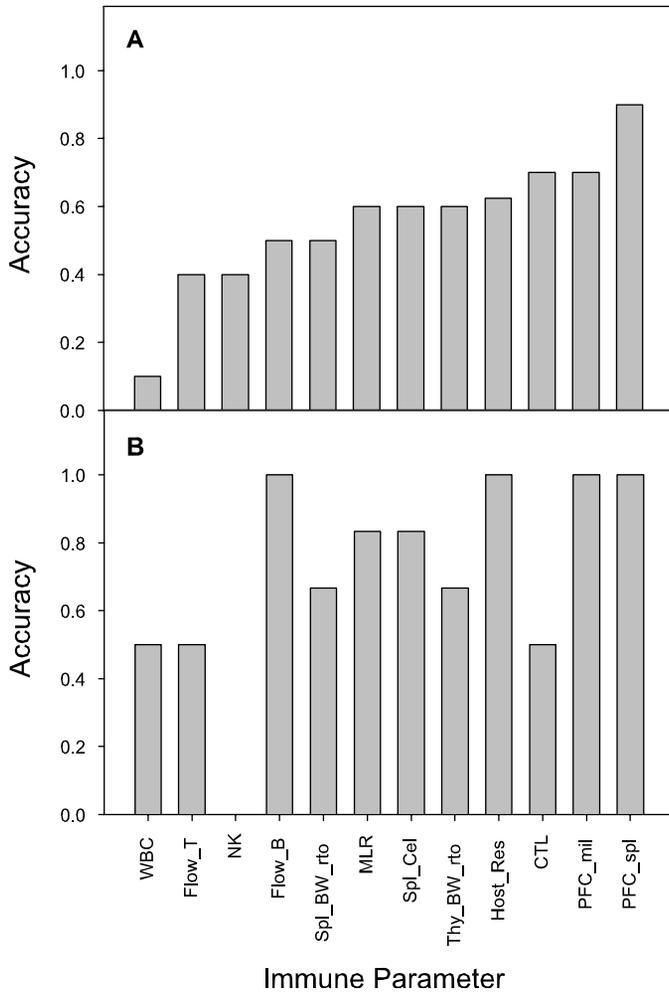
2 = Mild immunotoxic call – When the histopathology score does not exceed 6.

3 = Moderate immunotoxic call – When the histopathology score does not exceed 9.

4 = Marked immunotoxic call – When the histopathology score is more than 9.

The average histopathological scores obtained for each histological endpoint are shown in Figure 5 for the vehicle control, experimental data set (as a function of dose), and positive control data set. This analysis provides an indication of the relative sensitivity of the various histopathology parameters across

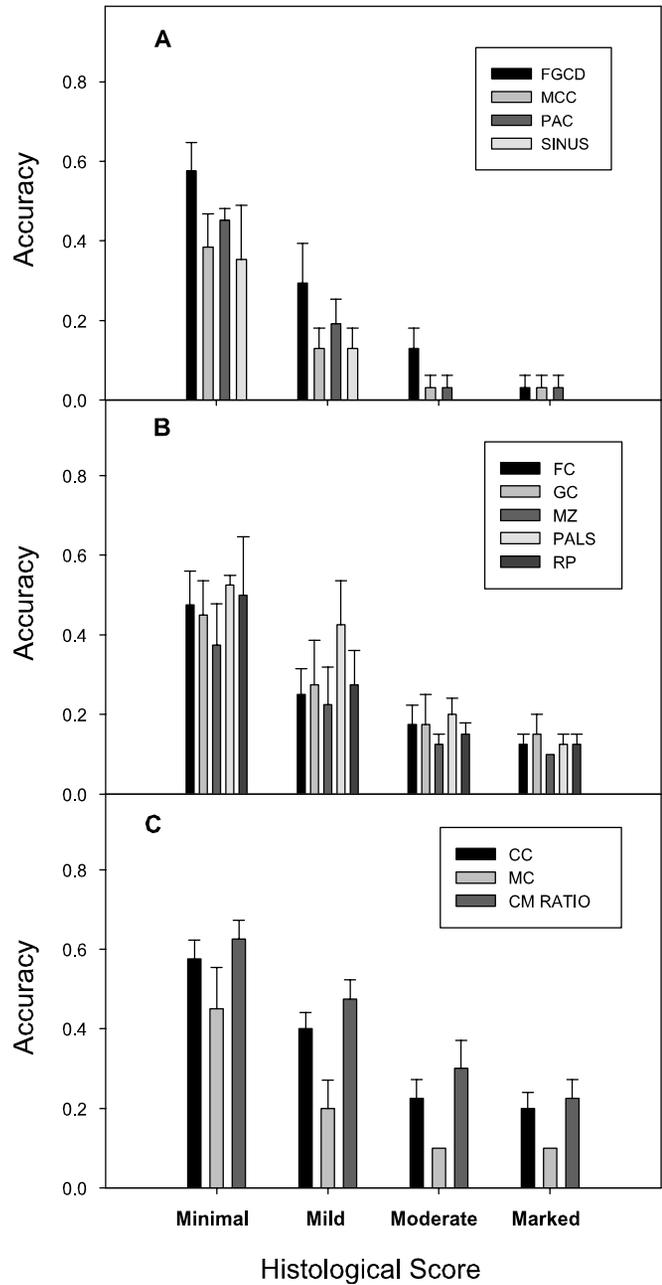
dose. While almost all endpoints demonstrated dose-response trends, differences in scores were not readily discernible between the vehicle-control treated groups and those from the low dose treatment groups and there were only minimal lesions observed in the medium dose group. The histological scores in



**FIG. 1.** Accuracy of various immune tests to identify an immunotoxic agent. Accuracy levels derived from the (A) test chemical data set or (B) positive control data set.

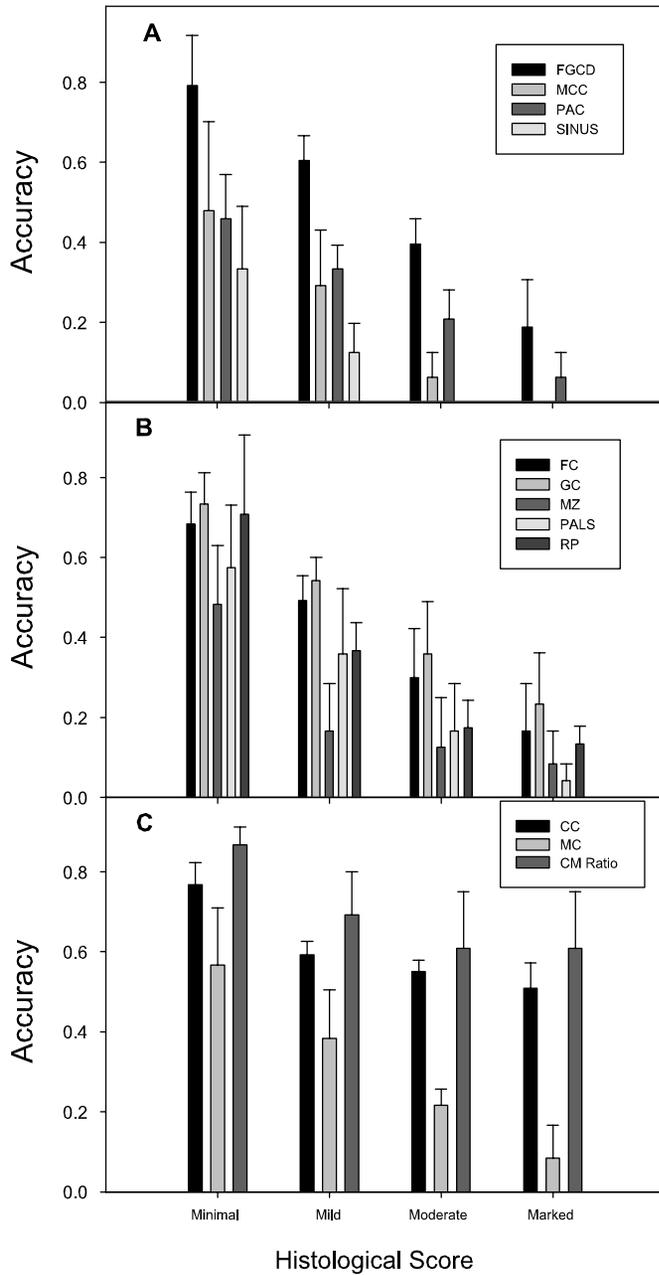
the high dose treatment groups, while never approaching the values obtained in the positive control groups, showed clear differences in lesion severity compared to the lower treatment groups. Within the specific parameters examined, thymus cortical cellularity and cortico-medullary ratios appeared to be the most sensitive or readily detected indicator of immunotoxicity (Fig. 5C).

To help address potential differences in the relative sensitivity between extended histopathology and classical immune tests, it was useful to compare the levels of accuracy that could be achieved between the histopathological scores and immune tests across the three dose levels that were evaluated in the test chemical data set (Fig. 6). For comparative purposes, the immune tests are considered individually while extended histopathology results are presented as a single variable i.e., “*in toto*.” For the immune tests, the accuracy level ranged from a high of 90% for the PFC response to a low of 20% for the WBC using results obtained from the high dose



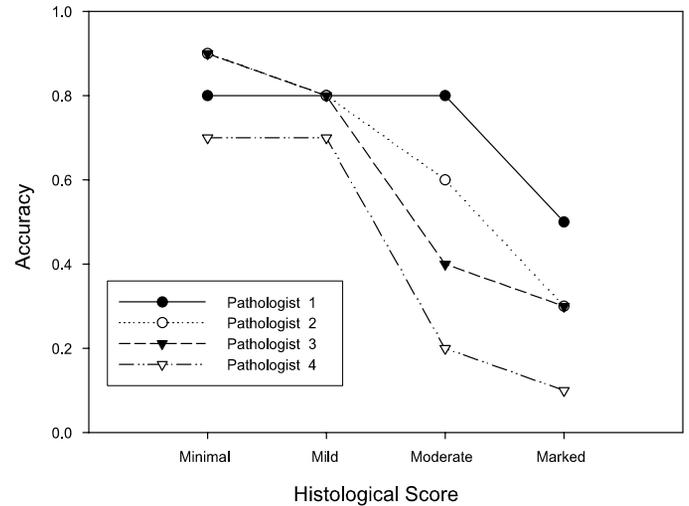
**FIG. 2.** Accuracy values as a function of the histopathology score obtained from a test chemical data set ( $n = 10$ ). Data are presented as mean  $\pm$ SEM of four independent pathologists for each histopathological endpoint examined. Endpoints measured included: lymph node cellularity in the follicles (FGCD), medullary cords (MCC), paracortical areas (PAC), and sinuses; spleen cellularity of lymphoid follicles (FC), marginal zone (MZ), periarteriolar lymphoid sheaths (PALS), red pulp (RP), and the total number of germinal centers (GC); and thymus cortex cellularity (CC), medullary cellularity (MC), and cortico-medullary ratio (CM-ratio).

treatment groups, to ~50% for the PFC response and 10% for WBC when considering results from the low dose treatment groups. Using the *in toto* histopathology data from the high dose treatment group, the accuracy values ranged from



**FIG. 3.** Accuracy values as a function of the histopathology score obtained from the positive control data set ( $n = 6$ ). Data are presented as mean  $\pm$  SEM of four independent pathologists for each histopathological endpoint examined. Endpoints measured included: lymph node cellularity in the follicles (FGCD), medullary cords (MCC), paracortical areas (PAC), and sinuses; spleen cellularity of lymphoid follicles (FC), marginal zone (MZ), periarteriolar lymphoid sheaths (PALS), red pulp (RP), and the total number of germinal centers (GC); and thymus cortex cellularity (CC), medullary cellularity (MC), and cortico-medullary ratio (CM-ratio).

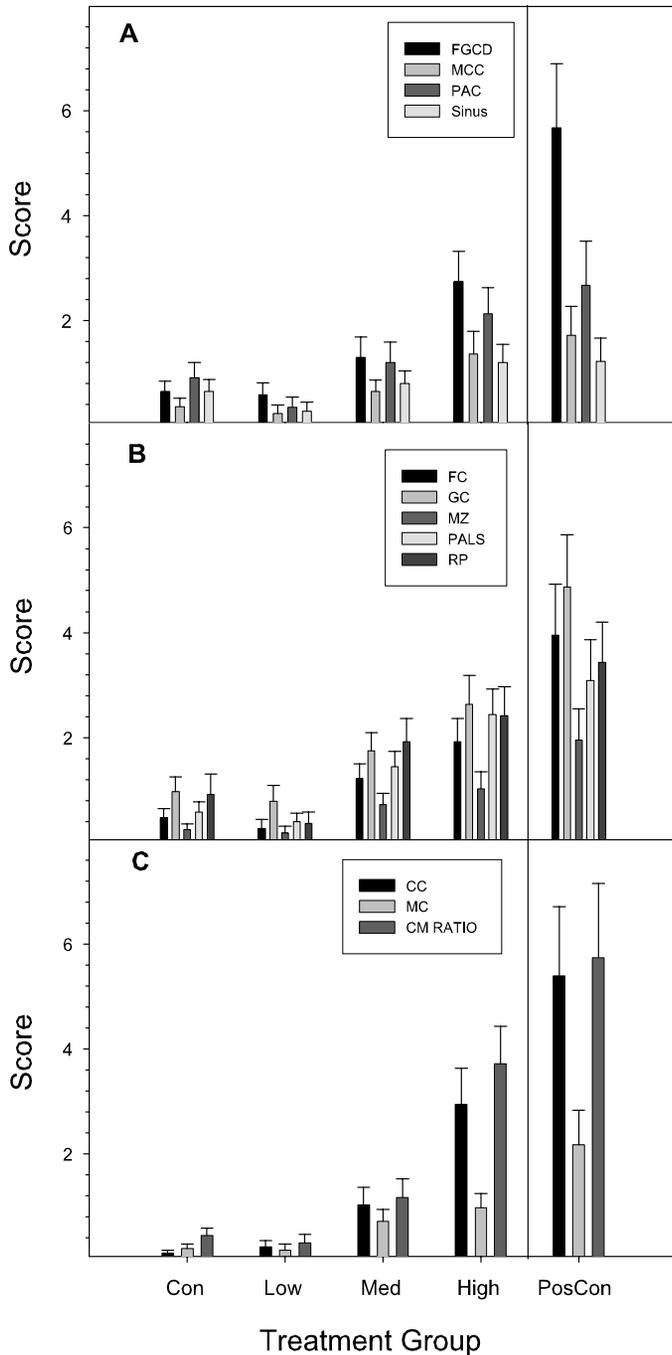
approximately 80%, using data obtained in the minimal and mild categories, to 30% when considering results from the marked category. Analysis of data obtained from animals treated with lower doses of chemicals suggested



**FIG. 4.** Accuracy values as a function of the histopathology score obtained by four independent pathologists when considering histological lesions *in toto*.

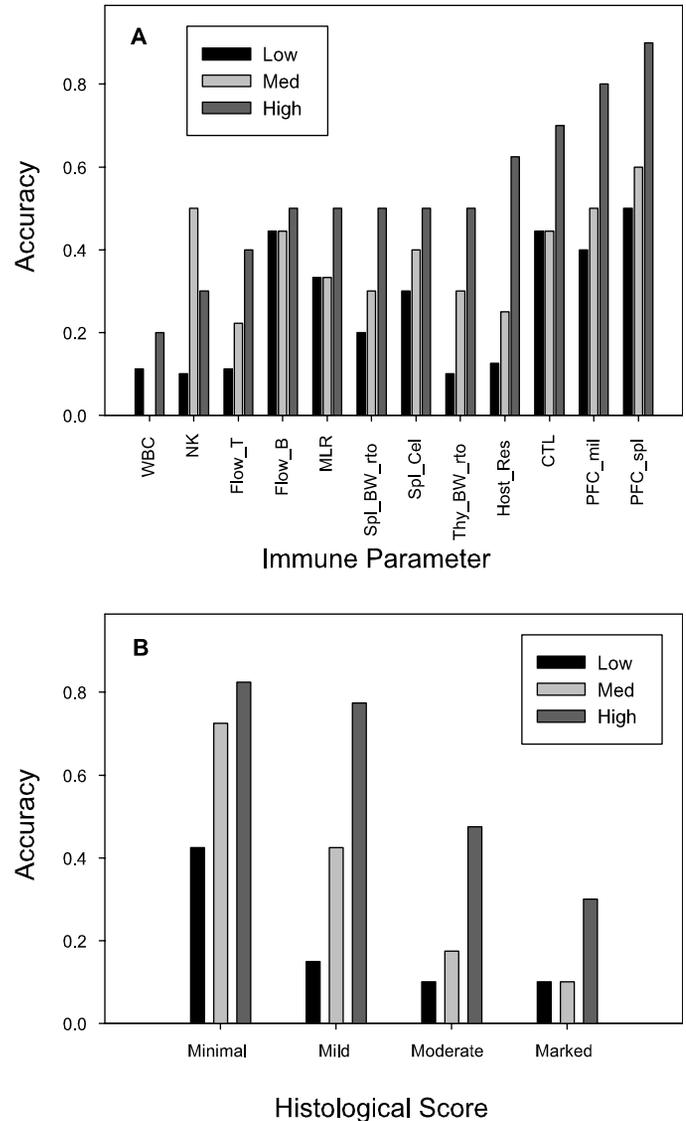
that histological scores, as did immune tests, generally followed dose-response trends. For the most part, accuracy levels comparable to the PFC response were obtained using *in toto* histopathology when limited to using the minimal, and in some instances mild criteria. Using the moderate or marked classification provided a limited degree of accuracy.

Table 3 provides the correlations obtained between the individual immune tests and mean histological scores from the four pathologists. While the composite, rather than individual, histopathological, and functional parameters are used to identify immunotoxic agents, this analysis provided an opportunity to determine which immune tests were more closely associated with specific histological changes and insight into the immunologic changes underlying the pathology. Although there was an overall tendency for low correlations, due to the small sample size and the large number of variables examined, several noteworthy associations were observed. The highest correlations were found between histological endpoints in the thymus, particularly those associated with cellularity, such as cortico-medullary ratios and thymus organ weights, and T-cell numbers. These histological endpoints also correlated well with host resistance tests. Host resistance tests in these data sets usually involved measuring resistance to challenge with the bacterium *Listeria monocytogenes* or with the B16F10 or PYB6 tumor cell lines, both of which evoke a strong cell mediated immune response. The lowest associations between the individual immune parameters and histological endpoints were observed in the lymph node. Table 4 provides the correlations between the individual immune tests and the histological scores from the pathologist with the most experience in evaluating extended histopathology of lymphoid tissues. Similar to that demonstrated in Table 3, the highest correlations were observed between endpoints in the thymus and the host resistance tests. In addition,



**FIG. 5.** Average histopathological scores obtained for each histological endpoint are shown for the vehicle control, experimental (as a function of dose), and positive control data sets. Each value represents the mean  $\pm$  SEM of four pathologists.

fairly strong associations were observed between quantitative measurements in the spleen (e.g., weight and cellularity) and histological endpoints in this tissue. Surprisingly, strong correlations between germinal center development in the spleen and lymph node with functional measures of antibody production were not observed.



**FIG. 6.** Accuracy as a function of dose for each of the immune tests (A) and *in toto* histopathology scores (B). Solid, low dose; lined, medium dose; open, high dose. Histopathology values represent means from the scores of the four pathologists.

DISCUSSION

At the most fundamental level, any laboratory incorporating a new test in toxicological assessment must establish the analytic sensitivity and specificity of the assay; its ability to detect the toxic endpoint when present and not detect it when absent (Grody, 2003). With this in mind, the present project was undertaken to help determine the level of accuracy expected to be achieved using extended histopathology to identify immunotoxic chemicals in female B6C3F1 mice at doses which do not induce overt toxicity. As there is no agreement on the magnitude of histopathological changes that would constitute an immunotoxic or biologically relevant effect, a continuous

TABLE 3  
Correlations between Individual Immune Tests and Mean Histological Scores from All Participating Pathologists

	L_FGCD	L_MCC	L_PAC	L_Sinus	S_FC	S_GC	S_MZ	S_PALS	S_RP	T_CC	T_MC	T_CMRA
CTL	0.407	0.235	0.331	0.122	0.100	0.208	0.067	0.208	0.189	0.281	0.060	0.185
FLOW_B	0.133	0.116	0.199	0.212	0.224	0.282	0.207	0.199	0.215	0.110	0.162	0.318
FLOW_T	0.329	0.227	0.433	0.175	0.365	0.060	0.292	0.443	0.190	0.534	0.306	0.423
HOST_RES	0.213	0.081	0.226	0.141	0.337	0.212	0.340	0.295	0.245	0.460	0.482	0.590
MLR	0.454	0.243	0.377	0.100	0.377	0.130	0.376	0.471	0.275	0.427	0.298	0.528
NK	-0.010	0.095	-0.055	-0.006	0.102	0.373	0.011	0.149	0.138	-0.001	0.057	0.244
PFC_MIL	0.330	0.285	0.422	0.322	0.367	0.324	0.393	0.372	0.419	0.481	0.351	0.433
PFC_SPL	0.326	0.237	0.377	0.274	0.355	0.371	0.375	0.373	0.365	0.486	0.328	0.479
SPL_BW_RTO	0.411	0.263	0.431	0.274	0.440	0.193	0.410	0.475	0.272	0.564	0.545	0.687
SPL_CEL	0.169	0.034	0.151	-0.046	0.281	0.057	0.348	0.334	0.104	0.329	0.195	0.398
THY_BW_RTO	0.352	0.186	0.420	0.266	0.335	0.130	0.426	0.331	0.193	0.541	0.478	0.556
WBC	-0.100	-0.090	-0.096	-0.089	-0.020	0.166	-0.097	-0.119	0.049	-0.046	-0.080	0.055

TABLE 4  
Correlations between Individual Immune Tests and Mean Histological Scores from Pathologist #2

	L_FGCD	L_MCC	L_PAC	L_sinu	S_FC	S_GC	S_MZ	S_PALS	S_RP	T_CC	T_MC	T_CMRA
CTL	0.358	0.205	0.081	0.064	-0.233	0.167	0.136	0.040	0.234	0.195	0.079	0.015
FLOW_B	0.112	0.224	0.420	0.371	0.229	0.498	0.055	0.299	0.291	0.060	-0.045	0.714
FLOW_T	0.233	0.507	0.327	0.233	0.162	-0.226	0.721	0.443	0.012	0.555	0.630	0.271
HOST_RES	0.175	0.175	0.391	0.175	0.526	0.328	0.358	0.299	0.452	0.447	0.590	0.806
MLR	0.305	0.205	0.031	0.024	0.244	0.287	0.454	0.396	0.167	0.303	0.389	0.504
NK	-0.208	0.019	-0.243	0.081	0.189	0.641	-0.210	0.282	0.329	-0.245	-0.160	0.510
PFC_MIL	0.138	0.426	0.478	0.384	0.309	0.368	0.372	0.402	0.394	0.557	0.385	0.664
PFC_SPL	0.108	0.395	0.435	0.355	0.287	0.368	0.347	0.412	0.393	0.528	0.338	0.756
SPL_BW_RTO	0.318	0.462	0.351	0.318	0.423	0.373	0.544	0.557	0.238	0.475	0.668	0.768
SPL_CEL	-0.021	0.138	-0.025	-0.021	0.339	-0.020	0.612	0.403	0.159	0.421	0.345	0.510
THY_BW_RTO	0.369	0.525	0.500	0.369	0.265	0.089	0.608	0.362	0.171	0.539	0.746	0.604
WBC	-0.083	-0.093	-0.096	-0.083	-0.081	0.265	-0.081	-0.135	0.189	-0.105	-0.062	0.285

scale was developed in which the association between functional criteria, as defined previously (Luster *et al.*, 1992), and histological changes could be established at any magnitude of histopathological change. Evaluating the histopathological endpoints individually, the level of accuracy in identifying immunotoxic chemicals, using the experimental data set, ranged from a high of approximately 60%, for several endpoints in the thymus when using the least stringent criteria to classify chemicals as immunotoxic, to less than 10% when using histological endpoints in the lymph node compartment under the most stringent criteria. When analyzing the positive control data set using the least stringent criteria, accuracy levels approaching 80% could be achieved. The fact that higher accuracy levels were obtained with the positive control data set would imply that the more potent an immunotoxic chemical (i.e., the more severe the histological lesion), the more likely it would be to correctly identify an agent. Relatively low levels of accuracy were obtained when histological endpoints were analyzed individually. However, accuracy levels for both the test chemical and positive control data sets were significantly improved when the histological

endpoints were considered *in toto* (i.e., combined), at least in the minimal to mild range, where only two of the immunosuppressive chemicals examined (2,4 diaminotoluene and Ribivarin) were not correctly identified. The interpretation of the histological lesions for some chemicals, such as gallium arsenide, presented particular challenges, as several of the pathologists observed a negative dose response, but in different endpoints. It should be noted, however, that the increased accuracy using the “*in toto*” analyses compared to the analyses of specific lesions, could be misleading, as there did not have to be agreement between the pathologists in the tissues in which lesions were observed. The degree of consistency obtained among the pathologists in identifying and grading specific lesions is shown in Table 2 and was specifically addressed in an earlier publication (Germolec *et al.*, 2004).

Of the three lymphoid organs examined, the most consistent and discernible histological lesions were observed in the thymus, specifically the cortical region. This is not surprising, as the susceptibility of the thymus to toxicity is well established (e.g., Schuurman *et al.*, 1992). Thymic alterations were

manifested as both a decrease in cortical cellularity and cortico-medullary ratios. As might be expected, the presence of these lesions was strongly associated with decreases in thymus:body weight ratios (see Tables 3 and 4). As a "stand alone" test, thymus:body weight ratios were previously shown to provide 68% concordance in identifying immunotoxic chemicals using a large data set (Luster *et al.*, 1992), which is comparable to the 60% accuracy obtained with thymus histological parameters in this study. It should be noted, however, that the same correlation is often observed as a nonspecific effect of stress in toxicology studies, so changes in both the histological and quantitative weight measurements should be interpreted within the context of other effects to help avoid misidentification of compounds as immunotoxic (Levin *et al.*, 1993). In addition to thymus:body weight ratios, histological lesions in the thymus were also associated with decreases in the antibody PFC response, previously shown to be a good indicator of immunotoxicity (Luster *et al.*, 1992, 1993). In contrast to thymus endpoints, lymph node measurements provided the least association with immune tests. The apparent sensitivity for the thymus, compared to the spleen and lymph node is unknown. From a biological standpoint, it can be argued that immunotoxic chemicals that operate by altering antigen recognition or antigen-dependent responses would most likely manifest histopathology in secondary lymphoid organs (i.e., spleen, lymph node), coinciding with an active immune response. In contrast, agents that operate through non-specific cytoreductive or antiproliferative processes would be expected to present histopathology in both primary (thymus) and secondary lymphoid organs.

These studies highlight that central to the successful application of extended histopathology in risk assessment for immunotoxicology is the necessity to determine an appropriate level of stringency (histological score) to be applied when assessing lesions. As a majority of the immunotoxic agents in the test chemical data set did not produce severe lesions (see positive vs. test chemical data set, Table 2), applying criteria that are too stringent, represented in our scale as moderate to marked histological scores, would not allow identification of the majority of immunotoxic agents in the test chemicals data set. Accepting minimal to mild histological scores as significant may provide an acceptable level of accuracy (~80%), but raises the possibility of increasing misclassifications (i.e., false positives). This was problematic with employing the minimal classification criteria, as this classification would have identified many of the negative controls as positive. A minimal score can be obtained by applying a minimal grade (1) to any of the 12 histopathological endpoints examined, and such minimal effects could be the result of variables such as diet, type of housing, and sex or strain of the test animal. Allowing for a mild rather than minimal classification as significant, however, would appear to limit false positives yet provide a reasonable level of accuracy (>80%) when considering lesions *in toto*, and >50 percent when considering lesions individually. Analyses of data obtained from animals treated with the low doses of chemicals, to help assess the relative

sensitivity between the immune and histological endpoints, did not significantly affect the outcomes.

In conclusion, it has been suggested that inclusion of extended histopathology of lymphoid organs as part of routine toxicology testing can be used as predictive indicator for immunotoxic agents. However, due to the paucity of available data and limited validation efforts, it is not currently possible to determine the predictive value of extended histopathology evaluation, either alone or as an adjunct with functional endpoints, as a screening test. Nonetheless, the present studies suggest that incorporation of extended histopathology into standard toxicological assessment has potential utility. While this study was conducted in female B6C3F1 mice at doses lower than the MTD, similar conclusions with regard to the need for training and the association between severity of lesions and the ability to evaluate tissue lesions were reached by the ICICIS investigators using inbred F344 and outbred Wistar rats and the standard OECD 407 28-day testing protocol (ICICIS, 1998). As a prerequisite for successful implementation, however, it will be necessary to adopt standardized histological scoring and quality assurance and controls, similar to that described in this and other studies, to ensure that subtle histopathological lesions can be consistently identified (Harleman, 2000; ICICIS, 1998). Comparable recommendations for standardization of controls, endpoints, and methods have been made for immune function tests (Luster *et al.*, 1988; van Loveren *et al.*, 1996). While acceptable levels of accuracy may be achievable using extended histopathology, in our studies, it was not equivalent to that which can be obtained by functional tests and, within the test chemical data set used for this exercise, two of nine immunotoxic chemicals would probably been classified as negative. This is in contrast to findings by Schulte *et al.* (2002) who reported that extended histopathology was more accurate and consistent than functional tests when used to evaluate the immunomodulatory compounds cyclosporine A and hexachlorobenzene. Furthermore, our analyses indicate that inclusion of nonfunctional endpoints (e.g., lymphoid organ weights and WBC) with extended histopathology would not significantly increase the predictive value (data not shown). As a final consideration, we did not monitor for persistence and currently the relative persistence of histological changes, as compared to functional immune effects is unknown. This may be of some concern as developmental immunotoxicology studies with estrogens may induce both thymic atrophy and alterations in functional immune parameters in the neonate; however, while changes in functional endpoints persist long into adulthood, changes in thymic cellularity appear to resolve rather quickly (Forsberg, 1984).

#### ACKNOWLEDGMENTS

The authors would like to extend their warmest appreciation to Dr. Robert Maronpot for his encouragement and facilitation of the study, as well as his thoughtful comments on the manuscript. We would also like to thank

Drs. Michael Andrew at NIOSH and Robert Luebke at USEPA for their insightful reviews of this work.

## REFERENCES

- Brown, L. D., Wilson, D. E., and Yarbrough, J. D. (1988). Alterations in the hepatic glucocorticoid response to mirex treatment. *Toxicol. Appl. Pharmacol.* **92**, 203–213.
- Burns, L. A., Bradley, S. G., White, K. L., Jr., McCay, J. A., Fuchs, B. A., Stern, M., Brown, R. D., Musgrove, D. L., Holsapple, M. P., Luster, M. I., and Munson A. E. (1994). Immunotoxicity of mono-nitrotoluenes in female B6C3F1 mice. I. Para-Nitrotoluene. *Drug Chem. Toxicol.* **17**, 317–358.
- Cao, W., Sikorski, E. E., Fuchs, B. A., Stern, M. L., Luster, M. I., and Munson, A. E. (1990). The B lymphocyte is the immune cell target for 2'3'-dideoxyadenosine. *Toxicol. Appl. Pharmacol.* **105**, 492–502.
- Clement, J. G. (1985). Hormonal consequences of organophosphate poisoning. *Fundam. Appl. Toxicol.* **5**, S61–S77.
- Forsberg, J. G. (1984). Short-term and long-term effects of estrogen on lymphoid tissues and lymphoid cells with some remarks on the significance for carcinogenesis. *Arch. Toxicol.* **55**, 79–90.
- Germolec, D. R. (2004) Sensitivity and predictivity in immunotoxicity testing: Immune endpoints and disease resistance. *Toxicol. Lett.*, **149**, 109–114.
- Germolec, D. R., Nyska, A., Kashon, M., Kuper, C. F., Portier, C., Kommineni, V., Johnson, K. A., and Luster, M. I. (2004). Extended histopathology in immunotoxicity testing: Interlaboratory validation studies. *Toxicol. Sci.* **78**, 107–115.
- Grody, W. W. (2003). Quests for controls in molecular genetics. *J. Mol. Diagnostics* **5**, 209–211.
- Harleman, J. H. (2000). Approaches to the identification and recording of findings in the lymphoreticular organs indicative for immunotoxicity in regulatory type toxicity studies. *Toxicology* **142**, 213–219.
- Hastings, K. L. (2002). Implications of the new FDA/CDER immunotoxicology guidance for drugs. *Int. Immunopharmacol.* **2**, 1613–1618.
- ICICIS Group Investigators. (1998). Report of validation study of assessment of direct immunotoxicity in the rat. *Toxicology* **125**, 183–201.
- Karrow, N. A., McCay, J. A., Brown, R., Musgrove, D., Munson, A. E., and White, K. L., Jr. (2000a). Oxymetholone modulates cell-mediated immunity in male B6C3F1 mice. *Drug Chem. Toxicol.* **23**, 621–644.
- Karrow, N. A., McCay, J. A., Brown, R. D., Musgrove, D. L., Pettit, D. A., Munson, A. E., Germolec, D. R., and White, K. L., Jr. (2000b). Thalidomide stimulates splenic IgM antibody response and cytotoxic T lymphocyte activity and alters leukocyte subpopulation numbers in female B6C3F1 mice. *Toxicol. Appl. Pharmacol.* **165**, 237–244.
- Kunimatsu, T., Kamita, Y., Isobe, N., and Kawasaki, H. (1996). Immunotoxicological insignificance of fenitrothion in mice and rats. *Fundam. Appl. Toxicol.* **33**, 246–253.
- Kuper, C. F., Harleman, J. H., Richter-Reichelm, H. B., and Vos, J. G. (2000). Histopathologic approaches to detect changes indicative of immunotoxicity. *Toxicol. Pathol.* **28**, 454–466.
- Liebtrau, A. M. (1983). *Measures of Association, Quantitative Application in the Social Sciences*, Vol 32. Sage, Beverly Hills, CA.
- Levin S., Semler, D., and Ruben, Z. (1993). Effects of two weeks of feed restriction on some common toxicologic parameters in Sprague-Dawley rats. *Toxicol. Pathol.* **21**, 1–14.
- Luster M. I., Munson, A. E., Thomas, P. T., Holsapple, M. P., Fenters, J. D., White, K. L. Jr., Lauer, L. D., Germolec, D. R., Rosenthal, G. J., and Dean, J. H. (1988). Development of a testing battery to assess chemical-induced immunotoxicity: National Toxicology Program guidelines for immunotoxicity evaluation in mice. *Fundam. Appl. Toxicol.* **10**, 2–19.
- Luster, M. I., Portier, C., Pait, D. G., White, K. L., Jr., Gennings, C., Munson, A. E., and Rosenthal, G. J. (1992). Risk assessment in immunotoxicology. I. Sensitivity and predictability of immune tests. *Fundam. Appl. Toxicol.* **18**, 200–210.
- Luster, M. I., Portier, C., Pait, D. G., Rosenthal, G. J., Germolec, D. R., Corsini, E., Blaylock, B. L., Pollock, P., Kouchi, Y., Craig, W., and *et al.* (1993). Risk assessment in immunotoxicology. II. Relationships between immune and host resistance tests. *Fundam. Appl. Toxicol.* **21**, 71–82.
- National Toxicology Program (1988a). Immunotoxicity of 2,4-Diaminotoluene (DAT). Final gavage report in Female B6C3F1 mice. NTP study number IMM87034.
- National Toxicology Program (1988b). The Immunotoxicity of Aldicarb Oxime in Female B6C3F1 mice. NTP study number IMM89025.
- National Toxicology Program (1989). Immunotoxicity of Ribavirin in Female C57Bl/6 Mice. NTP Study number IMM90010.
- OECD (1995). OECD Guideline for the Testing of Chemicals 407: Repeated Dose 28-day Oral Toxicity Study in Rodents.
- Phillips, K. E., McCay, J. A., Brown, R. D., Musgrove, D. L., Meade, B. J., Butterworth, L. F., Wilson, S. White, K. L., Jr., and Munson, A. E. (1997). Immunotoxicity of 2'3'-dideoxyinosine in female B6C3F1 mice. *Drug Chem. Toxicol.* **20**, 189–228.
- Pruett, S. B., Collier, S., Wu, W. J., and Fan, R. (1999). Quantitative relationships between the suppression of selected immunological parameters and the area under the corticosterone concentration vs. time curve in B6C3F1 mice subjected to exogenous corticosterone or to restraint stress. *Toxicol. Sci.* **49**, 272–280.
- Pruett, S. B., Ensley, D. K., and Crittenden, P. L. (1993). The role of chemical-induced stress responses in immunosuppression: A review of quantitative associations and cause-effect relationships between chemical-induced stress responses and immunosuppression. *J. Toxicol. Environ. Health* **39**, 163–92.
- Pruett, S. B., Fan, R., Myers, L. P., Wu, W. J., and Collier, S. (2000). Quantitative analysis of the neuroendocrine-immune axis: Linear modeling of the effects of exogenous corticosterone and restraint stress on lymphocyte subpopulations in the spleen and thymus in female B6C3F1 mice. *Brain Behav. Immun.* **14**, 270–287.
- Schulte, A., Althoff, J., Ewe, S., Richter-Reichelm and the BGVV Group Investigators (2002). Two immunotoxicity ring studies according to OECD TG 407 – Comparison of data on cyclosporin A and hexachlorobenzene. *Regul. Toxicol. Pharmacol.* **36**, 12–21.
- Schuurman, H. J., Van Loveren, H., Rozing, J., and Vos, J. G. (1992). Chemicals trophic for the thymus: risk for immunodeficiency and autoimmunity. *Int. J. Immunopharmacol.* **14**, 369–375.
- Sikorski, E. E., McCay, J. A., White, K. L., Jr., Bradley, S. G., and Munson, A. E. (1989). Immunotoxicity of the semiconductor gallium arsenide in female B6C3F1 mice. *Fundam. Appl. Toxicol.* **13**, 843–858.
- United States Environmental Protection Agency, U. (1998). Helth Effects Test Guidelines. *OPPTS 870.7800 Immunotoxicity*.
- van Loveren, H., Vos, J. G., and De Waal, E. J. (1996). Testing immunotoxicity of chemicals as a guide for testing approaches for pharmaceuticals. *Drug Information J.* **30**, 275–279.