International Harmonization of Toxicologic Pathology

Nomenclature: An overview and review of basic principles.

Peter C. Mann, John Vahle, Charlotte M Keenan, Julia F. Baker, Alys E Bradley, Dawn G. Goodman, Takanori Harada, Ronald Herbert, Wolfgang Kaufmann, Rupert Kellner, Thomas Nolte, Susanne Rittinghausen and Takuji Tanaka

Prior Harmonization Projects

Harmonization of nomenclature and diagnostic criteria in toxicologic pathology, especially for rats and mice, has been a goal of pathologists working in the profession for many years. In the latter part of the twentieth century, several initiatives were undertaken by the Society of Toxicologic Pathology (STP) in the United States and by the RITA data base group (Registry of Industrial Toxicology Animal – data) in Europe. Their efforts resulted in a number of internationally recognized publications: SSNDC (Standardized System of Nomenclature and Diagnostic Criteria): Guides for Toxicologic Pathology (www.toxpath.org/ssndc.asp) and the WHO/IARC International Classification of Rodent Tumors.

INHAND Nomenclature Project

Beginning in 2005, the STP and European Society of Toxicologic Pathology (ESTP), in conjunction with RITA, developed a collaborative process to review, update, and harmonize existing nomenclature documents and databases. In 2006, the British Society of Toxicologic Pathologists (BSTP) and the Japanese Society of Toxicologic Pathology (JSTP) joined the initiative, so that the project has become truly global. The goal of the project is to produce publications for each organ system that provide a standardized nomenclature and differential diagnosis for classifying microscopic lesions observed in laboratory rats and mice in toxicity and carcinogenicity studies. The project is referred to as INHAND (International Harmonization of Nomenclature and Diagnostic Criteria for Lesions in Rats and Mice).
Briefly the INHAND project is organized as follows: The Global Editorial and Steering Committee (GESC) oversees the activities of the project. The GESC is composed of toxicologic pathologists from all of the participating societies. In addition there are several technical consultants for online and print publishing support.

The Organ Working Groups (OWG) are the core of the project. Each Group is composed of expert toxicologic pathologists from each of the participating societies and are responsible for developing preferred nomenclature and diagnostic criteria. Once the draft nomenclature has been developed, the GESC completes an initial review, followed by a period during which all members of participating societies are requested to review the proposed nomenclature. The OWG then finalizes the nomenclature based on comments from the GESC and general membership. An important feature of the INHAND project is the use of the goRENI (global open Registry Nomenclature Information System, www.goreni.org) as a web-based platform to both review draft nomenclature and publish final nomenclature. Originally developed by RITA, goRENI provides access to members of all STPs as well as government regulators.

In addition to goRENI, the completed nomenclature for each organ system will be published in one of the official journals of the participating Societies: Toxicologic Pathology (STP and BSTP), Experimental and Toxicologic Pathology (ESTP), or Journal of Toxicologic Pathology (JSTP). The first systems to be published were the respiratory (Renne et al, 2009) and hepatobiliary (Thoolen, 2010) systems, with the remaining organ systems targeted to be published by the end of 2013.

General Concepts and Approaches

The individual organ system publications provide information on specific diagnostic entities and differential diagnoses; however, certain key principles and processes are common across the various organ systems. The purpose of the following sections is to provide guidance on general approaches to recording microscopic lesions in toxicity studies. The following comments are complimentary to the approaches and perspective
While the diagnosis of proliferative lesions is fairly straightforward (a mass either meets the criteria for diagnosis or it does not), non-proliferative lesions present a more formidable challenge. Different grading scales, whether a change should be graded, and the use of modifiers and thresholds all provide opportunities for significant variation between pathologists. In 2007, Greaves emphasized the importance of consistency in diagnosis, writing, “it is salutary to remember that toxicologists and physicians in government regulatory agencies usually read the text relating to pathology findings with extreme care. In addition the tabulated summaries of pathology are often reviewed with equal attention. Unclear language, inappropriate, misleading or unexplained terminology, conclusions not justified by data, any discrepancy between text and tables may all raise unnecessary questions. Thus, clarity of the report and explanation of all findings are essential” (Greaves, 2007). Since the purpose of the INHAND project is to generate (so far as possible) standardized nomenclature for both proliferative and non-proliferative changes, the more important of these variables are discussed below.

**Diagnostic versus Descriptive Terminology**

Most pathologists are initially trained in interpretive diagnostic pathology with the goal of rendering a definitive diagnosis through the use of morphologic, etiologic and disease diagnosis. In rendering a final disease diagnosis, the pathologist is often incorporating a wide range of data including history, clinical signs, and laboratory data. In toxicologic pathology, the goal is to determine if the test article produces changes through a comparison of treated animals with control animals. As such it is important that microscopic observations be recorded in a consistent, objective manner that readily allows tabulation and comparison of group effects. In this setting, the use of descriptive, rather than diagnostic, terminology is preferred. In many cases a disease diagnosis implies a particular pathogenesis or impact on organ function based on what is known about the spontaneous disease; this may be misleading in the experimental setting of a toxicity study. For example, a pharmacologic agent may cause an increase in trabecular
and cortical bone. If this observation was recorded as osteopetrosis, a specific disease state (osteopetrosis), mechanism (decreased osteoclast function) and etiology (genetic mutation) could be implied, when in fact the effect was due to a pharmacologic process. Another issue frequently encountered in toxicologic pathology is determining how to document vacuolation in multiple tissues due to the process of phospholipidosis. In recording the observations rendered from the routine hematoxylin and eosin stained section, the pathologist should document the morphologic diagnosis (cytoplasmic vacuolation). In the pathologist’s report, additional data (ultrastructural or biochemical data) can be integrated to render the interpretation that the vacuolation is due to the process of phospholipidosis. In some studies, pathologists are confronted with novel treatment-related changes which may require unique terminology; these need to be treated on a case-by-case basis. Consistently using descriptive, rather than diagnostic terminology in tabulating anatomic pathology data will decrease confusion and misconceptions. Additional description and interpretation on the likely process can be most appropriately provided in the pathologist’s report (narrative).

Modifiers
In addition to severity grades, several additional modifiers may be included in a diagnosis in order to provide more information and clarity. It is important to remember that the next person to review a diagnosis may not have access to the slide, and likely is not a pathologist. For these reasons, accuracy and consistency in diagnostic terminology are crucial. Modifiers that are commonly used include organ-specific topography, distribution, character of the change and duration (Frame and Mann, 2008).

Organ specific topography varies with the anatomic complexity of the organ; topography may also vary due to the prevalence of treatment-related changes in the organ. For the liver, examples include acinar, portal, periportal, midzonal, centrilobular, hilar, ductal, peri-ductal, peri-canicular or sub capsular – all of which may be used to indicate the specific location within the organ where the change has occurred. For the kidney and thymus, it is often useful to differentiate cortical from medullary changes, while in the
central nervous system the anatomical location affected is often crucial to understanding the pathogenesis of the change.

Distribution modifiers are utilized to indicate the pattern of change. Commonly used distribution modifiers include: focal, multi-focal, and diffuse. Based on the formal definition, a focal lesion refers to one specific area, or focus, whereas multi-focal refers to more than one focus (foci). However, some pathologists use focal for both focal and multifocal referring to the nature of the lesion rather than its actual distribution and using grading to reflect the extent of the lesion (Thoolen, et al 2010). Diffuse lesions affect the majority of the section examined, while focally-extensive may be used to describe a large lesion that affects a significant portion of the section, but which is still localized in nature.

A modifier for character of change becomes important when the general diagnostic term is quite broad. An example is necrosis. Character modifiers might include apoptotic, caseous, coagulative, single-cell or bridging. The use of character modifiers is often predicated by the organ where the changes occur (i.e. single-cell or bridging necrosis may be important indicators in the liver).

Terms for duration of a change have been used in both natural disease processes and treatment-related changes. Commonly used terms for duration include acute, subacute, chronic, and chronic-active, which is used to identify a lesion of chronic duration which also has ongoing acute changes. These terms are generally associated with a particular type of inflammatory cell: acute (neutrophilic), subacute (neutrophilic mixed with mononuclear cells) lymphocytes), chronic (lymphocytes, plasma cells, macrophages) mononuclear cells or macrophages) and chronic-active (both mononuclear cells and neutrophils). For routine toxicology studies, the recommendation is to forego chronicity modifiers and rely on cell type or process (fibrosis, hyperplasia, etc), especially since toxicology studies are a set number of days and the morphology represents a single point in time. For natural disease generally
there is more known about the time course of effects, so acute, chronic-active, etc have more meaning.

Not all of these modifier types are used by every pathologist for all situations. However, having the terms in our armamentarium and using them consistently will help immensely in producing clear tables and reports. A model of how a pathologist might construct a descriptive diagnosis using modifiers in presented in Figure 2.

**Combination Terms**

For some common findings which can involve a combination of morphologic responses, compound term diagnoses have been created to reduce the number of terms which need to be selected in pathology data entry systems. For example, degeneration may be present with concurrent evidence of regeneration in various tissues, and therefore the two processes are often combined into a single morphologic diagnosis “degeneration/regeneration” at the discretion of the pathologist. Another approach has been to use terms such as nephropathy, cardiomyopathy for organ-specific changes summarizing a constellation of histologic features. Combination terms may also be appropriate for longer term studies, including carcinogenicity studies, where there might be progression in toxicity (i.e. progressing from vacuolation to degeneration to necrosis to repair) over time. These compound terms are used with locators and modifiers in the usual way.

**Grading Systems/Severity Scoring**

Most toxicologic pathologists use a grading system to document lesion severity. In general, the grade of severity assigned to a diagnosis should be chosen to reflect a combination of the extent of the process (how many of its subordinate components are present), the distribution (focal to diffuse), and the actual degree of severity. Grading systems vary between laboratories, often based on differences in the computerized pathology data capture systems. These systems may differ in how they incorporate distribution, stage, and extent of lesions. Within a given grading system, it is important to allow the individual study pathologist flexibility in grading to allow them to accurately
demonstrate a dose response and characterize the toxicity of the test article. The problem of harmonization as it relates to lesion severity has been recognized and discussed in some detail (Hardisty and Eustis 1990, WHO 1978). Given this complexity and need for flexibility, a universal grading system would be difficult to achieve.

The most commonly used grading systems use either 4 or 5 distinct semi-quantitative grades. An example of a 4-point severity scale is shown in Table 1. The grades can either be expressed as numbers (where 0 indicates no change, and 1-4 or 1-5 indicate increasing severity) or as words indicating the degree of severity, which correlate to the numerical score. Some systems have been suggested where the percentage of the organ affected is utilized to assign a severity score (Shackelford, 2002), but this can vary according to organ involved. In addition, there may be a difference in the relative degree of severity assigned to a lesion depending on the length of the study (some changes may be graded as less severe in a chronic study than in a short-term study). Although the generation of ordinal data using a scoring system may allow statistical analysis for effects and trends (Gad and Rousseaux, 2002), routine statistical analysis of severity grades is strongly discouraged because the grading scale often is not linear, the assigned number is only semi-quantitative and is based on variable criteria. Determination of test-article effects and dose response is most appropriately done by the study pathologist, rather than by statistical analysis.
Table 1: Illustration of a 4-point scoring system

<table>
<thead>
<tr>
<th>Numerical Score</th>
<th>Description</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Within normal limits</td>
<td>Tissue considered to be normal, under the conditions of the study and considering the age, sex, and strain of the animal concerned. Alterations may be present which, under other circumstances, would be considered deviations from normal.</td>
</tr>
<tr>
<td>1</td>
<td>Minimal</td>
<td>The amount of change present barely exceeds that which is considered to be within normal limits.</td>
</tr>
<tr>
<td>2</td>
<td>Slight</td>
<td>In general, the lesion is easily identified but of limited severity.</td>
</tr>
<tr>
<td>3</td>
<td>Moderate</td>
<td>The lesion is prominent but there is significant potential for increased severity.</td>
</tr>
<tr>
<td>4</td>
<td>Severe</td>
<td>The degree of change is as complete as possible (occupies the majority of the organ).</td>
</tr>
</tbody>
</table>

Consistency in assigning severity grades within a study or between studies on the same compound is a key quality attribute of the pathology evaluation and inconsistency or “drift” (the tendency for diagnostic criteria to vary over time) in assigning diagnostic grades within a study can either compromise the ability to detect a test-article effect or lead to the appearance of a test-article effect when none is actually present. In some cases, toxicologic pathologists may attempt to modify or “split” their criteria for severity scores in a study in order to convey a treatment-related effect or the dose-responsiveness of an effect. For example, a pathologist may be able to reliably separate individual lesions into multiple groups, despite the fact that all the lesions represent variations of minimal. While this process may be appropriate in some situations, it creates challenges in inter-study consistency and may create the impression that a change is more severe than the glossary-defined severity terms would suggest. If this approach is taken, the toxicologic pathologist should clearly specify the process and criteria in the methods section of the pathology report. During the peer review process, the reviewing
pathologist should evaluate the consistency of grading within the study. As long as severity grades were consistently applied across a study, a one point difference of opinion regarding severity is acceptable, given the semi-quantitative nature of severity scoring.

**Non-graded Changes**
For some lesions pathologists may not assign a severity grade as no additional information would be gained. These changes are typically recorded as Present (P), rather than having a severity grade assigned. Examples include: neoplasms, cyst(s), autolysis and congenital anomalies. However, some pathology data entry systems require entry of grades for all lesions.

**Thresholds**
In addition to recording treatment-related alterations and notable spontaneous lesions, the pathologist must determine whether or not to document minor variations in normal tissue morphology. In many cases these variations are due to minor age-related changes, while in other cases they represent normal anatomic variability within a population of animals. Thresholding refers to the practice of determining which variations in normal morphology will be recorded and which variations are below a threshold and will not be recorded. In a population of untreated animals there are normal subtle variations in the morphology of tissues, so some degree of thresholding is needed to provide a meaningful compilation on microscopic pathology data. Factors the pathologist must consider in determining appropriate diagnostic thresholds include the morphology noted in the concurrent control animals and their understanding of normal morphology for the age, strain, and source of the test species. Great care must be taken in determining thresholds; indeed, some pathologists may employ limited, if any, thresholds. If a common spontaneous lesion is not recorded, the potential for a treatment-related effect on the incidence or severity of the change may be more difficult to detect. Alternatively, setting an appropriate threshold can aid in streamlining the number of diagnoses produced in a study, so that treatment-related changes are clear. Similar to other areas, the process of a pathology peer review will assist in ensuring appropriate diagnostic thresholds.
**Synonyms**

For all systems, the preferred term is listed at the start of each morphologic change. For many lesions, a list of synonyms (generally previously used terms) follows the preferred term. In toxicologic pathology, the name of a lesion may evolve over time although the morphology stays the same. Advancements in biology and understanding of how tissue alterations develop support the rationale for changing the terminology. In performing risk assessment, it is important that interpretation of data be based on current terminology and understanding of a particular lesion’s biological significance. However, interpretation may also require review of historical data, which may have been collected and published years or decades earlier. The evaluation of older literature must take into account changes in terminology and the current knowledge and thinking on the significance of a lesion.

**Pathology Processes Common to Organs**

**Vascular-based Changes**

Processes related to vascular changes, including congestion, hemorrhage, edema and degeneration and/or inflammation of blood vessels, will be discussed in the cardiovascular system manuscript.

**Infiltrate vs. Inflammation**

As discussed earlier the use of descriptive, rather than diagnostic, terminology is preferred. In many cases a diagnosis of inflammation, such as hepatitis, implies a particular pathogenesis, impact on organ function or clinical syndrome. It is recommended that terms describing the type of cellular infiltrate are used rather than terms which describe the type of inflammation, particularly if the response is predominantly aggregation of cells without other features of inflammation. Thus it is preferred to describe a change in the kidney as consisting of infiltrate, neutrophilic rather than either acute inflammation or acute nephritis, since these latter terms are more appropriately used in clinical medicine. Most toxicologic pathologists would not use organ specific disease descriptors (hepatitis, pneumonia, orchitis) but would build a
diagnosis using the descriptors described above to accurately describe the change. The character of the descriptors may be more fully described in the pathology narrative.

On the other hand, if the constellation of changes that characterize active inflammation including one or more of the following: increased blood flow (congestion), microvascular exudation of plasma fluids and proteins (edema), and/or margination and emigration of leukocytes are present, then a base term of inflammation followed by suitable descriptors may be appropriate. At the other end of the spectrum, more prolonged or chronic-active inflammation would be characterized by the collective features of ongoing tissue destruction, infiltration of a mixed inflammatory cell population and variable extent of granulation tissue and/or fibrosis. As noted earlier in this discussion, it is best to avoid use of terms that may have a clinical or disease connotation such as “hepatitis”.

**Intracellular Accumulations**

The recommendation is to start with a base term of vacuolation followed by modifiers as appropriate or more specific diagnosis following suitable special stains.

*Lipids*

Lipids such as cholesterol, triglycerides and phospholipids can accumulate in cells. This accumulation appears as clear vacuoles of various sizes and shapes and must be differentiated from accumulation of water or glycogen which may sometimes produce vacuoles. Cholesterol accumulation can also appear as crystalline or cleft-like spaces. Since the process of fixation with paraffin embedding removes lipid, special stains such as oil-red-O are evaluated on frozen sections. More specific descriptions and differential diagnoses are discussed in organ systems.

*Glycogen*

Excessive amounts of glycogen can be seen with glucose or glycogen metabolism abnormalities. It can also be observed in the liver of animals not fasted prior to necropsy. Excess glycogen can appear as vacuoles or more indistinct clear spaces. It is best
detected following fixation with alcohol and staining via PAS reaction. More specific
descriptions and differential diagnoses are discussed in organ systems.

**Extracellular Accumulations**

*Hyaline Change*

Extracellular hyaline change or hyaline substance refers to an alteration in which
homogeneous, glassy, eosinophilic material accumulates in tissue spaces, between cells
or along basement membranes or within the cytoplasm. Specific terminology and
descriptions are addressed in organ systems as appropriate.

*Amyloid*

The most common “hyaline” change is the accumulation of amyloid, a diverse group of
glycoproteins. It is generally extracellular, compressing adjacent parenchymal cells,
resulting in cell death from ischemia and eventual atrophy. Congo red stains amyloid
orange to orange-red. In polarized light the presence of amyloid is shown by apple green
birefringence of the Congo red-stained sections.

Amyloid may also be demonstrated by immunohistochemistry. Specific organ responses
and appearance will be reviewed in each organ system.

*Cholesterol crystals*

Cholesterol crystals are cleft-like deposits that may be present in areas of inflammation
and represent a by-product of hemorrhage and necrosis; routine processing generally
removes the deposit, leaving a clear angular space. Occasionally these crystals can incite
a granulomatous inflammatory response.

*Mineralization*

The recommendation is to start with a base term of mineralization (calcification)
followed by modifiers as appropriate. Tissue mineralization may be a result of several
distinct processes.

*Dystrophic calcification*
Dystrophic calcification can occur in areas of necrosis; the calcified material may have a basophilic, granular or clumped appearance. Progressive accumulation of layers can create a lamellate appearance (psammoma bodies). It can be intracellular or extracellular. Special stains to confirm include von Kossa and alizarin-red-S.

**Metastatic calcification**
Metastatic calcification occurs in normal tissue following syndromes that produce hypercalcemia such as hyperparathyroidism, destruction of bone secondary to neoplasms or genetic disease, renal failure and vitamin D toxicosis. It can occur throughout the body but gastrointestinal mucosa, kidneys, lungs, systemic arteries and pulmonary veins are the most common sites. The morphologic appearance is similar to dystrophic calcification.

**Osseous metaplasia**
Osseous metaplasia appears as distinct foci of immature woven bone or more dense lamellar bone. The pathogenesis of this change is not understood. It can occur throughout the body but in rodents most commonly found in lung, brain, adrenal gland, heart and eye.

**Pigments**
The recommendation is to use a base term of pigment, followed by appropriate modifiers (may be supported by special stains).

**Exogenous pigment**
Since most laboratory rodent species are bred and housed in a controlled environment with set humidity and filtered air circulation, exogenous pigments such as carbon are not common. Rodents that are identified with tattoos may have pigment within dermal macrophages and local lymph nodes.

**Endogenous pigment**
Melanin is a pigment normally present in epidermis, retina, and iris of the pigmented rodent strains. Melanin may occur normally in the spleen and meninges of mice. Lipofuscin pigment is a fine, yellow to brown, granular pigment present in the cytoplasm (frequently peri-nuclear). It has been noted as increasing with age and secondary to increased free radical injury and lipid peroxidation. Hemosiderin is formed from ferritin and is an intracellular golden-yellow to golden-brown and granular to globular pigment. Local or systemic excesses of iron result in hemosiderin accumulation. It can be easily visualized by Prussian blue stain.

Conclusions
Anatomic pathology is a descriptive and interpretive science and provides critically important information on the toxicity of environmental and biopharmaceutical agents. Toxicologic pathologists strive to provide clear and concise data and interpretations to toxicologists, physicians, regulatory reviewers and others involved in risk assessment. The inherent variability of spontaneous “background” changes in our biological models, the complexity and variability of responses to test-article, and differences in diagnostic terminology, severity grading and thresholds present challenges in consistently delivering clear interpretations. Application of the principles described in this paper and the consistent use of the harmonized nomenclature recommended in the INHAND publications will continue to increase the clarity and quality of anatomic pathology data in toxicity studies.

There are a number of possible grades and terms that may be used to describe non-proliferative lesions encountered in a toxicology study. Although there will always be some difference between the criteria utilized by individual pathologists, it is hoped that the descriptions given in this paper will help achieve some degree of international agreement and consistency as to the use of terminology to produce clear, concise data. It is hoped that the international character and review of the INHAND documents ensures that use by pathologists and national regulatory agencies will aid in the safety assessment of drugs, biological and chemicals
References


Figure 1 – Organization of INHAND
Figure 2: Organization of a Final Diagnosis
Various modifiers and severity grades can be added to a basic key term or diagnosis to describe and categorize microscopic observations. In blue is an example of the fields possible in a diagnostic term and in red are examples.