

The Society of Toxicologic Pathology's Recommendations on Statistical Analysis of Rodent Carcinogenicity Studies

The Society for Toxicologic Pathology (STP) is a nonprofit organization dedicated to improving the discipline of toxicologic pathology through education and professional interactions. The Society's membership includes over 800 pathologists and toxicologists involved worldwide in the nonclinical assessment of toxicity and carcinogenicity of chemicals, pharmaceutical candidates, and medical devices. Many of the study pathologists who interpret rodent carcinogenicity studies are members of the STP. The STP, in conjunction with representatives of the U.S. Food and Drug Administration, recognized that there were issues in the application of statistical methods to carcinogenicity studies (2). While most pharmaceutical companies used the Peto test (5), some did not categorize neoplasms as Fatal or Incidental (no contexts of observations were assigned). Some categorized neoplasms as Fatal or Incidental based solely on the type of neoplasm (diagnostic terminology) rather than on an animal-by-animal basis. Others categorized neoplasms as Fatal or Incidental based on the gross and microscopic findings for each animal. There were no guidelines for consistent assignment of Fatal and Incidental categories. The STP created an *ad hoc* Peto Working Group to review statistical methods for rodent carcinogenicity studies and make recommendations on how microscopic observations from these studies could be consistently and reliably categorized for analysis using existing statistical methods.

The STP Peto Working Group has reviewed the FDA draft *Guidance for Industry—Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals* (4) and has provided written comments and suggestions to FDA regarding this draft guidance. At the time these comments were provided, the STP was still gathering information and evaluating the use of pathology data in statistical analysis of rodent carcinogenicity studies. The STP has since solicited and published explanations and recommendations of best statistical practices from biostatisticians familiar with rodent carcinogenicity studies, and some positions provided by the STP to the FDA have been modified. The Peto Working Group also sponsored a discussion of the Peto test and the Poly-3 test with the STP membership during the 2001 STP annual meeting. The opinions of a broad segment of pathologists involved with rodent carcinogenicity studies varied and included:

1. Many pathologists believe that experienced pathologists can accurately and reliably determine if a neoplasm contributed to death or euthanasia of an individual animal, based on a review of clinical, necropsy, and microscopic findings for that individual. While most pathologists are

willing to classify neoplasms as Fatal if the neoplasms appeared to have contributed to death and Incidental if the neoplasm did not contribute to death, others are not convinced that this classification can be done consistently and accurately.

2. Pathologists cannot reliably or consistently determine the time of onset or duration of a neoplasm first found at necropsy based on gross and microscopic findings alone. Pathologists should not be requested to make such subjective judgments.
3. Pathologists cannot reliably determine whether a neoplasm contributed to death of an animal rapidly or not rapidly (slowly) based solely on gross and microscopic observations. Pathologists should not be requested to make such subjective judgments.
4. For cutaneous, subcutaneous, mammary, and other superficial neoplasms that generally can be palpated while quite small, the time of first palpation provides a better estimate of the time of onset of a neoplasm than the date of death. Pathologists believe that the time of first palpation of superficial neoplasms should be used as an estimate of the time of tumor onset when palpation data are available.

The STP offers these comments and recommendations regarding statistical analysis of rodent carcinogenicity studies. STP recommendations are noted in italics.

1. The FDA draft *Guidance for Industry—Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals* recommends that both a trend test and a pairwise comparison test be performed routinely for each study and that the results of both tests should be presented to regulatory officials. The FDA draft document also states that a trend test is generally more powerful except under specific conditions.

The STP recommends that only one statistical method should be required routinely and that the preferred method usually would be a trend test. Pairwise comparison tests should be performed on a case-by-case basis when evidence indicates that a pairwise comparison test may be more appropriate than a trend test. Analysis of study data using multiple methods increases the probability of false positive results and creates significant issues for interpretation when data are statistically significantly different using one method and not significantly different using a second method. Pathologists and other scientists should be aware that regulatory agencies may analyze data independently using more than one method.

2. Statisticians agree that assessment of "tumor incidence (the rate of tumor onset of a specific neoplasm in tumor-free animals among the previously tumor-free population) is the most appropriate measure of tumorigenesis" (4). *The STP concurs with this strategy and recommends that*

Address correspondence to: Dr. Daniel Morton, Pharmacia Corporation, 4901 Searle Parkway, H-153A, Skokie, IL 60077, USA; e-mail: dan.g.morton@pharmacia.com.

incidence rates (rather than prevalence rates) should be analyzed whenever possible. Evaluation of tumor incidence requires an estimate of the onset time of each neoplasm if this information is available. The Peto test, the Poly-3 test, and other methods require various assumptions regarding time to tumor onset and the tumor-free population over time.

3. Difference in survival between treated and control groups can bias statistical analyses of tumor incidence data by altering the relative number of animals at risk. Statisticians and pathologists involved with analysis of carcinogenicity studies agree that adjustments for intercurrent mortality should be performed, but there is not unanimous agreement regarding which method is most appropriate (3). Multiple methods exist for adjusting tumor incidence data for the effects of intercurrent mortality. The most common statistical method to adjust for intercurrent mortality is the Peto test. The Peto test requires the pathologist to classify as Fatal or Incidental all neoplasms first found at necropsy in animals that die or are euthanized in moribund condition prior to scheduled sacrifice. Alternative methods, such as Poly-k or Poly-3 tests, do not require that neoplasms be classified as Fatal or Incidental. Each statistical method requires assumptions about carcinogenicity study data and the biological behavior of neoplasms, and each has strengths and weaknesses. Dr. Peter Lee, a coauthor of the original paper describing the Peto test, and other statisticians agree that the original *ad hoc* run algorithm that Dr. Peto derived for determining time intervals for the prevalence portion of the Peto test has weaknesses and is no longer recommended (3). Assignment of time intervals is discussed in the FDA draft *Guidance for Industry—Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals*. An alternative method is used by the National Toxicology Program.

The STP recommends that statistical methods should be employed to control bias in statistical analyses related to differences in lifespan and survival among control and treatment groups. The STP does not recommend a preferred method for assigning time intervals for statistical analysis. Several options are available to companies interpreting study data that will support regulatory submissions. These include using the methods described in the FDA draft guidance, using the *ad hoc* run method until clear and final guidance is available, or consulting with regulatory agencies prior to statistical analysis.

4. The Peto Test

The original Peto test combines three statistical methods to analyze differences in tumor incidence (time to tumor onset) between treated and control groups (5). In the original Peto test, the rate-onset method is used to analyze all neoplasms that were observed before the animal came to necropsy. These neoplasms are classified as Mortality-Independent, and the time of first palpation is used to estimate of the time to tumor onset. The time of first palpation is a reasonable estimate of tumor onset time for superficial neoplasms of the skin, subcutis, mammary tissue, eyes, and distal extremities because these neoplasms usually are observed while still small. When neoplasms arise in the deep peripheral tissues or viscera, the neo-

plasms are usually large when first discovered and the time of first palpation often is an inappropriate estimate of time of onset.

For neoplasms not observed prior to necropsy, it is impossible to accurately determine or estimate the real time of onset. The Peto test requires classification of all neoplasms occurring in animals that died or were euthanized prior to scheduled sacrifice as Fatal or Incidental if the neoplasms were not observed during life. Neoplasms that contributed to death (Fatal neoplasms) in animals necropsied prior to scheduled study termination are analyzed by the death-rate method. In the death-rate method, the time of death (rather than time of first palpation) is used to estimate the time of tumor onset relative to neoplasms of the same diagnosis in other animals. There is concern among some pathologists that the death-rate method analyzes Fatal neoplasms as instantly fatal, however a majority of statisticians surveyed did not support this conclusion (3). Although the date of death is an imprecise estimate of the time of tumor onset, it is generally more appropriate to use the death-rate method for Fatal neoplasms than to use the prevalence method.

The third category of neoplasms includes those that were not observed during life and were not believed by the study pathologist to have contributed to death. All neoplasms first observed at scheduled necropsy are considered to be Incidental. These Incidental neoplasms are analyzed by the prevalence method and no estimates are made regarding the time of onset of these neoplasms.

The Peto test is required for product registration in Europe and perhaps in other locations. *Based on current regulatory requirements, the STP recommends that the Peto test should be performed whenever the study pathologist and the peer review pathologist can consistently classify neoplasms as Fatal or Incidental. Pathologists must be permitted to use their own professional judgment to determine if Fatal and Incidental classifications can be applied appropriately for most neoplasms across an entire study. If Fatal and Incidental classifications cannot be applied consistently and accurately, those responsible for interpreting the study should be permitted to collect the microscopic observations without making Fatal and Incidental classifications. If Fatal and Incidental classifications are not applied, then an alternative to the Peto method should be used in the statistical analysis (1). In some cases, Fatal and Incidental classifications will have been assigned, but later will be determined to be too inaccurate or unreliable to permit valid statistical analysis because of severe toxicity or other conditions unique to the study. In these cases, the STP recommends that the Poly-3 test or another alternative to the Peto test should be employed for statistical evaluation. These recommendation may change in the future as alternatives to the Peto test gain greater acceptance by regulatory agencies. The STP believes that is generally not advisable to carry out two sets of statistical tests in a single study, one by the Peto method and one by Poly-3 or another method, since this increases the probability of false positive results and can substantially complicate statistical interpretation.*

The reader is referred to the FDA draft *Guidance for Industry—Statistical Aspects of the Design, Analysis, and*

Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals for a discussion of alternatives to the Peto test. The FDA draft guidance suggests that the Bieler-William Poly-3 test is the most appropriate alternative (4).

If the Peto test is selected to analyze rodent carcinogenicity study data, the STP offers the following observations and recommendations for accurately and consistently classifying neoplasms as Mortality-Independent, Fatal or Incidental:

A. Neoplasms Palpated or Observed during Life in Animals Sacrificed Prior to or at Scheduled Termination

- 1) Superficial neoplasms of the skin, subcutis, mammary gland, eyes, body orifices, and distal extremities *The time of first observation or palpation is the best estimate for the time of tumor onset for all cutaneous, subcutaneous, mammary, and other superficial neoplasms that are observed during life. These neoplasms should be classified as Mortality-Independent and analyzed using the rate-onset method as described in the original Peto reference (5).*
- 2) Visceral and deep neoplasms *Fatal and Incidental classifications should be applied to palpable deep or visceral neoplasms in animals that come to necropsy prior to scheduled termination. Deep or visceral neoplasms should not be classified as Mortality-Independent even when they have been palpated while the animal was alive. Neoplasms located deep within the soft tissues or within visceral organs generally cannot be palpated in life until quite large. For neoplasms that generally are large when first observed, the date of first observation is an inappropriate estimate for the time of tumor onset.*

B. Neoplasms That Were Not Palpated Prior to Necropsy in Animals that Died or Were Euthanized Prior to Scheduled Sacrifice

- 1) *Each occult neoplasm (a neoplasm first discovered at necropsy) in animals found dead or euthanized prior to scheduled sacrifice should be classified as Fatal or Incidental.*
- 2) *Pathologists should use their best professional judgment to categorize each neoplasm first found at necropsy and deep and visceral neoplasms palpated prior to necropsy as Fatal or Incidental, utilizing all appropriate in-life, necropsy, and microscopic observations. It is inappropriate to classify all neoplasms of a certain diagnosis (eg, all malignant lymphomas) as either Fatal or Incidental a priori based on diagnostic terminology, since such a decision would not involve consideration of other lesions contributing to the morbidity or mortality of each animal.*
- 3) *On rare occasions, an animal may have more than one Fatal neoplasm if the pathologist determines that more than one neoplasm contributed to early death or euthanasia.*

4) *Neoplasms that otherwise might or would have contributed to death had the animal lived longer, but did not actually do so because the animal died first from a different cause or set of causes (another potentially fatal neoplasm, chronic renal failure, traumatic injury, etc.), should be classified as Incidental.*

5) *All Fatal neoplasms should be analyzed using the death-rate method. Incidental neoplasms should be analyzed using the prevalence method.*

C. Neoplasms First Observed at Scheduled Termination

- 1) *All neoplasms that were not observed during life in animals that survive until scheduled sacrifice should be considered Incidental and analyzed by the prevalence method, since these neoplasms did not contribute to the death of the host.*

D. Rapidly Fatal and Not Rapidly Fatal Classifications

In the Peto Working Group's previously published draft comments, there was concern that the Peto test treats Fatal neoplasms as instantly fatal, ie, that the time of death is used as a surrogate for the time of tumor onset (6). This assumption is important in estimating the number of animals at risk (the tumor-free population). Classification of Fatal neoplasms as instantly fatal seemed inappropriate to the Peto Working Group in most cases, since many Fatal neoplasms (such as pituitary neoplasms) were known to be present for several to many weeks before death or euthanasia, and the number of animals with occult neoplasms that will become Fatal cannot be estimated accurately at any given time. The Peto Working Group suggested that classifying Fatal neoplasms as Rapidly Fatal or Not Rapidly Fatal and handling each subclassification using different statistical approaches might be helpful. Most statisticians surveyed by the Peto Working Group did not believe that the Peto test modeled Fatal neoplasms as instantly fatal. Those statisticians who believed that the Peto test uses the time of death as a surrogate for the time to tumor onset did not consider this to be a significant concern provided that pathologists could reliably and consistently classify neoplasms as Fatal or Incidental. Most statisticians believed that the death-rate method is appropriate if pathologists can consistently and accurately classify as Fatal those neoplasms that significantly contributed to death. Although not ideal, the time of death is the best available indicator of the time of tumor onset for occult Fatal neoplasms. Consultation with many pathologists has indicated that pathologists cannot reasonably classify Fatal neoplasms as Rapidly Fatal or Not Rapidly Fatal. *Rapidly Fatal and Not Rapidly Fatal classifications of neoplasms are not practical and should not be performed.*

Note : Identification of the cause of death in individual animals (not to be confused with classification of neoplasms as Fatal or Incidental) adds significant value to the interpretation of carcinogenicity studies. Causes of death may include neoplasms, cardiomyopathy, chronic renal disease, trauma, and infectious conditions. Whenever possible, pathologists

should identify cause of death for animals dying or euthanized before scheduled sacrifice as a means to interpret causes of differential mortality and co-morbidities among groups. This classification cannot be construed to imply a date of onset or a rate of progression of lesions and cannot be specifically designed to monitor lesion onset. In some animals, the cause of death will not be apparent and would be recorded as undetermined. Cause of death data for each animal are evaluated qualitatively and are independent of standard statistical analyses in evaluation of the overall study results.

STP Peto Working Group Participants

Michael Elwell
William Fairweather
Xavier Fouillet
Kevin Keenan
Karl Lin
Gerald Long
Lori Mixson
Daniel Morton
Terry Peters
Colin Rousseaux
Darrell Tuomari

REFERENCES

1. Bieler GS, Williams RL (1993). Ratio estimates, the Delta method, and quantal response tests for increased carcinogenicity. *Biometrics* 49: 793–801.
2. Gad SC, Rousseaux CG (2002). Use and misuse of statistics in the design and interpretation of studies. In: *Handbook of Toxicologic Pathology*, Haschek WM, Rousseaux CG, Wallig MA (eds). Academic Press, San Diego, 2nd Edition, Volume 1, Chapter 15, pp 327–418.
3. Lee PN, Fry JS, Fairweather WR, Haseman JK, Kodell RL, Chen JJ, Roth AJ, Soper K, Morton D (2002). Current issues statistical methods for carcinogenicity studies. *Toxicol Pathol* (in press).
4. Office of the Federal Register (2001). Guidance for industry—statistical aspects of the design, analysis, and interpretation of chronic rodent carcinogenicity studies of pharmaceuticals (draft guidance). *Federal Register* 66(89), May 8, 2001, 23266–23267.
5. Peto R, Pike M, Day NE, Gray RG, Lee PN, Parish S, Peto J, Richards S, Wahrendorf J (1980). Guidelines for simple, sensitive significance tests for carcinogenic effects in long-term animal experiments. In: *IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Humans, Supplement of Long-term and Short-term Screening Assays for Carcinogens: A Critical Appraisal*. International Agency for Research on Cancer, Lyon, France, Supplement 2, pp 311–426.
6. STP Peto Working Group (2001). Draft recommendations on classification of rodent neoplasms for Peto analysis. *Toxicol Pathol* 29: 265–268.