DEPARTMENT OF HEALTH AND HUMAN SERVICES

Food and Drug Administration

[Docket No. 76F-0392]

Cyclamate (Cyclamic Acid, Calcium Cyclamate, and Sodium Cyclamate), Commissioner's Decision

AGENCY: Food and Drug Administration.

ACTION: Final decision following a formal evidentiary public hearing.

SUMMARY: The Commissioner of Food and Drugs is issuing his Final Decision concerning the food additive petition for the artificial sweetener cyclamate. The Commissioner has determined that cyclamate has not been shown to be safe for the proposed use as a food additive and is denying approval of the petition. The Commission has based this decision on two independent grounds: (1) cyclamate has not been shown not to cause cancer; and (2) cyclamate has not been shown not to cause heritable genetic damage. Accordingly, the Initial Decision of the Administrative Law Judge are affirmed, with supplementation and modification as contained herein.

EFFECTIVE DATE: December 15, 1980.

ADDRESS: The transcript of the hearing, evidence submitted and all other documents listed in this decision may be seen in the Office of the Hearing Clerk (HFA-305), Food and Drug Administration, Rm. 4-65, 5600 Fishers Lane, Rockville, MD 20857, from 9:00 a.m. to 4:00 p.m., Monday through Friday.

FOR FURTHER INFORMATION CONTACT: Ted Herman, Compliance Regulations Policy Staff (HFC-10), Food and Drug Administration, Department of Health and Human Services, 5600 Fishers Lane, Rockville, MD 20857 301-443-3480.

SUPPLEMENTARY INFORMATION: The purpose of this proceeding is to decide whether cyclamate has been shown to be safe under Section 409 of the Federal Food, Drug, and Cosmetic Act ("the act"), 21 U.S.C. 348.

Table of Contents—Cyclamate

I. Background

A. History [1]

The Food and Drug Administration ("FDA") first approved cyclamate for commercial use in 1951, when Abbott Laboratories, Inc. ("Abbott") filed a new drug application for use of cyclamate as a table top sweetener under the trade name "Sucaryl." Sucaryl was recommended for use in treatment of obese patients and by individuals with diabetes.

The regulatory status of cyclamate was changed as a result of the enactment of the Food Additives Amendment of 1958, 21 U.S.C. 348. This amendment was added to the act to require that food additives be tested to establish their safety prior to marketing. An exception to this premarket approval system was made for substances generally recognized as safe ("GRAS"). [2]

[1] The statement of the history of this proceeding is taken in part from a September 29, 1971, written statement by then Commissioner Charles C. Edwards which was presented to a subcommittee of the House Committee on Judiciary. *Cyclamates:* hearings on HR 4264, HR 4180, HR 4265, HR 4870, HR 4912, HR 4858, HR 5862, HR 6163, HR 6155 before Subcommittee No. 2 of the Committee on the Judiciary, House of Representatives, 92nd Cong., 1st Sess. 75-113 (1971).

[2] The term "generally recognized as safe" and its abbreviation, GRAS, are shorthand for the language in Section 201(s) of the act (underlined below) which has the effect of exempting GRAS substances

Footnotes continued on next page

To implement the Food Additive Amendment, FDA compiled an advisory "GRAS list" of substances already on the market. The final list of November 20, 1959, included cyclamate. In 1961, FDA advised Abbott that sodium cyclamate was no longer considered to be a drug, and was considered to be generally recognized as safe as a food ingredient.

In the early 1950's combinations of cyclamate and saccharin gained wide use in fabricated foods. To determine whether a mixture of cyclamate and saccharin gave results different from those reported in earlier experiments where cyclamate or saccharin was tested alone, Abbott, in 1967, contracted with the Food and Drug Research Laboratories, a private research institution, to conduct a study. In this study, eight of 60 rats fed a 10:1 mixture of sodium cyclamate and sodium saccharin for two years developed bladder tumors. (For a further discussion of this study, see Section IV.B.3.c ("Oser study"), below.)

Because the Food and Drug Research Laboratories study implicated cyclamate as a possible carcinogen, the then Commissioner of Food and Drugs, Herbert L. Ley, removed calcium cyclamate, magnesium cyclamate, potassium cyclamate and sodium cyclamate from the GRAS list (then 21 CFR 121.101) and limited the marketing of those cyclamate compounds to therapeutic uses as drugs (34 FR 17063, October 21, 1969). On August 27, 1970, FDA concluded that there was no substantial evidence of effectiveness of cyclamate compounds at any level for treatment of obese patients and individuals with diabetes and therefore prohibited continued sale of cyclamate-containing products with drug labeling (35 FR 13644). This action was based on the advice of a Medical Advisory Group established by the Assistant Secretary for Health and Scientific Affairs, Department of Health, Education and Welfare. The Medical Advisory Group also endorsed a prohibition, based on safety grounds, of cyclamate in beverages for general use and in the future processing of general purpose food (*id.* at 13645).

B. Administrative Proceedings

On November 15, 1973, Abbott filed a food additive petition (FAP 4A 2975) pursuant to Section 409(b) of the act seeking approval for the use of cyclamic acid, calcium cyclamate and sodium cyclamate (hereinafter collectively referred to as "cyclamate") [3] as sweetening agents in food and for technological purposes [4] in food. It is this petition which is the subject of this proceeding. FDA published a notice of filing of Abbott's petition in the Federal Register of February 8, 1974 (39 FR 4935). After reviewing Abbott's food additive petition to determine whether it met the criteria for approval of such a petition set forth in Section 409(c) of the act, the then Commissioner A. M. Schmidt concluded that the supporting data did not establish that cyclamate is safe for its intended use. The food additive petition was therefore denied by order in the Federal Register of October 4, 1976 (41 FR 43754).

Abbott and the Calorie Control Council, an industry trade group, filed objections to, and a request for hearing on, the October 4, 1976 order; only Abbott, however, made particularized objections. In the Federal Register of March 4, 1977 (42 FR 12515), the then Acting Commissioner, Sherwin Gardner, granted Abbott's request for a hearing pursuant to Section 409(f) of the act.

The formal evidentiary hearing began with a prehearing conference held on April 20, 1977. The issues considered at the hearing, as set forth by the Administrative Law Judge at the Prehearing Conference, were as follows:

(1) Whether the evidentiary record establishes to a reasonable certainty that cyclamate does not induce cancer when ingested by man or animals.

(2) Whether the evidentiary record establishes to a reasonable certainty that cyclamate does not cause genetic damage and is not mutagenic.

(3) Apart from the issues in Numbers 1 and 2 above, what does the evidentiary record show as an acceptable daily intake level for cyclamate?

(4) Whether apart from the issues in Numbers 1 and 2 above, because of the probable consumption patterns, safe conditions of use of cyclamate can be prescribed.

The parties in the hearing were the Bureau of Foods of the Food and Drug Administration ("Bureau") and Abbott. See 21 CFR 10.3(a).[5] The Bureau contended that Abbott's food additive petition for cyclamate should be denied. Abbott, of course, contended that its petition should be approved.

Testimony concerning the issues in the hearing was submitted in written form. Oral cross-examination was completed and briefs submitted to the ALJ by January 23, 1978.

On August 4, 1978, the Administrative Law Judge issued an Initial Decision in which he found that cyclamate has not been shown to be safe. Specifically, the ALJ found (ID at 38–39): [6]

(1) Cyclamate has not been shown to be safe as required by Section 409 of the Federal Food, Drug, and Cosmetic Act (21 U.S.C. 348).

(2) It has not been shown to a reasonable certainty that cyclamate does not cause cancer in man or animals.

(3) It has not been shown to a reasonable certainty that cyclamate is not a mutagen.

(4) In the event that the carcinogenicity and mutagenicity questions are resolved, the record in this proceeding would support a finding that the acceptable daily intake is five mg cyclamate/kg body weight/day or less.

(5) Even if the carcinogenicity and mutagenicity questions were to be subsequently resolved, the record in this proceeding does not establish probable consumption patterns of cyclamate to the extent necessary to establish safe conditions of use.

On June 26, 1979, the then Commissioner, Donald Kennedy, issued an interlocutory order remanding the case to the ALJ to develop the evidence further on certain issues relating to the safety of cyclamate. This order was published in the Federal Register of August 14, 1979, with minor non-substantive changes (44 FR 47620).

Footnotes continued from last page from the definition of "food additive:" The term "food additive" means any substance the intended use of which results or may reasonably be expected to result, directly or indirectly, in its becoming a component or otherwise affecting the characteristics of any food (including any substance intended for use in producing, manufacturing, packing, processing, preparing, treating, packaging, transporting, or holding food; and including any source of radiation intended for any such use), if such substance is not *generally recognized, among experts qualified by scientific training and experience to evaluate its safety, as having been adequately shown through scientific procedures (or, in the case of a substance used in food prior to January 1, 1958, through either scientific procedures or experience based on common use in food) to be safe under the conditions of its intended use;* * * * *21 U.S.C. 321(s) (emphasis added).*

[3] These three entities are being referred to simply as cyclamate because, in the gastrointestinal tract of animals fed any one of these three compounds, the actual form of cyclamate will be the same. For this reason, all three entities are considered to be chemically and biologically equivalent.

[4] Food additives are used for a variety of technological purposes, examples of which are set forth in 21 CFR 170.3(o).

[5] Dr. Michael Sveda, the discoverer of cyclamate, also appeared as a non-party participant *(id.)*. His appearance was subsequently stricken for failure to participate. See 21 CFR 12.45(e).

[6] The following abbreviations have been used in citing material in the record: Initial Decision: ID; Transcript: Tr.; Briefs to the ALJ: Brief; Exceptions to the Initial Decision: Exceptions; Replies to Exceptions: Reply; Briefs to the ALJ following reopened hearing: Remand Brief; Initial Decision following the reopened hearing: IRD; Transcript of hearing following the remand: R. Tr.; Exceptions to IRD: Remand Ex.; Replies to Remand Ex.: Remand Reply. This decision refers to the exhibits submitted to the record, including written direct testimony, by the following: Bureau: G; Abbott: A.

The order identified several areas in which the evidence needed further development. First, Commissioner Kennedy found data in the record concerning lung, liver, lymphoid tissue and mammary tumors in a number of studies which involved direct feeding of cyclamate to animals. Because these data could have had an impact on the final outcome of the proceeding, but were not fully analyzed or addressed by the parties, Commissioner Kennedy asked the parties to consider them. Second, Commissioner Kennedy asked that the evidence pertaining to the criteria for the evaluation of carcinogenicity data be further developed. The parties were asked to elaborate on their positions concerning what constitutes a "negative" study and the concept of "statistical significance." Third, the parties were asked eleven specific questions concerning the animal studies designed to determine the possible carcinogenicity or mutagenicity of cyclamate. The parties submitted stipulations on these eleven specific questions on September 18, 1979. On October 22, 1979, the parties submitted written testimony and written statements of position to the ALJ. Oral cross-examination was held on November 5, 1979. The parties submitted briefs on December 3, 1979.

Following the consideration of all the data submitted at the reopened hearing, on February 4, 1980, the ALJ issued an Initial Decision on Further Hearing. The ALJ concluded that "it is apparent that the reevaluation of the evidence presented on further hearing tends to increase the likelihood that cyclamate is a carcinogen" and that "[c]onsideration of the entire record in this proceeding requires the finding that petitioner has failed to sustain its statutory burden of establishing to a reasonable certainty that the proposed use of cyclamate will be safe * * *" (IRD at 23–24).

On February 25, 1980, Abbott and the Bureau submitted exceptions to the Initial Decision on Further Hearing. In its exceptions, Abbott requested oral argument before the Commissioner (Abbott's Remand Ex. at 32). Because I do not find oral argument necessary, I am denying that request. See 21 CFR 12.125(e).

Before proceeding further, a few words need to be said about Abbott's contentions that Commissioner Kennedy's Remand Order was "completely specious, consisting of inconsequential and artificially contrived questions none of which needed further evidentiary development prior to a final determination on Abbott's petition" (Remand Brief at 2–3). See also Abbott's Remand Ex. at 2–3. I find this contention to be without merit, for my own review of the full record reveals that the further analyses of evidence undertaken pursuant to the Remand Order have materially improved the quality of the record.

Significantly, the reopened hearing established that the Kroes study, which was previsously believed by Abbott and the Bureau to be negative, in fact contained data that, when analyzed, showed a statistically significant incidence of lymphosarcomas (G–139 at 7). This study, discussed in more detailed below (Section IV.B.1.c.), plays an important role in my final decision. It also became clear on remand that other important data had previously been overlooked, see *e.g.* finding of statistical significance for total tumors in the Rudali study, Section IV.B.1.a.(3).

The remand also gave Abbott a further opportunity to submit additional evidence and argument on important and complex issues raised by Commissioner Kennedy. The record reflects that Abbott took full advantage of this opportunity. Abbott submitted the testimony of three witnesses totalling sixty pages and an eight page stipulation. Some of Abbott's comments submitted at the reopened hearing, such as the use of certain statistical corrections, have been adopted in this decision.

It is true that some of the questions raised by the Remand Order, standing alone, might not ordinarily warrant reopening a hearing. However, once it became necessary to reopen the hearing because the record contained potentially significant but unanalyzed data, it was only prudent to include less significant inquiries in the Remand Order.

II. Statutory Requirements For Approval of a Food Additive Petition

Section 409 of the act sets up a premarket approval system for food additives.[7] It declares that the presence of an unapproved food additive renders a product adulterated, and therefore unlawful. 21 U.S.C. 409(a). It also provides a mechanism by which the sponsor of a food additive may seek approval from the Food and Drug Administration.

This premarket approval system represented a considerable departure from the prior system. Before passage of the Food Additive Amendment of 1958, food additives could be marketed without any advance demonstration of safety. In order to prohibit sale of a food additive prior to 1958, FDA was required to show, through its own testing, consumer injuries, or other means, that the food additive posed a hazard to health. The Amendment thus reflects a Congressional response to the need in contemporary society for a scientifically and administratively sound basis for determining the safety of food additives prior to their marketing. *Cf. Certified Color Mfg. Ass'n.* v. *Mathews*, 543 F.2d 284, 286–87 (D.C. Cir. 1976).

Section 409 of the act provides that a regulation approving a food additive petition shall not issue if a fair evaluation of the data

> [F]ails to establish that the proposed use of the food additive, under the conditions of use to be specified in the regulation, will be safe: Provided, that no additive shall be deemed to be safe if it is found to induce cancer when ingested by man or animal, or if it is found, after tests which are appropriate for the evalution of the safety of food additives, to induce cancer in man or animal * * *

21 U.S.C. 348(c)(3)(A). The proviso to this subsection of the act (*i.e.*, the language after the word "Provided") is the so-called "Delaney clause." It prohibits the marketing of any food additive that has been found to induce cancer when ingested by man or animal. While the Delaney clause is often the subject of considerable attention, see, *e.g.*, Merrill, *Regulating Carcinogens In Food: A Legislator's Guide To the Food Safety Provisions of the Federal Food, Drug, and Cosmetic Act*, 77 Mich. L. Rev. 171 (1978), it is not being invoked in this proceeding because the evidence submitted does not conclusively establish that cyclamate is a carcinogen. My analysis, therefore, will be conducted under the first clause of the above-quoted provision (the language before the word "Provided"). This clause is known as the "general safety clause." The general safety clause applies to a wide range of adverse health effects, including the potential carcinogenicity and mutagenicity of a food additive, the two issues to be addressed in this decision.

Under the general safety clause, a food additive regulation permitting use of a substance can be issued only if "the data" submitted to the agency in a food additive petition "establish" that the proposed use of the food additive "will be safe." 21 U.S.C. 348 (c)(3)(A). Two aspects of this statutory standard deserve attention: the locus of the burden of proof, and the meaning of the word "safe."

By requiring that the data in support of a food additive petition "establish" safety, Congress has put the burden of proof on the petitioner. *Monsanto* v. *Kennedy*, 613 F.2d 947, 955 (D.C. Cir. 1979). FD&C Act Red No. 2; Denial of

[7] The definition of "food additive," 21 U.S.C. 321(s), is set forth at footnote 2.

Petition for Permanent Listing; Final Decision, 45 FR 6252 (January 25, 1980); Benylin; Denial of Approval of Supplemental New Drug Application; Final Decision; 44 FR 51512 (August 31, 1979). See 5 U.S.C. 556(d); 21 CFR 12.87(d).

In determining whether petitioner has met its burden, the agency must, as a logical matter, arrive at one of three possible conclusions. First, it may find that the evidence establishes that the additive is "safe." Second, the agency may find that the evidence establishes that the additive is unsafe. Third, the agency may find that the evidence is such that the safety of the additive is unknown or uncertain. By allocating the burden of proof to the petitioner, Section 409 authorizes FDA approval of a food additive petition only in the first situation. Confronted with either the second or third situation, the agency must deny the petition.

Although the term "safe" is not defined in Section 409 of the act, the legislative history of the Food Additives Amendment of 1958 makes clear that the term "safe" was not intended to require absolute proof of safety. The House Report states that:

> * * * Safety requires proof of a *reasonable certainty that no harm will result* from the proposed use of an additive. It does not—and cannot—require proof beyond any possible doubt that no harm will result under any conceivable circumstance.
>
> This was emphasized particularly by the scientific panel which testified before the subcommittee. The scientists pointed out that it is impossible in the present state of scientific knowledge to establish with complete certainty the absolute harmlessness of any chemical substance.

(H.R. Rept. 2284, 85th Cong., 2d Sess., pp. 4–5, 1958.) (Emphasis added.)

The Senate Report agreed with the assessment of the term "safety" contained in the House Report, noting:

> * * * Conscious of the fact that any substance or, for that matter, any particular food known to be good for the health of human beings can be deleterious to the health of an individual who insists on consuming inordinate amounts of it, the committee agrees with the Food and Drug Administration that, instead of insisting on proof beyond any possible doubt that no harm will result under any conceivable circumstances from the use of a particular additive * * * the test which should determine whether or not a particular additive may be used in a specific percentage of relationship to the volume of the product to which it might be added should be that of reasonable certainty in the minds of competent scientists that the additive is not harmful to man or animal, subject to the procedural safeguards provided in the bill which assure the right to hearing and judicial review.

(Senate Report No. 2422, reprinted in [1958] U.S. Code Cong. and Admin. News 5301.)

FDA's interpretation of the term "safe" used in section 409 of the act is consistent with the act's legislative history. FDA's regulations provide that a food additive is "safe" if "there is a reasonable certainty in the minds of competent scientists that the substance is not harmful under the intended conditions of use" 21 CFR 170.3(i).

Taken as a whole, then, Section 409 means that Abbott has the burden of proving that the data in the record establish that there is a reasonable certainty of no harm from use of cyclamate. There is considerable disagreement, however, about how to apply that principle.

Abbott contends that the Bureau's witnesses did not base their opinions on presently accepted scientific methods and that therefore the Bureau is advocating a standard of "emotional certainty" rather than "reasonable certainty" (Abbott's Brief at 2–7). I recognize that Congress did not intend to impose a burden higher than "reasonable certainty." At the same time, it must be understood that what must be proved to a reasonable certainty is "no harm." That burden may be hard to meet, for credible proof of some harm will undercut efforts to prove no harm, even if there is not enough proof to make out a certain case of harm. That is the way Congress intended it, and for good reason. The Food Additives Amendment of 1958 protects against carcinogens, mutagens, and other dangers in our food supply. By allocating the burden of proof as it did, Congress asked FDA to be conservative in deciding whether to approve food additives.[8]

Abbott also contends that, for a scientist to conclude that cyclamate has not been shown to be safe, there must be an "objective basis for the evaluation of the data presented" (Abbott's Brief at 2–7). I agree.

It is not possible, however, to provide a formula specifying precisely the quantity and quality of evidence an applicant is required to submit in order to meet its burden. But the lack of a precise formula does not mean that the process lacks objectivity. Nor does a lack of certainty mean a lack of objectivity, especially where the subject matter is complex and the science evolving. The requirement of objectivity is met, I believe, if the agency reviews the evidence carefully, conducts a fair evaluation of the evidence, states its reasons for crediting or not crediting a piece of evidence, weighs all the evidence, applies the correct statutory standards, and decides.

As the discussion in Section IV below demonstrates, the evidence submitted in this proceeding does not provide a reasonable certainty of no harm from cyclamate. Many of the studies contain deficiencies and are, therefore, simply inadequate, whether to prove safety or lack of safety. Of the studies in the record entitled to weight, a significant number suggest, though they do not prove, that cyclamate is a carcinogen and a mutagen. As a scientific matter, one can imagine studies which would negate these suggestions of carcinogenicity and mutagenicity. But no such studies are included in the record. Those studies in the record in which no carcinogenic or mutagenic effect was found are either too insensitive to rely on as proof of safety or do not detract sufficiently from the studies which suggest that cyclamate is a carcinogen or mutagen. In these circumstances, the petition must fail, for the evidence supporting it does not establish the safety of cyclamate.

III. Carcinogenicity: The Scientific Framework

A. Criteria for the Evaluation of Carcinogenicity Studies

Beginning in Section IV, I examine the carcinogenicity studies contained in the food additive petition for cyclamate. Two major issues recur in that discussion. One is "statistical significance." The other is "biological significance." These two concepts are applied to interpret the results of animal studies in which one or more groups of animals [9] are fed a test substance and

[8] In any event, as discussed in Sections IV and V below, the Bureau's witnesses did not hold the evidence in this proceeding to a standard higher than "reasonable certainty," but rather evaluated it in light of presently accepted scientific methods.

[9] Both parties rely on their interpretations of results from tests conducted on laboratory animals. Indeed, one of the underlying premises of this proceeding is that results from such tests can be used as a basis for determining the safety or carcinogenic potential of a test substance in humans, a principle generally recognized by scientists. This principle was expressly recognized in section 409(c)(3)(A) of the act (the Delaney Clause) which commands the denial of a food additive petition if the food additive in question . . . is found to induce cancer when ingested by man or *animal*, or if it is found, after tests which are appropriate for the evaluation of the safety of food additive, to induce cancer in man or *animal*. . . . [emphasis added]. 21 U.S.C. 348(c)(3)(A). That the Delaney Clause is not being invoked in this proceeding does not preclude reference to it for purposes of ascertaining Congressional intent with respect to use of animal data.

Courts have consistently upheld government regulatory actions against carcinogens or suspected carcinogens based, at least in part, on results from tests on laboratory animals. *Environmental Defense*

Footnotes continued on next page

one or more control groups are fed the same diet and handled in the same manner as the treated groups except that the controls do not receive the test substance. The incidence of tumors in the treated group is then compared to the incidence of tumors in the control group. A finding that a test result has "statistical significance" involves the use of statistical methodology to determine the probability that the observed difference, if any, in the incidence of tumors in treated animals compared to controls is associated with the test substance rather than a chance occurrence. A finding that a test result has "biological significance" involves consideration of certain biological factors which provide information about the proper interpretation of the results. Together, these two criteria help scientists to decide what, if any, conclusions can be drawn from the results of a study.

1. *Statistical Significance.* The term "statistical significance" is generally understood to refer to a conclusion that there is a small probability that the observed difference between control and treated animals is due to chance. This probability is expressed as a decimal, *e.g.*, P=.1. The smaller the P-value, the less the probability that the effect is associated with chance and hence the greater the likelihood that the effect is associated with treatment. The larger the P-value, the greater the probability that the result is due to chance and hence the less the likelihood that the effect is associated with treatment.

For example, assume that a study is performed in which both treated and control groups consist of 100 animals and five tumors are found in treated animals and none in controls. In this hypothetical study, the probability (P) that the observed difference in tumor incidence between treated and control animals is due to chance is P=.03. A p-value of .03 means that the probability of the observed difference in tumor incidence being due to chance alone is 3 in 100 (3 percent) and therefore the probability of the observed difference being associated with treatment is 97 in 100 (97 percent).[10] If the number of tumors found in treated animals were four instead of five, and none was found in controls, the P-value would be P=.061 (rather than P=.03) and there would thus be a greater likelihood that the result was due to chance (6.1 percent rather than 3.0 percent). In contrast, if ten tumors instead of five were found in treated animals and none in controls, the P-value would be P=.001 (rather than P=.03) and there would thus be less likelihood that the result was due to chance (0.1 percent rather than 3.0 percent).

2. *Biological Significance.* "Biological significance," as its name implies, involves consideration of biological factors. Some of the factors typically considered are the methodology of the study involved, the existence of a dose response relationship, the rarity of tumors, and the presence of similar results in other studies (G–139 at 5).

For example, there may be an observed difference in tumors between treated and control animals. If it is determined, however, that due to a mistake those treated animals with tumors did not receive the test substance, then, obviously, the tumor difference in the experiment cannot be attributed to the test substance. Similarly, if there is no difference in tumor incidence between treated and control animals, but there is a substantial defect in the design or conduct of the study, the results of the study would be considered biologically insignificant.

The methodology of a study includes consideration of factors such as whether animals in the study are randomly allocated to treated and control groups, whether treated and control animals are handled in the same way, whether all control animals receive the same feed, whether all treated animals receive the same test substance, and the manner in which the test substance is administered. Each of these factors can have an effect on the outcome, and must be considered in deciding how much weight to give a study. Suppose, for example, that treated animals are administered the test substance through a tube which irritates their throats. The better practice would be to insert the same tube in the control animals, so that their throats are subjected to the same irritation as the treated animals. If this is not done, one cannot be as sure as the statistical significance might suggest that any resulting throat cancers are due to the substance (rather than to the irritation).

The relationship between increasing dose of the test substance and the effect observed is known as the dose response relationship. Dose response relationship is another consideration involved in a determination of biological significance. Carcinogens are known to exhibit dose response relationships.[11] The presence of a dose relationship is looked for in studies employing more than one dose level of the test substance. If the effect observed increases as the dose level of the test substance increases, it is more likely that the effect observed is due to the test substance and more weight can be given to the results of the study. Conversely, the absence of a dose response relationship in studies where such a relationship would be expected to occur, may detract from the weight to be given a study.

Another consideration involved in a determination of biological significance is whether or not the same effect occurs in more than one study. If it does, the significance of the studies may be enhanced.

As noted above, there is an interrelationship between statistical significance and biological significance. Scientists view the statistical and the biological data together to determine what, if any, conclusions can be drawn from the results of the study.

It should be emphasized, however, that neither statistical significance nor biological significance supplies formulaic answers to questions about the meaning of data. They are very useful tools—analogous to canons of statutory construction in assessing legal problems—but that is all they are. They must be used, as Commissioner Kennedy has said, with "the purposes of the scientific enterprise" for which they are being applied in mind (44 FR 47622).

3. *Position of the Parties and Findings of the ALJ on Statistical Significance.* Abbott equates statisical significance with P<.05. In other words, Abbott contends that only when the P-value for the incidence of cancer in cyclamate-treated animals is less than or equal to .05 can a study be considered positive and therefore serve as a basis for denying approval of a food additive petition (Abbott's Remand Brief at 15).

Footnotes continued from last page

Fund v. *E.P.A.*, 598 F.2d 62, 87–89 (D.C. Cir. 1978); *Environmental Defense Fund* v. *E.P.A.*, 548 F.2d 998, 1006–10 (D.C. Cir. 1976), *rehearing denied*, 548 F.2d 1012 (D.C. Cir. 1977); *Synthetic Organic Chemical Manufacturers Ass'n.* v. *Brennan*, 506 F.2d 385, 387 (3d Cir. 1974), *cert. denied*, 420 U.S. 973 (1975). This principle is also recognized throughout the record (*see, e.g.*, G–97 at 1; A–647 at 3–8).

It should also be noted that use of animal data serves an important ethical purpose as well: it obviates the need for routine testing in humans of potential carcinogens.

[10] Instead of using the decimal which expresses the likelihood that the effect is due to chance (here, .03), some statisticians refer to a confidence level that the effect is due to the treatment (here, 97%). The two expressions are different ways of saying the same thing.

[11] Lowering the dosage of carcinogens known to follow a dose response curve can result in a "noncarcinogenic" effect, *i.e.*, a dosage at which the carcinogen will not produce a statistically significant increase in tumors (see *e.g.* Section IV.B.3.b.(3) below). It is important to note, however, that such a "noncarcinogenic" dosage of a known carcinogen would not be considered safe because thresholds for carcinogens have not been established (see Tr. at 1068–69).

Abbott contends that where the P-value for the increased incidence of cancer in cyclamate-treated animals compared to control animals is greater than .05, the study must be treated as negative and therefore can provide a basis for approving a food additive petition (Abbott's Remand Brief at 18).

Abbott argues that use of the .05 confidence level is standard and is supported by traditional usage (Abbott's Remand Brief at 14). Abbott further contends that if carcinogenic effects that are not significant at $P<.05$ are used to conclude that cyclamate is potentially a weak carcinogen, "science is done a disservice and any hearing is an exercise in futility" (Abbott's Exceptions at 9). In Abbott's view, consideration of any carcinogenic effect that is not statistically signifiance at $P<.05$ as a basis for concluding that cyclamate has not been shown to be safe is a "subjective and arbitrary treatment [that] has never been the established practice of the Agency" (Abbott's Remand Brief at 14). In support of the latter statement, Abbott relies on a Bureau of Foods strategy document which it claims shows that the Bureau will not label a finding "positive" unless that finding has a P-value of less than .05 (Abbott's Remand Ex. at 23–25). Abbott thus contends that the Bureau is advocating in this proceeding a higher standard than it ordinarily uses in reviewing food additive petitions.

The Bureau recognizes that "out of convention $P<.05$ continues to serve as a benchmark for statistical significance" and that statistical significance at $P<.05$ may well be a prerequisite to labeling a study unequivocally positive (Bureau's Position Paper at 8; Bureau's Remand Reply at 5). The Bureau contends, however, that effects which are not statistically significant at $P<.05$ may nevertheless be relied upon as a basis for denial of a food additive petition (Bureau's Remand Reply at 5–6).

In support of its position, the Bureau has adopted the following observations made by Commissioner Kennedy in his Remand Order with respect to statistical significance:

> The use of "statistical significance" in the scientific community has not had the degree of inflexibility that the parties in these proceedings have assumed it has. Although the ".05" confidence level has often been used in the scientific literature to determine whether a result is positive, there is no fixed convention on the matter, * * *
>
> * * * * * *
>
> There is always a temptation to adopt the highest possible confidence level, particularly in the scientific community where a very high value is given to the avoidance of a false positive result. Especially high reliance is placed on reports of positive results because they are used to construct new hypotheses and theories and will be incorporated into the body of assumed scientific knowledge. But no particular value of significance constitutes a law of nature; it is a matter of scientific custom, reflecting human value judgments about the purposes of the scientific enterprise. And in some contexts we are especially troubled by the prospect of mistakenly declaring that the results of a study are negative, i.e., of mistakenly concluding that a study demonstrates safety. Such a decision, if incorrect, could result in the widespread marketing of a carcinogen. A regulatory agency may therefore have less reason than scientists do to insist on a very high degree of certainty before concluding that a study is positive. Similarly, there may be reason for a regulatory agency to require greater stringency than other scientists require before concluding that a study is negative.

(44 FR 47622; Bureau's Position Paper at 7; see G–139 at 3–6.)

The Bureau further contends that the strategy document Abbott relies on is not the official position of the Bureau, is not in evidence, and therefore should not be considered (Bureau's Remand Reply at 4). The Bureau also asserts that the use of statistical criteria discussed in the strategy document is not inconsistent with the position the Bureau has advocated in this proceeding (Bureau's Remand Reply at 5–6). The ALG adopted the Bureau's position that effects which are not statistically significant at $P<.05$ may nevertheless support the conclusion that a food additive has not been shown to be safe (IRD at 12–13).

4. *Commissioner's Findings on Statistical Significance. Although $P<.05$ has in the past been used as a standard, this usage is grounded in history, not in science (G–139 at 4; A–859 at 3–4) or law. Before the advent of computer technology, statisticians relied on statistical tables to determine statistical significance (G–139 at 4). These tables generally reported only three significance levels: .01, .05, and .1 (id.).* The use of $P<.05$ as a reference point evolved from the use of these tables. Indeed, Abbott's witnesses seem to recognize the lack of scientific basis for use of $P<.05$. One of these witnesses, Dr. Smuckler, stated that "it is true that (the use of the .05 confidence level) is an arbitrary decision, and, from a strictly mathematical standpoint, the selection of this limit could be criticized * * *" (A–859 at 4). Dr. Oser, another Abbott witness, could say only that the .05 confidence level is "commonly used" (A–858 at 24). Dr. Carlborg, a third Abbott witness who is a statistician, did not articulate any rationale for use of $P<.05$, but rather stated that "NCI regularly uses the .05 level" (A–857 at 9). Traditional usage of a scientific method is not necessarily, however, a valid reason for usage of that method in a particular case.

Moreover, although use of the $P<.05$ as a standard is grounded in tradition, it is no longer the method used by most statisticians. Most statisticians, with the use of computers, now can and do report to the precise P-value for an observed result and allow toxicologists and other scientists to make a judgment for themselves on whether or not the level of statistical significance obtained is sufficient for them to reach a conclusion that the effect seen is the real effect of the substance tested (G–139 at 4; see also G–140 at 13).

In deciding how to apply the concepts of statistical and biological significance in proceedings under Section 409 of the act, we do well to keep in mind the fact, adverted to earlier, that evidence not conclusive enough to confirm harm may yet be probative enough to harm to negate safety. Consider this example. Suppose the data tell us there is a 90 out of 100 chance that cancer is associated with ingestion of the test substance (that is $P=.1$). If the rule of decision is that we will not conclude that a substance causes cancer unless we think the chances are 95 out of 100 that it does (*i.e.*, $P=.05$) then the data do not "prove" the substance is a carcinogen. But to say we lack proof of cancer is scarcely to say we have proof of safety. It is that distinction which is mandated by the statute. We are commanded to seek proof of safety, not merely to accept as proof of safety anything falling minutely short of proof of harm.

Commissioner Kennedy put it another way in pointing out that one's choice of a P value may depend on the purpose to which it will be put. In some cases, the consequences of a false positive are very serious. Suppose, for example, that we are testing a new component for a rocket to be used in a moon shot, and that that component's survival is critical to success of the mission. In such a circumstance, we would want to be virtually 100% certain that the new component is more reliable than the component it is replacing. Thus, a P-value of .000001 might be desirable.

Where, however, it is a false negative that presents a problem, a test with a P-value higher than .05 may supply important information. In this proceeding, there is good reason to be seriously concerned about an incorrect finding of safety, for the consequence is the marketing of a carcinogen. Using this principle, there is a valid reason for FDS to consider effects tht are not significant at $P<.05$ even though scientists or regulators engaged in different

endeavors may not.[12] In so doing I emphasize that the difference between a confidence level of P=.05 and P=.06 is merely a matter of the degree of certainty. In the former case, one is 95 percent certain that the observed result is not due to chance. In the latter case, one is 94 percent certain. There is no valid scientific rationale for concluding that there is a substantial difference between these two confidence levels. In the latter case, one is a little less certain about whether the carcinogenic effect is associated with treatment. I cannot, however, ignore such an effect. It may not be conclusive, but it is at least suggestive of a carcinogenic effect and therefore supports the conclusion that the tested substance has not been shown to be safe. Such suggestive results are especially important where they recur in a number of studies, for as a scientific matter, several inconclusive but suggestive studies containing similar results increase the likelihood that the effect observed is real (G–139 at 5; G–140 at 13). Adopting Abbott's suggested use of P≤.05 for all studies would preclude consideration of such inconclusive but suggestive results and therefore would be both scientifically and legally inappropriate. *Ethyl Corp.* v. *EPA*, 541 F.2d 1, 28 n. 58 (D.C. Cir.) (*en banc*) *cert. denied*, 426 U.S. 941 (1976); *Environmental Defense Fund* v. *EPA*, 598 F.2d 62, 89 (D.C. Cir. 1978); *Color Mfg. Ass'n* v. *Mathews, supra*, 543 F.2d at 297.

I also reject Abbott's argument that in evaluating other food additive petitions, the Bureau of Foods always uses P≤.05 as a standard. There is no evidence in this record to that effect. Indeed, even the internal Bureau working paper which Abbott cites as support for its position is to the contrary.[13] The memorandum does state that the incidence of a tumor should be significant at P≤.05 before a study will be found to be positive (Abbott's Remand Ex.; Exhibit 21 at 2). The memorandum further states, however, that "(i)f the data in a study indicate a trend of increased tumor incidence that is not statistically significant at P≤.05, doubts about the safety of the additive will be raised which will warrant further testing. This testing would in all probability require a chronic feeding study with a 'higher power of test' *e.g.* more animals per group, higher doses etc." (*id.* at 2–3). Thus, it is plain that the memorandum upon which Abbott relies, recognizes that effects which are not statistically significant at P≤.05 and therefore not conclusively positive, may nevertheless raise a doubt as to the possible carconogenicity of a food additive. It is therefore clear that the Bureau of Foods customarily considers effects that are not significant at P≤.05 where such effects raise uncertainty as to the safety of a food additive.

Moreover, even if the Bureau had in the past used P≤.05 as a standard, the Bureau's past practice is not controlling because the Bureau does not set the *agency's* standards for approval of food additive petitions. As the Court in *Abbott Laboratories* v. *Harris*, 79C 3732 (N.D. Ill., decided June 12, 1980) made clear, the function of the Bureau of Foods' staff is to serve as advisors to the Commissioner (Slip Opinion at 3). The Commissioner makes all final decisions and is in no way bound by the advice he receives from the Bureau of Foods.

Finally, it is important to note that, although I find that it is appropriate to rely on effects that are not significant at the P≤.05 level, I am not relying solely on such effects in denying approval of the food additive petition for cyclamate. The incidence of lung tumors in one strain of female mice in the Rudali study (discussed below) is significant at P=.003 and the incidence of total tumors in the same strain of female mice and second strain of mice in the Rudali study is also statistically significant at P≤.05. Moreover, the incidence of lymphosarcomas in three combined generations of mice in the Kroes study (discussed below) are statistically significant at P=.0036. Finally, the dose response relationship between cyclamate and the incidence of lymphosarcomas in the Brantom study (discussed below) is statistically significant at P=.008. These studies strongly suggest that cyclamate is a carcinogen and therefore are sufficient to raise a serious doubt concerning the carcinogenicity of cyclamate. Thus, even if I were to use P≤.05 as a standard, as Abbott has suggested, I would nevertheless find that Abbott has failed to show that cyclamate is safe.

5. *Position of the Parties, Findings of the ALJ and Commissioner's Findings On Biological Significance.* Abbott agrees that "evaluating effects for their biological significance, if any, is a valid scientific and regulatory exercise" (Abbott's Remand Brief at 15). Moreover, it is undisputed that to determine whether a tumor incidence is biologically significant, the consideration of biological factors, such as methodology of the study involved, chemical structure, length of use, dose response, rarity of tumors, and the presence of similar results in other studies is involved (Abbott's Remand Brief at 16; G–139 at 5). The ALJ found that "biological significance must be attached to study findings where borderline statistically significant effects occur (e.g. P=.06), but additional factors exist" (IRD at 13).

Abbott contends, however, that the concept of biological significance can be applied only to reject effects that are statistically significant at P≤.05 (Abbott's Remand Brief at 15–15), but cannot be applied to attribute significance to effects that are not statistically significant at P≤.05. I find Abbott's "one way" test to be untenable, for it would operate only to prove safety, not to disprove it. Scientifically, it is just as appropriate to rely on biological factors to conclude that an effect has biological significance, even though it is not statistically significant at P≤.05, as it is to rely on biological factors to reject effects that are significant at P≤.05 (G–139 at 4–6; G–140 at 13).

Consideration of biological factors can add further credence to or detract from the weight that would normally be given to findings with a particular P-value. For example, two different types of tumors may occur at the same P-value in a particular study. If only one of these tumor types recurs in other studies, the recurring tumor type will be considered to have greater biological significance than the tumor type that does not recur in other similar studies. (The latter tumor type may be found to be

[12] In my decision denying approval of a color additive petition for Red No. 2, I addressed an issue similar to that raised by Abbott here. I emphasized there, as I do here, the importance of using methods that are most likely to detect a carcinogenic effect because of the consequences of mistakenly concluding that a food additive is safe:

In reviewing the adequacy of the existing studies, I have, in accordance with the philosophy of the color additive law adopted a conservative approach in order to be sure that the public health will be adequately protected * * * I have used methods that are valid and are also the ones most likely to detect any carcinogenic effect that may be present, * * * When a study is used to evaluate the safety of a substance to be widely used by the public, the risk of a false negative—of incorrectly failing to detect an adverse effect that is present—is of greater concern than the risk of a false positive—of incorrectly reporting an adverse effect when none exists. * * * I am not, however, imposing an absolute standard of safety for evaluation of safety studies * * * I would not use a procedure, even if it were the most conservative, if the procedure were not a valid one. If the questions about a substance or the defects in a study are insubstantial, they do not preclude approval of the substance. However, when uncertainty remains about safety, after a fair evaluation of the record in accordance with scientific principles of evaluation, then, under applicable law, the importance of protecting the public health must guide the final decision. FD&C Red No. 2; Denial of Petition for Permanent Listing; Final Decision; Docket No. 76C–0033 (January 25, 1980, 45 FR 6253).

[13] Although the Bureau correctly notes that this memorandum is not in evidence and is not the official position of the Bureau, I have considered it because it helps to resolve this issue.

insignificant it it does not recur in any studies.)

Similarly, an effect may occur at P-value that, when viewed by itself, does not appear to be significant. However, consideration of biological factors may result in a conclusion that the effect has biological significance. For example, in a number of direct cyclamate feeding studies in rats (see Section IV.B.2. below) more bladder tumors occurred in cyclamate treated rats than occurred in controls. The occurrence of these tumors in each of the individual studies is not statistically significant at $P<.05$. However, because bladder tumors are historically rare in the strains of animals used in these studies, because the occurence of these tumors in cyclamate-treated animals is consistent with a small treatment effect, because the occurence of these tumors in control animals is consistent with the incidence of these tumors in historical controls, and because these bladder tumors have recurred in a number of studies involving different strains of rats, these bladder tumors are biologically significant.[14]

To summarize, the concepts of statistical significance and biological significance should be viewed together in determining the significance of a treatment related incidence of tumors. The closer the P-valve is to $P<.05$ the greater the confidence that can be placed in the results of the study. The factors to be considered in determining biological significance may increase or decrease that confidence. This evaluation results in a decision as to how much, if any, weight a study should be given (see G–139 at 3–6; G–140 at 13).

Moreover, each study is not only considered independently, but also is considered as part of the totality of the evidence. An individual study, standing alone, may not raise a serious question as to the safety of a substance. When that study is viewed with other similar studies, a trend of a particular effect may become apparent. Where several studies, viewed together, point in the direction of carcinogenicity, those studies, even though inconclusive, are a valid and objective basis for concluding that a food additive has not been shown to be safe. This is particularly true when the inability to demonstrate a statistically significant treatment effect in the individual studies is a result of the insensitivity of the studies.

Courts have consistently upheld decisions made by federal agencies where those decisions have been based on evidence that was inconclusive but suggestive. In *Ethyl Corp.* v. *EPA, supra,* the court stated that:

> * * * we need not seek a single dispositive study that fully supports the Administrators' determination. Science does not work that way; nor, for that matter, does adjudicatory fact-finding. Rather, the Administrator's decision may be fully supportable if it is based, as it is, on the inconclusive but suggestive results of numerous studies. By its nature, scientific evidence is cumulative: the more supporting, albeit inconclusive, evidence available, the more likely the accuracy of the conclusion.

541 F.2d at 37.[15]

The District of Columbia Circuit Court of Appeals recently reaffirmed the opinion in *Ethyl Corp.* and further recognized that a regulatory agency could not carry out its statutory mandate to protect the public from incompletely understood dangers such as cancer if the agency could not rely on suggestive results:

> * * * [R]egulations [prohibiting marketing of a suspected carcinogen] may jeopardize plants or whole industries, and the jobs depending on them. In such circumstances, the temptation to demand that the agency furnish conclusive proof of carcinogenicity as support for the regulations is great. However, the decision to delegate authority to an agency to control suspected carcinogens is a legislative judgment that is not open to question in this court. Congress's direction to EPA to protect against incompletely understood dangers could not be carried out if we were to adopt the proof requirements advocated by industry petitioners.

Environmental Defense Fund v. *EPA, supra,* 598 F.2d at 89. *Accord, Color Mfg. Ass'n* v. *Mathews, supra,* 543 F.2d at 297. *See Hercules* v. *EPA,* 598 F.2d 91, 110 (D.C. Cir. 1978).

B. Classification of Carcinogenicity Studies

Classifications for carcinogenicity studies are simply terms used to reflect the conclusions drawn from a study. Studies submitted in this proceeding can be classified as (1) positive, (2) inconclusive but suggestive of a positive effect, (3) negative, or (4) deficient. These classifications reflect whether a study supports the conclusion that the test substance causes cancer (positive), suggests that the test substance causes cancer (inconclusive but suggestive of a positive effect), supports the conclusion that the test substance is safe (negative), or is inadequate for drawing any conclusions as to the safety of the test substance (deficient). These classifications are discussed below.

1. *Positive.* A positive study is a study with contains results that establish that a test substance causes cancer. Such a study would result in a conclusion that the food additive is unsafe under the general safety clause, and, under the Delaney clause of section 409 of the act, would require that the food additive be banned. There does not seem to be much disagreement among the parties concerning the definition of a study which contains results which are positive. Abbott contends that to be positive a finding must be statistically significant at $P<.05$ and biologically significant as well (Abbott's Remand Ex. at 24; Abbott's Remand Brief at 15). The Bureau seems to agree with this assessment (Bureau's Remand Reply at 5; see Bureau's Position Paper at 2).

Although I agree that the level of statistical significance for determining that a study is conclusively positive should be at or near $P=.05$ and that the study should be biologically significant as well, I am not deciding in this proceeding whether the confidence level need be $P=.05$. Although the Rudali, Kroes and Brantom studies contain results that are statistically significant at well below $P=.05$ and suggest that cyclamate is a carcinogen, I find that, in light of questions raised about the biological significance of these studies, they are not conclusively positive (see Section IV B. below). In view of the fact that the precise P-value for determining that a study is conclusively positive is irrelevant to this proceeding, I will not resolve that issue here, but rather will resolve it when it is presented in the context of an administrative proceeding in which it is relevant.

2. *Inconclusive But Suggestive of a Positive Effect.* As discussed above in Section III.A.3., Abbott contends that all studies that are not positive should be considered as negative and cannot be relied upon to deny approval of a food additive petition (Abbott's Remand Brief at 12–18). The Bureau contends that an inconclusive study may raise serious questions as to the safety of cyclamate and thus support the conclusion that the additive has not been shown to be safe (Bureau's Remand Reply at 5–6).

As also discussed above, I find that a study which is inconclusive because of questions about its statistical or biological significance may nevertheless raise a serious doubt as to the safety of a food additive and be relied on by the

[14] It should be emphasized that the great majority of substances do not cause cancer when tested in the types of animal studies contained in this record. Attention in therefore properly paid to such studies whenever cancerous tumors are found.

[15] The Court in *Ethyl Corp.* was reviewing EPA's decision under the arbitrary and capricious standard of the Administrative Procedure Act, 5 U.S.C. 706(2)(A) (1976). Although the cyclamate decision is subject to review under the substantial evidence standard of 5 U.S.C. 706(2)(E) (1976), the proposition stated above is nevertheless applicable here because it does not relate to the applicable standard of review but rather to the application of scientific principles to administrative factfinding.

agency as a basis for denial of a food additive petition.

3. *Negative.* A negative study is a study that supports the conclusion of a reasonable certainty of no harm. As with positive studies, negative studies are attributed various weights depending on the statistical and biological significance of the study. It is important to bear in mind, however, that in view of the serious consequences of mistakenly finding that a negative study proves safety, a flawed negative study may be entitled to little or no weight whereas a positive study with a similar flaw may well be entitled to some weight.

One issue that reoccurs with respect to a number of studies that Abbott considers negative is the sensitivity of a study. Abbott recognizes that although a study may not detect any effect, it may be entitled to little or no weight if the size of the study is so small that the study is too insensitive to detect an effect even if one is in fact present (Abbott's Remand Brief at 18). This issue can best be understood by considering a scientist use of a microscope. A scientist may be unable to observe an object with a microscope because the microscope is not powerful enough to sufficiently magnify the object to make it visible. Similarly, a small study may be too insensitive to detect a carcinogenic effect, even though one is present. In evaluating carcinogenicity studies, statistical methodology is used to determine the likelihood that a real effect is present even though the study did not detect any effect. In the Remand Order, Commissioner Kennedy asked the parties to further explain their positions on this issue:

> Another issue that needs further development by the parties concerns criteria for determining proof of safety. This determination involves an assessment of the quality of a study which in turn involves two main considerations: the minimum difference that a study can detect between effects on control animals and effects on treated animals, and the frequency with which this difference can be detected. Abbott appears to argue that any study not significant at the ".05 confidence level is negative and should be considered as proof of safety regardless of the sensitivity of the test or the frequency with the which the study would detect a specified difference.

(44 FR 47622).

The terms referred to by Commissioner Kennedy are used to describe the sensitivity of a study. The term "minimum difference" refers to the minimum difference between treated and control animals that a study is capable of detecting at a specified confidence level and frequency. The "power" of a statistical test or false negative error rate is the probability (frequency) that the test will detect, at a specified confidence level, a specific minimum difference between treated and control animals, if the difference is present. For example, the Plank study had only a 50% chance of detecting, at the 95% confidence level, a true difference in tumor incidence of approximately 33% between the controls and the high dose treated animals. The "33%" figure in this example is the minimum detectable difference that this study is capable of detecting at the 95% confidence level. The power of this study is 50%. This statement tells us that even if a true difference in tumor incidence of 33% between cyclamate-treated and control animals existed in the Plank study, the study would have only a 50/50 chance of detecting that difference at the $P \leq .05$ confidence level.

The minimum detectable difference, the power of a study and what constitutes a statistically significant result are dependent on one another and on the number of animals in a study (G–120 at 4). Generally, the larger the number of animals in a study, the more sensitive the study will be, *i.e.*, the lower the minimum detectable difference the study can detect at a specified power and confidence level.[16]

Abbott contends that "if no statistically significant ($P \leq .05$) effects are observed in a study then it is negative; however, all negatives are not of equal value" (Abbott's Remand Brief at 18). Abbott does not, however, articulate what it considers to be the criteria for determining whether the sensitivity of a study is adequate. Abbott states only that "commonly accepted scientific standards for determining safety are well known and understood" (Abbott's Remand Brief at 18; A–858 at 25). Abbott also lists the Schmaehl, Kroes, Taylor, Gaunt and Carson studies as examples of negative studies providing proof that cyclamate is safe. Although I agree that the Gaunt and Carson studies are negative, I disagree with the remainder of that statement. My reasons are discussed below in Section IV.

[16] The Bureau also notes that the power of a test "depends on how exaggerated the highest dose studied is compared to the estimate of human consumption" (Bureau's Position Paper at 5 n. 1). This statement is incorrect. The statistical power of a study will remain constant even though the dose studied may vary. If the Bureau means to suggest that a study may have an adequate statistical power but nevertheless be inadequate because the highest dose studied is too low, I agree. However, the Bureau has not criticized any of the dose levels employed in the cyclamate carcinogenicity studied as being too low. Nor, for that matter, has Abbott criticized any of the dose levels studies as being too high.

In response to the specific question asked in the above-quoted languag of the Remand Order, the Bureau referred to a statistical review in the Temporary Committee Report (G–41 App. V at 19–20). That statistical review reports the minimum detectable difference between cyclamate-treated and control animal for each cyclamate carcinogenicity study reviewed by the Temporary Committee (Bureau's Position Paper at 5 n. 1). The Bureau also cites the Interagency Regulatory Liaison Report, which I have not considered because the admission of that report into evidence was properly denied by the ALJ (see Section VII. F. below).

I find that the power of a study and the minimum detectable difference a study can detect are important criteria for determining what, if any, weight should be attributed to a study that fails to detect a statistically significant effect. This method of analysis provides an objective means of comparing the relative sensitivity of the cyclamate carcinogenicity studies. In my analysis of the cyclamate carcinogenicity studies of questionable sensitivity (see Section IV.B.2.a. (2)–(4); IV.D.) I have therefore reported and considered the findings of the Temporary Committee concerning the minimum detectable difference each cyclamate carcinogenicity study is capable of detecting.

It should be noted that, in determining the minimum detectable difference between cyclamate-treated animals and control animals in the cyclamate carcinogenicity studies, the Temporary Committee (1) assumed that the power of each study was 50%, (2) assumed that statistical significance was $P \leq .05$, and (3) reported the resulting minimum detectable difference for each study. For example, the Temporary Committee reported that the Ikeda study (discssed below) had only a 50% change of detecting, at the 95% confidence level, a true difference in tumor incidence of approximately 13% between the controls and the high dose treated animals (G–41 App. V at 20). I do not consider the 50% power utilized by the Temporary Committee to be an especially high one. It means that 50% of the time, when the specified minimum detectable difference is actually present it will not be declared significant at the $P \leq .05$ level. Given the consequences of incorrectly declaring that a study is negative, I do not find a potential false negative error rate of 50% to be very reassuring. I find, however, that even assuming that a 50% power is adequate, the minimum detectable difference in the cyclamate carcinogenicity studies of questionable sensitivity is unacceptably high.

I recognize that it is impossible to prove a negative to an absolute certainty and I am not asking Abbott to do so. However, I disagree with Abbott as to what weight, if any, should be attributed to many of the studies that Abbott considers negative (see Sections IV.B.2.a.(1)–(5) and IV.D.). For the reasons discussed below, I find that many of the studies that Abbott contends are negative, do not provide a basis for any valid conclusions as to the safety of cyclamate because of the low sensitivity of those studies. Indeed, I have found that there are only two studies (Gaunt and Carson) submitted by Abbott that are properly classified as negative (see Section IV.C.). These two negative studies are not, however, entitled to sufficient weight to meet Abbott's burden of proving to a reasonable certainty that cyclamate does not cause cancer nor do they rebut the safety questions raised by other studies (see Section IV.C).

4. *Deficient.* A study may be deficient because of defects in the design or conduct of the study. The parties do not dispute that where a study contains a significant defect it should not be given any weight. The parties also agree that even where the conduct of a study is not defective, that study may be entitled to no weight because it is too insensitive to provide any useful information about the safety of the test substance (Abbott's Remand Brief at 18). Abbott contends that inadequately sensitive studies should be classified as negative, although entitled to little or no weight (*id.*). The Bureau contends that a study of inadequate sensitivity or an otherwise deficient study should be classified "inconclusive but uninformative" (Bureau's Position Paper at 7). I find that there is no substantive difference in these approaches, but only a question or nomenclature. I have decided to classify such studies as deficient.

IV. Carcinogenicity: The Evidence

With the principles discussed in Sections II and III in mind, I will now discuss the evidence submitted in this proceeding. One piece of evidence that was the subject of much dispute was the Report of the Temporary Committee For The Review of Data On Carcinogenicity of Cyclamate (G–41). Because this report is cited by both parties as part of their discussion of most of the carcinogenicity studies, I will discuss it first.

A. The Review of the Temporary Committee of the National Cancer Institute

On March 14, 1975, then Commissioner A.M. Schmidt requested that the National Cancer Institute ("NCI") establish an advisory committee of experts to review the carcinogenicity evidence concerning cyclamate and advise the agency as to whether or not cyclamate is a carcinogen (G–41, App 1). The National Cancer Institute thereafter established a Temporary Committee for the Review of Data on Carcinogenicity of Cyclamate ("Temporary Committee") to advise NCI concerning its scientific review on all available data on the carcinogenicity of cyclamate. The Temporary Committee consisted of a number of distinguished scientists, including oncologists, pathologists, medical doctors and doctors of veterinary medicine. In addition, four working groups were established to provide staff support and additional expertise to the Temporary Committee. These working groups included the NCI Epidemiology Working Group, the NCI Experimental Design and Toxicology Working Group, the NCI Pathology Working Group and the NCI Statistics Working Group.

In February, 1976, the Temporary Committe submitted its report to the Director of the National Cancer Institute. The Temporary Committee concluded that:

> 1. The present evidence does not establish the carcinogenicity of cyclamate or its principal metabolite, cyclohexylamine, in experimental animals.
> 2. No conclusions can be made regarding the question of cyclamate's potential carcinogenicity in humans due to the short post-exposure observation time, the insensitivity of epidemiologic studies to detect relatively small changes in cancer incidence, and other factors.
> 3. The Committee is concerned over the implications of the increased incidence of tumors in the urinary tract of cyclamate-fed animals from several studies, even though those increases were not statistically significant. It is not clear whether this represents a weak carcinogenic response or random variation.
> 4. An additional concern is the carcinogenic responses obtained in cyclamate-treated animals from studies employing unconventional procedures or in which the specificity of the response is questionable. The bladder implantation study done by Bryan et al. was considered to be inappropriate for assessing carcinogenicity of a human dietary constituent. Of particular concern is the Food and Drug Research Laboratories' study (Oser et al.) in which a statistically significant increase in bladder tumors occurred in animals treated with a mixture of cyclamate and saccharin. The cocarcinogenicity system used by Hicks et al. has yet to be validated as a bioassay for carcinogenicity. Although the dose-dependent increase in lymphosarcomas in cyclamate-treated mice (Brantom et al.) was statistically significant, there is the likelihood that this reflects a nonspecific response in the strain of mice employed.
> 5. Short-term or in vitro test systems cannot now be used to establish carcinogenicity. However, the results from such systems are useful for determining the need for appropriate carcinogen bioassay studies, as well as for enlarging the mutagenicity-carcinogenicity correlative data base. In this regard, the Committee notes that in several studies cyclamate or cyclohexylamine has been found to produce chromosomal damage in human and rodent cells.

(G–41 at 48).

The advice of the Temporary Committee is, of course, not controlling in this proceeding. The Temporary Committee's conclusions are, however, evidence in this proceeding and should be considered as such. Abbott contends that the Temporary Committee could not have made "a more definitive statement regarding cyclamate's safety" (Abbott's Remand Ex. at 9). I disagree. A much more definitive statement could have been written, namely that cyclamate has been shown to be safe. The Temporary Committe did not make such a finding. Indeed, as is apparent from paragraphs 3 and 4 above, the Temporary Committee expressed substantial uncertainty about the safety of cyclamate. In addition to the statements in paragraphs 3 and 4, the Temporary Committee stated that

> None of those studies (referring to the Bryan, Oser, Hicks and Friedman studies) satisfy the Committee's criteria for concluding that cyclamate is a carcinogen. They do, however, create a sense of uncertainty.

G–41 at 46. Moreover, the Experimental Design and Toxicology Working Group of the Temporary Committee found that "the studies thus far conducted have been inadequate to assess the carcinogenicity of cyclamate in animals" (G–41, App. V at 55).

The Temporary Committee also described a study designed to resolve the Committee's uncertainty about the safety of cyclamate. Abbott contends that requests for additional testing result in a "never-satisfied posture" in view of the Temporary Committee's statement that "[c]yclamate has pushed the technology of carcinogenicity testing to its limit" (G–41 at 47). I disagree. The statute places on Abbott the burden of proving that cyclamate is safe. Congress has thus decided that where the evidence is uncertain the petition must be denied, regardless of whether additional testing could resolve that uncertainty. The fact that the "uncertainty [about cyclamate's safety] does not appear to be easily resolvable by currently available bioassay technology" (G–41 at 46) does not lessen Abbott's burden.

In the case of cyclamate, it is certainly possible that further adequate testing,

such as the study proposed by the Temporary Committee, could resolve the current questions about cyclamate's possible carcinogenicity. If such testing is done, it may yet be possible for FDA to conclude that there is a reasonable certainty that cyclamate does not cause cancer.

B. Inconclusive but Suggestive Studies Raising a Serious Question as to the Possible Carcinogenicity of Cyclamate

1. *The Occurrence of Lung and Liver Tumors and Lymphosarcomas in Mice.* The Rudali, Brantom, Kroes and Hardy studies all involve the direct feeding of cyclamate to test animals and suggest that cyclamate is a carcinogen. In the Rudali study, one strain of cyclamate treated female mice was found to have a statistically significant incidence at the P≤.05 level of lung tumors and of total tumors combined. A different strain of cyclamate treated male mice in the Rudali study was found to have an increased incidence (P=.07) of liver tumors and a statistically significant incidence at P≤.05 of total tumors combined. The Brantom, Kroes and Hardy studies all resulted in increased levels of lymphosarcomas (a malignant tumor) in treated animals. In the Brantom study, there was a statistically significant dose response relationship (P=.008) between cyclamate and lymphosarcomas for female mice and the total incidence of reticuloendothelial sarcomas (P=.06) was biologically significant for female mice. In the Kroes study, the incidence of lymphosarcomas for three generations of male mice combined was statistically significant (P=.0036). Finally, in the Hardy study, although the incidence of lymphosarcomas was not statistically significant at the P≤.05 level, it is important because the study used the same mouse strain as the Brantom study. Thus, the increased lymphosarcoma levels in the Hardy study enhance the credibility of the results of the Brantom study. Each of these studies is discussed in detail below.

a. *Rudali, et al. (G–43).* (1) *Study Design:*[17] The ALJ described the Rudali study as follows:

> Sodium cyclamate was placed in the drinking water of several strains of mice at a concentration of 6 gm/liter. A breakdown of the test animals is as follows: 30 male mice of the RIII strain and an equal number of male controls; 20 mice and 20 female mice of the C3H strain and an equal number of controls; 30 female mice of the XVII/G strain and an equal number of female controls; and 40 male laboratory-bred mice of the F1 (C3H x RIII) strain and an equal number of controls. The study was conducted for the lifetime of the animals. The animals were examined grossly but special attention was not given to the bladders nor were bladders examined histopathologically.

(ID at 10).

(2) *Study Results:* The authors of the Rudali study concluded that cyclamate is a weak carcinogen (A–412 at 3). The ALJ found that in the first Rudali study, an increased incidence and shortened latency was seen for lung tumors in XVII/G reated female mice (ID at 10). The incidence of these lung tumors was statistically significant at P=.0003. In the F1 (C3H x RIII) treated male mice, an increased incidence of hepatomas was seen (*id.*). The incidence of these liver tumors was significant at P=.07. In addition, the incidence of total tumors combined in XVII/G treated female mice and F1 (C3H x RIII) male mice were statistically significant at P≤.05. Most of the liver and lung tumors found were multiple (A–412 at 2–3).

In the Initial Decision, the ALJ noted that, "[e]ven though an incidence of tumors was seen, the study is deficient in that not all the animals and all their organs were subjected to histopathologic examination" (ID at 11).[18] The Temporary Committee reached the same conclusion (G–41 at 22).

Following the reopened hearing, the ALJ found that "[t]he lung tumor incidence in the Rudali study tends to indict cyclamate as a carcinogen" but that due to lack of histopathology, "possible microscopic tumors present in the control animals could have been missed" and that therefore "the biological significance of the Rudali lung data is compromised" (IRD at 20).

The ALJ further found that:

> In order to accept the overall statistical significance of the lung and liver tumor levels, data from different biological systems and different mouse strains must be combined. The controversy over the propriety of such combinations would not allow labeling the overall data biologically significant. Thus, the borderline significance level for liver tumors is the only biologically significant effect.

(*id.*)

(3) *Analysis:* The Bureau takes exception to the ALJ's finding that the lack of histopathology compromised the Rudali study (bureau's Remand Ex. at 2–3). The Bureau contends that large lung lesions visible (without histopathology) in treated animals but not in controls are at the least an indication of a more rapid onset of the effect seen and are evidence of a greater chance for metastasis (spread of cancer) (*id.*). The Bureau argues that even if histophathology revealed some tumors in the control group, those tumors would have been smaller and later in developing. These factors, the Bureau concludes, make the lung tumor findings in the Rudali study toxicologically significant, even though no histopathology was performed (*id.* at 3). Abbott, relying on the testimony of Dr. Smuckler, contends that due to the lack of histopathology the Rudali study contributes nothing to the assessment of the potential carcinogenicity of cyclamate (Abbott's Remand Reply at 4–5). Abbott further contends that the possibility of metastasis is purely speculative (*id.*).

The Bureau also takes exception to the ALJ's finding that lung and liver tumors in the Rudali study cannot be combined. (Dr. Frankos combined total tumors, which included lung an liver tumors, in his analysis of the Rudali study and found them to be significant (R. Tr. at 193–96).) The Bureau contends that combining data on lung and liver tumors is permissible and that the resulting data are biologically significant (Bureau's Remand Ex. at 4). In response, Abbott cites testimony of Dr. Carlborg who states that the National Cancer Institute has not adopted the practice of combining tumors from different biological systems and that he has confirmed this fact with a Ken Chu of NCI (Abbott's Remand Reply at 5–6).

Abbott takes exception to the ALJ's finding of a borderline significant effect for liver tumors (Abbott's Remand Ex. at 19) Abbott contends that (1) finding is limited to one sex and one strain; (2) the effect is not significant at the P≤.05 level and therefore a higher standard is being applied to cyclamate than any other food additive; and (3) lack of histopathology compromises this finding (*id.* at 19–20).

I find that the statistically significant (at the P=.003 level) incidence of lung tumors in the XVII/G female mice in the Rudali study is a key finding suggesting a possible carcinogenic effect of cyclamate. This finding is sufficient by itself to raise a serious question about the carcinogenicity of cyclamate. The finding of increased incidence of liver tumors (significant at the P=.07 level) in cyclamate treated F1 (C3H x RIII) male mice and the statistically significant at P≤.05 increase in total tumors combined

[17] This and other descriptions of Study Designs are taken essentially without change from the ALJ's Initial Decision. They are included here to assist the reader in understanding the analysis of the study results.

[18] Histopathlologic examination refers to the process by which tissues are dried, sectioned, stained, placed on slides, and examined under a microscope.

in cyclamate treated F1 (C3H x RIII) male mice and XVII/G female mice are also important because they reinforce the concerns about the carcinogenicity of cyclamate arising from the key finding of lung tumors in the XVII/G female mice (G–140 at 10–11). This total tumor finding is appropriately relied on as part of the overall basis for concluding that a serious question has been raised as to the possible carcinogenicity of cyclamate.

I agree with Abbott that the Bureau's argument concerning the possibility of metastasis of the tumors found in the Rudali study is speculative. I do not, however, find persuasive Abbott's argument that the lack of histopathology invalidates the findings of liver and lung tumors in this study. Although histopathology may have revealed tumors in control animals, it is equally possible that it would also have revealed more tumors in treated animals. Moreover, lung and liver tumors that were found macroscopically were examined microscopically (A–412 at 2–3). The Site Visit Committee stated that histologic confirmation of all tumors is essential (G–41 App. III, Foundation Curie at 4). Although the Site Visit Committee stated that the quality of slides available was generally poor, the Site Visit Committee did confirm a number of the lung and liver tumors found in the Rudali study from a sample of the slides (*id.*). This microscopic confirmation of the tumor findings in the Rudali study supports the validity of the macroscopic examinations of lung and liver tumors found in the Rudali study.

Finally, I agree with the Bureau's contention that the large lung and liver lesions, visible without histopathology, found in cyclamate treated animals but not in controls are an indication of a more rapid time of onset of the tumors found in cyclamate treated animals (R. Tr. at 188; see also G–140 at 10; R. Tr. at 190; A–412 at 3). This factor, by itself, is supportive of a finding of carcinogenicity. Thus, even if histopathologic examination of the lungs and livers of the mice in the Rudali study revealed an equal incidence of lung and liver tumors in treated and control mice, the more rapid time of onset of the tumors in the cyclamate treated mice would still raise a serious question as to the carcinogenicity of cyclamate.

Abbott also relies on the testimony of Dr. Smuckler, who in addition to questioning the lack of histopathology, stated that "[s]ince mice are notorious for the appearance of spontaneous disease, the absence of lymphoma and the absence of critical analysis of the type of pulmonary tumor found need clarification" (A–859 at 9–10). Dr. Smuckler, however, does not even suggest why the absence of lymphomas, even if unusual, would negate the observed significant difference in the evidence of lung and liver tumors between the treated and control groups. As to the second part of Dr. Smuckler's statement, the type of pulmonary tumor present is irrelevant so long as that tumor is malignant. I find that the macroscopic examination of tumors confirmed in part by histopathologic examinations fully supports the conclusion that the lung and liver tumors found by Rudali were malignant. Accordingly, I reject Dr. Smuckler's criticism of this study.

Abbott also attacks the credibility and reliability of Dr. Frankos, a Bureau witness. Although I agree with Abbott that Dr. Frankos' opinion regarding the possible occurrence of metastasis in the Rudali study was speculative, I emphatically reject Abbott's contentions that "Dr. Frankos' testimony is brought into question in virtually every answer during his cross-examination"; "that Judge Davidson accorded little weight" to Dr. Frankos' testimony and that Dr. Frankos is "inexperienced" (Abbott's Remand Reply at 4). A careful review of the testimony of Dr. Frankos and his curriculum vitae reveals that Dr. Frankos has substantial experience in the evaluation of carcinogenicity studies and that the cross-examination of Dr. Frankos, if anything, enhanced his credibility.[19]

(*Id.* at 96–97).

Dr. Frankos received a Ph.D. in 1977 from the University of Maryland School of Pharmacy, Department of Pharmacology and Toxicology, where he had experience in the area of experimental toxicology of drugs (R. Tr. at 79–81). This experience is relevant to the evaluation of the safety of other chemicals (*id.* at 81). From 1977–79, Dr. Frankos worked as a toxicologist in FDA's Division of Toxicology, Bureau of Foods. In that capacity, Dr. Frankos reviewed a total of approximately 100 toxicity studies (including carcinogenicity studies) submitted in support of compounds for which industry firms sought FDA approval (G–140 at 2; R. Tr. at 99–100). Dr. Frankos has also participated in the design of toxicology studies (R. Tr. at 81–86; 96–98). Dr. Frankos demonstrated a detailed knowledge of the type of studies that the Bureau of Foods receives in support of food additive petitions (R. Tr. at 88–93; 100–101). Abbott surely cannot be suggesting that experience gained by a scientist serving in a federal regulatory agency is of no value.

Finally, the ALJ did not make any finding that Dr. Frankos lacked credibility and did make a number of findings that were supported by Dr. Frankos' testimony: *e.g.*, the ALJ found that the borderline significant level for liver tumors in the Rudali study is biologically significant (IRD at 20) and that the findings of lymphosarcomas in the Brantom study are biologically significant (IRD at 22). Accordingly, I reject Abbott's criticism of Dr. Frankos and find that his testimony is entitled to substantial weight.

Abbott further contends that since the incidence of liver tumors found in the F1 (C3H x RIII) male mice in the Rudali study is not significant at the $P \leq .05$ level, a higher standard is being applied to cyclamate than is applied to other food additives. I do not find this argument convincing. When the incidence of liver tumors in the treated mice is compared to the incidence in controls, the P-value is .07. Thus, there is a 935 probability that the increased incidence in liver tumors in treated animals is a result of cyclamate treatment rather than a result of chance. I would have more confidence that these results were not a random occurrence if they were significant at the $P \leq .05$ level, a higher standard is being applied to cyclamate than is applied to other food additives. I do not find this argument convincing. When the incidence of liver

[19] For example, Dr. Frankos testified that:

At the Bureau of Foods I spent a number of years helping to design the protocols for studies that will be considered adequate for submission in the cyclic review that is going to be initiated in the Bureau of Foods. This was one of my prime jobs there, writing quality assessment factors for the protocols; also writing up, designing the protocols that we are going to require the petitioners to submit to us under cyclic review.

* * * And people would come to me and ask me how would you design this experiment. And I would custom design things * * * (R. Tr. at 82–83); and

Q. But why does it require innovative thinking if every study consists of 50 rats of each species at each of four levels?

A. Well, it is not that simple. When you design a study you have to look at—well, how much of this am I going to have to feed in the study to establish a level that is going to be usable in the human population? * * *

When you evaluate that data you could get toxicological effects that weren't due to the compound because you designed the study improperly. You have to design a study that takes into consideration the nutritional requirements of that animal. You have to consider the palatability. You have to consider the findings from subchronic studies or actue studies because those findings will indicate to you, hey, there is an effect in the liver. I had better look very specifically at the liver in this study. Then you have to incorporate the proper enzymatic assays that might be needed, the proper histopathologic studies that will be needed to zero in on that organ. So those are the more innovative types of studies that I am talking about.

tumors in the treated mice is compared to the incidence in controls, the P-value is .07. Thus, there is a 93% probability that the increased incidence in liver tumors in treated animals is a result of cyclamate treatment rather than a result of chance. I would have more confidence that these results were not a random occurrence if they were significant at the P≤.05 level. I do not, however, consider these results to be insignificant. I agree with the Bureau and the ALJ that these results are important (G–140 at 10–11; IRD at 20) and are at least supportive of the conclusion that cyclamate has not been shown to be safe. Moreover, my consideration of carcinogenic effects, such as the liver tumors found in the Rudali study, which are not significant at P≤.05 level, does not impose a higher standard on cyclamate than the agency has imposed on other food additives. As the discussion in Section III establishes, effects may have biological significance even though they are not statistically significant at the P≤.05 level. Although such effects are not entitled to as much weight as effects which are significant at the P≤.05 level, they are nevertheless entitled to some weight especially when considered together with other statistically significant results.

Abbott also argues that the liver tumors found in the F1 (C3H x RIII) male mice are not significant because they were found in only one strain and one sex. Presumably, Abbott would make the same contention with respect to the lung tumors found in females of the XVII/G strain of mice in the Rudali study. I do not find these contentions convincing. The significance of a tumor finding in one strain and sex of a species is not reduced where that effect does not occur in other strains or sexes of the same species. In order to negate tumor findings in a particular strain and sex of a species, it is necessary to conduct further studies in the same strain and sex of the species in which the tumor finding was made. Such testing is necessary because it is not unusual to find more of an effect in a particular sex or a particular strain (R. Tr. at 107–08; G–140 at 11–12). The fact that other strains or sexes of mice tested by Rudali did not exhibit the same lung and liver tumor effect does not lessen the significance of the liver tumors found in F1 (C3H x RIII) strain of male mice and lung tumors found in the XVIII/G strain of female mice. Even within the same species, strains or sexes can vary in sensitivity (G–140 at 11; R. Tr. at 107–08). Thus, to negate the lung and liver tumor findings in the Rudali study, further testing must be done in the F1 (C3H x RIII) strain of male mice and the XVII/G strain of female mice.

It is appropriate to use the most sensitive strain of a species for detecting a toxic effect (G–140 at 11–12), because the induction of cancer in any strain or species is a good indication that the chemical will probably cause cancer of some type in humans (*id.*). Even though a tumor finding may be limited to a specific species, strain, sex and organ, that finding cannot be dismissed as being irrelevant to humans (*id.*). Absent data indicating what species or strain is most like man insofar as similarity of carcinogenic response to cyclamate is concerned, I have to assume in the interest of public safety that the response in the most sensitive species, strain and sex is most like that of man (*id.* at 12).

Moreover, it is not entirely true that the liver tumors in F1 (C3H x RIII) male mice and lung tumors in XVII/G female mice were found in only one sex and only one strain. The combined incidence of total tumors, which consisted primarily of lung and liver tumors, in F1 (C3H x RIII) male mice and XVII/G female mice, were statistically significant at the P≤.05 level. Thus, there is evidence that Rudali found an increased incidence of liver tumors in two different strains and sexes of mice and an increased incidence of lung tumors in two different sexes and strains of mice. Finally, it should be noted that Rudali did not test F1 (C3H x RIII) female mice or XVII/G male mice. Thus, it is possible that, if tested, the male XVII/G mice and the female F1 (C3H x RIII) mice would have exhibited the same response as their counterparts.

I further disagree with Abbott's contention and the ALJ's conclusion that it is inappropriate to combine total tumors (which consisted primarily of lung and liver tumors) found in the same strain of mice in the Rudali study, for the purpose of obtaining additional information about the potential carcinogenicity of cyclamate. Combining tumors from different organ sites is appropriate in order to evaluate cyclamate's overall carcinogenic potential (R. Tr. at 193–96; see G–118 at 18–19). This approach is particularly valid where, as here, a statistically significant tumor increase is seen in one organ (lung) in one strain of mice (XVII/G) and borderline significant tumor increase is seen in the same organ and a second organ (liver) in a second strain of mice (F1 (C3H x RIII)) (R. Tr. at 193–96). The finding of statistical significance for total tumors in the strain of mice with two borderline effects increases the confidence to be placed on the biological significance of those two borderline effects. The finding of an increased incidence of a specific type of malignant tumor in a specific location (such as the lung tumors found in XVII/G female mice) is more definitive than findings of generalized increased malignancies (such as the combined total tumors in the Rudali study), but the generalized finding is still entitled to some weight.

One of Abbott's witnesses, Dr. Carlborg, a statistician, contends that the National Cancer Institute ("NCI") has rejected the practice of combining tumors from different biological systems in its bioassay program (A–857 at 6). The only support Dr. Carlborg provided for this statement was an experience he had in which he combined rumors from different biological systems in a study of toxaphene (R. Tr. at 48). Dr. Carlborg testified that his analysis of the combined tumors resulted in a finding of no effect, *i.e.*, "the tumor rates in the control and all the treated groups were exactly the same" (*id.*). Dr. Carlborg stated that his practice in the case of toxaphene was rejected by NCI (*id.*). The example provided by Dr. Carlborg is, however, distinguishable from the procedure employed with the lung and liver tumor data in the Rudali study. As Dr. Frankos testified, it is invalid to combine all tumors to obliterate an effect (R. Tr. at 194–95). Thus, it is not surprising that NCI rejected Dr. Carlborg's combination of tumors where it resulted in a finding of no effect.

Even if NCI does not accept the practice of combining tumors from different organ sites where a statistically significant (at the P≤.05 level) effect is found, I find that the method used to analyze the data from the Rudali study is valid. I recognize that this method does not provide conclusive evidence of cyclamate's carcinogenicity. However, it does contribute to the assessment of cyclamate's carcinogenicity and raises a serious question as to the possible carcinogenicity of cyclamate.

In sum, I find that the Rudali study suggests, but does not prove, that cyclamate is a carcinogen.

b. *Brantom, et al. (G–3).* (1) *Study Design:* This study involved groups of 30 male and 30 female mice fed .7, 1.75, 3.5 or 7.0% sodium cyclamate. A control group of 60 mice of each sex was maintained. The study was continued for 80 weeks, after which survivors were sacrificed.

(2) *Study Results:* In the Initial Decision, the ALJ found that "a statistically significant increase of lymphosarcomas was found in the Brantom study" (ID at 31). Following the

reopened hearing, the ALJ made the following finding with respect to the Brantom study:

* * * the Bureauy found a biologically significant effect for cyclamate in the total incidences of lymphosarcomas and reticulum cell sarcomas in the female treated groups when compared to the controls (Ex. No. G–140 at 7). Abbott challenges this data because the Bonferroni multiplier was not applied. Even if this multiplier is used, however, two figures remain of borderline statistical significance [Linear trend for lymphosarcomas and the incidence of lymphosarcomas and reticulum cell sarcomas combined]. One of these two effects is also challenged for failing to properly use the Armitage test (Ex. No. A–857 at 13). But even assuming the validity of this challenge, a borderline statistically significant effect of the remaining figure, for the total reticuloendothelial sarcoma rates, exists. When considered in conjunction with the dose related increase in lymphosarcomas for female treated animals, this trend renders the Brantom data biologically significant.

(IRD at 21–22).

The Temporary Committee made the following finding with respect to the Brantom study:

* * * the Committee agrees that the test material did not induce a carcinogenic response in the urinary bladders of the treated animals. Although the increased incidence of lymphosarcomas in the cyclamate-fed female mice requires close evaluation, the nonspecific nature of this response makes its significance questionable with respect to establishing carcinogenicity.

(G–41 at 16).

(3) *Analysis:* In its exceptions, Abbott contends that the two findings which the ALJ found to be at "borderline statistical significance" (if the statistical corrections insisted on by Abbott are applied) are negative based on established Bureau criteria (Abbott's Remand Ex. at 26).

The borderline findings to which Abbott refers are the increased incidence of combined lymphosarcomas and reticulum cell sarcomas (P=.06) and the linear trend for lymphosarcomas (P=.076) (linear trend is a statistical test used to test for presence of a dose response relationship). Abbott further contends that the lymphosarcoma and reticulum cell sarcoma finding in the Brantom study "was a chance occurrence such as is bound to arise in such a vast amount of data" (*id.* at 27).

The Bureau contends that the key finding in the Brantom data is the dose response relationship between cyclamate and lymphosarcomas for female mice which was statistically significant at the P=.008 level (Bureau's Remand Reply at 5; G–139 at 6). The Bureau also argues that the linear trend test for lymphosarcomas, which was significant at the P=.076 level, and the incidence of lymphosarcomas and reticulum cell sarcomas, for female mice, which was significant at the P=.06 level, are biologically significant (Bureau's Remand Reply at 5).

I find that the incidence of lymphosarcomas and the incidence of lymphosarcomas and reticulum-cell sarcomas combined are key findings that suggest that cyclamate is a carcinogen. There is a statistically significant (P=.008) dose response relationship between cyclamate and the incidence of lymphosarcomas in female mice in the Brantom study (G–139 at 6). Moreover, even accepting Abbott's statistical analysis of the data, the incidence of lymphosarcomas and reticulum cell sarcomas combined is significant at the P=.06 level. I agree with the ALJ that the dose response relationship in female mice, when viewed with the borderline statistically significant incidence for all reticulonendothelial sarcomas, renders the Brantom data biologically significant.

I reject Dr. Carlborg's statement that "when the multiplier of 4 is applied to [the P-value for lymphosarcomas and reticulum-cell sarcomas combined], the P-value is .060 (4×.015), and any significance vanishes." (A–857 at 13). Even assuming that the use of this Bonferroni multiplier is valid, there is no basis in science for the proposition that the potential carcinogenic effect "vanishes" simply because the P-value is greater than .05. There is no qualitative difference between a P-value of .05 and .06. The difference is merely quantitative. To suggest that the relavitely small quantitative difference between a P-value of .05 and .06 renders the resulting data meaningless is to ignore the scientific realities of the situation.[20]

I find that the strong dose-response relationship between cyclamate and the incidence of lymphosarcomas (P=.008) and the linear trend for lymphosarcomas and reticulum cell sarcomas combined (P=.045) support the conclusion that the incidence of lymphosarcomas and reticulum cell sarcomas combined are biologically significant. In addition, the findings of lymphosarcomas in the Kroes and Hardy studies also support the conclusion that the lymphosarcomas and reticulum cell sarcomas in the Brantom study are biologically significant (G–139 at 9–10; see G–140 at 7–8). The occurrence of the same finding in more than one study is a factor that should be considered in determining the biological significance of a borderline significant effect (G–140 at 13). I therefore conclude that the incidence of lymphosarcomas and reticulum-cell sarcomas combined found in the Brantom study are biologically significant.

It is important to note that the P-value of .076 cited by the ALJ for the linear trend for lymphosarcomas is erroneous. This figure was arrived at by applying the Bonferroni correction to the P-value for the linear trend for lymphosarcomas. However, as Abbott's witness, Dr. Carlborg conceded, the Bonferroni correction is applied only to individual comparisons and not to trend tests and dose responses (R. Tr. at 33). Thus, the Bonferroni multiplier of four was improperly applied to the linear trend for lymphosarcomas and the correct P-value is .019. Although Dr. Carlborg criticizes this result because it was achieved by use of the Armitage test, which he claims is inappropriate for the lymphosarcoma finding, Dr. Carlborg does *not* state that the result would be any different if the method he claims is correct were used. Moreover, although Dr. Carlborg identified all linear trend tests which he thought were inappropriate (A–857 at 13), he did not state that the Armitage test was inappropriate for analyzing the linear trend for lymphosarcomas and reticulum-cell sarcomas combined (*id.*). That trend test was statistically significant at P=.045. Finally, Dr. Gaylor found that the dose response relationship between cyclamate and lymphosarcomas in the Brantom study was significant at P=.008 (G–139 at 6) and his statistical methodology was not challenged.

Abbott also contends here, as it does with respect to the Kroes study (in which a statistically significant incidence of lymphosarcomas was found), that the incidence of lymphosarcomas and the dose response relationship are artifacts, *i.e.*, chance occurrences. Abbott contends that this result is due to the "infinite number of comparisons [that] can be made" (Abbott's Remand Ex. at 26). Abbott also relies on the fact that the chance of an arithmetic decrease in lymphosarcomas in male mice in the Brantom study is 1 in 120 (exactly the opposite of the increase found in female mice) and a statistical analysis of liver tumors in the Brantom study indicates that cyclamate is a carcinogen in females and an "anticarcinogen" in males (*id.*). Abbott claims that there is no known scientific rationale to support

[20] I have assumed for the sake of argument, without deciding upon its intrinsic merits, that the Bonferroni correction should be used in analyzing data such as that in the Brantom study.

the validity of these inconsistent conclusions (*id.*).

I do not find Abbott's arguments persuasive. First, there is a scientific explanation for what Abbott has characterized as an "anticarcinogenic" effect. An apparent decrease in tumors with increase in dose may be a result of competing risks of deaths from other diseases which obscure the presence of cancer at high doses (R. Tr. at 53). Thus, what Abbott claims is cyclamate's "anitcarcinogenic" effect on lymphosarcomas and liver tumors may not be an artifact but may be due to mortality from other causes (see, *e.g.* G–41, App. VII, British Industrial Biological Research Association at 1).

Even if there were an "anticarcinogenic" effect in male mice in the Brantom study, it would not negate the biological significance of the lymphosarcoma and reticulum-cell sarcoma findings in the female mice (R. Tr. at 182). This is particularly true in view of the occurrence of lymphosarcomas in the Kroes and Hardy studies (discussed below). The occurrence of lymphosarcomas in the Hardy and Kroes study adds credence to the lymphosarcoma finding in the Brantom study (G–139 at 6–7; G–140 at 9–10) and tends to negate Abbott's argument that the Brantom findings are artifacts.

It is hard to understand how Abbott can argue that the reticuloendothelial sarcoma findings in the Brantom study are artifacts resulting from the infinite number of possible comparisons in view of Abbott's application of the Bonferroni correction. The purpose of the Bonferroni correction is to adjust for the increased false positive error rate that can result from multiple comparisons. As the above discussion establishes, however, even applying the Bonferroni correction where Dr. Carlbourg contends it should be applied, the effect on the reticuloendothelial system is significant at P=.06 and the dose response relationship is significant at P=.008. Abbott cannot have it both ways. If Abbott wants to correct for multiple comparisons, it cannot complain that the resulting figures are nevertheless invalid because of the multiple comparisons that have been employed.

The fact that one or more artifacts is likely to occur in a study such as the Brantom study does not prove that a particular effect, such as the lymphosarcomas, is an artifact. I cannot disregard a potential carcinogenic effect based on such a speculative argument. In order to rebut such a finding, it is necessary to adequately study the same sex/strain/species under the same experimental conditions and obtain valid negative results (R. Tr. at 186–87). Absent such evidence, mere speculation is insufficient to support a conclusion that the findings of lymphosarcomas and reticulum cell-sarcomas combined in the Brantom study are artifacts.

I recognize that my conclusion with respect to the Brantom study is contrary to the finding of the authors of the study and the Temporary Committee. The authors of the study concluded that "the incidence of lymphoma was not affected by the feeding of cyclamate" (G–3 at 744). The Temporary Committee found that the significance of the lymphosarcomas was questionable of the nonspecific nature of the response (G–41 at 16).

The conclusions of the authors of a study that the test results are negative is not dispositive (R. Tr. at 157). That conclusion can be rebutted by other evidence, for example, a statistical analysis showing some positive results that need further investigation, or evidence of a defect in the execution of the study. In the case of the Brantom study, two statistical analyses (nonparametric dose-response and linear trend) show a statistically significant effect and an analysis of lymphosarcomas shows biologically significant effect. This evidence rebuts the conclusion of the authors of the study and the Temporary Committee and, as discussed above, has not been adequately refuted by Abbott. As to the Temporary Committee's finding that lymphosarcomas were not site specific, I agree with Dr. Samuel Epstein, a Bureau witness, who stated that "* * * the comments of the [Temporary Committee] Report that lymphosarcomas are inconsequential because they are 'nonspecific tumors' appears incomprehensible. A lymphosarcoma is a malignant tumor * * *" (G–121 at 6; see G–118 at 19).

c. *Kroes, et al. (G–76; A–734).* (1) *Study Design:* The ALJ described the Kroes study as follows:

> This study employed SPF-derived swiss mice in groups of 50 animals of each sex. The groups were fed 2 or 5% sodium cyclamate, 2 or 5% cyclamate-saccharin in a 10:1 mixture, or 0.2 or 0.5% saccharin or 0.5% CHA. A control group of equal size was also maintained.

(ID at 10.)

(2) *Study Results:* In the Initial Decision, the ALJ found that "[b]oth parties agree that the study is negative, but the Bureau contends that its sensitivity is severely reduced because of the large number of animals lost to autolysis" (ID at 10). (Autolysis is a decay of tissue that begins shortly after death, thus preventing meaningful histopathological examination.)

Following the reopened hearing, the ALJ found that "[a] statistically significant effect for lymphosarcomas exists in the Kroes study if all three treated male generations are compared with the sum of their control counterparts" (IRD at 22). The ALJ further found that "[o]nly if the worst case against cyclamate is assumed, however, does the data withstand Abbott's criticism [that combining the three generations is inappropriate]" (*id.*).

The Temporary Committee found the study "* * * to have been well designed and conducted, although its significance was reduced somewhat as a result of a substantial number of mice lost from autolysis . . . [N]one of the test materials displayed carcinogenicity." (G–41 at 26.)

(3) *Analysis:* In its exceptions to the Initial Decision, Abbott contended that the significance of the Kroes study was reduced by autolysis, but that the study is not insignificant as a negative study (Abbott's Exceptions at 29–30). With respect to the ALJ's findings after the reopened hearing, Abbott concedes that the lymph system sarcomas in the three combined generations of the male mice in the Kroes study are statistically significant at the $P \leq .05$ level (Abbott's Remand Ex. at 27). Abbott contends, however, that (1) it is inappropriate to combine these generations because this method has not been emloyed elsewhere; (2) the effect is sex specific for males, but a sex specific effect is not confirmed by other studies; (3) the high spontaneous incidence of lymphosarcomas easily explains this finding; and (4) the effect is an artifact because the treated males in another study, the Brantom study, experienced fewer tumors than their controls (Abbott's Remand Ex. at 27–29).

The Bureau's reply is that (1) Dr. Frankos' testmony on the appropriateness of combining generations is uncontradicted; (2) the alleged high spontaneous incidence of lymphosarcomas in other studies is irrelevant because there is no testimony that the control incidence of lymphosarcomas in the Kroes study is unusually low; and (3) the finding of lymphosarcomas in the Brantom and Hardy study negate the possibility that the Kroes finding is an artifact (Bureau's Remand Reply at 6–7). The Bureau also contends that autolysis limited substantially the detectability of effects in the Kroes study, thus limiting the sensitivity of the study (Bureau's Brief at 18; G–121 at 9; G–126 at 12; G–113 at 7; G–112 at 15).

The Bureau also takes exception to the ALJ's criticism of combining generations in the Kroes study. The Bureau contends that the uncontradicted testimony establishes that combining the data from generations is appropriate (Bureau Remand Ex. at 2; R. Tr. at 159–60, 164–65). The testimony cited by the Bureau is that of Dr. Frankos who testified that he approves of the combination of generations because it increases the sensitivity of the study and is very analogous to the human situation of many generations being exposed to a compound (R. Tr. at 164–65).

I find that the data generated from the three generations of mice fed cyclamate in the Kroes study were properly combined and analyzed and that the statistically significant (P = .0036) lymphosarcoma finding is a key finding that suggests that cyclamate is a carcinogen. Moreover, the finding of lymphosarcomas and reticulum cell sarcomas in the Brantom study reinforces the concerns about the carcinogenicity of cyclamate arising from the lymphosarcomas found in the Kroes study (G–140 at 9–10).

I reject Abbott's argument that the combination of generations in the Kroes study is inappropriate. Dr. Carlborg, who is Abbott's witness and who raised every conceivable criticism of the statistical analyses contained in the Remand Order, did not criticize the combination of generations (A–857). Indeed, Dr. Carlborg performed his own statistical analyses of the data utilizing all of the adjustments and types of tests he deemed appropriate, and concluded that when the three generations were combined the evidence of lymphosarcomas for control vs. male mice treated with 5% cyclamate was significant at the P=.031 level, that lymphosarcomas for control vs. all cyclamate treated male mice was significant at the P=.017 level and the linear trend for male mice was significant at the P=.036 level (A–857, Exhibit 2 at lines 17–18). Although Dr. Carlborg dismisses these statistically significant results as "artifacts" he does not dispute the validity of combining generations. Indeed, no Abbott witness disputes the validity of this method. In view of the lack of evidence to the contrary, the combining of generations by Abbott's own witness, that witness's conclusion that the results were statistically significant and thus the method implicitly valid, and Dr. Frankos' testimony and the testimony of Dr. Gaylor (G–139 at 7) acknowledging the validity of this method, I conclude it is a valid method.

Abbott contends, however, that Dr. Frankos' testimony concerning the use of this method in other studies is equivocal and should be given no weight (Abbott Remand Ex. at 27–28; Abbott Remand Reply at 2–3). Dr. Frankos testified that the combining of generations was employed as a method of analyzing data on the possible carcinogenicity of xylitol (R. Tr. at 168). Abbott contends that just prior to giving this testimony Dr. Frankos was uncertain about his answer (Abbott's Remand Ex. at 28). However, Dr. Frankos' second answer is emphatic and I find it has probative value.

Moreover, Dr. Frankos also testified, in response to a question about whether the FDA permits reviewers to combine generations for review of a multigeneration study, that "[o]ur statisticians have done it * * *. We have your statistician, Dr. Carlborg and Dr. Gaylor and other statisticians, they all have done that" (R. Tr. at 160). Thus, I find tht Dr. Frankos' testimony, when read in its entirety, is credible and supports the conclusion that the combining of generations in the Kroes study was appropriate.

I also reject Abbott's argument that the lymphosarcoma finding in the Kroes study is not biologically significant. Abbott contends: (1) it is only sex specific in males (not in females and not in males and females combined), and (2) this sex specificity of lymph system sarcomas is not confirmed by other studies. Abbott's argument misallocates the burden of proof. The burden is not on the Bureau to submit an additional study confirming the finding in the Kroes study, but rather the burden is on Abbott to produce negative results in the same sex, species and strain of mice as in the Kroes study. The absence of increased lymphosarcomas in female mice in the Kroes study may have been due to the fact that the survival of the females was significantly less than the survival of the males (see G–41, App. VII, National Institute of Public Health, Netherlands at 1). Moreover, the fact that a cancer is found only in a specific sex and a specific strain does not mean that it can simply be dismissed as being irrelevant to humans (G–140 at 11). This issue is discussed in detail in my discussion of the Rudali study. For the reasons given there, I reject Abbott's argument that the lymphosarcomas in the Kroes study are not biologically significant.

Finally, I reject Abbott's argument that the historical spontaneous incidence of the particular type of tumor in the animal strain in the Kroes study easily explains the finding (Abbott's Remand Ex. at 28). The only evidence of the spontaneous incidence of lymphosarcomas of the strain of mice in the Kroes study is the Temporary Committee Report. (Abbott also cites an exhibit submitted by the Bureau which reports the spontaneous incidence of leukemia-lymphomas as being between 1.6 and 6.8% (G–141 at 962). However, this report does not involve the same strain of mice as that used in the Kroes study.) The Temporary Committee report states that the spontaneous diseases for the strain of mice used in the Kroes study "includes a 5–10% incidence of leukemia (primarily lymphocytic)" (G–41, App. III, National Institute of Public Health, at 2). However, the Kroes study reported leukemias separately from lymphosarcomas. Therefore, it is unclear whether the spontaneous incidence of lymphosarcomas in the strain of mice in the Kroes study is in fact 5–10%. The three separate generations of mice in the Kroes study had a zero incidence of lymphosarcomas (R. Tr. at 180). This would indicate that the historical incidence of lymphosarcomas in this strain of mice is low (*id.*). Even if the spontaneous incidence of lymphosarcomas is 5–10%, there is no testimony that the incidence of lymphosarcomas in the control mice in the Kroes study was *unusually* low. Thus, the evidence does not establish that the spontaneous incidence of lymphosarcomas in the strain of mice used in the Kroes study is 5–10% or that the incidence of lymphosarcomas in the control mice is unusually low.

Finally, even assuming that the incidence of lymphosarcomas in the Kroes study control mice was unusually low, the results of the study were nevertheless statistically significant. Moreover, there was a dose response relationship between cyclamate and the incidence of lymphosarcomas (P=.036) (A–857, Exhibit 2 at "Linear trend" for males). If the tumor difference between cyclamate-treated and control animals were due to the spontaneous occurrence of tumors in treated animals, the effect seen would not be expected to have a dose response relationship. I cannot conclude that such results are biologically insignificant absent sufficient additional data in the same strain and sex of mice showing negative results, Abbott here relies on the testimony of Dr. Carlborg who allegedly found other effects that were artifacts (Abbott's Remand Ex. at 29). This issue is discussed in detail in my discussion of the Brantom study. For the reasons given there, I reject Abbott's argument that the lymphosarcomas found in the

Kroes study are artifacts and therefore not biologically significant.

The Bureau also contends that autolysis limited substantially the detectability of effects in the Kroes study (Bureau's Brief at 18). I agree. Autolysis is a decaying of tissue that begins shortly after death and that makes examination of tissue more difficult. Abbott contends that autolysis reduced the significance of the negative results in the Kroes study, but argues that the study is not insignificant as a negative study. It should first be noted that, with respect to lymphosarcomas, I have found the results of the Kroes study to be suggestive of carcinogenicity, not negative. To the extent that the Kroes study did not reveal a significant difference between cyclamate-treated and control animals in the incidence of tumors other than lymphosarcomas, I agree with the Bureau that the autolysis in the Kroes study substantially reduced the detectability of effects in that study and thus reduces the sensitivity of the study (G–121 at 9; G–126 at 12; G–113 at 7; G–112 at 15).

d. *Hardy, et al. (A–690).* (1) *Study Design:* The ALJ described this study as follows:

> This study employed 48 male and 50 females ASH/CS1 (SPF) strain mice. The mice were fed CHA–HCL at concentrations of 300, 1,000 or 3,000 ppm. A control group of 48 males and 50 females was maintained. The study was conducted for 80 weeks, after which the survivors were sacrificed.

(ID at 16).

(2) *Study Results:* The ALJ found that: This increase * * * was not statistically significant. The Hardy data is important because the treated female group reflecting the increased lymphosarcoma levels was the same mouse strain which showed an effect in the Brantom study (HSH–CS1) [sic] mice.) Although this data enhances the Brantom data's credibility, taken alone, it is too tenuous to warrant declaring cyclamate a carcinogen.

(IRD at 22).[21]

(3) *Analysis:* Abbott contends that, applying the Bonferroni inequality multiplier of 3 to the P-value reported for lymphosarcomas in the Hardy study, the resulting P-value is .384. Abbott further notes that there was no dose response relationship exhibited in the Hardy study (Abbott's Remand Ex. at 29–30). The Bureau agrees with the ALJ that, taken by itself, the Hardy study is not positive. The Bureau argues, however, that because an increased incidence of lymphosarcomas were found in the same sex and in the same strain as in the Brantom study, the increased incidence is biologically significant (Bureau's Remand Reply at 6–7; G–139 at 7).

I find that the results of the Hardy study do add to the weight to be given the finding of lymphosarcomas in the Brantom study which employed the same strain of mice as the Hardy study. Two factors support this conclusion. Brantom used dose level of cyclamate of 0, .7, 1.75, 3.5, and 7.7% of the diet. Hardy used dose levels of cyclohexylamine that were considerably lower (0, .03, .10, and .30%) than the levels of cyclamate used in the Brantom study. Allowing for the differences in dose levels and metabolism of cyclamate to cyclohexylamine, the responses in the female strain of mice used in both the Hardy and Brantom studies are consistent.

I recognize that when the Bonferroni multiplier is applied to the P-value for lymphosarcomas in the Hardy study that the resulting P-value is .384. I do not, however, find that the results of the Hardy study are statistically significant. Thus, the precise P-value selected is not that important. The important aspect of the Hardy study is that it is biologically significant in that it supports the finding of lymphosarcomas in the same mouse strain in the Brantom study.

e. *The Significance of Lymphosarcomas.* The ALJ made the following statement concerning the significance of lymphosarcomas generally:

> Evidence was also submitted regarding the potential effects lymphosarcomas have upon different body organs and systems. Both parties agreed tht because the lymph system is crucial to an organism's immunological defenses, any assault upon its smooth functioning threatens that organ's viability. However, the parties did not agree that cyclamate was a carcinogen. Only if cyclamate was a cancer promoter would these factors be relevant to its safety.

(IRD at 22).

The Bureau agrees with the first three sentences of the above-quoted statement, but takes exception to the final sentence (Bureau's Remand Ex. 5). The Bureau contends that there is no evidence of record to support the statement that evidence of cyclamate's causing damage to the lymph system would be relevant only if cyclamates were a cancer promoter (*id.*). Abbott did not reply to this exception.

It is unclear what the ALJ was referring to in the last sentence of the above-quoted statement. Several matters are, however, clear. Lymphosarcomas are malignant tumors (G–121 at 6). This evidence is uncontradicted and I do not believe that the fact that lymphosarcomas are a form of cancer can be seriously disputed (G–140 at 8).

A separate issue, and perhaps the issue that caused confusion for the ALJ, concerns the role played by the lymph system in immunological defense. An effect on the lymph system could reduce an animal's immunological defenses to an infectious disease causing the animal to die from that disease or to be sick for longer periods of time than it might ordinarily (G–140 at 9). This adverse health effect is, however different from cancer, and is not being relied on to support my finding that cyclamate has not been shown to be safe, from a carcinogenicity standpoint. (I note, however, that one study in the record in this proceeding examined the effect of calcium cyclamate on the humoral immune response of rabbits and found that "cyclamate given to rabbits for 150 days increased the period required for the immune system to respond to stimulation by BSA [bovine serum albumin]" (G–54 at 53). Thus, there is evidence in the record to support the theory that cyclamate may also have an adverse effect on the immune system.

2. *The Occurrence of Bladder Tumors in Direct Feeding Studies in Rats.* The occurrence of bladder tumors in a number of strains of cyclamate-treated rats in a number of cyclamate direct feeding studies raised a serious question about the safety of cyclamate. Bladder tumors in these strains of rats are rare. Their occurrence, even in small numbers that are not statistically significant at $P \leq .05$ within each study, is biologically significant.

The method employed by the Bureau to evaluate the possible carcinogenicity of cyclamate in these studies was to combine a number of studies involving a specific strain of rats and compare the occurrence of bladder tumors in the cyclamate-treated rats to the background rate for that type of tumor obtained from historical controls. This method revealed that the difference in tumor incidence between cyclamate-treated animals and historical controls is statistically significant at $P < .05$. As the subsequent discussion establishes, this method is a valid and scientifically

[21] CHA–HCL (Cyclohexylamine hydrochloride) is a metabolite of cyclamate. When humans ingest cyclamate, enzymes in the body may transform (metabolize) some of the cyclamate to cyclohexylamine (G–41 at 36). Thus, exposure of a human to cyclamate may result in exposure to cyclohexylamine also. Several forms of the cyclamate metabolite, cyclohexylamine, were used in studies that comprise the record of this proceeding. These forms are cyclohexylamine (CHA), cyclohexylamine hydrochloride (CHA–HCL) and cyclohexylamine sulfate (CHS). In the stomach all three of these compounds will be present in the same form. For this reason, all three forms of cyclohexylamine are considered to be biologically equivalent, and studies using them are relevant in this proceeding.

acceptable means of evaluating the possible carcinogenicity of a test substance and raises a serious question as to the carcinogenicity of cyclamate.

a. *The Occurrence of Bladder Tumors in Sprague-Dawley and Wistar Rats:* The Schmaehl, Homberger, Taylor, Ikeda, and Hicks (direct feeding) studies involved the direct feeding of cyclamate to Sprague-Dawley or Wistar rats to determine whether cyclamate is a carcinogen. (All of these studies are discussed in detail below.) Although cyclamate treated animals in most of these studies did develop tumors, when comparing treated and control animals within each study, the tumor incidences were not statistically significant at the $P<.05$ level. The Bureau contends, however, that the three bladder tumors found in cyclamate-treated Sprague-Dawley rats in the Homberger, Schmaehl, and Taylor studies combined is statistically significant at the $P=.02$ level when compared to the spontaneous rate of bladder carcinomas in Sprague-Dawley rats (approximately .23%) based on historical data (G–120 at 10; Tr. at 601–604). Moreover, the Bureau notes that the one bladder tumor found in the control animals in these three studies is not inconsistent [22] with the low background rate based on historical data *(id.)*. The Bureau further contends that the three bladder tumors found in cyclamate-treated Wistar rats in the Ikeda and Hicks direct feeding studies combined is statistically significant at the $P=.002$ level when compared to the background rate for bladder carcinomas (approximately .116%) developed from historical data of the National Cancer Institute for all species of rats combined (G–120 at 10–11). Moreover, the absence of any bladder tumors in the control groups in these two studies is consistent with the low background rate based on historical data *(id.)*.

I agree with the Bureau's analyses of these data. Although a comparison to historical controls would not ordinarily be accepted as a basis for contradicting the results of a comparison to concurrent controls within a study, where, as here, the individual studies are of low sensitivity and the tumor in question has a very low background rate, such a comparison has validity. What is significant about these studies is that in a number of studies involving different strains of rats we are seeing the occurrence of the same rare tumor in treated animals and fewer in controls (G–121 at 8; see G–120 at 16; G–139 at 6). The importance of the occurrence of such tumors in rats was recognized by the Temporary Committee (G–41 at 20–21; 25). Moreover, the occurrence of these tumors in cyclamate treated animals is consistent with a small treatment effect and the occurrence of these tumors in control animals is consistent with the incidence of these tumors in historical controls. The fact that a similar effect is present in two separate strains of rats adds credence to the conclusion that these effects are important (Tr. at 613–14). These findings, by themselves, are biologically significant.

The statistical method employed by the Bureau confirms that these findings are biologically significant. It provides an objective means of evaluating the significance of these rare tumors. The results of the application of this method to the Wistar and Sprague-Dawley rat strain studies, when viewed together with the results of one of the Friedman studies (discussed below), which involved Osborne-Mendel rats, led one Bureau witness, Dr. Charles Brown, who was the head of the statistics working group for the NCI-Temporary Committee, to conclude that "for rats, the evidence of positive carcinogenicity is not overwhelming, but it is suggestive that they are sensitive to carcinogenic insult by cyclamate" (G–120 at 16). I agree with Dr. Brown's conclusion.

In its exceptions, Abbott does not contest the propriety of combining these studies, but contends that the results within each study were not statistically significant at the $P<.05$ level and that there was no dose response relationship (Abbott's Exceptions at 27). Even thought Abbott is correct in its characterization of the individual studies, I do not believe this argument affects the overall significance of the bladder tumors found in these studies as a group. The sensitivity of most of these studies is low (see G–41, App. V at 19–20; see discussion below). Low sensitivity is important because, if cyclamate is a weak carcinogen, it would not be expected to produce tumors significant at the $P<.05$ level or to exhibit a dose response relationship in such small studies (see G–120 at 7–8). Indeed, even in the most sensitive of these studies, the Schmaehl study, there was a reasonable chance that the study would fail to detect a true difference in tumor incidence of 4% between control animals and those treated at the 5% feeding level (G–120 at 6–7).

Accordingly, the lack of a statistically significant effect in each of these studies when considered alone does not rebut the question about cyclamate's safety raised by the comparison between the combined incidence of bladder tumors found in cyclamate treated Sprague-Dawley and Wistar rats and the background rate for such tumors based on historical data.

I recognize that the validity of combining the results of different studies and comparing it to historical controls can be questioned on the ground that the studies being combined were conducted in a different manner (G–120 at 11–12). I find, however, that the method of combining these studies used by the Bureau was appropriate for two reasons. First, only studies involving the same species and strain of rats were combined (G–120 at 12). This eliminates the possibility that a strain difference in the sensitivity of these animals to cyclamate would complicate the analysis. Second, the tumor findings in the studies that were combined are not inconsistent (see Tr. at 628). The incidence of bladder tumors in the control animals in the combined Sprague-Dawley and Wistar rat studies were consistent with each other and with the incidence of bladder tumors found in control animals based on historical data. Thus, the combination of the data from these studies is valid.

It should be emphasized that these findings of bladder tumors in rats do not conclusively establish that cyclamate is a carcinogen. Moreover, these findings do not provide the same degree of confidence that one would have if the results were statistically significant when compared to controls in each study. These findings do, however, raise a valid and serious question as to cyclamate's safety. It is therefore necessary that cyclamate be tested further to resolve this issue. In reaching this conclusion, I am not requiring that Abbott prove a negative, which is, of course, impossible. I am, however, holding that Abbott cannot escape the force of these studies unless it submits additional evidence in the form of sufficiently sensitive studies that demonstrate to a reasonable certainty that cyclamate does not cause bladder tumors in rats.

A detailed discussion of the study design, study results and my analysis of the Hicks (direct feeding), Ikeda, Taylor, Homberger, Schmaehl, and Plank studies follows. Abbott contends that, when viewed individually, the Hicks (direct feeding), Ikeda, Taylor, Homberger, Schmaehl and Plank studies are negative. As the discussion below establishes, however, Abbott's contention is without merit.

(1) *Hicks, et al. (direct feeding study) (G–2, A–832).* (a) *Study Design:* This study was conducted in conjunction with the Hicks MNU study (discussed separately below). The study involved

[22] The incidence of 1 bladder tumor in 225 total control animals is approximately .44%.

Wistar rats fed cyclamate alone. The study was continued for 24 months, after which time the animals were sacrificed.

(b) *Study Results:* Of the 84 surviving rats fed cyclamate alone, five males were found to have tumors (3 bladder, 2 kidney). No control animals were found to have tumors. The Pathology Working Group of the Temporary Committee confirmed three malignant bladder and two malignant kidney tumors (G–41 at 25). Although the incidence of these tumors was not statistically significant at $P \leq .05$ ($P = .2$), the Temporary Committee found that "their occurrence in a low incidence must be evaluated with respect to the reported absence of these tumors in matched and historical control animals" (*id.*). Dr. Hicks stated that the background bladder and kidney tumor rate for these rats in her lab was zero (G–114 at 21).

(c) *Analysis:* Abbott contends that the results of this study are not statistically significant at the $P \leq .05$ level and that therefore the study is negative. I disagree. The total tumor incidence in this study is significant at the $P = .2$ level (G–114 at 21). There is thus an 80% probability ($P = .2$) that the results of the Hicks direct feeding study are due to cyclamate instead of a 95% probability necessary for statistical significance at the $P \leq .05$ level. Obviously, I would be more certain of the importance of these results if the incidence of bladder tumors were significant at the $P \leq .05$ level. I do not, however, consider these results to be negative, particularly in view of the biological factors present in the study. The significance of these results is enhanced by the fact that Dr. Hicks testified that she had never seen such tumors in untreated animals in her laboratory (G–114 at 21). This factor led Dr. Hicks to conclude that the total tumors found in this study were "pathologically * * * very significant" (*id.* at 20). Moreover, as previously noted, Dr. Brown testified that the "probability of observing three or more tumors in the 217 treated animals in [the Hicks and Ikeda] studies combined is .002" (assuming a background rate of .116% tumor incidence as obtained from the NCI data on all species of rats combined) (G–120 at 11). Although this evidence does not conclusively establish that cyclamate is a carcinogen, the study cannot be considered proof of safety and indeed raises a question as to the potential carcinogenicity of cyclamate.

(2) *Ikeda, et al. (G–79).* (a) *Study Design:* This study involved groups of 54–56 male Wistar rats fed sodium cyclamate or a sodium saccharin plus sodium cyclamate mixture. The concentration given was 2% for the first 20 days, 3% for days 21–60, 4% for days 61–150, 5% after 150 days and 6% at one year. The study was continued for 28 months. At the time of the Temporary Committee Report, only 40% of the microscopic examinations had been performed.

(b) *Study Results:* The Temporary Committee found that "none of the test materials induced tumors of the urinary bladder in any of the treated animals" (G–41 at 26). The ALJ stated that "[n]o bladder tumors were observed in the animals so far examined. However, testicular degeneration and urinary calculi were observed in treated animals and appeared to be treatment related" (ID at 13–14).

(c) *Analysis:* Abbott relies on the report of the Temporary Committee and contends that the study is negative (Abbott's Brief at 24). The Bureau concedes that no bladder tumors were found in the animals in the Ikeda study, but notes that "histopathology had not been performed on other animal organs at the time of the report's publication" (Bureau's Brief at 18). The Bureau contends that the Ikeda study is therefore inconclusive.

I find that the histopathology in the Ikeda study is insufficient for classifying this study as negative, particularly in light of the evidence of lymphosarcomas in the Brantom, Hardy and Kroes studies and the evidence of lung and liver tumors in the Rudali study. These studies support the conclusion that tumors at sites other than the bladder may have been present in the Ikeda study, but were not detected, since only bladder histopathology was done.

Moreover, the Ikeda study has only a 50% chance of detecting, at the 95% confidence level, a true difference in tumor incidence of approximately 13% between the controls and the high dose (2.5 gm/kg) treated animals (G–41, App. V at 19). This study is therefore unlikely to detect a small treatment effect. This lack of sensitivity is especially important in view of the findings of bladder tumors in Sprague-Dawley rats combined and Wistar rats combined (discussed above). Accordingly, I cannot consider this study to be proof of cyclamate's safety.

(3) *Taylor, et al. (G–13).* (a) *Study Design:* The Taylor study involved 48 male and 48 female Sprague-Dawley strain rats fed a diet containing 5% calcium cyclamate. The animals were derived from parents who were also administered cyclamate from the time of mating through delivery and weaning of the test generation. The study was continued for 114 weeks.

(b) *Study Results:* One bladder tumor was found in a control animal and none in cyclamate treated animals. The Temporary Committee reported that the study was "* * * particularly good in that animals were exposed *in utero* and continued on treatment for their lifetimes" and that "the test material did not display carcinogenicity" (G–41 at 21). The ALJ found that the Taylor study "employed an unacceptably small number of animals per group" (ID at 31).

(c) *Analysis:* Abbott takes exception to the ALJ's finding on study size (Abbott's Exceptions at 29). Abbott, relying on the Temporary Committee Report, contends that the Taylor study is negative (Abbott's Brief at 18, 25). The Bureau contends that "the relatively small number of animals examined microscopically (for the bladder 49 controls and 53 treated) reduced the sensitivity of the study" (Bureau's Brief at 15–16; G–41, App. VII, Taylor and Friedman 1974 at 2).

I find that the Taylor study has only a 50% chance of detecting, at the 95% confidence level, a true difference in tumor incidence of approximately 9% between the controls and the high dose (5%) treated animals (G–41, App. V at 19). Moreover, the presence of one bladder tumor in a control animal makes the detection of a positive effect more difficult because the difference between the number of animals with tumors in the treated and control groups needs to be greater in order for that difference to be statistically significant (G–112 at 15; G–113 at 8; see G–41 App. VII, Taylor and Friedman 1974 at 2). This study therefore is unlikely to detect a small treatment effect. This lack of sensitivity is especially important in view of the findings of bladder tumors in Sprague-Dawley rats combined (discussed above). These tumor findings suggest a low treatment effect and thus emphasize the need for studies of greater sensitivity than the Taylor study. Accordingly, I cannot consider this study to be proof of cyclamate's safety.

(4) *Homberger et al. (A–348).* (a) *Study Design:* The Homberger study involved groups of 25 Charles River CD–1 Sprague-Dawley male rats which were fed 0, 1 or 5% sodium cyclamate. The bladders of at least 12 animals per group were examined microscopically. They were started on test at approximately six weeks of age and continued on treatment for two years.

(b) *Study Results:* The authors of the study concluded that:

> On the basis of these experiments, it cannot be concluded that * * * cyclamate [is] carcinogenic. This may be considered of significance since for smaller doses of other compounds under similar conditions were unquestionably carcinogenic for liver,

bladder, subcutaneous, vascular and other tissues of rats and/or mice.

(A–348 at 9). Two carcinomas of the bladder were found in cyclamate-treated animals (one in the high dose group and one in the low dose group) and none in control animals (Tr. at 602–604). The ALJ found that the Homberger study employed an "unacceptably small number of animals per group" (ID at 31).

(c) *Analysis:* Abbott takes exception to the ALJ's finding (Abbott's Exceptions at 29). Abbott, relying on the authors' conclusion, contends that the study is negative (Abbott's Brief at 17, 22).

The Bureau relies on the Temporary Committee Report. The Report stated that:

[a] number of questions were raised regarding the experimental design and conduct of this study. The small animal group size and the possibility of cross-contamination of the cyclamate-treated animals with other chemicals being tested in the same room, including one later found to be a bladder carcinogen, limit the value of this study in assessing the carcinogenicity of cyclamate.

(G–41 at 18; Bureau's Brief at 14). The Bureau concludes that the Homberger study is inconclusive (Bureau's Brief at 19).

I agree with the Bureau and the ALJ that the sample sizes employed in the Homberger study were unacceptably small. The Homberger study had only a 50% chance of detecting, at the 95% confidence level, a true difference in tumor incidence of approximately 26% between the controls and the high dose (5%) treated animals (G–41, App. V, at 19). This study therefore is unlikely to detect a low treatment effect. This lack of sensitivity is especially important in view of the findings of bladder tumors in cyclamate-treated rats in this study (G–121 at 7) and in Sprague-Dawley rats combined and Wistar rats combined (discussed above). Accordingly, the study cannot be considered proof of cyclamate's safety.

The author's remark that smaller doses of other compounds under similar conditions were unquestionably carcinogenic under similar conditions does not alter my conclusion. At best, that finding only tends to show that cyclamate is not a strong carcinogen in this species. However, because of the poor sensitivity of the study, it does not provide any reliable insight into whether cyclamate is a weak carcinogen.[23]

As to the possibility of cross-contamination of the animals in the Homberger study with other chemicals, including a bladder carcinogen, I find that the likelihood of cross-contamination is too speculative to be relied upon, especially where I have found no tumors occurring in control animals (see Section IV.B.2.b.(1)(c) below).

(5) *Schmaehl (A–555).* (a) *Study Design:* This study involved groups of 104 Sprague-Dawley rats fed either sodium cyclamate, sodium cyclamate and saccharin, or CHA. Animals were started on study between 70–90 days of age and continued on treatment for their lifetimes.

(b) *Study Results:* One bladder tumor was found in a cyclamate treated animal. The authors of the study concluded, however, that "[i]n spite of the high dosages and the duration of the experiments over an entire lifetime, no evidence was found of chronic toxic or carcinogenic activity of the substances tested" (A–555 at 6). The Temporary Committee found that the reported extremely rare occurrence of spontaneous bladder tumors in the rat strain used "must be taken into consideration when evaluating the significance of the one bladder transitional cell carcinoma found in a cyclamate-treated animal" (G–41 at 23–24).

(c) *Analysis:* Abbott, relying on the conclusion of the authors of the study, contends that the Schmaehl study is negative (Abbott's Brief at 23–24). The Bureau does not criticize the conduct or design of this study, but considers the one bladder tumor found in this study together with the tumors found in the Homberger and Taylor studies and contends the results are biologically important. The Bureau's contention concerning the analysis of this study with other similar rat studies and Abbott's exception to this analysis is discussed above. As noted there, the occurrence of a bladder tumor in the Schmaehl study is consistent with a small treatment effect (G–120 at 6–7; see G–126 at 11–12), even though it is not significant at the $P \leq .05$ level. Moreover, as the Temporary Committee noted, this finding must be viewed in light of the extremely rare occurrence of spontaneous bladder tumors in this rat strain (G–41 at 23–24). Accordingly, I cannot consider the Schmaehl study to be proof of cyclamate's safety.

(d) *Other matters:* As part of the Remand Order, the parties were asked to comment on the apparent failure of the Schmaehl study to report the results of the study separately by sex. Abbott contends that a number of effects were reported by sex and that even if the reported incidences of tumors all occurred in one sex, none of the reported findings would be statistically significant at $P \leq .05$ (A–858 at 23). (Findings other than tumor findings, such as water intake and body weight gains, were reported by Dr. Schmaehl with information about the sex of the animals. However, this information is not relevant to the question raised by the Remand Order. That question was intended to inquire whether certain *tumors* may have been statiscally significant if they occurred only in one sex.)

The Bureau argues that tumor findings statistically significant at $P \leq .05$ may be present in the Schmaehl study because, if the tumors of the reticuloendothelial system (reticular cell sarcomas, lymphosarcomas and leukemia combined) all occurred in the same sex, their incidence would in fact be statistically significant at $P \leq .05$ when compared to controls (Bureau's Remand Brief at 5–6).

I agree with Abbott that the incidence of lymphosarcomas, reticular cell sarcomas, and leukemias occurring in the Schmaehl study, if examined independently, would not be statistically significant, even if occurring in one sex (R. Tr. at 215). I find, however, that by combining either reticulum cell sarcomas and lymphosarcomas or these two effects and leukemias, a result statistically significant at $P \leq .05$ would be achieved, if these effects occurred all in one sex (G–140 at 3–4; R. Tr. at 161; 210–11; 215–16). Such a combination of the data is appropriate because lymphosarcomas, reticular cell sarcomas, and leukemias all involve cells derived from reticulum cells (R. Tr. at 113–19). Without a report of these tumor findings by sex, this issue cannot be conclusively resolved.

Abbott's position is that a detailed report of the tumor findings by sex is nevertheless unnecessary. Abbott contends that a statement contained in the Schmaehl report and a conversation between Dr. Oser and Dr. Schmaehl are sufficient to resolve this issue (Abbott's Remand Ex. at 21). The Schmaehl report states that "No greater incidence regarding either sex could be detected with reference to the benign or the malignant tumors" (A–555 at 5). Additionally, in a conversation with Dr. Oser, Dr. Schmaehl is alleged to have said that he would have reported significant differences as to sex if they

[23] The terms "strong" and "weak" carcinogen are used here to differentiate between compounds which respectively cause relatively high and relatively low incidences of tumors when tested in experimental animals. Even a "weak carcinogen", however, by this distinction can cause important and unacceptable incidences of cancer in the human population as a whole. Indeed, even a 1% increase in tumor incidence would be unacceptable in the human population (Tr. at 102).

were present (A–858 at 20). It is unclear, however, whether Dr. Schmaehl's statements refer solely to specific tumor findings, *e.g.*, lymphosarcomas, or whether they also refer to combined tumor findings, *e.g.*, effects on the entire reticuloendothelial system. This question is particularly important in view of Abbott's questioning of the propriety of combining effects on the reticuloendothelial system (R. Tr. at 117–18). If Dr. Schmaehl shared Abbott's skepticism about the combining of effects, he probably would not have analyzed combined effects on the reticuloendothelial system. Accordingly, without a report of lymphosarcomas, reticular cell sarcomas and leukemias by sex, the precise meaning of Dr. Schmaehl's statements remain uncertain. As noted previously, it is not unusual for scientists to disagree with the conclusions of the author of a study as to the significance of the results of a study. I therefore agree with the ALJ's conclusion that "The only conclusion that can be drawn from the author's failure to report this data separately by sex is that it is uncertain whether a true sex specific effect occurred" (IRD at 12).

It should be noted that Dr. Oser's conversation with Dr. Schmaehl was stricken as hearsay and Abbott took exception to this ruling. Although I agree with the ALJ's ruling, I have nevertheless considered the statement and found that it does not resolve the issue because of the ambiguity contained in the statement.

(6) *Plank, et al. (A–401–404).* (a) *Study Design:* This study involved Charles River CD–1 Sprague-Dawley albino rats, in groups of 25 of each sex, fed the following concentration of cyclohexylamine sulfate: 0.15 mg/kg/day, 1.5 mg/kg/day, or 15 mg/kg/day. A control group of 25 of each sex was also maintained. The study was conducted for two years, after which the survivors were sacrificed.

(b) *Study Results:* A single bladder carcinoma was found in one male from the high dose treatment group (G–41 at 24). The Temporary Committee found that "[t]he value of this study is limited by its poor sensitivity. It is thus considered to be of minimal value in assessing the carcinogenicity of cyclohexylamine"(*id.*). The ALJ found that "[b]ecause of the extreme rarity of spontaneous bladder tumors in this strain, the positive finding raises questions concerning CHA's carcinogenic potential. Furthermore, the study's sensitivity was limited due to the small number of animals used" (ID at 16).

The ALJ also grouped the Plank study with the Hicks, Friedman, and Schmaehl cyclamate direct feeding studies and found that:

> In the rat studies, seven transitional cell carcinomas of the bladder, two of the kidney, three bladder papillomas, five hyperplastic lesions, and a bladder proliferative lesion were found in rats treated solely with cyclamate [Hicks, (Ex. No. G–2), Plank (Ex. No. A–146), Friedman (Ex. No. A–195) and Schmaehl (Ex. No. A–0555) [Studies]].

(ID at 31).

(c) *Analysis:* Abbott takes exception to this study's being grouped with other studies involving rats fed cyclamate rather than CHA (Abbott's Exceptions at 27). Abbott also contends that in grouping the Plank Study with other studies (1) the ALJ erroneously grouped together carcinomas and noncarcinomas, such as papillomas, and (2) that the three bladder papillomas cited by the ALJ were not confirmed by the Pathology Working Group (Abbott's Exceptions at 27). The Bureau does not dispute these points (Bureau's Reply at 14–15). I agree with Abbott that a study of CHA should not be grouped with cyclamate studies, although it can, by itself, provide important information about the safety of cyclamate because it is a metabolite of cyclamate in humans. I also agree with Abbott that the ALJ erroneously lumped noncarcinomas together with carcinomas. I find, however, that the Plank study does not prove the safety of CHA. The sensitivity of the Plank study is unacceptably low. The Plank study had only a 50% chance of detecting, at the 95% confidence level, a true difference in tumor incidence of approximately 33% between the controls and the high dose (15 mg/kg) treated animals (G–41, App. V at 20). As a result, statistical significance at the $P \leq .05$ level is difficult to demonstrate unless the test substance causes an exceptionally high tumor incidence (G–121 at 8). The single bladder tumor found in a cyclohexylamine treated animal has biological significance because the occurrence of bladder tumors in the strain of rats employed in the Plank study is rare. This single bladder tumor may be an indication of a weak carcinogenic effect which might have been statistically significant if the study had been larger. The single bladder tumor found in the Plank study thus has biological significance (*id.* at 8–9). Accordingly, I cannot consider this study to be proof of cyclamate's safety.

b. *The Occurrence of Bladder Tumors in Holtzman and Osborne-Mendel Rats:* Two studies submitted in the cyclamate hearing were conducted by Friedman, *et al.* and involved Osborne-Mendel or Holtzman rats. These studies conducted independently, were published together and are discussed below.

(1) *Friedman, et al. (A–388).* (a) *Study Design:* The first of the two Friedman studies (hereafter "first Friedman study") was conducted using seven male and seven female Osborne-Mendel rats per group. These rats were fed sodium cyclamate or calcium cyclamate at 0.4%, 2.0%, or 10% of their chow diet for 101 weeks. A group of 14 controls per sex fed a standard chow diet was maintained. The animals were started as weanlings and the study continued for 101 weeks.

The "second Friedman study" was conducted using male Holtzman rats. A group of twenty of these rats were fed a semisynthetic diet containing calcium cyclamate at 1% level plus 20% casein, and 2% level plus 20% casein, and 2% level plus 10% casein. An equal number of controls were fed the semisynthetic diet with 20% casein.

(b) *Study Results:* Three transitional cell carcinomas (two at the low dose and one at the high dose) and two papillomas of the bladder were found in the calcium cyclamate treated animals in the first Friedman study (ID at 11). Three papillomas were found in the sodium cyclamate treated rats in the first Friedman study (*id.*). One papilloma was found in a calcium cyclamate treated animal in the second Friedman study (G–41, App VII, Food and Drug Administration L. Friedman et al. 1972 at 4).

The Pathology Working Group of the Temporary Committee confirmed the three bladder carcinomas found in the calcium cyclamate treated animals in the first Friedman study, but did not confirm the papillomas. The Temporary Committee found that:

> The small number of rats used is considered to be a major deficiency in this study. Although the incidence of bladder tumors was not statistically significant, their importance, even in small numbers, must be evaluated with respect to the reported rarity of spontaneous bladder tumors in the rat strain used.

(G–41 at 20–21). The calcium cyclamate portion of the first Friedman study was among the studies that the Temporary Committee found create a "sense of uncertainty" about the safety of cyclamate (*id.* at 46).

The ALJ, who recognized the deficiencies in the first Friedman study cited by Abbott (discussed below), concluded that with respect to the first Friedman study the "incidence of tumors is important, even though not statistically significant, because spontaneous tumors are extremely rare in the rat strains employed" (ID at 12).

(c) *Analysis of the Calcium Cyclamate Portion of the First Friedman Study:* Abbott contends that (1) the three carcinogens found in calcium cyclamate treated animals in the first Friedman study were neither statistically significant nor dose related; (2) the small number of rats used in this portion of the first Friedman study is a major deficiency; (3) the tumor findings are complicated because they appeared only in the first Friedman study which utilized a chow diet and did not appear in the second Friedman study which utilized a semisynthetic diet; and (4) the first Friedman study is complicated by the presence of calculi and bladder parasites (Abbott's Brief at 23).

The Bureau contends that the three bladder tumors found in the calcium cyclamate portion of the first Friedman study are biologically significant, notwithstanding the lack of statistical significance at the $P<.05$ level, because the spontaneous bladder cancers in mice and rats are rare (G–121 at 8). The Bureau further contends that the lack of dose response might be attributable to the small size of the study (Bureau's Brief at 24).

I find that the three bladder tumors found in the calcium cyclamate treated animals of the first Friedman study add to the doubt about the safety of cyclamate that was raised by the findings of bladder tumors in Sprague-Dawley and Wistar rats discussed in subsection B.2.a. above.

I recognize that the three bladder tumors found in this study were not significant at the $P<.05$ level ($P=0.29$; G–41, App. VII, Food and Drug Administration Friedman et al. at 2). These tumors are nevertheless biologically significant because (1) the sensitivity of this portion of the Friedman study is low (this portion of the study had only a 50% chance of detecting, at the 95% confidence level, a true difference in tumor incidence of approximately 34% between the controls and the high dose treated animals (G–41, App. V at 19)) and (2) the spontaneous or background rate for bladder tumors in Osborne-Mendel rats is reported as being low (G–41 at 20–21; G–121 at 8). Thus, the occurrence of three bladder tumors in the calcium cyclamate treated rats is consistent with a small treatment effect even though they are not significant at the $P<.05$ level. The similar findings of bladder tumors in Sprague-Dawley and Wistar rats reinforces the conclusion that these three bladder tumors are biologically significant.

The calcium cyclamate portion of the first Friedman study is, however, further questioned by Abbott because there was no clear dose response relationship. If there were some correlation between the increased dose levels and an increase in tumor production, I would have greater confidence in the results of the study. However, the lack of such a response may have been due to the small sample size (G–120 at 13). A small sample size makes the finding of a dose response more difficult, because of random fluctuation (*id.*).

I reject Abbott's argument that the results of this portion of the first Friedman study are unreliable because tumors appeared only in animals on a chow diet (used in the first Friedman study) but did not appear in animals on a semisynthetic diet (used in the second Friedman study). Abbott's argument might have merit but for the fact that all control animals in the first Friedman study received the same chow diet (absent cyclamate) as the treated animals and there were no tumors found in the control animals. The use of concurrent controls in which no tumors were found negates the possibility that tumors found in treated animals were due to the chow diet (see Tr. at 1049–50). Thus, the study design ensured that the results of the study would not be biased by the type of diet received by the cyclamate treated animals.

Moreover, Abbott's only citation for this contention is the report of the study (A–195; Abbott's Brief at 18). This reference does not state that the results of the first Friedman study are complicated by the chow diet, but rather describes the results of the histopathology for the two studies (A–195 at 755–56). The only other support for Abbott's contention that could be found is the report of the Statistics Working Group to the Temporary Committee which speculates that the tumors in animals on a chow diet "may" have been due to contamination (G–41, App. VII, Food and Drug Administration L. Friedman et al. 1972 at 4). However, there is not evidence or other explanation supporting the suggestion that the chow diet may have been contaminated. Moreover, the report of the full Temporary Committee did not state that the chow diet was a complicating factor and recognized the potential importance of the three bladder tumors found in the calcium cyclamate portion of the first Friedman study (G–41 at 20–21). I therefore conclude that there is no basis upon which to attribute the three bladder tumors found in the cyclamate treated animals in the first Friedman study to the chow diet.

I also reject Abbott's argument that the calcium cyclamate portion of the Friedman study is deficient in that it utilized a small number of animals. Although this portion of the Friedman study utilized a small number of animals, the small sampler size is not a valid reason for discounting the three bladder tumors found in the calcium cyclamate-treated animals. A study, such as the Friedman study, which because of its insensitivity is unlikely to detect a carcinogenic effect, may nevertheless detect a carcinogenic effect in some cases. There is nothing inconsistent in finding that a study is too small to yield reliable negative results yet is sufficiently sensitive to raise serious doubts as to the safety of the tested substance (see Tr. at 630–31). Thus, the lack of sensitivity of the sodium cyclamate portion of the Friedman study is not a valid reason to criticize the finding of three bladder tumnors in the calcium cyclamate portion of the study, even though both portions of that study employed the same number of animals.

Finally, Abbott contends that the three tumors in the calcium cyclamate treated group may have been due to bladder calculi[24] or bladder parasites. The evidence on the relationship between bladder calculi and tumors is at present inconclusive (G–41, App V at 48–49). Moreover, in a related context, Abbott contends that "if a study is to have relevance on whether parasites cause bladder tumors, the length of exposure to parasites must be known" (Abbott's Exceptions at 31). This comment would seem to apply equally to bladder calculi. Abbott has not cited any evidence as to each animal's length of exposure to bladder calculi or bladder parasites. It may be that the bladder tumors in this portion of the first Friedman study were caused by bladder calculi. However, the evidence submitted is insufficient to establish that the bladder tumors were cause by bladder calculi.

Since the randomly selected control group presumably had an equal chance to develop such calculi, the observed bladder calculi may be treatment related in which case cyclamate might be producing a carcinogenic response, albeit an indirect one. Thus, even if there were definitive evidence that the bladder tumors in this study were caused by bladder calculi (which there is not), it would not resolve the question of the safety of cyclamate.

In sum, the three bladder tumors found in the calcium cyclamate portion of the first Friedman study do not conclusively establish that cyclamate is a carcinogen. Moreover, this finding

[24] Calculi are concretions usually of mineral salts around organic material found in the bladder.

does not provide the same degree of confidence that one would have if the results of the study were significant at the $P \leq .05$ level. This bladder tumor finding does, however, add to the doubt raised by the bladder tumors found in Sprague-Dawley and Wistar rats.

(d) *Analysis of the Sodium Cyclamate Portion of the First Friedman Study:* The Bureau contends that the sodium cyclamate portion of the first Friedman study should not be given any weight as a negative study because of the small number (fifteen) of animals treated (Bureau's Brief at 24). The Bureau also notes that the dose levels were rather low in this portion of the first Friedman study (G–120 at 12).

I agree with the Bureau. The size of this study is too small to permit reliable conclusions concerning the safety of cyclamate. This portion of the first Friedman study had only a 50% chance of detecting, at the 95% confidence level, a true difference in tumor incidence of approximately 34% between the controls and the high dose (10%) treated animals. (G–41, App. V at 19). This degree of sensitivity is unacceptably low (G–120 at 12). This portion of the study is therefore too insensitive to be considered proof of safety.

(e) *Analysis of the Second Friedman Study:* I find that the size of the second Friedman study is too small to permit reliable conclusions concerning the safety of cyclamate. The study had only a 50% chance of detecting at the 95% confidence level, a true difference in tumor incidence of approximately 20.5% between the controls and the high dose (2%) treated animals (G–41, App. V at 19). This degree of sensitivity is unacceptably low. This study is therefore too insensitive to be considered proof of safety.

(f) *Other matters.* As part of the Remand Order, the parties were asked to comment on the reported increased overall mortality in the Friedman study and the author's report that a small number of animals in the study were unaccounted for. The parties stipulated that these events result in a smaller pool of animals from which to measure biological effects and that therefore, the ability of the study to detect biological effects is decreased (Stipulation dated September 17, 1979 at 3).

3. *The occurrence of Bladder Tumors In Rats In Studies Other than Cyclamate or Cyclohexylamine Direct Feeding Studies.* The three studies discussed in detail below involve (1) the implantation of a pellet consisting of cyclamate and cholesterol in the bladders of mice (Bryan, G–1); (2) the direct feeding of cyclamate to animals which have a potent carcinogen (MNU) instilled in their bladders (Hicks, A–832, G–2); and (3) the direct feeding of a cyclamate/saccharin mixture to rats (Oser, G–81). It is undisputed that the incidence of bladder tumors in the treated group in all three of these studies is statistically significant. Abbott argues that, even if properly conducted, the techniques employed in the Hicks and Bryan study are invalid for assessing the carcinoginicity of a substance. Abbott further argues that the presence of saccharin in the Oser study makes that study inappropriate for assessing the carcinogenicity of cyclamate.

Although these studies are not as reliable as direct feeding studies, such as the Rudali study, I find that the results of these studies give rise to a high degree of suspicion concerning the possible carcinogenicity of cyclamate and add support to the bladder tumor findings in the cyclamate direct feeding studies discussed in Section IV.B.2. I recognize that the significance for human health of the findings in the Hicks and Bryan studies can not yet be fully evaluated. We do know, however, that the difference between a low incidence of cancer and no incidence of cancer (as in the Hicks controls) is the presence or absence of cyclamate (G–112 at 20). The suspicions raised by these studies could be negated by valid and convincing negative direct feeding studies or evidence that the carcinogenic response is unique to this mode of administration and could not result from ingestion of cyclamate. However, as the discussion in Sections IV.C. and D. establishes, Abbott has failed to submit such studies.

I note that these three studies do not play a major role in my decision. Indeed, even in the absence of these studies, I would reach the same conclusion, i.e., that cyclamate has not been shown to be safe. I have, however, given these studies some weight because, although the Hicks and Bryan techniques and the Oser study may not involve the methods of choice and should not be relied on primarily to screen food additives for carcinogenicity, these methods have shown biological effects cannot be ignored (G–112 at 20). The scientific basis for this conclusion is discussed below.

a. *Bryan, et al. (G–1).* (1) *Study Design:* The ALJ described the Bryan study as follows:

> Cholesterol pellets containing 20% sodium cyclamate were surgically implanted in the bladders of 100 female swiss mice. A control group of 100 mice with cholesterol pellets in their bladders was maintained. The mice were permitted to survive 55 weeks after which they were sacrificed and given a histologic exam.

(2) *Study Results:* The ALJ described the results of the Bryan study as follows:

> * * * the incidence of bladder tumors in the animals implanted with cholesterol and cyclamate pellets was 78% whereas the incidence was 13% in the controls. In a duplicate experiment conducted by Dr. Bryan the incidence was 61% in the test animals and 12% in the controls.
>
> A positive control group was also maintained. Mice were implanted with cholesterol pellets containing 8-methyl ether xanthurenic acid, a compound previously found to be carcinogenic in mouse bladders. The incidence of bladder tumors was 35%, as expected.

(ID at 18). The ALJ concluded that "* * * although there are questions as to whether [the Bryan technique] is still an appropriate procedure . . . the results cannot be totally disregarded. The results represent a major biological and statistically significant effect which has not been satisfactorily explained and which increases doubt concerning cyclamate's safety" (ID at 32).

The Temporary Committee, however, concluded that "the rout of administration [in the Bryan study] is inappropriate for assessing the carcinogenicity of a human dietary constituent" (G–41 at 27). The Bryan study was nevertheless among those studies referred to by the Temporary Committee which created a "sense of uncertainty" about the safety of cyclamate (G–41 at 46).

(3) *Analysis:* Abbott takes exception to the ALJ's finding with respect to the Bryan study (Abbott's Exceptions at 24). Abbott contends that the Bryan study contributes nothing to the evaluation of the carcinogenicity of cyclamate (Abbott's Brief at 32). Abbott relies on the testimony of Dr. Bryan who stated that "* * * It's a technique that is not at all replicative of normal human experience, experimental variables may be difficult to control * * * [and] utilization of this technique really has diminished substantially in the last several years * * *" (G–113 at 17–18; see also G–120 at 16).

The Bureau contends that the Bryan study is positive (Bureau's Brief at 19). The Bureau notes that Dr. Bryan conducted two replicate experiments one year apart. Both of the cyclamate treated replicate groups developed augmented incidences of bladder tumors when compared to the respective control groups. In both instances the statistical evaluation revealed a highly significant difference between treated and control groups (G–113 at 15–16).

I find that the Bryan study does support the conclusion that cyclamate

has not been shown to be safe. I recognize that the method used raises the possibility that the carcinogenic effect seen may be caused in part by the instillation technique or the cholesterol pellet or both. However, I disagree with the Temporary Committee's finding that the route of administration in the Bryan study was inappropriate because the animals which were exposed to the cyclamate incorporated in a cholesterol pellet and surgically implanted in the animals' bladder were compared to control animals exposed to the same type of surgically implanted cholesterol pellet. The only difference between the treatment and control groups is the exposure to cyclamate (G–112 at 20). The statistically significant difference between treatment and control thus shows that cyclamate is the sole or the primary cause of this tumor production. I do not find Abbott's attempts to explain these tumor findings, which are discussed below, persuasive. I agree with the ALJ that the Bryan study has shown "a major biological effect" which adds to the doubt concerning cyclamate's safety.

Although Dr. Bryan admitted that the technique he used is not replicative of human experience and that its use has declined substantially in the last several years, Dr. Bryan did testify that the reliability of the technique is supported by concordance of results, both positive and negative, between laboratories where careful studies have been conducted (G–113 at 19). Dr. Bryan further testified that the correlation between the results of studies in which the pellet implantation technique is employed and those utilizing direct feeding studies is "remarkably high" (Tr. at 828). Finally, Dr. Bryan testified that his technique has validity where large enough population samples and adequate controls are utilized (Tr. at 823).

It is also important to note that the bladder implantation technique is currently being used by three other groups, which suggests the continuing vitality of the technique (Tr. at 823). Moreover, the usage of different techniques is not due to a lack of confidence in the Bryan technique, but rather is largely due to the ability of different techniques to generate bladder tumors more quickly, in a higher yield, and with less expense than the Bryan technique (G–113 at 18).

It is possible that the surgical procedure used to implant the pellet, or the implanted pellet itself, acted synergistically with cyclamate to produce the tumors found in the cyclamate treated animals in the Bryan study. The Bureau argues that the presence of a foreign body in the urinary tract is a condition that occurs in human pathology (G–112 at 19). The Bureau further argues that surgical procedures on the bladder do occur in people, some of whom might be exposed to cyclamate both before and after surgery (Tr. at 818). Finally, the Bureau argues that stone formation can occur after a surgical procedure is performed on the bladder (*id.*). I do not find these arguments totally convincing. The tumors found in cyclamate treated animals may have been due in part to the unique circumstances of this test. Abbott's *unsupported* argument that these unique circumstances are responsible for the tumors found in cyclamate treated animals in the Bryan study is, however, insufficient by itself to rebut the suspicion raised by the Bryan study. I find that the Bryan technique is sufficiently analogous to human experience to require that valid and convincing negative direct feeding studies (using the conventional route of administration) or studies which prove that the carcinogenic response is unique to this technique and will not occur as a result of direct feeding of cyclamate, be submitted to rebut the suspicion raised by the Bryan study.

The Bureau also argues that the total duration of exposure of the bladder to cyclamate in the Bryan study is less than a day, whereas human exposure to cyclamate as a food additive, while involving considerably lower levels per day, would involve exposure for a much larger period of time (25,000 days in a lifetime) (Bureau's Brief at 39; G–112 at 18–19). The Bureau concludes that because of the longer exposure, the carcinogenic effect seen in the Bryan study "could be potentiated many, many fold in an exposure continued for thousands of days in a human population" (Bureau's Brief at 39). I am not persuaded by this argument. The animals in the Bryan study are exposed to a single brief (short elution time) but intense and highly localized exposure, to the unmetabolized agent, directly at a target site. Although human exposure to cyclamate ingestion would be long term, it would also be systemic exposure with relatively lower concentrations at any given tissue. Thus, because of the differences in the nature of the exposure, the longer term of the human exposure would not necessarily result in a greater carcinogenic effect in humans than was found in the animals in the Bryan experiment. It is, however, unnecessary for the Bureau to show that the effect in humans would be greater than that shown in animals, because the latter effect was statistically significant ($P<.05$).

(4) *Request for Rebuttal:* Abbott also contends that rebuttal testimony it attempted to submit on the Bryan technique was wrongfully excluded. I disagree. The purpose of rebuttal testimony is to allow the party with the burden of proof to adduce evidence on matters the relevance or existence of which were unknown or could not be reasonably foreseen at the time of the presentation of its case in chief. Abbott conceded that it "has long been familiar with the Bryan study" (Amended Motion for Leave to Adduce Rebuttal Testimony at 8). Abbott nevertheless attempted to justify its rebuttal testimony on the ground that it did not anticipate that the Bureau would rely heavily on the Bryan study (*id.* at 8). This reason is insufficient. If allowed, it could permit rebuttal testimony on almost any topic. The Bryan study is one of eight studies that the Bureau characterized as positive and does not appear to have been given any more reliance than the other seven studies. Abbott thus had ample opportunity to submit testimony challenging the Bryan study, and therefore was properly precluded from submitting rebuttal testimony on this issue.

Moreover, I agree with the ALJ that the proposed rebuttal testimony of Dr. Clayson (which would have argued that the Bryan technique was invalid and involved conditions which do not occur in humans) is "largely in the nature of argument and should best be presented as argument on brief" (Order of September 12, 1977 at 1). These arguments were presented by Abbott in their brief and are fully considered above.

Abbott also argues that the ALJ erred in refusing to allow it to submit a published scientific journal article by Jull, *et al.* (A–853). The procedural regulations governing the submission of such articles required Abbott to submit all documentary data and information upon which it sought to rely by June 15, 1977 (21 CFR 12.85(b)). The Jull article was published in 1975, but was not offered into evidence until November 3, 1977, some five months after the date for its submission and two months after cross-examination was completed. The procedural regulations allow a participant to supplement its submission under § 12.85 where "the material contained in the supplement was not reasonably known or available when the submission was made" 21 CFR 12.85(c). Abbott claimed that in spite of its efforts to locate this information, the article was unknown to it prior to

October 31, 1977, "primarily because the study was not published under its own title, but rather within the book whose title gives no specific indication of its existence." (Motion to Add A Document to the Administrative Evidentiary Record at 3). The ALJ ruled that this reason is insufficient. I agree. In view of the fact that this article was not offered until five months after it was due, two months after the completion of cross examination and one month after Abbott's request for a ruling on rebuttal testimony and that this article was available two years before the date for submission of such articles, I find that it was properly excluded from the evidentiary record.

Although the Jull article and the proposed testimony of Dr. Clayson were properly excluded, I have nevertheless decided to consider the main argument contained in these submissions. The Bryan technique is criticized on the ground that if the control animals are kept alive for their normal lifespan the tumor incidence is so high that valid conclusions concerning the carcinogenicity of a test compound cannot be drawn. This criticism lacks merit. Dr. Bryan testified that he had conducted an experiment (unrelated to his cyclamate experiment) in which a cholesterol pellet was left in the mouse bladder for 110 weeks (the normal lifespan of a mouse) and found that the incidence of bladder tumors under that circumstance is only about 12, 13, or 14 percent (Tr. at 803). Indeed, the Jull article recognized that the high tumor incidence Jull found in the strain of mice he subjected to the Bryan technique and kept alive for a normal lifespan could be unique to the strain of mice used in the Jull study and might be different for other strains (A–853 at 388–89). This strain variation could explain why Dr. Bryan found a 12, 13 or 14% tumor incidence in the mice he kept alive for 110 weeks. Thus, the Jull article does not rebut Dr. Bryan's findings.

b. *Hicks, et al. (A–832).* (1) *Study Design:* This study involved rats whose bladders were stripped of the epithelium by instillation of methylnitrosourea (MNU), a potent carcinogen, and then fed sodium cyclamate. A control group, given an MNU injection was also maintained.

(2) *Study Results:* When animals were exposed to MNU and fed a cyclamate containing diet for two years, a very high incidence of tumors developed in these animals (G–114 at 14). Dr. Hicks characterized this response as "very dramatic" (*id.*). In a subsequent experiment, utilizing a more potent dose of MNU, the MNU treated animals had a 20 percent tumor rate, but the MNU plus cyclamate group again produced a 50 percent incidence of bladder cancer (G–64). The results in both of these experiments were highly statistically significant ($P<.001$) (G–41, App. VII).

The Temporary Committee found that the "MNU-plus-cyclamate regime resulted in a clear carcinogenic response in the treated animals" (G–41 at 25). The Temporary Committee further found that the Hicks technique "may very well become an important screening method for substances suspected of being a urinary bladder carcinogen" but that "[i]t has not yet . . . been validated for this purpose" (*id.* at 45–46). The Hicks study was among the studies that the Temporary Committee found "create a sense of uncertainty" about the safety of cyclamate (*id.*).

The ALJ found that the Hicks study raised serious questions concerning the carcinogenicity of cyclamate (ID at 20, 32). The ALJ found further, however, that several factors raise questions as to the validity of the Hicks study: (1) the feed was not analyzed for pesticides; (2) no separate control group was anesthetized or instilled with an innocuous material; and (3) the lack of a formal randomization might have introduced additional bias (*id.* at 20).

(3) *Analysis:* Abbott takes exception to the ALJ's finding that the Hicks study raises serious questions concerning cyclamate's carcinogenicity. Abbott contends that (1) the technique used by Dr. Hicks is unlike any human circumstance and therefore suspect; (2) Dr. Hicks was uncertain as to the technique she actually used; (3) attempts to replicate her work have failed; (4) there were no anesthetized controls or controls instilled with innocuous materials; (5) no formal randomization was used; and (6) the animal facilities and environmental control were below the optimum standards. Abbott further contends that the ALJ erred in refusing to allow Dr. Deutsch Wenzel to appear and testify in rebuttal (Abbott's Exceptions at 22).

The Bureau takes exception to the ALJ's criticisms of the Hicks study, contending that (1) although Dr. Hicks did not use a table of randomization, she did use an appropriate system of randomization; (2) analyzing feed for pesticides is unnecessary for such a study and, in any event, both treated and control animals received in the same feed; and (3) there is no evidence that the anesthetization or instillation in treated animals was any different than that for controls.

(a) *Abbott's Exceptions.* The exceptions raised by Abbott and the Bureau are discussed below:

(i) *Relevance to Human Experience:* The theory underlying Dr. Hicks' technique is important to an understanding of its relevance to the human experience. Dr. Hicks utilizes MNU as part of her technique because there is a very good dose response relationship between MNU and the incidence of bladder cancer (G–114 at 8). A single intravesicular dosage of either 1.5 or 2. mg of MNU has been shown to be "noncarcinogenic" in the bladder (G–2 at 226; Tr. at 991). A second similar dosage (at either 1.5 or 2 mg) of MNU will cause tumors (G–2 at 226–27). The Hicks method involves substituting the test substance for a second dose of the known carcinogen MNU. The underlying theory of the Hicks method is that if the test substance does produce tumors, it is either initiating the tumor production and thus is a carcinogen or is promoting the effect by acting synergistically with the MNU (G–114 at 8). The Hicks methodology was explicity or implicitly endorsed by five leading oncologists or toxicologists (G–113 at 9–10; G–116 at 11; G–112 at 17; G–121 at 9; Tr. at 1173).

The only difference between the treatment and control groups in the Hicks studies was the feeding of cyclamate (G–112 at 20). Thus, it is reasonable to attribute the high tumor incidence found in the cyclamate treated animals to cyclamate. Although the precise mechanism by which this tumor incidence was caused is unknown, a convincing explanation has not been provided as to why these results, which are highly statistically significant, cannot be attributed solely to cyclamate. Moreover, the fact that a known carcinogen, MNU, would produce a similar increase in tumor production if substituted for cyclamate in the Hicks model supports the conclusion that cyclamate is producing a carcinogenic effect under the circumstances of this test model. I therefore find that it is reasonable to attribute the tumors produced in the Hicks study to cyclamate.

Abbott contends that the instillation of MNU in the bladder of the test animal is unlike any human experience and thus renders the Hicks model totally inappropriate. In support of this contention, Abbott notes that "the basis of the (Hicks) model is to initiate neoplastic changes with MNU" and that "in focal areas, the epithelium is stripped" (Abbott's Exceptions at 23). The lack of a completely analogous human experience does not, in itself, invalidate the Hicks model. It is equally plausible that the increase in tumor production is not due to the action of MNU, but rather is caused solely by

cyclamate. The Hicks model may thus be a valid technique for determining the possible carconogenicity of a test substance. In light of the "dramatic results" found using the Hicks model and the possible ultimate validation of the Hicks technique for the detection of carcinogens, the increase in tumor production found in this study raises considerable suspicion as to the safety of cyclamate.

It is possible that the presence of MNU in the bladder of the test animals plays a role in the tumor production found in the Hicks study. The Bureau argues that that circumstance is not totally unlike certain human experiences. The Bureau contends that it is reasonable to expect that, if ingested by humans, cyclamate will interact with carcinogens or suspect metabolites in the bladder (G–114 at 14–15). MNU is a nitrosamide which breaks down spontaneously to a carcinogen which is thought to be identical to a metabolite of dimethylnitrosamine (G–114 at 7; G–65). Dimethylnitrosamine in turn can be produced in the urine of people with bladder infections (G–114 at 7; G–60; G–61). The Bureau further contends that other carcinogens or suspect metabolites may also be found in the human bladder (G–114 at 14–15). I do not find these arguments totally convincing. The tumors found in cyclamate treated animals may have been due in part to the unique circumstances of this test. Abbott's *unsupported* argument that these unique circumstances are responsible for the tumors found in cyclamate treated animals is, however, insufficient by itself to rebut the suspicion raised by the Hicks study. The Hicks study is sufficiently analagous to human experience, when considered with the lack of a convincing explanation negating the strong results found by Dr. Hicks, to cause me to conclude that the results of the Hicks study cast doubt upon the safety of cyclamate. Valid and convincing negative direct feeding studies are required to rebut this doubt.

(ii) *Criticisms of the Hicks Technique:* Abbott makes two crtiticisms concerning the technique employed by Dr. Hicks. First, Abbott alleges that Dr. Hicks was uncertain as to how much MNU she used (Abbott's Brief at 29–30). In one report (G–2), Dr. Hicks refers to the usage of a 2 milligram (mg.) dosage of MNU. In a second report (G–64), use of a 1.5 mg. dosage of MNU is reported. A review of these reports and Dr. Hicks' testimony shows that the reference to a 2 mg. dosage of MNU in the first report appears to refer to a pilot experiment which preceded the two cyclamate studies that are the subject of this hearing (Tr. at 1040). The reference to a 1.5 mg. dosage of MNU in the second report appears to refer to the dosage of MNU used in both treated and control animals in the two cyclamate MNU studies that are the subject of the cyclamate hearing (Tr. at 1046, 1048). In any event, as Dr. Hicks explained, it is irrelevant whether a 1.5 or 2 mg. dosage of MNU was used in the cyclamate-MNU studies that are the subject of the cyclamate hearing, because both dosages represent a "noncarcinogenic dose". The results obtained from preliminary studies using the Hicks model and either of these dosages produced identical "noncarcinogenic" results (Tr. 991–92; 1036–37). Moreover, whenever Dr. Hicks employed this method a control group was utilized with the identical amount of MNU as the treated group (Tr. at 992). So long as the difference between treated and control animals is statistically significant, the increase can be attributed to cyclamate. Thus, Abbott's exception is without merit.

Abbott further contends that Dr. Hicks admitted that she did not follow the technique described in her publication (A–804 at 3). However, Abbott's contention is based upon a misinterpretation of a discussion Dr. Hicks had with a Dr. Moore in a round table discussion in Geneva. In that discussion, Dr. Hicks was not referring to the cyclamate experiments which are at issue in the hearing, but rather to a different experiment in which Dr. Hicks used a batch of MNU which caused tumors (Tr. at 990–91). Thus, Abbott has failed to establish that Dr. Hicks did not follow the technique described in her publication.

(iii) *Failure to Analyze Feed:* Abbott contends that Dr. Hicks' failure to analyze the feed for pesticides and other contaminates is a deficiency in the study (Abbott's Brief at 30). The ALJ agreed with Abbott and the Bureau took exception to the ALJ's finding. The Site Visit Report of the Temporary Committee did state that "no analysis of the feed was made for pesticides, mycotoxins, or other contaminants" (G–41, App. III, Hicks, et al. at 3). However, the Site Visit Report concluded that "the facilities, environmental controls, experimental design, and conduct of the study were thought to be adequate to warrant the consideration of the experimental results" (*id.* at 10). Furthermore, as was the case with the chow diet utilized in the calcium cyclamate portion of the Friedman study, the same feed was used in both the treated and the control animals and there have been no specific allegations that the feed may have contained a carcinogenic contaminant. Moreover, there was a statistically significant difference between the incidence of tumors in cylamate-treated animals and controls. As is the case with many factors which complicate carcinogenesis bioassays, the presence of a carcinogenic contaminant in the feed cannot negate a positive finding of carcinogenesis as long as both the control and treated animals consume the same feed. Thus, the failure to analyze feed does not invalidate the results of the Hicks study. It should be noted, however, that such contamination can compromise a negative result by causing such a high tumor incidence in the treated animals becomes statistically insignificant.

(iv) *Alleged Failure to Replicate Dr. Hicks' Work:* Abbott alleges that Dr. Hicks' model cannot be accepted until it has been replicated and that attempts to do so "have been unavailing" (Abbott's Brief at 30). Although it is true that Dr. Hicks' study has not been replicated, the attempts of Dr. Mohr to do so were incomplete at the time of the hearing (A–842 at 2–3). Thus, no final conclusions can be drawn from his work. Until such time as valid efforts to replicate Dr. Hicks' findings are unsuccessful, her work cannot be dismissed on this basis. Of course, if Dr. Hicks' work is replicated, greater confidence can be placed in her methodology.

Abbott also contends that Dr. Mohr found 2 mg. of MNU to be carcinogenic whereas Dr. Hicks found the same dosage to be "noncarcinogenic" (Abbott's Brief at 30). Abbott claims that this discrepancy "means that something went wrong with Dr. Hicks' [sic] MNU." This contention is incorrect. Dr. Mohr used a more active bath of MNU than Dr. Hicks used (Tr. at 990–91). As a result, even though Dr. Mohr used the same dosage of MNU as Dr. Hicks, his batch produced tumors whereas Dr. Hicks' batch did not (*id.*). In any event, as long as the treatment and control groups receive the same dosage of MNU from the same batch and there is a statistically significant difference between the two groups, the data are acceptable (*id;* Tr. at 587).

Additional support for the validity of the Hicks' technique is found in the work of other scientists who have employed methods analogous to those of Dr. Hicks and obtained favorable results. Dr. Gilbert Friedell, head of the American National Bladder Cancer Program, has employed Dr. Hicks' method with a dosage of nitrofuran, instead of MNU, as the initiating

carcinogen (Tr. at 989). Animals administered this dosage of nitrofuran were then fed a saccharin containing diet (*id.*). The incidence of tumors in the treated group was approximately 50% (*id.*). Dr. Hicks also testified that another scientist, Dr. Bryan, who appeared as a witness for the Bureau, conducted an experiment in which MNU was used to initiate a carcinogenic response in the epithelium (Tr. at 989). Using this method, Dr. Bryan demonstrated a synergistic effect between a tryptophan derivative and MNU (*id.*). (Abbott attacks this part of Dr. Hicks' testimony on the ground that Dr. Bryan never mentioned this experiment in his written or oral testimony (Abbott's Brief at 30, n.1). However, Dr. Bryan was never asked about this experiment, nor did Abbott seek to have this question posed to Dr. Bryan following Dr. Hicks' testimony.)

The Bureau also notes that Dr. Hicks achieved negative results with the known noncarcinogens coffee and cyclophosphamide (Bureau's Brief at 33). Although the experiments involving (1) nitrofuran (rather than MNU) and a tryptophan derivative and (2) MNU and coffee or cyclophosphamide (rather than cyclamate) cannot be considered true replications of Dr. Hicks' experiment, they lend some support to the validity of her method.

In sum, I find that there are no unsuccessful attempts to replicate Dr. Hick's experiment. In view of the fact that Dr. Hick's technique may yet be validated, and the fact that the technique has been successfully used with other carcinogens and noncarcinogens, I do not consider the lack of a successful replication a significant deficiency.

(v) *Alleged Lack of Formal Randomization:* Abbott contends that Dr. Hicks failed to randomize the animals in her study (Abbott's Exceptions at 23). Dr. Hicks explained, however, that although she did not use a table of randomization, she did use a system of randomization (Tr. at 1072). The Site Visit Report found that "It is unlikely that any biases were introduced by the lack of a formal randomization method being used to assign the animals to the experimental groups" (G–41, App. III, Hicks, et al. at 8). As noted above, the Temporary Committee concluded that the experimental design and conduct of the study were adequate (*id.* at 10). I agree with the conclusions of the Temporary Committee.

(vi) *Alleged Inconsistency in Tumor Findings:* Abbott questions the validity of Dr. Hicks' technique on the ground that no tumors were found in rats receiving 2 mg. of one preparation of MNU whereas tumors were found in 20% of the rats receiving 1.5 mg. of a different preparation of MNU. This difference in tumor production is attributable to the differing potency of different preparations of MNU (Tr. at 990–91). This alledged inconsistency thus does not render the experiment invalid, provided, as was the case in both of Dr. Hicks' experiments with cyclamate, that both treatment and control groups in each experiment receive the same dosage of MNU from the same preparation, and that there is statistically significant difference between the treatment and control groups.

In support of this argument, Abbott attempted to present, by way of rebuttal, the testimony of Dr. Deutsch-Wenzel. Dr. Deutsch-Wenzel would have testified that the "potency" of MNU does not vary from preparation to preparation. Although this testimony was properly excluded by the ALJ, I have nevertheless considered it. Even assuming, *arguendo,* that the potency of MNU does not vary, it is irrelevant to the validity of the Hick's experiment so long as both the treatment and control groups received the same preparation of MNU and it was handled the same way. If the procedure is followed, the difference in tumor incidence between treated and control groups can be attributed to cyclamate. Dr. Hicks repeatedly testified that the same MNU was given to treatment and control groups and it was handled the same way (Tr. at 991–93; 1020; 1030; 1070–72). Thus, the alleged constant potency of MNU is irrelevant.

(vii) *Failure to Start All Animals at the Same Time and the Alleged Uncertainty as to the Number of Animals Started at Various Times During the Test:* Abbott contends that these factors are significant in evaluating the validity of the Hicks' experimental procedure (Abbott's Exceptions at 23–24).

The animals in the Hicks' study were entered into the study over a period of months (Tr. at 1053). Whenever vacancies for storing the animals became available, a paired group of control and treated animals were added to the study (Tr. at 1052–53). Although Dr. Hicks could not recall the number of animals started at various times during the study, she testified that her records would reflect when each animal was entered into the study (Tr. at 1054–55). Many of these records were examined by the Temporary Committee's Site Visit Team (Tr. at 1055). The Site Visit team concluded that the conduct of the study was adequate (G–41, App. III, Middlesex Hospital Medical School at 10).

Abbott does not even suggest why the staggered starting times should invalidate the study. Since equal numbers of control and treated animals were started together, even if the MNU was unstable and broke down during the course of the experiment, it would not affect the validity of the study. Such a study would evaluate the carcinogenic response of cyclamate under varying potencies of MNU. The difference between treated and control animals would still be attributable to cyclamate. More importantly, there is no reason to suspect that the MNU used by Dr. Hicks did break down during the course of the experiment. Dr. Hicks testified that precautions were taken to ensure the stability of the MNU preparation used in the cyclamate experiments (Tr. at 992). When a bulk batch of MNU was received by Dr. Hicks, it was immediately weighed into small aliquots (*id.*). Each one was sealed in a glass bottle, wrapped in foil to keep out light and stored at minus 20 degrees (*id.*). This procedure maintained the stability of each batch (Tr. at 993). The procedure followed for dosing the treated and control animals was the same in every case (Tr. at 1071; 1073). Thus, there is no reason to believe that the staggered starting times of animals entered in the study would diminish the reliability of the study's results.

(viii) *Alleged Differences in Numbers of Animals Examined at Histology:* One report of Dr. Hicks' cyclamate experiment (G–2) states that 54 animals were treated with MNU and cyclamate whereas a second report of the study (A–832) states that 69 animals were treated with MNU and cyclamate. The discrepancy is attributable to the fact that the first report was a preliminary report, whereas the second report, which lists a larger number of animals, is the final report. The discrepancy appears to be due to the fact that all animals had not been sorted or had not yet been examined at the time of the preliminary report (Tr. at 1064).

(ix) *Slide Examination:* Abbott questions whether the examination of the slides to verify the existence of tumors was totally blind, *i.e.,* whether the investigator could determine that the slide came from a treated or control animal (Abbott's Exceptions at 23). Dr. Hicks explained that although the slides examined for purposes of histology were numbered, they were not numbered in the same way as the animals were (Tr. at 1066). There was no way of determining from the number on the slide whether the animal came from the

control or treated group (*id.*). Thus, for all practical purposes, Dr. Hicks' examination of slides was blind. Even if there is some question about the histology employed by Dr. Hicks, that histology was confirmed by the Site Visit Team of the Temporary Committee (G–41, App. III, Middlesex Hospital Medical School at 10). The Temporary Committee Site Visit Team concluded that the histopathologic examinations were satisfactory (*id.* at 10). Thus, Abbott's criticism is groundless.

(x) *Cyclophosphamide:* Dr. Hicks conducted an experiment in which treated animals received MNU plus cyclophosphamide and control animals received MNU. The cyclophosphamide MNU did not produce any tumors in the treated animals (G–2 at 225). Dr. Hicks cited this evidence as support for the validity of her technique on the ground that "there is no evidence to establish that cyclophosphamide is a bladder carcinogen in man" (Tr. at 1067) or rats (Tr. at 1068). Two other Bureau witnesses also confirmed Dr. Hicks' opinion that cyclophosphamide is not a bladder carcinogen in rats (Tr. at 555; 965–66). Abbott contends that Dr. Hicks is incorrect about the noncarcinogenicity of cyclophosphamide and the the ALJ improperly excluded the testimony of Dr. Schmaehl on this issue. For the reasons discussed immediately below in Subsection (b), I find that Dr. Schmaehl's testimony was properly excluded. I therefore find that Dr. Hicks' cyclophosphamide experiment was properly considered by the ALJ and does lend support to the validity of the Hicks' technique.

Moreover, even if Dr. Hicks' cyclophosphamide experiment is excluded from consideration because of the possibility of that substance's carcinogenicity, Dr. Hicks also performed a study using her technique with coffee and obtained negative results (Tr. at 553). Thus, even without considering the results of the cyclophosphamide experiment, there is a study providing a negative correlation for Dr. Hicks' technique and supporting Dr. Hicks' conclusion that tumors found in animals receiving MNU followed by cyclamate should not be attributed solely to MNU.

(b) *Abbott's Request for Rebuttal.* Abbott claims that it was prejudiced because it was prohibited from introducing rebuttal testimony from Drs. Schmaehl and Deutsch-Wenzel, whom Abbott asserts would have testified as to the alleged deficiencies in the Hicks' model. In denying Abbott's motion to adduce this rebuttal testimony, the ALJ correctly found that Abbott could and should present the matters it sought to introduce as rebuttal testimony as argument in its brief (Order dated September 12, 1977). Abbott did present each of these arguments in its exceptions (Abbott's Exceptions at 23) and they have been considered above. Thus, Abbott has not been prejudiced by the exclusion of its rebuttal testimony. Moreover, as discussed above, that rebuttal testimony was properly excluded.

(i) *Testimony of Dr. Deutsch-Wenzel.* Dr. Wenzel was to testify concerning the properties of MNU, namely that its potency does not vary from batch to batch and as to its volatility. This evidence is irrelevant. As discussed above, the "potency" and "volatility" of MNU are irrelevant so long as both treatment and control groups receive MNU from the same batch and that MNU is handled the same way.

(ii) *Testimony of Dr. Mohr:* Dr. Mohr's rebuttal testimony concerning his attempts to replicate Dr. Hicks' works and how his results allegedly conflicted with those of Dr. Hicks were admitted into evidence. The ALJ only excluded Dr. Mohr's general criticisms of the Hicks' model. This testimony was properly excluded because Abbott did not properly state its intention ahead of time to introduce this testimony in rebuttal (See Motion For Leave to Adduce Rebuttal Testimony; Docket No. 117, p. 5). Since these general criticisms went beyond the scope of Abbott's request to adduce rebuttal testimony, the ALJ properly excluded pp. 4–6 (beginning with ¶1, p. 4 through the last sentence on p. 6) of Dr. Mohr's testimony (Order, Docket No. 149, November 21, 1977). Moreover, even if Abbott had properly requested leave to adduce this rebuttal testimony, these matters could have and should have been introduced as part of Abbott's written direct testimony.

(iii) *Testimony of Dr. Schmaehl:* Many of the alleged deficiencies which Abbott sought to have Dr. Schmaehl testify about were known to Abbott prior to the cross-examination of Dr. Hicks. The procedural regulations governing the cyclamate hearing required the Bureau to file with FDA's Hearing Clerk, at the time of publication of the notice of hearing, all documentary data and information upon which it relied. 21 CFR 12.85(a)(3).

The majority of issues for which Abbott sought to introduce rebuttal were contained in two reports (G–2, G–64) which were filed with FDA's Hearing Clerk pursuant to 21 CFR 12.85(a)(3) prior to the hearing, or were otherwise known to Abbott. These issues are: (i) Abbott's contention that there was a discrepancy in the amount of MNU used by Dr. Hicks. This contention is based on two publications (G–2, G–64) that were available to Abbott prior to the hearing; (ii) Abbott's allegation that the Hicks' study lacked formal randomization. This issue was mentioned in the Temporary Committee's Report and therefore was known to Abbott prior to the hearing; (iii) the alleged discrepancy in the number of bladders each of the two reports states were examined (G–2, G–64). This information was also known to Abbott prior to its cross-examination of Dr. Hicks; and (iv) the fact that Dr. Hicks used her technique with cyclophosphamide and considered cyclophosphamide a noncarcinogen. This information is contained in a report of a study (G–2) which was filed with FDA's Hearing Clerk prior to the hearing. Abbott thus had ample opportunity to submit testimony on these issues prior to the hearing and has no grounds to claim that that testimony was proper rebuttal.

Abbott also sought to introduce rebuttal on statements by Dr. Hicks that slides were examined blind "to a large extent" (Tr. at 1066); that the "basis of the model is to initiate neoplastic changes with MNU" (G–114 at 8); that "in focal areas the epithelium is stripped" (Tr. at 1067) and that all animals in the test were not started at the same time (Tr. at 1051–52). These statements, standing alone, do not entitle Abbott to rebuttal. Abbott failed to provide the ALJ with any explanation, let alone a convincing explanation, of why these factors are so important as to require expending the addtional time and resources to hear additional rebuttal testimony. Abbott did not provide any information as to how these factors affect the validity of Dr. Hicks' technique. In the absence of such an offer of proof, Abbott has failed to provide sufficient justification for rebuttal. To the extent that these statements by Dr. Hicks, on their face, indicate that her technique is invalid, Abbott does not need rebuttal witnesses to restate the obvious. These matters have been considered above and do not raise any meritorious questions as to the validity of Dr. Hicks' technique.

(c) *Bureau's Exceptions.* The ALJ questioned the validity of the Hicks' study on the ground that no separate control groups were anesthetized or instilled with an innocuous material. The Bureau takes exception to this finding. The Bureau contends that there is no evidence in the record that anesthetization or instillation could

affect tumor production. The Bureau further contends that the use of a control group that was anesthetized and received MNU is sufficient because there is no evidence that the treated and control animals were handled any differently (Bureau's Exceptions at 3). Abbott contends that these variables "can disrupt the orderly interpretation of a study's results, and perhaps make a study's results impossible to interpret meaningfully" (Abbott's Reply at 4–5).

Abbott further contends that the manner in which the animals were anesthetized and subjected to MNU was never adequately explained by Dr. Hicks (*id.*). This latter point is different from that of the ALJ. The ALJ did not find that the method of anesthetization or instillation of the MNU was different for controls or treated animals. Indeed, Dr. Hicks testified that "the controls were given exactly the same treatment with MNU prepared in exactly the same way, from the same batch, at the same pH, in the same medium, very often from the same individual solution [as the treated group]" (Tr. at 1020). Thus, Abbott's latter point is without merit.

The basis for the ALJ's criticism is apparently the statement in the Temporary Committee's Site Visit Report that "[n]o anesthetized control groups or ones instilled with an innocuous material were established" (G–41, App. III, Middlesex Hospital Medical School at 3). The Site Visit Report concluded, however, that "[t]he facilities, environmental controls, experimental design, and conduct of the study were all thought to be adequate to warrant consideration of the experimental results" (*id* at 10). No expert testified that these factors cast doubt on the validity of the Hicks experiment. More importantly, as noted above, the procedure for instilling the MNU in treated and control animals was the same. Thus, the anesthetization and instillation of MNU were controlled variables and these factors do not cast doubt on the validity of the study.

c. *Oser, et al. (G–81).* (1) *Study Design:* This study involved Wistar-derived FDRL strain rats in groups of 35 males and 45 females fed a diet containing a mixture of ten parts sodium cyclamate to one part sodium saccharin. After 78 weeks on study the animals were subdivided and some were additionally treated with cyclohexylamine hydrochloride.

(2) *Study Results:* Bladder tumors were reported in 12 of the 80 rats given the high dose of the treatment mixture. The Temporary Committee found that "no conclusion can be made as to whether cyclamate was the causative agent, acted in concert with saccharin, or was noncontributory" (G–41 at 21). The Temporary Committee further found that "[o]f particular concern is the [Oser study] in which a statistically significant increase in bladder tumors occurred in animals treated with a mixture of cyclamate and saccharin" (G–41 at 48–49).

The ALJ found that "the study might be relevant to showing cyclamate's cocarcinogenic potential * * *" (ID at 18). The ALJ further found that "there are many confounding variables (presence of saccharin, CHA, bladder parasites and calculi and the design) that prohibit relying on the study to show the carcinogenicity of cyclamate" (*id.*).

3. *Analysis:* In its exceptions, Abbott contends that the uncontrolled variables in the Oser study make it inappropriate for assessing either the carcinogenicity or cocarcinogenicity of cyclamate. Specifically, Abbott contends that (1) the study was not intended to examine carcinogenicity; (2) the presence of saccharin and CHA are uncontrolled variables; (3) there was no data on possible trace impurities in either the saccharin, the cyclamate, or the cyclamate/saccharin mixture; (4) the presence of bladder calculi in many of the rats with tumors complicated any finding of a direct causal connection between the tumors and the test substance; (5) spontaneous tumors could have been caused by bladder parasites; (6) the pathologists differed in their diagnosis of the tumors, agreeing unanimously on only 4 of the 12 rats diagnosed; and (7) attempts by Schmaehl, Ikeda and Kroes to replicate the Oser study have been unsuccessful (Abbott's Exceptions at 19, 20–22).

The fact that the Oser study was not specifically designed to determine the carcinogenicity of cyclamate is irrelevant except insofar as specific aspects of the design or conduct of the study can be shown to make it more difficult or impossible to draw conclusions concerning the carcinogenicity of cyclamate from the study. Abbott alleges that it has found two such defects. First, Abbott contends that the presence of saccharin and CHA complicate any findings with respect to cyclamate. The Bureau contends that it is highly unlikely that effects seen are due solely to saccharin because (1) the ratio of cyclamate to saccharin in the mixture used in the Oser study is ten to one and (2) saccharin has not shown a carcinogenic effect at the feeding level utilized in the Oser study Bureau's Brief at 35; G–120 at 14). Although it may be probable that cyclamate was the sole or primary cause of the carcinogenic effects seen in the Oser study, the Bureau's arguments do not eliminate the possibility that saccharin or possibly CHA contributed to or was the sole cause of the carcinogenic effects seen in the Oser study. The fact that studies of saccharin have not shown an effect at the feeding level of saccharin utilized in the Oser study (1%) does not eliminate the possibility that saccharin played a role in the carcinogenic effects found. Thresholds for carcinogens have not been established (see Tr. at 1068–69). Moreover, it is unknown whether the species, strains and conditions of the saccharin studies on which the Bureau relies are the same as the species, strains and conditions of the Oser study. The possible effects of saccharin in the Oser study therefore cannot be eliminated. The same, of course, holds true for cyclamate, particularly in view of Hicks and Bryan studies and recurrent findings of bladder tumors found in the direct feeding studies (Friedman, Schmaehl, Homberger and Hicks (direct feeding) studies). All of these findings add credence to the possibility that cyclamate was the sole or primary cause of the production of bladder tumors in the Oser study. The design of the Oser study is certainly not ideal for determination of the carcinogenicity of cyclamate, but it does raise a suspicion as to cyclamate's safety and requires a close examination of the studies involving the direct feeding of cyclamate.

The suspicion raised by the Oser study could be rebutted by valid and convincing negative studies. I do not find persuasive, however, Abbott's arguments that studies by Schmaehl, Ikeda and Kroes represent such studies. First, as previously noted, at a minimum, to disprove results that are inconclusive but suggestive of a positive effect, the test substance must be tested in the same strain of the same species as used in the experiment with positive results. The Oser study involved Wistar derived FDRL rats whereas the Kroes study involved mice and the Schmaehl study involved Sprague-Dawley rats. Moreover, the Kroes study was positive for lymphosarcomas and the Schmaehl study cannot be considered negative because there was one tumor found in the cyclamate treated group in that study (see Section IV.B.2.a.(5)). Although the Ikeda study involved the same strain and species of rats as the Oser study, the study is inconclusive because of the low sensitivity of the study and the fact that histopathology had not been performed on all organs (see Section IV.B.2.a.(2)).

I also disagree with Abbott's argument concerning the alleged impurities in the cyclamate and saccharin mixture used in the Oser study. I find persuasive the Bureau's contention that there is no evidence that such impurities are present or what significance they might have (Bureau's Brief at 37). Dr. Oser states that "The possibility of the presence of impurities and their effect cannot be overlooked in light of later developments particularly with respect to commercially produced saccharin" (A–803 at 4–5). Dr. Oser does not state the identity of these suspected impurities or whether they would be expected to have a carcinogenic effect. Moreover, Dr. Oser does not explain his reference to "later developments" concerning "commercially produced saccharin," nor does Dr. Oser state whether or not the saccharin used by FDRL was "commercially produced." The cyclamate/saccharin mixture used in the study was supplied by Abbott (G–81 at 4). Surely, if there was some reason to suspect the presence of impurities in the cyclamate of saccharin, Abbott would be in a position to provide more specific information as to the nature of those impurities. I find that the information provided by Dr. Oser is not sufficiently specific to justify questioning the validity of the study.

The record does not support Abbott's contention that bladder parasites or bladder calculi were responsible for the tumors found in the treated animals. Only three of the tumors in the treated group were associated with calculi (G–41, App. VII, Oser et al. at 3; G–120 at 15). This factor supports the conclusion that calculi are not necessary for tumor development (G–120 at 15). In addition, a dose response relationship between the cyclamate/saccharin mixture and the incidence of tumors was found (G–114 at 28). This does response relationship would tend to negate the likelihood that calculi or parasites caused the tumors found in the Oser study because calculi or parasites would be expected to cause tumors in all groups with approximately the same frequency (*id.*).

Finally, I do not find persuasive Abbott's criticism of the Oser study on the ground that there was a lack of unanimity on all tumor diagnoses. The NCI Pathology Working Group confirmed the diagnosis of all twelve tumors (G–41, App. III), all of which were originally reported by FDRL. I therefore see no reason to question the diagnoses of tumors in the Oser study.

In sum, I find that the validity of the Oser study was comprised somewhat by the presence of saccharin and CHA. I find, however, that cyclamate is a probable cause of carcinogenic effects in the study and that therefore, the study does raise a suspicion as to the safety of cyclamate.

C. Negative Studies

1. *Gaunt, et al. (A–706).* a. *Study Design:* This study involved groups of SPF Wistar rats, with 48 of each sex in a group, fed diets containing either 600, 2,000 or 6,000 ppm of cyclohexylamine hydrochloride. The stuy was conducted for 104 weeks. All major organs including bladders were microscopically examined.

b. *Study Results:* No tumors were found. The authors of the study concluded that "[t]here was no indication of a carcinogenic effect at any of the levels of treatment" (A–611 at 2). The Temporary Committee found that "the test material did not display carcinogenicity" (G–41 at 17).

c. *Analysis:* Abbott and the Bureau agree that the resuts of the study were negative in terms of carcinogenicity (Abbott's Brief at 19; Bureau's Brief at 19).

I find that under the conditions of this test, cyclohexylamine did not display any carcinogenicity. I note, however, that although studies of the cyclohexylamine metabolite of cyclamate are relevant to the safety of cyclamate, such studies, standing alone, are insufficient to establish the safety of other metabolites or of cyclamate itself. Moreover, the stuy is inadequate to rebut questions raised by other studies in other species or strains of animals (see Section IV.B.1.a.). (I also note that, I agree with the Bureau that pulmonary effects and testicular atrophy were strongly associated with dose level in this study.)

2. *Carson, et al. (G–4).* a. *Study Design:* This study involved five groups of 30 weanling FDRL Wistar rats of each sex 0, 15, 50, 100 or 150 mg./kg. cyclohexylamine per day. The study was conducted for 113 weeks, after which the survivors were sacrificed. Most of the bladders were examined microscopically.

b. *Study Results:* No tumors were found. The authors of the study stated that "[e] xamination of multiple sections (about 16 to 20) of the urinary bladder from each rat revealed no evidence of tumorigenesis * * * It is of particular interest to note the absence of any bladder carcinoma despite the intensive examinations that were carried out" (A–274 at 28).

The ALJ noted that only 32% of the animals survived the two years" of the study (ID at 16). The ALJ also stated that "[o] ccasionally the incidence of tumors is related to the presence of bladder parasites" and that although "parasites were found in every animal * * * no bladder tumors were found" (ID at 16).

c. *Analysis:* Abbott contends that the "32% survival figure" is incorrect; that it is nowhere mentioned in the published study and that it cannot be drawn from the tables on the report of the study (Abbott's Exceptions at 30). I agree with Abbott that the 32% survival figure is incorrect. The correct survival rate is, however, nevertheless unusually low. The authors of the Carson study made the following comment on the survival rates of the animals tested:

> Survival ranged from 80–90% up to the 78th week in all groups except the control females where it was 71%. Toward the end of the second year, mortality increased, the terminal survival rates for all test groups averaging 45.1 and 55.9% for the males and females, respectively (with no grading related to dose level), compared with 46% survival for each sex in the controls.

(G–4 at 50.) Thus, although the ALJ's "32% survival figure" is incorrect, the ALJ is nevertheless correct in that a large percentage, approximately 50%, of the animals in the study did not survive the first two years of the study. These average 45.1 to 55.9% survival figures represent a reduction in the sensitivity of the study because the animals which died prior to the termination of the study might have developed tumors had they survived. (Only the animals that survived were evaluated.) Even if all the animals had survived and had been evaluated, the study had only a 50% chance of detecting, at the 95% confidence level, a true difference in tumor incidence of approximately 9% between the controls and the high dose (150 mg/kg) treated animals (G–41 App. V, at 20). I find that this low sensitivity significantly reduces the confidence that can be placed on the results of this study. Moreover, although studies of the cyclohexylamine metabolite of cyclamate are relevant to the safety of cyclamate, as noted above, such studies, standing alone, are insufficient to establish the safety of cyclamate.

With respect to the presence of bladder parasites, Abbott contends that "if a study is to have relevance on whether parasites cause bladder tumors, the length of exposure to parasites must be known" (Abbott's Exceptions at 31). Abbott concludes that the Carson study is "meaningless on this issue" because the length of exposure to parasites is unknown. I agree with Abbott's exception. It is important to note, however, that the length of exposure to parasites is also a factor to consider in those studies where Abbott argues that the occurrence of tumors may be due to

bladder parasites or bladder calculi rather than the test substance.

D. Deficient Studies

The following studies do not contain results from which responsible conclusions as to the safety of cyclamate can be drawn because of deficiencies in the design or conduct of the studies.

1. *Altoff study (A–691).* a. *Study Design:* This study involved levels of .156, .312, .625 and 1.25% of sodium cyclamate or calcium cyclamate given in drinking water to groups of 30 male and 30 female hamsters. The study was continued for the lifetime of the hamsters.

b. *Study Results:* The authors of the study stated that "(t)he present experiment in Syrian golden hamsters adds to the volume of negative evidence on carcinogenicity of saccharin and cyclamate" (A–691 at 23).

The ALJ found that the study was negative (ID at 31), but noted that "(t)he sensitivity of the study was limited by the small group size and the poor survival rate. Less than 15% of the animals were alive after 74 weeks" (ID at 14).

c. *Analysis:* Abbott contends that the study is negative, relying on the conclusion of the authors of the study (Abbott's Brief at 18). In its exceptions, Abbott contends that the NCI Site Visit group found that the "small initial group size limited *somewhat* the sensitivity of the study A–647 App. III at 4", and that therefore the ALJ distorted the evidence and is prejudiced because he found that the sensitivity of the study was "limited" (Abbott's Exceptions at 27–28).

The Bureau notes that the Temporary Committee found that "(t)he small number of effective animals, resulting from high early mortality of the treated hamsters, reduced the sensitivity of the study" (Bureau's Brief at 14). The Bureau concludes that the study is inconclusive (*id.* at 19).

I disagree with Abbott's exception. The terms "limited" and "somewhat limited" are roughly synonomous. In any event, the record supports the ALJ's statement. This Altoff study had only a 50% chance of detecting, at the 95% confidence level, a true difference in tumor incidence of approximately 38% between the controls and the high dose treated animals (G–41, App. V, Table IV at 19). This study is therefore too insensitive to be considered proof of safety.

2. *Altoff et al. ("Second Altoff study") (G–41 at 19).* a. *Study Design:* The ALJ described this study as follows:

> * * * Syrian golden hamsters were given 1.5% sodium or calcium cyclamate in their drinking water as follows: to seven females for four weeks before mating, to five females between the time of mating and delivery, to five females after mating and continued for 25 days after delivery, and to four females for four weeks before mating and and continued until delivery.
> The study was continued for the lifetime of the F–1 generation, consisting of 13–35 hamsters per group. The study had not been completed at the time the NCI Temporary Committee made their report.

b. *Study Results:* After 80 weeks no tumors were found in the hamsters that had died (ID at 14–15). The Temporary Committee found that the study "has limited value in that none of the animals were continued on treatment for their lifetime" (G–41 at 19). The ALJ found that this study "has limited value with respect to carcinogenicity because of the low-dose level and that no F–1 group was treated beyond 25 days following birth" (ID at 15).

c. *Analysis:* Abbott argues with the ALJ's assessment that the second Altoff study was incomplete at the time the Temporary Committee wrote its report (Abbott's Exceptions at 28). Abbott also agrees with the Temporary Committee's finding that the second Altoff study has "limited value" (*id.* at 28–29). Abbott further states that, contrary to a statement in the Initial Decision (p. 15), Abbott does not contend that the second Altoff study reinforces other negative findings (*id.*). The Bureau states that the study is "worthless" because it was not a lifetime feeding study (Bureau's Brief at 15).

Abbott takes exception to the ALJ's statement that the "NCI Temporary Committee Site Visitors suggested that hamsters were not sensitive enough to detect weak carcinogens" (ID at 15). Abbott contends that this statement is not synonomous with that of the Temporary Committee and therefore prejudicial to Abbott (Abbott's Exceptions at 28). This contention is groundless. The Site Visitors stated that:

> The hamster has proven to be a good animal model to demonstrate the carcinogenicity of some bladder carcinogens. When beta-napthylamine (a relatively strong carcinogen) was initially tested in the hamster, no carcinogenic response was elicited, although the dose was rather large (0.1%). It was found that even a higher dose (1.0%) was needed to produce a carcinogenic response. *In view of these results, it must be questioned whether the hamster also is a good animal model for detecting relatively weak bladder carcinogens.*
>
> *The questionable sensitivity of the hamster to detect relatively weak bladder carcinogens, the rather poor sensitivity of the studies referring to both Altoff studies) and the low dose levels tested must all be considered in determining the value of the results obtained.*

(G–41, App. III, Site Visit Report for the Eppley Institute for Research in Cancer and Allied Sciences at 4–5; emphasis added).

I find that the above quoted language from the Site Visitor's Report fully supports the ALJ's statement. Moreover, I find that the questionable sensitivity of the hamster to detect relatively weak bladder carcinogens should be considered and further limits the value of both Altoff studies.

d. *Other matters:* As part of the Remand Order in this proceeding, the parties were asked to submit any data pertaining to this study. That was done on September 17, 1979. A review of the data did not, however, reveal any significant effects.

3. *Colston, et al. (A–207).* a. *Study Design:* This study involved a group of rhesus monkeys fed an oral dose of sodium cyclamate, 6 days per week since January 1968. As of June 1975, only three monkeys were still being studied.

b. *Study Results:* The Temporary Committee found that "(t)he value of this study to assess the carcinogenicity of cyclamate is severly limited as a result of the small number of animals used, the low dose level tested, and the relatively short portion of the monkeys' life span studied" (G–4 at 16).

c. *Analysis:* Abbott designates the Colston study as being negative (Abbott's Brief at 18). The Bureau agrees with the findings of the Temporary Committee (Bureau's Brief at 13; G–126 at 9–10). For the reasons stated by the Temporary Committee, I find that this study does not contain results from which responsible conclusions as to the safety of cyclamate can be drawn.

4. *Fitzhugh (A–192).* a. *Study Design:* This study involved Osborne-Mendel rats, in groups of ten males and ten females, fed 0, 0.01%, 0.1%, 0.5%, 1.0% or 5.0% sodium cyclamate. The study was conducted for 24 months. The bladders were not microscopically examined, except for those animals in the high dosage group.

b. *Study Results:* The Temporary Committee considered this study "deficient is that the bladders were examined only microscopically" (G–41 at 20).

The ALJ found that the failure to microscopically examine all bladders "seriously questions the validity of the negative results" (ID at 12).

c. *Analysis:* Abbott did not take exception to the ALJ's finding. The Bureau contends that the Fitzhugh study is deficient (Bureau's Brief at 15, 19).

The lack of microscopic examination of bladders in the Fitzhugh study raises the possibility that tumors may have been overlooked. Therefore, I cannot

draw any valid conclusions concerning the safety of cyclamate from this study.

It should be noted that my decision to place no reliance on the results of this study is not inconsistent with my finding that the Rudali study, in which no histopathology was performed, is inconclusive but suggestive of a positve effect. In the Rudali study, tumors were visible macroscopically. As I explained in Section IV.B.1.a., even if microscopic tumors were present in control animals in the Rudali study, the large lung and liver lesions visible in cyclamate treated animals without histopathology are an indication of a more rapid time of onset of the tumors and thus are supportive of a finding of carcinogenicity. In contrast, in the Fitzhugh study, tumors may have been present in cyclamate treated animals but overlooked due to lack of histopathology. Unlike the Rudali study, where the presency of microscopic tumors in control animals would not substantially alter the interpretation of the results, a finding of microscopic tumors in cyclamate treated animals and none in controls could result in a finding that the Fitzhugh study was positive. This possibility prevent me from concluding that the Fitzhugh study is negative.

d. *Other Matters:* As part of the Remand Order, the parties were asked to submit the data for the Fitzhugh study. A review of that data did not reveal any significant effects and confirmed that all bladders were not microscopically examined.

5. *Schmaehl et al. (A–386A).* a. *Study Design:* This study involved groups of Sprague-Dawley rats fed butylbutanol-nitrosamine (BBN) or BBN plus sodium cyclamate.

b. *Study Results:* There was a 100% incidence of bladder tumors in both treatment groups (G–41 at 23). The Temporary Committee concluded that the study "has limited value in providing evidence for or against cyclamate's potential carcinogenicity" (*id.*).

c. *Analysis:* Because of the 100% incidence of tumors in both groups, I find that the study is deficient.

6. *Bar (A–131).* a. *Study Design:* The ALJ states that "[t]his study employed rats that were laboratory-bred and were fed these [sic] doses of sodium cyclamate: 150 mg/kg/day, 300 mg/kg/day and 450 mg/kg/day in groups of males and females ranging from 30–55" (ID at 12).

b. *Study Results:* The Temporary Committee stated that "[v]ery few details on this study were available to the Committee. Thus, an evaluation of it cannot be made until after it is completed and details of its design, conduct and results are known" (G–41 at 19).

c. *Analysis:* In view of the lack of information on this study, it is of no use in determining the carcinogenicity of cyclamate.

d. *Other Matters:* As part of the Remand Order, the parties were asked to provide a report of this study. The only information provided was a review article which discussed the early studies on artificial sweeteners.

7. *Adamson (G–41 at 25).* a. *Study Design:* This study involved twenty treated monkeys (three groups of 5 of each sex) fed 0, 100 or 500 mg/kg/day of sodium cyclamate. The study had been ongoing for five years at the time of the Temporary Committee Report.

b. *Study Results:* The Temporary Committee found that "no conclusions regarding cyclamate's potential carcinogenicity in monkeys can be made until either a response is detected or the study is terminated" (G–41 at 25). The ALJ reached the same conclusion (ID at 15).

c. *Analysis:* Abbott did not take exception to this finding. I agree with the Temporary Committee's findings and conclude that valid conclusions as to the safety of cyclamate cannot be drawn from this study.

8. *Industrial Biotest (A–394–400).* a. *Study Design:* This study involved Beagle dogs in groups of each sex given cyclohexylamine sulphate (CHA–S) as a 25% mixture with lactose for the first 193 weeks, and as a 50% mixture with lactose for weeks 194–400. The first group was fed a concentration of 0.15 mg CHA–S/kg//week which was increased to 50 mg/kg/wk at the 194th week. For the second group, the concentrations were 1.5 mg/kg/wk and 100 mg/kg/wk, and for the third group 15 mg/kg/wk.

b. *Study Results:* The study is incomplete in that seven of the original sixteen animals were still being tested at the time of the Temporary Committee Report. The Temporary Committee found that "[u]nless a statistically significant carcinogenic response is demonstrated, this study has limited value in assessing the carcinogenicity of cyclohexylamine due to the small number of test animals" (G–41 at 24).

c. *Analysis:* Abbott agrees that the Industrial Biotest study was incomplete at the time the NCI Committee wrote its report. I find that this study does not contribute to the evaluation of the safety of cyclamate because the study is not completely reported and the number of animals involved is too small for the study to be of any value.

9. *Roe, et al. (A–286).* a. *Study Design:* This study involved Swiss albino mice in groups of 50 females each. One group was exposed to 5% sodium cyclamate pretreated with polyethylene glycol. The other group received 5% sodium cyclamate pretreated with polyethylene glycol and benzo(a)pyrene. The mice were not randomly allocated in that the older mice were placed in the control group.

b. *Study Results:* The Temporary Committee considered this study deficient because bladders were not examined microscopically (G–41 at 18).

c. *Analysis:* Abbott initially listed the Roe study as negative, but did not discuss it in its brief to the ALJ (Abbott's Brief at 18).

The Bureau listed the Roe study as "inconclusive" (Bureau's Brief at 10). One Bureau witness, Dr. Cranmer, stated that the lack of microscopic examination of animal bladders was a "major flaw" in the Roe study (G–126 at 10).

In the Remand Order, the parties were asked to comment on the maldistribution of weight or age in the various groups in the study. In response, the parties stipulated that "since complete histologic examination was not conducted, this study does not contribute to an assessment of the carcinogenic potential, if any, of cyclamate" (Remand Proceedings Stipulation of the Parties at 4). I agree with the parties, and, consequently, have not considered this study in my assessment of the safety of cyclamate.

E. Other Evidence

In addition to the animal studies discussed above, the record contains studies performed in test tubes (*in vitro*) and retrospective studies on the use of cyclamates by humans (epidemiological studies). The ALJ found that the *in vitro* tests "represent no more than a predictive tool, and in light of the results of the animal feeding studies, cannot be considered as determinative of the issue of carcinogenicity" (ID. at 33). As to the epidemiological studies, the ALJ found that the studies inquired about artificial sweetners and did not distinguish between saccharin and cyclamate *id.*). The ALJ further found that the sensivity of these studies was severely limited because at the time the studies were conducted cyclamate had only been on the market for five years and subjects were generally questioned within five years. Because the proposed use of cyclamate would result in lifetime exposure and because carcinogenic effects often do not manifest themselves for periods of 25 to 30 years after exposure, the ALJ concluded that "no conclusion concerning cyclamate's safety can be drawn on the basis of

these studies" (ID at 33). Abbott and the Bureau did not take exception to the ALJ's conclusion with respect to the *in vitro* and epidemiological studies. I concur with the conclusions of the ALJ with respect to these studies.

V. The Mutagenicity Issue

A. Introduction

1. *Issue Presented:* The second issue presented is: Whether the evidentiary record establishes to a reasonable certainty that cyclamate does not cause genetic damage and is not mutagenic. (ID at 4). In layman's terms, the question is whether cyclamate causes changes in the genetic code which could lead to an abnormal individual in future generations (G–124 at 7; see also A–800 at 1–3). One expert defined mutagenicity most succinctly as the induction of "heritable genetic damage" (G–121 at 11).

The basic genetic material in man and animals is called deoxyribonucleic acid (DNA). The DNA, which forms the genetic code, is distributed among the many "genes" that determine traits to be inherited. These genes are grouped in packages called "chromosomes." Chromosomes are physically much larger than genes and can be seen under a microscope. Humans have 46 chromosomes (23 pairs) each of which contains numerous genes (A–800 at 1–2).

Either chromosomes or genes can be harmed in such a way as to create an abnormal individual in future generations (G–123 at 3; G–124 at 14; A–800 at 1). The question before me in this proceeding is whether Abbott has produced evidence which proves that there is a reasonable certainty that cyclamate does not cause the type of genetic damage, either to chromosomes or to genes, which may lead to an abnormal individual in future generations.

From a medical standpoint, mutagenicity is an extremely significant issue. Dr. Legator, the Bureau's chief mutagenicity expert, explained as follows:

> There are a variety of genetic diseases—Down's syndrome, which is a product of chromosomal abnormalities, various neurological diseases, mental retardation, a host of inborn errors of metabolism. Genetic abnormalities in our population are probably the most significant health burden we now face. Indeed, it has been estimated that 25 percent of our overall health problems are due to genetic or genetically related diseases.

(G–124 at 7; see also G–122 at 7 and G–121 at 11). Other Government agencies, such as the Environmental Protection Agency, have also recognized the medical significance of mutagenicity. As Dr. Epstein, another Bureau witness, stated:

> EPA takes the position that mutagenesis is an extremely critical public health risk, because its effects may extend to a large number of generations to come.

(G–121 at 11). Genetic abnormalities are further significant in that they often affect an individual from the moment of birth onward, rather than merely during later life, as do many other diseases. Thus, great caution should be exercised in determining the mutagenic potential of cyclamate.

2. *Conclusion:* For the reasons stated below, I find that Abbott has not shown that there is a reasonable certainty that cyclamate does not cause heritable genetic damage.

3. *Summary of Evidence:* As discussed in detail in Section II. above, the statutory scheme governing the evaluation of a food additive petition provides that the petition shall be denied if a fair evaluation of the data fails to establish that the food additive will be safe under the specified conditions of use. 21 U.S.C. 348(c)(3)(A) and 21 CFR 170.3(i). In order for this determination to be made, the parties have introduced into the record 72 scientific studies designed to test the potential mutagenicity of cyclamate and its metabolites. Of these, 49 studies were performed on live animals or human beings (called *in vivo*), and the remaining 23 studies were performed in test tubes using plant, animal or human cells (called *in vitro*). The parties agree that the dispositive information must come from *in vivo* studies because only in these can the test compound be examined under conditions most closely approximating actual human use (Abbott's Brief at 44; Bureau's Brief at 73–74; G–124 at 30).

I find that the results from one group of *in vivo* studies, called cytogenetic studies, raise a serious question as to the potential mutagenicity of cyclamate. An *in vivo* "cytogenetic" study, as will be explained in greater detail in Subsection F.3. below, is designed to measure a test compound's effect upon chromosomes. Six *in vivo* cytogenetic studies each found a statistically significant increase in chromosome aberrations. These findings were obtained in five different species: Holtzman rats (G–9, both portions), Mongolian gerbils (G–26), fetal lambs (G–44), Chinese hamsters (G–45) and human beings (J–1); and in three different types of cells: bone marrow (G–9 and G–26), blood (G–44, G–45, and J–1), and spermatogonia (G–9). (See discussion of individual studies in Subsection F.4. below.)

The repeated nature of these findings across such a variety of species, cells, and laboratories greatly enhances their credibility (see Subsection D below). Although these chromosome aberrations were predominantly "breaks" which do not themselves directly cause inherited abnormalities, these findings are nevertheless biologically significant for two reasons: (a) because breaks may lead to another type of chromosome aberration, called "exchange figures," which do cause heritable genetic diseases; and (b) because breaks may be indicators of gene mutations which may also cause inherited abnormalities (see Subsection F.2.c. below). I would therefore categorize these six studies as being "inconclusive but suggestive" of mutagenicity. These findings, by themselves, require the denial of Abbott's food additive petition. *Certified Color Manufacturer's Ass'n* v. *Mathews*, 543 F.2d 284, 297 (D.C. Cir. 1976); *accord, Environmental Defense Fund* v. *E.P.A.*, 598 F.2d 62, 89 (D.C. Cir. 1978); *Ethyl Corp.* v. *E.P.A., 541 F.2d 1, 37–38 (D.C. Cir.) (en banc), cert. denied*, 426 U.S. 941 (1976); *see Hercules, Inc.* v. *E.P.A.*, 598 F.2d 91, 110 (D.C. Cir. 1978). Moreover, a number of cytogenetic studies performed *in vitro* provide additional support for this conclusion (see Subsection F.7. below).

The valid "negative" *in vivo* cytogenetic studies are insubstantial by comparison. Although four such valid studies (A–143, A–151, A–716 and A–811, App. 19) found no statistically significant increase in chromosome aberrations, each of these is readily distinguishable from the suggestive findings just described. For example, one negative study (A–151) used an entirely different animal species (mice). The remaining three studies, although using the same species as one suggestive study (Chinese hamster), analyzed a different type of cell: the negative studies using bone marrow (A–143) or spermatogonial cells (A–716 and A–811, App. 19) versus blood cells (G–45) for the suggestive study. (See discussion of individual studies in Subsection F.5. below). Thus, none of the negative studies directly rebuts any of the suggestive evidence.

I have eliminated from consideration as being "deficient" 15 additional *in vivo* cytogenetic studies because they do not meet the minimum criteria set forth in Subsections C.2 and 3. below in terms of statistical or biological significance. (See discussion of individual studies in Subsection F.6 below).

Mutagenicity studies on cyclamate were also performed in three other types

of *in vivo* studies: host-mediated assay; dominant lethal assay; and drosophila. Although the findings in each of these groups were predominantly negative, they do not outweigh the suggestive cytogenetic findings just described because known mutagens have been found to show mutagenic effects in some test methods but not in others (G–124 at 9–10 and 31; Tr. at 933–34; Tr. at 498–501; Tr. at 717–18 and 734). (See discussion of studies in Subsection G. below).

Finally, the record contains some additional *in vitro* findings. These types of studies, however, are never sufficient to outweigh suggestive *in vivo* findings because *in vitro* studies, being performed in test tubes, cannot take into account a live animal's metabolism (Tr. at 937–38; see also G–124 at 10; G–121 at 12). (See discussion of studies in Subsection G.5. below).

In sum, the repeated *in vivo* cytogenetic findings of breaks raise a serious question about cyclamate's mutagenic potential—specifically, its capacity to induce exchange figures and gene mutations, both of which are capable of producing genetic abnormalities in future generations. Indeed, as Dr. Legator concluded, "Any compound which shows the [mutagenic] effects cyclamate has shown should be considered a high risk agent" (G–124 at 26). Given this strongly suggestive mutagenicity evidence, the statutory scheme mandates that Abbott's food additive petition be denied on this additional ground.

B. The Statutory Scheme

I have already discussed the legislative history as well as the judicial and administrative interpretations of the general safety clause (see Section II. above), and I adopt that discussion here. Nevertheless, one point worth repeating here is that the general safety clause requires disapproval of a food additive petition if the evidence "suggests" lack of safety, even if that evidence is inconclusive. For example, in *Certified Color Manufacturer's Ass'n* v. *Mathews, supra,* which presented an analogous situation involving the act's Color Additive Amendments of 1960, the court stated:

> The information available to [the Commissioner of Food and Drugs] indicated a statistically significant relationship between high dosages of Red No. 2 and the occurrence of cancer in aged female rats. That relationship concededly did not establish conclusive proof that Red No. 2 was a carcinogen, but it was at least suggestive of it, . . .

Id. at 297. *Accord, Environmental Defense Fund v. E.P.A., supra,* 598 F.2d at 89; *Ethyl Corp. v. E.P.A., supra,* 541 F.2d at 37–38; *see Hercules, Inc. v. E.P.A., supra,* 598 F.2d at 110. Moreover, the effect of "inconclusive but suggestive" evidence on an agency's safety analysis applies with equal force to human health risks other than cancer. *Ethyl Corp. v. E.P.A., supra* (lead poisoning). Thus, Abbott's food additive petition must be denied if a fair evaluation of the evidence suggests that cyclamate may cause heritable genetic damage. For the reasons stated below, that is precisely the situation here.

C. Criteria for the Evaluation of Individual Mutagenicity Studies.

I have adopted several minimum criteria, involving both statistical and biological significance, necessary for a study to be considered valid. Based largely upon these criteria, mutagenicity studies may be clasified into four categories: (a) positive; (b) suggestive of a mutagenic effect; (c) negative; and (d) deficient. I first will define these terms and then set forth the minimum criteria.

1. *Classification of Mutagenicity Studies.* I have adopted the same classification terminology for the mutagenicity studies as I used for the carcinogenicity studies. Very briefly, these terms are defined as follows:

a. *Positive:* A "positive" study is one which conclusively demonstrates that cyclamate causes heritable genetic damage. There are no such studies on this record.

b. *Suggestive of a Mutagenic Effect:* A "suggestive" study is one which, although inconclusive, suggests that cyclamate may cause heritable genetic damage. The principal examples on this record are the *in vivo* cytogenetic studies which found a statistically significant increase in chromosome aberrations, predominantly breaks. These studies are suggestive rather than positive because breaks themselves are not inherited. Rather, as explained in Subsection F.2.c. below, breaks are biologically significant because they may: (a) lead to exchange figures; and/or (b) be indicators of gene mutations, both of which are capable of inducing heritable genetic abnormalities. (Note, however, that *findings* of breaks at $P<.05$ are termed "positive findings" even though those *studies* are termed "suggestive.")

c. *Negative:* A "negative" study is one where: (1) no statistically significant ($P<.05$) increase in genetic damage is found; and (2) the minimum criteria for biological significance set forth below are met. Although negative studies satisfying this definition are considered to be valid, they may still be considered inconclusive and entitled to differing weights depending upon the nature and extent of any internal flaws.

d. *Deficient:* A "deficient" study is one which does not meet the minimum criteria for either statistical or biological significance set forth below. Deficient studies are entitled to no weight at all.

2. *Statistical Significance.* In contrast to the sharp debate on statistical significance sparked by the carcinogenicity data, the parties are in general agreement as to the statistical significance of the mutagenicity studies. This is because the studies predominantly reported findings at the .05 confidence level; this was uniformly true, in fact, among the pivotal group of suggestive *in vivo* cytogenetic studies. Thus, the only issue relating to statistical significance of the mutagenicity studies is whether a statistical analysis had been performed on a given study. I believe that the performance of a statistical analysis is a minimum requirement necessary to demonstrate the validity of a study's results. Those studies which fail to give this critical information (and where the parties have not themselves performed a statistical analysis using reported data) have been eliminated from consideration as being "deficient" (see *e.g.,* A–217).

3. *Biological Significance:* I have also employed three minimum criteria necessary to establish the biological significance of a study. These involve: a) study size; b) reporting of data; and c) positive controls.

a. *Study size.* For a study to be considered valid, it must employ a sufficient number of animals to give the study an adequate degree of sensitivity. Dr. Legator testified, for example, that with respect to the *in vivo* cytogenetic studies, at least ten animals should be used per treatment group (G–124 at 18). This figure was based not only on Dr. Legator's own experience, but also upon the minimum protocol recommended by the Ad Hoc Committee of the Environmental Mutagenic Society (*id.*). Abbott produced no expert testimony to the contrary. I have therefore adopted the figure of ten animals per treatment group as a general guideline in determining the adequacy of an *in vivo* citogenetic study's sensitivity. (In so stating, however, I note that all studies that have been eliminated from consideration for this reason employed six or less animals per treatment group).

The parties did not elicit specific expert testimony regarding minimum study size for other types of *in vivo* or *in vitro* studies. Where questions have arisen as to the sufficiency of the population size of a particular study in one of these other categories, I have

resolved them on a study-by-study bais by considering the expert testimony on that particular study.

This criterion of minimum study size has a different effect on "negative" studies than it does on "suggestive" studies. The concept of adequate sensitivity means that the study must have been large enough (i.e., sensitive enough) so that, if a test compound is mutagenic, there is sufficient likelihood that the mutagenic effect will be detected. Thus, if an *in vivo* cytogenetic study using only three or four animals per group produced "negative" findings, no confidence can be placed in those results. For this reason, I have rejected as being "deficient" these so-called "negative" studies.

In contrast, if a study with only a few animals produces statistically significant results, it cannot be criticized for being too insensitive. Quite the contrary, what such a result suggests is that the test compound is sufficiently potent that it is capable of being detected by even an insensitive study (Tr. at 941). I therefore consider results from studies in this category (particularly G–44) to be facially valid, although perhaps entitled to slightly less weight than results derived from a larger test population. This approach is consistent with that taken in the carcinogenicity section of this decision. (See discussion of the calcium cyclamate portion of the first Friedman study, Section IV.B.2.b.(1)(c) above.)

b. *Reporting of Data.* The second criterion under the rubric of biological significance is that each valid study must contain an adequate presentation of data so as to enable a full evaluation of the study's results. For example, some of the studies are published only in "abstract" form, a mode which normally contains only a brief summary of the study's methods and results and virtually no presentation of data. Addtionally, other scientific papers contain results of several types of studies that were run concurrently; in these, the results of one portion (*e.g.*, the *in vivo* cytogenetic portion) were apparently of secondary interest to the investigators and therefore insufficiently presented. I have rejected as being "deficient" all of these studies, regardless of whether they reported positive or negative findings, which do not supply enough information to be assessed intelligently (see *e.g.*, G–124 at 19; see also Tr. at 490 and G–122 at 21). I have also not considered the results of one purportedly negative *in vivo* cytogenetic study (A–241) because it was submitted in a foreign language, was not translated, and is therefore impossible to evaluate.

c. *Positive Controls.* Ideally, every mutagenicity experiment (except those conducted on human beings) should have a positive control group, which is simply an additional treatment group dosed with a known mutagen, (see Tr. at 717). As the parties agree, the purpose of a positive control, is to serve as a check on the sensitivity of the test—*i.e.*, to ensure that the experiment is able to detect a mutagenic effect where one would be expected as to be present (Tr at 717; 975). In laboratories which specialize in the particular type of mutagenicity testing being performed, this same assurance can be gained from positive control data derived from previous experiments. Such data are called "historical controls" (Tr. at 960).

The following examples illusrate how results from positive controls either verify, weaken, or completely nullify a study's otherwise "negative" findings. First, if the positive control values are clearly positive, this verifies the negative results from the test compound and enhances their credibility because the experiment has been proven to be able to detect mutagenicity where it is expected to exist. In contrast, if the positive control values are positive but below their norm, the test compound's "negative" findings are of questionable significance because the experiment has been shown to be not as sensitive as it should be (see discussion of A–716 and A–811, App. 19 in Subsection F.5. below). Finally, if the positive control values are negative, this completely nullifies the test compound's negative results because the experiment has been shown to be too insensitive to detect a known mutagen (see discussion of bone marrow portion of A–177 in Subsection F.6.a. below).

Where a negative study contains no positive control data at all, I have taken the following approach. First, for negative studies with no internal flaws suggesting the experiment's lack of sensitivity (*i.e.*, A–143 and A–151 in Subsection F.5. below), I have not considered the absence of a positive control, by itself, to render the study "deficient." This is because the investigator may have had historical positive control data which was not reported in the published paper (for example, compare G–9 with Tr. at 960). In this situation, I have treated the lack of a positive control as reducing the weight to be given to a study, rather than as affecting its overall validity. In contrast, where other factors in a study suggest that the test is insensitive, I have considered the absence of a positive control to be the determinative factor in declaring the study to be "deficient" (see, *e.g.*, A–274 in Subsection F.6.a. below).

The lack of a positive control has a different effect upon a study with "positive findings" (*e.g.*, suggestive *in vivo* cytogenetic studies). Statistically significant ($P \leq .05$) findings in the test group are sufficient, by themselves, to domonstrate that the sensitivity of the experiment is adequate (Tr. at 975). As Dr. Legator testified "If one gets a[n] effect without a positive control, that again, as I said can be classified as a good experiment" (*id.*). Thus, I have considered these studies to be valid (see, *e.g.*, G–45 in Subsection F.4.e below).

D. Criteria for the Evaluation of Mutagenicity Evidence as a Whole

After each study has been reviewed and classified, the evidence as a whole must be evaluated to determine if Abbott has demonstrated that there is a reasonable certainty that cyclamate does not cause heritable genetic damage. This overall evaluation necessarily involves a judgmental process by the decision-maker, especially in making factual determinations such as whether certain negative studies outweight other suggestive ones. To objectify this process as much as possible, however, the record contains three criteria. These involve the necessity for using a battery of test methods, for testing different animal species, and for obtaining results from different laboratories.

1. *Battery of Test Methods:* The parties are in agreement that cyclamate must be tested in wide variety of experimental methods because know nutagens often produce positive results in some test methods but negative results in others (G–124 at 9–10 and 31; Tr. at 933–34; Tr. at 498–501; Tr. at 717–18 and 734). As Dr. Legator explained:

> At the present state of the art, all of our tests for describing or characterizing mutagenic agents have very serious drawbacks. Often, a particular type of test may miss a particular agent because of the insensitivity of the procedure, or the type of chemical being tested, the time of analysis, or many other factors. The great majority of compounds, with very few exceptions, do not give us a positive effect in all tests. Therefore, all responsible agencies in this area recommend that we use a battery of tests, that is, a number of tests, to study a single agent.

(G–124 at 9). Dr. Legator goes on to cite two examples of known mutagens (ionizing radiation and nitrogen mustard) which do not produce positive effects in one or more accepted

mutagenicity test methods (*id.* at 9–10). Dr. Green, another Bureau witness, cities five additional examples of this type (Tr. at 498–501). Abbott's chief mutagenicity witness, Dr. Hsu, also agrees with this general proposition (Tr. at 717–18 and 734). Thus, for Abbott to be able to establish safety, it must present a complete battery of mutagenicity tests which show negative results.

2. *Different Animal Species:* Mutgenicity tests must also be performed in a variety of animal species in order to prove safety. This is because a compound may produce negative results in one species but positive results in another (Tr. at 965; A–143 at 16; see also G–123 at 5). Although it would be impractical to require tests in every imaginable animal species, studies designed to rebut specific positive or suggestive findings should be performed using the same animal species and strain (*id.*) Thus, the species employed is an important factor to be considered in weighing the suggestive findings against the negative ones (see discussion in Subsection F.1. below).

3. *Different Labaoratories:* The need to obtain test results from different laboratories has recently been scrutinized and documented. According to Dr. Legator, eight laboratories performed a collaborative *in vivo* cytogenetic study on the compound triethylenephosphoramide (TEPA), a known mutagen. Before commencing the actual study, the investigators agreed on standardized procedures and on standardized definitions for scoring slides. Nevertheless, one of the eight laboratories produced results that were clearly different from the other labs (Tr. at 923–25). Thus, the extent to which the key data comes from the same or different laboratories is also a factor to be considered in weighing the evidence.

E. Credibility of Expert Witnesses

The credibility of the expert mutagenicity witnesses is very much in issue (see Abbott's Brief at 39–42 and Bureau's Brief at 106–110). In reaching my own conclusion as to the credibility of each expert, I have considered the following factors: (1) his training and experience; (2) the extent to which he has demonstrated a familiarity with the cyclamate studies of record; (3) the extent to which his testimony is corroborated or supported by other evidence in the record; (4) clarity or vagueness of his opinions; and (5) possible bias.

The parties' chief mutagenicity experts are Tao-Chiuh Hsu, Ph.D. for Abbott (A–800; Tr. at 715–734) and Marvin Legator, Ph.D. for the Bureau (G–124; Tr. at 894–976). I have reviewed each expert's curriculum vitae and relevant testimony and find that each holds outstanding qualifications in the field of mutagenicity testing. Both men have extensive experience in the design and execution of both *in vivo* and *in vitro* mutagenicity testing, using a variety of test compounds. Moreover, both men have special expertise in the area of cytogenetics which is of central importance in this proceeding.

Dr. Hsu gained his experience in academia, primarily at the University of Texas (Houston campus), but also at Baylor University, Brown University, Rice University, and Wayne State University. Dr. Legator served ten years (1962–72) as Chief of the Genetic Toxicology Branch of the Food and Drug Administration. Dr. Legator then moved on to academia, first at Brown University and most recently at the University of Texas (Galveston campus), which is the same university where Dr. Hsu teaches. Indeed, Dr. Hsu was one of the cytogenetic instructors at a toxicology course organized there by Dr. Legator (Tr. at 918). The two men therefore know each other well, and each has readily acknowledged the other's expertise (Tr. at 719; 919).

Rather than training and experience, the pivotal factor in the relative credibility of each man's testimony is the extent to which each has demonstrated a familiarity with the cyclamate studies of record in this proceeding. Dr. Hsu's entire direct testimony evaluating the cyclamate evidence consists merely of a brief, general summary which does not even mention the name or author of a single cyclamate study (A–800 at 7–8). Instead of criticizing these cyclamate studies, Dr. Hsu seems to rely on a review article which is not of record in this proceeding and whose completeness is in some doubt (Bureau's Brief at 106; Tr. at 919–21). Moreover, the fact that Dr. Hsu has not conducted any cyclamate studies himself (Tr. at 732) precludes another avenue by which he may have become familiar with all or part of the cyclamate evidence. In contrast, Dr. Legator described and evaluated in some detail many of the cyclamate studies, especially the key cytogenetic ones (G–124 at 16–25). Thus, Dr. Legator's testimony on specific studies is entirely unrebutted by Dr. Hsu (or any other Abbott witness).

In addition to Dr. Hsu's questionable familiarity with the record, his conclusions on the ultimate mutagenicity issue are not convincing. Dr. Hsu concluded that "there is no *decisive* evidence to show that cyclamates and their metabolites cause a *significant* amount of chromosome damage" (A–800 at 7) (emphasis added). This statement has two major flaws. First, Dr. Hsu's conclusion is vague. He talks in terms of a "significant amount" of chromosome damage without elaborating on how much or what kinds of chromosome damage would be needed before he would term it "significant." Second, Dr. Hsu's conclusion is incomplete. He mentions only the potential for chromosome damage without offering any opinion on the potential for gene mutations. This omission is somewhat surprising in light of his own introductory statement that a "(gene) mutation affecting an important gene can cause lethality; and a mutation affecting a less important gene can alter the organism's morphology or physiology" (A–800 at 1).[25]

I find these shortcomings of Dr. Hsu's testimony to be extremely significant and of far greater consequence than are Abbott's criticisms of Dr. Legator. Abbott suggests, for example, that Dr. Legator has "preconceived notions" about and an "emotional involvement" in the cyclamate issue and therefore is not able to render an objective, scientific opinion on this subject (Abbott's Brief at 40; Tr. at 907). Abbott bases this claim on certain letters to the press (G–136) and to FDA (A–828 and A–830) in which Dr. Legator advocated a cyclamate ban (Abbott's Brief at 40–41). I have reviewed these letters and decline to find the inferences which Abbott suggests. The first letter was based, from a mutagenicity standpoint, on objective scientific data—*i.e.*, the then recent laboratory findings of Dr. Legator (G–9) (see G–136 at Ref. 8). Moreover, Dr. Legator's views in this letter were shared by four other prominent scientists, including a Nobel Laureate (Tr. at 907). The thrust of the other two letters is that insufficient information was then known about the way cyclamate is metabolized to permit a responsible finding that cyclamate is safe. This opinion also was grounded in scientific facts which were stated in the letters themselves (A–828 and A–830).

Abbott also suggests that these letters reveal Dr. Legator's extra-scientific opinion that "cyclamate should not be

[25] In any event, Dr. Hsu's conclusion would not support a finding of safety. As discussed above, the general safety clause of the act requires disapproval of a food additive upon a showing of evidence "suggestive" of mutagenicity, even if that evidence is inconclusive. *Certified Color Manufacturers Ass'n* v. *Mathews*, *supra*, 543 F.2d at 297; *accord*, *Environmental Defense Fund* v. *E.P.A.*, *supra*, 598 F.2d at 89; *Ethyl Corp.* v. *E.P.A.*, *supra*, 541 F.2d at 37–38; *see Hercules, Inc.* v. *E.P.A.*, *supra*, 598 F.2d at 110. Thus, the fact that "decisive" evidence does not exist does not mean that Abbott should prevail.

approved as a food additive because it provides no benefits to society" (Abbott's Brief at 41). This is not the import of these letters. Dr. Legator was simply stating that cyclamate possesses no societal benefits capable of outweighing the public health risks which he perceives. Under the statutory scheme, however, possible societal benefits are not even to be considered. See 21 U.S.C. 348(c)(3)(A). Since Dr. Legator clearly distinguishes between the "risk assessment" and the "benefit assessment" in his analysis, I have considered the former but not the latter in my evaluation.

Finally, Abbott challenges Dr. Legator's scientific abilities by asserting that his positive findings with cyclamate (especially in G–9, his most important study) were artifacts since they could not be replicated by other investigators (Abbott's Brief at 40). I disagree with Abbott for two reasons. First, Dr. Legator is not the only investigator who has found an increased incidence of chromosome aberrations in an *in vivo* cytogenetic test. As illustrated in Subsection F.4. below, four other investigators, unaffiliated with Dr. Legator, have also reported a statistically significant increase in chromosome aberrations after using cyclamate or CHA. Moreover, the "negative" studies relied upon by Abbott as demonstrating lack of replicability of G–9 (*i.e.*, A–177, A–195 and A–297) are all "deficient" in terms of design or procedure and therefore do not detract from Dr. Legator's findings. (See dicussion in Subsection F.6. below.)

I therefore find that Dr. Legator has provided credible expert testimony which was based on objective facts, not personal bias. Moreover, the fact that he provided extensive detailed analyses of many of the specific studies at issue makes this testimony far more persuasive than the general, conclusory remarks offered by Dr. Hsu.

I need not go into detail on the expertise of the Bureau's other witnesses, Drs. Green, Epstein and Zimmering, (G–123, G–121 and G–122, respectively), as their credibility in not directly challenged by Abbott. I have reviewed each's curriculum vitae and relevant testimony and find that each qualifies as an expert in mutagenesis. Although each's testimony is limited in scope (Drs. Green and Epstein to certain *in vivo* cytogenetic and dominant lethal studies, and Dr. Zimmering to drosophila experiments), each of these experts demonstrated a familiarity with the specific studies evaluated. I therefore find that Drs. Green, Epstein and Zimmering are all credible witnesses.

I also need not discuss in detail the qualifictions of Abbott's other mutagenicity expert, Dr. Lorke (A–811T and A–827), but for a different reason. Unlike the other witnesses, Dr. Lorke did not present either written or live testimony in which he evaluated specific studies or the evidence as a whole. Instead, Dr. Lorke merely attached to his curriculum vitae a number of cyclamate mutagenicity studies which he performed. These studies were properly introduced into the record and have been reviewed along with the other experiments, and the general statements made in them are entitled to the same weight as those of any other investigator whose study is of record.

F. Evidence Raising a Serious Question as to the Mutagenicity of Cyclamate: The in Vivo Cytogenetic Studies.

1. *Summary of Evidence.* An *in vivo* cytogenetic study is an established type of mutagenicity test, carried out on live animals, used to examine a test compound's possible effects on chromosomes. The current administrative record contains 25 *in vivo* cytogenetic studies. A review of each of these studies has revealed that six are suggestive of mutagenicity, four are negative, and 15 are deficient.

The six suggestive studies (G–9 [both bone marrow and germ cell portions], G–26, G–44, G–45 and J–1) collectively present strong evidence that cyclamate or CHA causes chromosome aberrations. These findings were predominantly breaks, with some evidence of exchange figures. As fully explained in the following Subsection of this decision, breaks are biologically significant because they: (a) may lead to exchange figures; and (b) may be indicators of gene mutations. Both exchange figures and gene mutations are capable of causing heritable genetic defects.

The six suggestive studies may be summarized as follows. Legator, et al. (G–9) found a statistically significant increase in chromosome aberrations, predominantly breaks, in both the bone marrow and spermatogonia of Holtzman rats. Majumdar and Solomon (G–26) found similar results in the bone marrow of Mongolian gerbils. Turner and Hutchinson (G–44) found a statistically significant increase in both breaks and exchange figures (scored separately) in the blood cells of fetal lambs, while van Went-de Vries (G–45) found a statistically significant increase in chromosome aberrations (breaks and exchange figures being grouped together) in blood cells of Chinese hamsters. Finally, Bauchinger, et al. (J–1) found a statistically significant increase in breaks after analyzing the blood cells of human test subjects.

Abbott has challenged the validity of each of these six studies. In the study-by-study analysis which follows later in this decision, I have considered each of these alleged flaws and have concluded that each of the studies has strengths (*e.g.*, the dose response in G–9, both portions, G–26 and G–44) which outweigh any claimed weaknesses.

Moreover, the fact that findings of breaks were reported by five different laboratories using five different animal species and three different types of cells greatly enhances the studies' collective credibility and makes the evidence as a whole surpass the sum of its parts. Even more important, Bauchinger's findings from human beings lend confidence to the extrapolation of the animal study findings to potential human use.

The four valid negative studies are Brewen, et al. (A–143), Cattanach, et al. (A–151) and two studies by Lorke, et al. (A–716 and A–811, App. 19). All of these studies reported no statistically significant increase chromosome aberrations in treated animals over the negative controls. Brewen's study used the bone marrow of Chinese hamsters. Cattanach analyzed spermatocytes of mice, while both Lorke studies involved the spermatogonia of Chinese hamsters.

In comparing these negative studies with the suggestive ones, it is clear that the suggestive findings predominate. Cattanach's study is clearly distinguishable because he used an entirely different animal species (mouse). the three negative Chinese hamster studies are also distinguishable because they analyzed either bone marrow (A–143) or spermatogonia (A–716 and A–811, App. 19) cells whereas the positive Chinese hamster study (G–45) examined blood cells. Thus, none of the negative findings directly rebuts any suggestive study.

Moreover, each of the four negative studies has internal flaws which reduce the weight accorded to it. Brewen (A–143) did not specify the size of his test population, and neither he nor Cattanach (A–151) supplied positive control data. As to the Lorke studies (A–716 and A–811, App. 19), the positive control values for each were low, and the test population size for the latter one (A–811, App. 19) was too small (six per group) which lessened its sensitivity.

I therefore find that the *in vivo* cytogenetic evidence, when viewed as a whole, strongly suggests that cyclamate may be capable of inducing heritable genetic damage. This evidence alone in

sufficient to deny Abbott's food additive petition.

2. *Biological Significance of Different Types of Chromosome Damage.* The major issue surrounding the *in vivo* cytogenetic evidence involves the biological significance of three types of chromosome damage: breaks, gaps, and exchange figures.

a. *Types of Chromosome Damage:* As noted above, deoxyribonucleic acid (DNA) is the basic genetic material in man and animals. The DNA, which forms the genetic code, is distributed among the many "genes" that determine traits to be inherited. These genes are grouped in packages called "chromosomes." Humans have a total of 46 chromosomes; lower animal species, such as rats and mice, have fewer (A–800 at 1–2).

Physically, each chromosome (at the cell cycle stage called metaphase) contains two rods which are joined either at their centers or at one end, depending on the species. Each rod is called a "chromatid." Sometimes part of one of the chromatids cuts off and separates from the main rod. If the separation is greater than the width of a chromosome, the aberration is called a "break." If the separation is less than the width of a chromosome, it is called a "gap" (G–124 at 15; Tr. at 958). These are the definitions published by the Ad Hoc Committee of the Environmental Mutagenicity Society (*id.*), and the ones which I adopt.

Often chromosomes will repair themselves after breaks or gaps have occurred, or else die. Sometimes, however, two different chromosomes, each with breaks, join together at the site of those former breaks. When this happens, the resulting configuration is called an "exchange figure" (A–800 at 4; G–123 at 3; G–124 at 14). Exchange figures are also sometimes referred to as "reunion figures" (A–177 at Table III), "rearrangement figures" (A–297 at Table 1), "translocations" (A–716 at 243), or "major structural aberrations" (G–44 at Table 1). (For pictorial illustrations of breaks, gaps, and exchange figures, see, *e.g.,* G–45 at Figure 2 and A–239 at Figure 2).

Abbott questions the comparative biological significance of breaks versus gaps versus exchange figures (Abbott's Brief at 61). Relying upon A–239 at 350, Abbott claims that "'* * * gaps are the least conclusive criterion in determinating [sic] cytogenetic effects * * *', and chromatid breaks are only slightly better * * * the best criterion * * * are rearrangement figures * * *'" (Abbott's Brief at 61).

I agree with Abbott that A–239 rank-orders exchange fixtures, breaks and gaps. However, the comparison made in that study, when read in context, is slightly different from that presented by Abbott. A–239 is a collaborative *in vivo* cytogenetic study conducted by four independent laboratories. (The study did not involve cyclamate or its metabolites.) One purpose of the study was "* * * to test the variability in interpreting [jointly prepared] slide preparations by participants in their respective laboratories" (A–239 at 338). This type of study recognizes the fact that the cytogenetic analysis is, to some extent, a subjective art as much as an objective science. The investigators therefore sought to compare analyses by different laboratories of jointly prepared slides. These investigators, which included Dr. Legator, concluded:

> The results indicate that gaps are the least conclusive criterion in determining cytogenetic effects. The variability between laboratories was greatest for gaps; in addition, the values for gaps resulting from TMP and the different doses of DDT were not significantly different in the ip [injected] and oral parts of the experiment. Agreement between laboratories was close for the criterion of breaks, and was even closer for rearrangement figures * * *

(A–239 at 349–50). This passage, read as a whole, clearly demonstrates that the investigators rank-ordered the three types of chromosome damage in terms of the agreement/variability between laboratories. In other words, when four laboratories separately read jointly prepared slides, their findings were "close" for breaks, "even closer" for exchange figures, and varied the "greatest" for gaps. I interpret these results to mean that in an *in vivo* cytogenetic study, findings of breaks can be considered reliable, exchange figures even more reliable, but gaps not very reliable. In so finding, I note that this conclusion only reaches the issue of whether certain findings in *in vivo* cytogenetic studies can be considered to be accurate, not whether those findings, even if accurate, are biologically significant. For the remainder of this Subsection, I will consider the latter issue.

b. *Biological Significance of Exchange Figures:* The parties agree that exchange figures are biologically significant in that they can survive and pass on genetic defects to the next generation. As one Bureau witness, Dr. Green, stated: "It is generally thought that exchange figures are the type of abnormality that can be associated with heritable genetic damage" (G–123 at 3; see also G–124 at 14; Bureau's Brief at 100–02; and Abbott's Brief at 49). I agree and find accordingly.

I also find that one *in vivo* cytogenetic study in the record (Turner and Hutchinson, G–44) reported a statistically significant increase in exchange figures after dosing fetal lambs with CHA. (See general discussion of this study in Subsection F.4.d. below). Although this finding is not by itself sufficient to prove that cyclamate is mutagenic, I believe this evidence does cast doubt upon the safety of cyclamate in this regard.

c. *Biological Significance of Breaks:* The central cytogenetics question, and one on which the parties strongly disagree, concerns the biological significance of breaks. This issue is of central importance because breaks were the predominant finding throughout the six suggestive studies relied upon by the Bureau (G–9 [both portions], G–26, G–44, G–45, and J–1; see in Subsection F.4. below).

The Bureau's position is that breaks are biologically significant for three reasons: (a) they lead to exchange figures; (b) they are indicators of other types of genetic damage such as gene mutations; and (c) they can cause heritable genetic damage themselves (Bureau's Brief at 100–03).

Abbott does not directly respond to points (a) or (b), but does strongly disagree with the third, arguing that chromosomes with just breaks are not inherited because they either repair themselves or die (Abbott's Exceptions at 36; 47; Abbott's Brief at 47–52; 56–57; and 61).

The ALJ did not make specific findings on this issue but did note that statistically significant findings of breaks could not be disregarded (ID at 34).

A careful review of the testimony on this issue shows that it strongly supports the Bureau's position that breaks are biologically significant for the first two of the three reasons advanced by the Bureau. I will now discuss them in the order presented by the Bureau.

(1) *As Leading to Exchange Figures:* As Dr. Green explained:

> Exchange figures are produced as a result of the rejoining of chromosomes which possess chromatid breaks. It is, therefore, apparent that chromatid breaks are the necessary events which subsequently lead to exchange figures.

(G–123 at 3; see also G–124 at 14; A–800 at 4). This theoretical point is confirmed by the statistically significant findings of exchange figures in the Turner and Hutchinson study (G–44). Moreover, according to the experts, the fact that several studies found breaks without exchange figures is not unusual. As Dr. Legator explained: ". . . the frequency

here of exchange figures by our current techniques are always minimal. That is, we do see it, but it is the rare kind of event" (Tr. at 956; see also Tr. at 957 and G–124 at 29–30). Dr. Legator attributed the low frequency of exchange figure findings to the imperfect state of the art for this type of scientific test: ". . . when we do see breaks or gaps, if we readjust our techniques or timing, we probably could see exchange figures as well" (G–124 at 14; see also G–124 at 30 and G–123 at 3). As Dr. Legator concluded: "It is very difficult to find a compound that has been thoroughly investigated that does not cause exchange figures when breaks or gaps are found" (G–124 at 15; see also Tr. at 957).

On the basis of this evidence, which is unrebutted by Abbott's chief expert, Dr. Hsu, I find that one basis for the biological significance of breaks is that they will likely lead to the formation of exchange figures which, as the parties agree, cause heritable genetic damage.

In theory, Abbott could rebut this conclusion by presenting sufficient valid negative studies proving that there is a reasonable certainty that cyclamate or its metabolites do not cause exchange figures. Abbott believes that it has already done this (Abbott's Exceptions at 65, 72, 73 and 74–75; Abbott's Brief at 56–63), but I disagree. Although Abbott introduced into the record 18 studies which it labeled as being "negative," I have found 14 of these studies to be "deficient" because they do not meet the minimum criteria for a valid study (see discussion in Subsection F.6. below).

Of the four valid negative studies, each was internally flawed because it lacked validation by positive controls (A–143 and A–151), had unusually low positive control values (A–716 and A–811, App. 19), did not specify the size of (A–143) or did not employ a sufficiently large (A–811, App. 19) test population. The best evidence presented by Abbott attempting to demonstrate the lack of exchange figures is the Cattanach study (A–151). As explained in Subsection F.5.b. below, that study was conducted exclusively to look for exchange figures. However, Dr. Cattanach conducted his study on mice which was not one of the species in which evidence of breaks were found (*i.e.*, rats, gerbils, lambs, Chinese hamsters and humans).

For Abbott to prove to a reasonable certainty the lack of exchange figures, it would have to present a group of valid negative studies designed to detect exchange figures; these studies should have large test populations, several dose levels, and validation by positive controls. Moreover, these studies should include the animal species in which positive findings of breaks *in vivo* have already been reported. I recognize that this is a heavy burden to impose upon Abbott, but it is one that I believe is necessary to prove safety, as the statute requires.

(2) *As Indicators of Gene Mutations:* Breaks are biologically significant for a second reason, the Bureau argues, because they serve as indicators of other types of genetic damage, especially gene (or point) mutations. Dr. Hsu, Abbott's chief witness, described genes and gene mutations as follows:

> A cell of an organism contains numerous genes each of which determine a particular step of a biochemical process. It is a sequence of DNA with code which determines a particular protein. If the code changes in whatever manner or if the code is missing, the gene becomes 'mutated' or deleted respectively, and the gene cannot perform its designated function. A mutation affecting an important gene can cause lethality; and a mutation affecting a less important gene can alter the organism's morphology or physiology.

(A–800 at 1). Thus, in terms of heritable damage, gene mutations are as biologically significant as exchange figures (G–124 at 8).

Dr. Legator testified that there is an "extremely good" correlation in other compounds between chromosome abnormalties and gene mutations (Tr. at 931). "[I]n fact," he said, "I cannot think of more than perhaps one exception—[of compounds] that cause chromosomal damage that do not also cause gene mutations" (*id.;* see also G–124 at 8; G–123 at 3). Dr. legator has also emphasized this correlation between chromosome damage and gene mutations in a context completely independent of cyclamate: "Although there is no proven quantitative relationship between point mutation and chromosomal changes, the correlation between either physical agents or chemical agents that can cause both types of alteration has been well established [reference omitted]" (A–239 at 349). At least one other investigator of record in this proceeding agrees: "* * * minor chromosomal lesions which cannot be regarded as permanent breaks, may well be indicators of submicroscopic damage, pinpoint mutations" (Schoeller, G–18 at 3).

Therefore, on the basis of this unrebutted expert testimony, I find that breaks are also biologically significant because they serve as indicators of gene mutations which, as the parties agree, can cause serious heritable damage.

In theory, Abbott could rebut this expert testimony with valid negative gene mutation studies demonstrating that cyclamate is the exception, rather than the rule, with respect to the correlation between breaks and gene mutations. On this record, however, there is only one gene mutation study conducted on mammalian cells, and that was an incompletely reported *in vitro* study conducted by Chu, et al. (A–699; G–47). Chu reported negative findings with cyclohexylamine (CHA, a metabolite of cyclamate) and positive findings with N-hydroxycyclohexylamine (N–OHCHA, another metabolite) both with Chinese hamster cells. These findings, however, were reported only in a biref abstract, with no supporting data, and are therefore entitled to little, if any, weight (see general discussion on abstracts in Subsection C.3.b. above and specific discussion of the Chu study in Subsection G.5. below.) I therefore conclude that Abbott has not shown that there is reasonable certainty that cyclamate or its metabolites do not cause gene mutations.

(3) *As Heritable Genetic Damage Themselves:* Finally, the Bureau contends that breaks "can cause serious genetic damage themselves" (Bureau's Brief at 102). Specifically, the Bureau argues that chromosomes with breaks can be replicated and passed on to progeny, and, relying on Dr. Hsu's testimony, this constitutes genetic damage because it represents a change in the genetic code (*id.* at 102–03).

Abbott responds that breaks themselves are not biologically significant because they will either repair themselves or die; thus, a single broken chromosome, when not part of an exchange figure, will not be inherited by future generations (Abbott's Exceptions at 36; 44; Abbott's Brief at 47–52; 56–57).

I agree with Abbott that chromosomes with breaks themselves do not constitute heritable genetic damage. A review of the passages from the testimony which the Bureau quotes in its brief shows that, when read in context, they so not support the Bureau's position. For example, Dr. Legator stressed: "The intriguing thing about exchange figures *as opposed to breaks* is that these rearrangements can survive and multiply" (G–124 at 14) (emphasis added). This statement clearly suggests that chromosomes with just breaks themselves cannot "survive and multiply." I read the language quoted by the Bureau to mean that chromosomes with breaks which do not die (or correctly repair themselves) are significant because they may lead to exchange figures. The Bureau's reference to Dr. Hsu's description of gene mutations is also, I believe, not supportive. As noted above, Dr. Legator

was quite clear that breaks are merely "indicators" (as opposed to direct causes) of possible gene mutations. I therefore conclude that, on the basis of the current record, breaks themselves do not constitute heritable genetic damage.

d. *Biological Significance of Gaps:* The biological significance of gaps is a less important issue for the purposes of this proceeding because the suggestive *in vivo* cytogenetic findings relied upon by the Bureau were based primarily upon breaks rather than gaps (G–9 at 1140; "single chromatid breaks predominated"; G–26 at 191, 193: breaks scored separately; G–44 at 409: breaks scored separately; G–45 at 417: exchange figures, breaks and gaps all grouped together; and J–1 at Table 3: breaks scored separately.)

Nevertheless, the Bureau argues that gaps are entitled to some weight (Bureau's Brief at 103–04). The Bureau relies upon: (1) the fact that Dr. Hsu describes breaks and gaps without distinguishing between them in terms of biological significance (A–300 at 4); and (2) Dr. Legator's testimony that certain other, unnamed scientists recently stated that gaps were as significant as breaks (G–124 at 15). The Bureau admits, however, that gaps may indeed be entitled to less weight than breaks (Bureau's Brief at 103–04). In response, Abbott contends that gaps are entitled to no weight at all because: (a) they may be quickly repaired by cellular mechanisms; (b) they are difficult to detect; and (c) in any event, are not heritable (Abbott's Exceptions at 36; Abbott's Brief at 48–52).

As seen from the definitions of breaks and gaps, the difference between the two is one of degree rather than kind—*i.e.*, breaks are wider separations than are gaps, but both are separations nonetheless. It logically follows that gaps, like breaks, are entitled to some weight, for gaps may develop into breaks.

However, as Dr. Legator himself found in his collaborative *in vivo* cytogenetic study (A–239) described in Subsection F.2.a. above, ". . . gaps are the least conclusive criterion in determining cytogenetic effects. The variability between laboratories was greatest for gaps . . ." (A–239 at 350). Based upon this conclusion, I find that gaps are entitled to considerably less weight than are breaks. I would therefore classify gaps in terms of "additional support" rather than as "primary evidence" of potential mutagenicity.

3. *Conduct of an In Vivo Cytogenetic Study.* An *in vivo* cytogenetic study is carried out in three principal steps: (1) dosing; (2) obtaining and preparing cell specimens; and (3) analyzing cell specimens.

In the dosing stage, the test group of animals or humans is given the test compound, either by feeding or injection, for a specified period of time. Concurrently, both negative and positive control groups are usually identified and dosed by the same means and for the same duration. The negative control group is given a placebo, and the positive control group is given a known mutagen.

At the conclusion of the dosing period, cell specimens from each group are obtained for microscopic analysis. Three types of cells were used in the cyclamate experiments: bone marrow, blood, and sperm cells. Obtaining bone marrow cells sometimes requires that the test animal be sacrificed. This procedure, therefore, is usually reserved for smaller animals, such as rodents. In human beings, and often in larger animals, blood cells are used instead which are obtained by simple, well-known procedures.

Once the cell specimens are obtained and prepared, they are examined microscopically for the type and frequency of chromosomal aberrations. If the results from the test group are "positive," they are compared to the negative control for statistical significance. In contrast, if the results from the test group are "negative," the results are compared to the positive control to ensure that the experimental environment was conducive to obtaining a positive response.

4. *Suggestive Studies.* As noted above, the administrative record contains six studies whose findings are inconclusive but suggestive of mutagenicity. Each study's findings were statistically significant at the $P<.05$ level. The reason these studies are "suggestive" rather than "positive" is that the findings were primarily of breaks rather than exchange figures (see discussion in Subsection C.1.b. above).

a. Legator, et al. (G–9) (bone marrow portion)

(1) *Study Design:* This study was performed on Holtzman strain albino male rats using CHA as the test compound. Five test groups of 20 to 30 rats each were formed, and each group was given daily CHA intraperitoneal injections of 1, 10, 20, 40 or 50 mg/kg, respectively, for a period of five days. A similar-sized negative control group was also established and given daily injections of distilled water over the same period of time. The animals were sacrificed 24 hours after the last injection, and slides were prepared and analyzed for 625 cells at each dose level. (These cells are called "metaphases" because the cells are at the metaphase stage of the cell cycle.) Although the published report of the study does not so state, these slides were coded so that the persons reading them had no knowledge of whether they came from treated or control groups (Tr. at 960).

(2) *Study Results:* Analyses of the bone marrow cells revealed a statistically significant increase ($P<.01$) over the control group in the percentage of cells with breaks in each of the four highest dose groups. Moreover, a linear dose-response was observed throughout these four groups. The authors also noted "infrequent exchange figures" (G–9 at 1140), but presumably, these alone did not reach statistical significance.

The ALJ found that this study produced "positive results" with a "dose-response" trend (*id.* at 25–26, 35). He also noted that the test compound was tested for impurities and that none were found (*id.* at 26). However, the ALJ emphasized that two other investigators (Ford, A–279 and Dick, A–177) failed to replicate Legator's results, despite "appear[ing] to have used the exact protocol used by Dr. Legator." (*id.* at 25). The ALJ concluded, therefore, that "the inability of replicating (Legator's) results puts them in doubt." (*id.* at 26).

(3) *Analysis:* Abbott takes no exception to this aspect of the ALJ's decision. The Bureau's position on the replicability issue is that "the Dick study was not an exact duplication and has problems of its own" (Bureau's Brief at 77). The Bureau also suggests possible bias in the Dick study because Dr. Dick was an Abbott employee at the time the study was performed (*id.* at 76–77; 87). The Bureau has made no comments on the Ford study.

As to the quality of the Legator study itself, Abbott argues: (a) that Legator's findings of breaks do not constitute "permanent" breaks (Abbott's Brief at 56–57); and (b) that Legator did not use positive controls (*id.* at 57). The Bureau's response is that breaks do constitute significant genetic events (Bureau's Brief at 100–04); and that positive controls are not necessary to validate positive results, only negative ones (*id.* at 76).

Viewed by itself, I would characterize the Legator study as a very well-designed experiment which produced clear positive findings of breaks. By the phrase "very well-designed," I mean that Dr. Legator employed a sufficient number of test animals (20–30 per dose group) and analyzed a sufficient number of cells (625 per dose group). Additionally, as the ALJ noted, he tested the CHA for impurities and none were found. He also coded the slides to prevent possible bias. Finally, he used a

dose range (five dose levels) which permitted dose-response information to be obtained. By the phrase "clearly positive findings," I refer both to the statistically significant findings at four out of five dose levels, and to the consistent dose-response trend which was observed. The dose-response, especially, adds credence to the positive findings (G–124 at 17; see G–18 at 2).

I have already considered and dismissed Abbott's first contention, regarding the biological significance of breaks (see Subsection F.2.c. above.) Abbott's second criticism, regarding the lack of positive controls, is also without merit. I agree with the Bureau that positive controls are not always necessary to validate positive findings (Tr. at 975; see discussion in Subsection C.3.c. above), especially where, as here, the testing laboratory has historical control data (Tr. at 960).

Furthermore, I disagree with the ALJ's finding that "the inability of replicating [Legator's] results puts them in doubt" (*id.* at 26). I disagree because the two allegedly "replicate" studies (Ford, A–297 and Dick, A–177) are each "deficient." Ford administered CHA to only one group of three animals and then analyzed a total of only 150 metaphases. This test population is simply too small for any weight to be given to its "negative" results. It certainly in no way detracts from Legator's positive findings which were derived from an experiment employing five groups of 20–30 animals per group with the total number of metaphases examined exceeding 3000.

The Dick study (A–177) is deficient for a different reason. Although it is true that Dick reported no statistically significant increase of breaks in the treated over the negative control group, neither did Dick find such an increase of breaks in the positive control, triethylenemelamine (TEM), as compared to the negative control (A–177 at Table III). This absence of breaks in the positive control demonstrates that Dick's study was so insensitive that she could not even find breaks in a compound known to be mutagenic. *A fortiori,* no conclusion can be drawn from Dick's failure to observe breaks in the test compound, CHA.[26]

(4) *Other Matters:* Former Commissioner Kennedy, in his Remand Order, requested the parties to provide certain underlying data to this study because it would be "helpful" in evaluating the study "more fully" (44 FR 47623). The parties have since stipulated, however, that the requested data could not be located (Stipulation dated September 17, 1979 at 8). Although I agree with the former Commissioner that the requested information would have been helpful, I find that I can adequately evaluate this study on the basis of information currently in the record.

b. Legator, et al. (G–9) (spermatogonial cell portion)

(1) *Study Design:* The design of the spermatogonial cell portion of this study was identical to the bone marrow portion discussed above, except that 500 metaphases were analyzed for each dose level (rather than 625).

(2) *Study Results:* The investigators observed a statistically significant increase ($P \leq .05$) over controls in the percentage of cells with breaks in each of the five dose groups. The investigators also found a linear dose-response throughout these five groups. Finally, as with the bone marrow portion, the authors noted "infrequent exchange figures" which presumably did not reach statistical significance.

The ALJ characterized this portion of the Legator study as showing an "adverse effect" (*id.* at 35), and, more specifically, as demonstrating a statistically significant increase in "breaks" in the treated animals over the controls which was found to be dose-related (*id.* at 27–28). The ALJ also noted that "infrequent exchange figures" were observed (*id.* at 28). The ALJ distinguished the Friedman study (A–195) because the investigator there used a different test compound (cyclamate rather than CHA) and a different method of administering that test compound (feeding rather than intraperitoneal injection) (*id.* at 28). However, the ALJ did note that the Ford study (A–297) "appear[ed] to have used the same protocol as Dr. Legator but obtained negative results" (*id* at 28).

(3) *Analysis:* Abbott agrees with the ALJ that Legator's findings should be discounted because they could not be replicated by Ford (Abbott's Exceptions at 65). Other than that, Abbott has not raised any additional criticisms directed towards the design of the spermatogonial cell portion of this study. Moreover, Abbott explicitly does not take exception to the manner in which the ALJ distinguished the Friedman study (see *id.* at 65). The Bureau makes no additional comments on these issues.

I would adopt here, by reference, my earlier evaluation of the design and results of Legator's bone marrow portion. Moreover, positive findings in spermatogonial cells carry special significance from a mutagenicity standpoint. While positive findings in somatic cells (*e.g.,* bone marrow cells and blood cells) help us learn whether a certain compound causes chromosomal aberrations at all, positive findings in spermatogonial cells give us the added information that those chromosomal aberrations occur in the very cells that determine heredity (G–124 at 16–17). Thus, this study presents important evidence that genetic abnormalties caused by cyclamate or its metabolites may be passed on to future generations.

With respect to the Friedman study (A–195), I agree with the ALJ that this study is distinguishable from Legator's because of the difference in test compounds and routes of administration. Moreover, for the additional reasons stated in Subsection F.6.a.(4) below, I find that this study is deficient and therefore entitled to no weight at all.

With respect to the Ford study (A–297), however, I strongly disagree with the ALJ that is was a replicate of Legator's study. As I explained in my analysis of the bone marrow portions of these two studies (the design of each investigator's sperm cell portion being virtually identical to the design of his own bone marrow portion), Ford simply used too small a test population (1 dose level; 3 animals; 124 metaphases) for it effectively to rebut Legator's findings (which were based on 5 dose levels; over 100 animals, and 2500 total metaphases). Moreover, Ford's study size is so small that I have classified it as "deficient" and have attributed no weight to it at all (see Subsections F.6.a. and d. below).

(4) *Other Matters:* My earlier comments (in the bone marrow section) regarding the former Commissioner's request for additional data are equally applicable here.

c. Majumdar and Solomon (G–26)

(1) *Study Design:* This study was performed on Mongolian gerbils using calcium cyclamate as the test compound. The design of the study was similar to that of Legator, et al. (G–9) in that the test population consisted of five dose groups (ten animals per group) that were given daily injections of 10, 30, 50, 70 or 100 mg/kg, respectively, for a period of five days. A negative control group of ten animals received injections of distilled water over the same period of time. All animals were sacrificed on the fifth day. Cells selected for analysis were from the bone marrow and numbered 300–350 per dose group.

(2) *Study Results:* In the four highest dose level groups, the investigators

[26] I do note, however, that, contrary to the suggestion raised by the Bureau, the mere fact that Dr. Dick was an Abbott employee at the time her study was conducted is, by itself, simply not relevant to the issue of investigator bias.

reported a statistically significant increase ($P<.001$) over controls in the percentage of cells with breaks as well as in the percentage of cells with gaps and fragments. Moreover, across these four dose levels, a slight but consistent dose-response was found.

The ALJ characterized this study as having "positive results" due to the findings which I have just summarized (*id.* at 35).

(3) *Analysis:* Abbott makes no specific exceptions to this study. I agree with the ALJ that this study produced clear, statistically significant positive findings, especially in light of the dose-response obtained (see G–124 at 17). Moreover, the investigators employed a sufficient number of animals (ten per dose group) and analyzed a sufficient number of cells (300 per dose group) to give added credence to the results. These findings tend to confirm the results of the bone marrow portion of Legator, et al. (G–9) (see G–121 at 13), and therefore enhance the credibility of both studies. Although the number of chromosomal aberrations observed in this study was somewhat lower than that reported by Legator, et al. (G–9), this could be explained either by the fact that Majumdar and Solomon used calcium cyclamate rather than its metabolite, CHA (G–124 at 20), or by the fact that they used a different animal species (see discussion in Subsection D.2. above). In any event, this study stands virtually unimpeached. This evaluation is further supported by expert testimony (G–124 at 20).

d. Turner and Hutchinson (G–44)

(1) *Study Design:* This study was carried out on fetal lambs using CHA as the test compound. Each treated animal received one dose of CHA of either 50, 100, 200, or 250 mg/kg. The animals were given the CHA *in utero* by intravenous injections over a period of either five or 18 hours. Eight treated animals were used in all, one for each dose level and dose period. In addition, one control animal was used for each dose period. Cells obtained for analysis were peripheral blood cells that were drawn from each fetus. Results were based on the analysis of a combined total of 500 cells.

(2) *Study Results:* The investigators reported a statistically significant increase over the control (in each time period) in three different categories: (a) percentage of cells with major structural aberrations (i.e., exchange figures); (b) percentage of cells with breaks; and (c) percentage of cells with total aberrations. In addition, a linear dose-response was found.

The ALJ characterized this study as a "positive" one, showing both "chromatid and chromosome aberrations" (*id.* at 27, 35).

(3) *Analysis:* Abbott takes exception to the low number of animals utilized in the study, and, relying on alleged conclusions by the authors, claims that: (a) CHA is a "clastogen" only; (b) breakage was not dose-related; and (c) CHA does not induce translocations (Abbott's Exceptions at 60). The Bureau responds only to the issue of study size, arguing that where positive results are found in a small study, this is an indication that the test compound is quite potent (Bureau's Reply at 22).

The major strength of this study is the statistically significant findings of major structural aberrations, which include exchange figures. As noted previously, these are the types of chromosomal abnormalities which the parties agree cause heritable genetic damage (Abbott's Brief at 49); Bureau's Brief at 100). Moreover, because a lamb is a much larger animal than the conventionally used rodent, these findings have a more direct applicability to man (Tr. at 964). Finally, the demonstration of a dose-response, as noted in the previous studies greatly enhances the credibility of the positive results (G–124 at 17). Once again, this analysis is supported by the relevant expert testimony (*id.* at 17–18).

I have reviewed in general terms Abbott's criticism that this study employed too few animals (see discussion in Subsection C.3.a. above), but I will elaborate here. It is true that Turner and Hutchinson employed only one animal per dose group per treatment period. It is also true that this is far less than the ideal number of animals to use. For example, had this study produced negtive results, it could have been justly criticized for being too insentitive because there would not have been a sufficient likelihood of detecting mutagenic effects, even if present. On the other hand, where, as here, a study with few animals produces positive results, it cannot be critized for being too insensitive. Quite the contrary, what such a test suggests is that the test compound is sufficiently potent that it is capable of being detected by even an insensitive test (Tr. at 941). Therefore, although Turner and Hutchinson's findings would have been stronger if their test population size had been larger, the findings based upon the population used are nevertheless valid.

Abbott's other exceptions regarding the author's alleged conclusions are totally without merit. Although the authors do conclude that CHA may be a "clastogen," this in no way advances Abbott's cause. The company's own witness, Dr. Hsu, defined a clastogen as a mutational agent which causes chromosomal aberrations (Tr. at 718). Second, contrary to Abbott's assertions, the authors do not conclude that the breakage was not dose-related. In fact, the authors concluded that they observed a "dose-effect correlation in the frequencies of both major and minor aberrations . . ." (A–725 at 411; G–44 at 411). This would include both exchange figures and breaks, respectively. Finally, Abbott's counsel apparently misread page 410 of this study. Nowhere on that page does the word "translocations" appear, although the word "tranformation" does. This exception, therefore, requires no response.

e. Van Went-de Vries (G–45)

(1) *Study Design:* In this study, Chinese hamsters were given CHA through oral intubation (forced feeding). Twenty hamsters were each dosed with 200 mg/kg for three successive days. The cells analyzed were peripheral blood cells. Blood was drawn both before and at the conclusion of the dose period; thus, each animal served as its own negative control. A total of 1,000 metaphases were analyzed for both treated and control groups. The slides to be analyzed were coded so that the persons reading them had no knowledge of whether they came from treated or control animals.

(2) *Study Results:* The investigator found a statistically significant increase ($P<.005$) over controls in the total number of structural chromosome abnormalities. The findings included several exchange figures, one ring, and numerous breaks and fragments.

The ALJ characterized this study as "positive" with an "increase in structural aberrations" (*id.* at 26, 35).

(3) *Analysis:* Abbott criticizes this study in three ways: (a) absence of positive controls; (b) difficulty in evaluating findings since all types of aberrations were grouped together; and (c) allegedly small increase in total aberrations in treated over controls (Abbott's Brief at 58–59). In response, the Bureau simply emphasizes the positive findings (which included exchange figures) and the extra controls employed by the investigator, such as the precautions taken to ensure purity of the test compound (Bureau's Brief at 75–76).

The strength of this study lies in the statistically significant increase over controls in terms of total chromosomal aberrations found, and the fact that this included exchange figures (see Tr. at 921). Unfortunately, due to the design of the study which called for only one dose level, no dose-response information could be obtained.

There is some merit to Abbott's criticism regarding the small increase in the amount of total chromosomal aberrations in treated over controls. Although the findings were statistically significant at a high confidence level (P<.005), the actual number of aberrations observed was relatively low (see G–45 at table 2). This reduces the biological significance of the findings. Indeed, this study presents a perfect example of how the concept of biological significance can reduce the weight otherwise accorded to statistically significant results. While this does not mean the findings of this study should be completely discounted, the weight given is not as great as in the previous studies discussed.

Abbott's complaint regarding the grouping of the chromosomal aberrations has less merit. Although the findings would have been somewhat stronger if exchange figures had been grouped separately and found to be statistically significant (as was true with the Turner and Hutchinson study), so, too, would the findings have been somewhat weaker if no exchange figures had been found at all. Thus, the findings simply are what they are.

Finally, Abbott's third comment regarding the absence of a positive control is without merit. As noted above in Subsection C.3.c. above, positive controls are not necessary to confirm positive findings (see TR. at 975).

f. *Bauchinger, et al. (J–1):* (1) *Study Design:* This was the only suggestive study conducted on human beings. Cyclamate was the test compound. The treatment group consisted of 11 persons, all of whom suffered from liver or kidney disease(s). Each was fed either 2 grams or 5 grams of cyclamate per day for periods ranging from 310 to 1160 days. In addition, two control groups were established. The first ("Control I") consisted of 10 persons with the same or similar diseases as the treated group. Control I was given fructose instead of cyclamate. The second control group ("Control II") consisted of 52 healthy persons. The authors made special mention of the fact that none of the participants in this experiment received therapeutic radiation or thereapy with alkylating drugs. The blood cells used for analysis were peripheral lymphocytes. Approximately 100 cells were analyzed from each individual from the treated and Control I groups, and 55 cells from each member of the Control II group.

(2) *Study Results:* The results of this study are most clearly presented in Table 3 (J–1). Here the investigators reported a statistically significant increase in the treated group (diseased individuals on cyclamate) over the Control I group (diseased individuals on placebo) in terms of: (a) percentage of cells with chromosomal aberrations (*i.e.*, breaks, gaps and exchange figures grouped together) (P=.032); (b) total number of breaks (P=.05); and (c) aberrations with open breaks (P=.038). No significant difference was found for the fourth category labeled "restructurings," which would include exchange figures.

The investigators also found, however, a statistically significant increase in the Control I group (diseased individuals on placebo) over the Control II group (healthy individuals on placebo) in terms of: (a) total number of breaks (P=.024); and (b) restructurings (P<.003). No significant difference was found in the other two categories.

(3) *Analysis:* The strong points of this study are that: (a) it was conducted on humans; (b) it used dose levels "that are frequently encountered in individuals who are consuming cyclamate" (G–124 at 22); and (c) the investigators found a statistically significant increase in breaks in the treated over the Control I group. Dr. Legator termed this study "probably the most relevant piece of information we have right now to be expanded on" (Tr. at 974). The ALJ apparently agreed (see extended discussion in at 27; 35). (See also G–121 at 13; G–123 at 4; G–124 at 21–22.)

Although Abbott takes numerous exceptions to the ALJ's findings on this study, only one of these requires a detailed discussion. This relates to the possibility that a synergistic effect could have been at work between the cyclamate and the diseases. In this regard, there are two ways in which to interpret Bauchinger's findings. The first is to say that sinec the cyclamate was the only differing factor between the treated and the Control I group, it was the cyclamate which caused the increased incidence of breaks. This view was adopted by the ALJ (*id.* at 27). The second possible interpretation is that the increased incidence of breaks was caused by a synergistic effect, *i.e.*, the combination of the cyclamate and the diseases. This theory is somewhat supported by the increased incidence of breaks found in the diseased control group over the healthy control group. However, even were this second interpretation the proper one, the results of the Bauchinger study would still be "relevant because, if approved, cyclamate would be ingested by a broad segment of the population, including those with kidney and/or liver disease" (*id.* at 35).

Consistent with the cautious approach taken throughout this decision, which is aimed at maximizing protection of the public health, I am interpreting the positive results of this study as having been caused by the cyclamate (*i.e.*, the first option just discussed). This interpretation is at least as likely to be correct as the synergistic effect interpretation, and Abbott has not satisfactorily shown the contrary to be true.

Abbott's remaining exceptions can be dealt with briefly. First, Abbott challenges the small population size of the treated (11 persons) and Control I (10 persons) groups (Abbott's Exceptions at 62), as well as the disparity in size between those two groups and Control II (52 persons) (*id.*). Abbott also questions the validity of analyzing only about half as many cells per person in Control II (55) as in the other two groups (100) (*id.*). However, there is unrebutted testimony that the size of the test population was adequate (Tr. at 508) and that an ample number of cells was analyzed (Tr. at 515). I agree with this testimony, and I would again emphasize that the actual size of a test population is less important where positive findings are obtained.

Abbott next seeks to clarify the exact nature of the positive findings by stating: "the only chromosome aberrations that were significantly increased were open breaks." (Abbott's Exceptions at 61). Abbott is generally correct on this point. For a more precise statement of Bauchinger's findings, see my statement of the Study Results, *supra.* Abbott also states that "similar kinds of chromosomal aberrations and frequencies were observed in both the treated group and control group I" (Abbott's Exceptions at 63). Abbott is also correct here in so far as "chromosomal aberrations" refers to Bauchinger's "Restructurings" category, which would include exchange figures (J–1 at Table 3). Abbott's point, I believe, is that the only statistically significant findings were in terms of breaks and not exchange figures. This is true, as I have already pointed out. To the extent that the ALJ's decision suggests anything to the contrary, I would modify it accordingly.

Given these findings of breaks, Abbott criticizes their validity because they were not related to dose or duration of exposure (Abbott's Exceptions at 61). However, Dr. Legator testified that he would not expect to find a dose-response relationship in a human study of this size (Tr. at 971–72). I agree with the Bureau's position on this point. In reaching this conclusion, I note that Abbott has not produced any expert

testimony to lend scientific credence to its theory.

Finally, Abbott raises two concerns regarding possible confounding variables. First, Abbott questions the validity of using patients in the treated and Control I groups with "similar" rather than "identical" diseases (Abbott's Exceptions at 61). The Bureau's expert testimony, however, dismisses this Abbott concern (Tr. at 970–71), and I agree with this unrebutted evidence. Second, Abbott claims that the patients' exposure to diagnostic radiation and to non-alkylating drugs confounded the study's results (Abbott's Exceptions at 62), and that the ALJ's finding that cyclamate was the "causative factor" of chromosome damage (breaks) (ID at 27) rested upon assumptions unsupported by the record (Abbott's Exceptions at 63). A review of the record, however, has shown there also to be unrebutted expert testimony in the Bureau's favor on this issue of confounding variables (Tr. at 526), and I agree. The very purpose of using the Control I group (having similar diseases as the treated group and therefore similar exposure to diagnostic radiation and non-alkylating drugs) undoubtedly was to eliminate the very kind of confounding variables which Abbott is raising. Moreover, the ALJ's conclusion that cyclamate was the "causative factor" of chromosome breaks is more than adequately supported by the record (G–121 at 13; G–123 at 4; G–124 at 21–22).

I therefore conclude that the Bauchinger study presents statistically significant findings of breaks which are strongly suggestive of cyclamate's mutagenic potential.

5. *Negative Studies:* The administrative record also contains four *in vivo* cytogenetic studies which I have classified as "negative"—*i.e.*, the studies meet the minimum criteria set forth in Subsections C.2. and 3. above and found no statistically significant increase over controls in the types of chromosome aberrations which were scored. These studies were obtained from the bone marrow of Chinese hamsters (A–143), the spermatocytes of mice (A–151), and the spermatogonia of Chinese hamsters (A–716 and A–811, App. 19). As explained in Subsection F.1. above, however, the findings are insufficient to outweigh the suggestive *in vivo* cytogenetic findings just described.

a. *Brewen, et al.* (A–143). (1) *Study Design:* This study was performed on Chinese hamsters using CHA as the test compound. Three test groups of unspecified size were injected daily for three consecutive days with 50, 150 or 450 mg/kg body weight. Negative control animals were given identical regimens of distilled water. No concurrent positive controls were used. After sacrifice, cells were obtained from the bone marrow for analysis.

(2) *Study Results:* The authors reported no statistically significant increase in the treated animals over controls in terms of chromosome aberrations. These findings were based on analyses of either 200 or 400 cells per treatment group. The authors did note, however, that half of the treated animals at the highest dose level died before completion of the experiment.

The ALJ found that this study "evidenced no chromosomal damage" (*id.* at 26).

(3) *Analysis:* Abbott takes no exception to the ALJ's finding (Abbott's Exceptions at 57), and asserts that the study is important because it used a dose level five times that used by Legator (Abbott's Brief at 57). The Bureau criticizes the study because: (a) the size of the treated groups was not specified; and (b) no positive controls were used (Bureau's Brief at 85–86). The Bureau stressed that positive controls are especially necessary where, as here, a test animal whose sensitivities are not well known is used (*id.* at 86).

I agree with the ALJ and with Abbott that Brewen, et al. did not find a statistically significant increase in chromosome aberrations. The study is therefore "negative" in the general sense. However, I also agree with the Bureau that the study has shortcomings which reduce the weight to be accorded to it.

First, the authors' failure to specify the number of animals used raises a question which is not answered by the current record. Although I might be justified in rejecting this study altogether as "deficient" due to this shortcoming (since it is Abbott's burden to establish its proof to a reasonable certainty), the fact that 400 cells were analyzed at the middle dose level suggests that at least at that dose level the study size may have been sufficient. (Note, for example, that in the Majumdar and Solomon study, G–26, 10 animals and 300–350 cells were used per treatment group.) Thus, I consider this unknown fact to affect the weight but not the overall validity of this study.

Second, Brewen's failure to use a positive control also reduces the study's weight. While it is true that Brewen may have had adequate historical control data, none was presented, and, again, it is Abbott's burden to ensure this information is presented. I also note, however, that no other evidence exists suggesting that this study was insensitive (the Bureau's claim that the test animal is insensitive being unsubstantiated, especially in light of van Went-de Vries' positive findings with Chinese hamsters (G–45)). Therefore, consistent with the approach outlined in Subsection C.3.c. above, I find that Brewen's lack of a positive control reduces the study's weight but does not undermine its overall validity.

Finally, I consider it important to note that in the highest dose group (1350 mg/kg total dose), half of the animals died before the end of the experiment. I interpret this to mean that cells from these animals were not examined for mutagenicity. This inference is supported by the fact that only half as many cells (200) were examined in this dose group as compared to the middle dose group (400 cells). I therefore consider the results from this one dose level to be invalid because mutagenic effects could have been masked by the fatality of the dose.

I also note that having eliminated this upper dose level, Brewen's highest valid dose level (450 mg/kg total dose), although greater than Legator's (150 mg/kg), was well below van Went-de Vries' (600 mg/kg) highest total dose which also involved CHA. Thus, Brewen's findings do not pre-empt the suggestive studies in terms of dose size.

In conclusion, the Brewen study presents inconclusive evidence that CHA does not cause chromosomal aberrations in the bone marrow of Chinese hamsters. I would have more confidence in these results if Brewen had specified an adequate test population size, and if he had presented adequate positive control data.

b. *Cattanach, et al.* (A–151). (1) *Study Design:* This study was conducted on mice using CHA as the test compound. Different sized test groups were used. The first group of four mice were given daily CHA injections of 50 mg/kg body weight for five days. The second group consisted of eight mice which received daily CHA injections of 100 mg/kg body weight, also for five days. The negative control group, consisting of eight mice, received distilled water. No concurrent positive controls were used. The cells examined were spermatocytes and numbered 200 per animal (which total 800 for the first group and 1600 for the second and control groups). Only translocations (*i.e.*, exchange figures) were scored.

(2) *Study Results:* The investigators reported no statistically significant increase of translocations in either treated group over controls. In fact, no translocations at all were observed in the treated groups, although one was seen in the control group.

The ALJ found this study to have produced "negative" results (*id.* at 28).

(3) *Analysis:* Neither party has taken exceptions or commented in any detail on this study. The purpose of the study was to determine whether CHA induces exchange figures in mice. This type of study is important in evaluating the mutagenicity of a compound because, as the parties themselves agree, exchange figures are one means by which genetic abnormalities can be transmitted to future generations (Abbott's Brief at 49; Bureau's Brief at 100). One of the Bureau's hypotheses is that breaks are significant because they may lead to exchange figures. (Bureau's Brief at 101; see general discussion in Subsection F.2.c.(1) above). The Cattanach study, in essence, is designed to test that hypothesis with respect to cyclamate.

I agree with the ALJ that this study is negative, but I find that it has two internal shortcomings which must be noted. The first shortcoming is that the lower dose treatment group contained only four mice, and these findings must therefore be rejected due to that group's insensitivity. (See discussion in Subsection C.3.a. above.) This shortcoming, however, does not substantially reduce the weight given to the study as a whole because the findings of the higher dose group are valid. That group contained eight animals. This number is sufficiently close to the guideline to ten which I have followed (see Subsection C.3.a. and G–124 at 18) to be considered adequate, especially in light of the large number of cells (1600) analyzed.

The second shortcoming is the lack of a positive control. For the same reasons discussed immediately above in connection with the Brewen study, I find that this shortcoming reduces the weight but not the overall validity of Cattanach's findings.

I therefore find that this study presents inconclusive evidence that CHA does not induce exchange figures in mice. I emphasize the inconclusiveness of these findings because exchange figures are rare events that are difficult to detect (G–124 at 30; Tr. at 956). As Dr. Cattanach himself admitted:

> . . . but here a word of caution must be introduced. For the induction of translocations at least two breaks must occur in the same cell at the same time and the broken chromosomes must rejoin in such a way that each rearranged chromosome possesses I centromere. A failure to detect translocations is not therefore incompatible with spermatogonial chromosome breakage. It is clear that more work is needed.

(A–151 at 474). Given the apparent difficulty in detecting exchange figures, I would have more confidence in the results of this study if positive control data were available to validate the negative findings.

c. *Lorke, et al.* (A–716). (1) *Study Design:* This study was conducted on Chinese hamsters using CHA as the test compound. The CHA was administered orally to one group of eight hamsters at a dose of approximately 100 mg/kg body weight/day for five consecutive days. Negative and positive control groups of eight animals each were run concurrently. The negative controls were not dosed. The positive controls were given cyclophosphamide at a dose of 100 mg/kg body weight/day for five days. 100 spermatogonial cells from each animal were then analyzed.

(2) *Study Results:* The investigators scored the cells for three types of chromosomal aberrations: (a) cells containing aberrations, including gaps; (b) cells containing aberrations, without gaps; and (c) cells with translocations. No statistical difference between the treated and negative controls was found in any of these categories. In fact, no translocations were observed in either of these two groups. A statistically significant increase in all categories was observed in the positive controls when compared to the negative controls.

The ALJ found this study to produce "negative" results (*id.* at 28).

(3) *Analysis:* The Bureau criticizes this study on two principal grounds: (a) that the positive control values were unusually low, suggesting an insensitivity in the test (Bureau's Brief at 86–87; Bureau's Reply at 26); and (b) that Dr. Lorke employed an inappropriate method for statistical analysis (Tr. at 853). Abbott counters the Bureau's first argument by stating that the sensitivity of the Chinese hamster spermatognia were validated by earlier studies, and that the statistically significant positive control values validated the sensitivity of this particular study (Abbott's Brief at 62; Abbott's Exceptions at 66). As to the appropriateness of Dr. Lorke's statistical method, Abbott contends: (a) the chi-square method is appropriate; (b) the only testimony elicited by the Bureau is based on hearsay; and (c) the Bureau did not show that the results of the study would be any different if any statistical test had been used (Abbott's Brief at 62–63).

I agree with the ALJ and with Abbott that this is a "negative" study, but I also agree with the Bureau that the postive control values raise a question about the study's sensitivity that reduces the weight I would otherwise have given to it. The Bureau adduced testimony from two expert witnesses that the positive control values for cyclophosphamide reported by Lorke were well below the norm (Tr. at 851; G–124 at 19–20). Specifically, Dr. Green testified that cyclophosphamide dose levels employed by Lorke normally produce chromosomal aberrations in 20% to 40% of cells examined, whereas Dr. Lorke's values did not exceed 9% (Tr. at 851). I find the Bureau's testimony persuasive on this point, especially since Abbott produced no expert testimony to the contrary. Arguments by Abbott counsel miss the point for two reasons. First, even if the positive control values are statistically significant, those values can still be lower than would be anticipated. Second, even if Dr. Lorke validated the sensitivity of the Chinese hamster in earlier tests, it is still quite possible that something in the current study reduced the sensitivity of the results at issue. I therefore attribute less weight to this study than would Abbott.

I do not, however, find merit in the Bureau's criticism of Dr. Lorke's statistical analysis. As Abbott correctly points out, the Bureau did not present any evidence demonstrating that Dr. Lorke's findings would not have reached statistical significance if another method had been used. Moreover, the Bureau's witness on this point, Dr. Green, is not himself a statistician and instead based his testimony on his conversations with other, unnamed persons (Tr. at 853) who were not available for cross-examination.

I therefore conclude that this study presents inconclusive evidence that CHA does not produce chromosomal aberrations in the spermatogonia of Chinese hamsters. I would have more confidence in these results if the positive control data had been within the normal range for that compound.

d. *Lorke, et al.* (A–811, App. 19). (1) *Study Design:* This study was also conducted on Chinese hamsters but used sodium cyclamate (rather than CHA) as the test compound. Six hamsters were orally given 2,000 mg/kg body weight/day of sodium cyclamate for five days. A negative control group of six hamsters were not dosed. Two sets of positive controls were also run concurrently. In the first, six hamsters received 1,000 mg/kg body weight/day of trimethylphosphate orally for five days. In the second, six other hamsters received 250 mg/kg body weight/day of cyclophosphamide orally, also for five days. 600 spermatogonial cells (100 per hamster) were analyzed in both the sodium cyclamate group and negative control group. The total number of cells analyzed varied for the positive controls.

(2) *Study Results:* The cells were scored in the same three categories described above in connection with the other Lorke study (A–716). The

investigators found no statistically significant increase in any category in the treated group over the negative controls. Analyses of the positive control groups yielded statistically significant increases in all categories when compared to the negative controls.

The ALJ found these results to be "negative" (*id.* at 28).

(3) *Analysis:* The parties apparently directed their comments made in connection with the other Lorke study (A–716) to this study as well. In addition, the Bureau criticizes this particular study for only using six animals per group (Bureau's Brief at 86–87).

I adopt here by reference my earlier analysis of the other Lorke study (A–716) with respect to the positive control and statistical methodology issues.

In addition, I agree with the Bureau that the number of animals per group in this study (six) is too small. Under the criteria set forth in Subsection C.3.a. above, I would be justified in totally rejecting this study as deficient because its sensitivity is too low. However, this low sensitivity is partially offset by the extremely high dose used, 2,000 mg/kg body weight/day for five days. (Compare, for example, Majumdar and Solomon (G–26) which used doses of calcium cyclamate on Mongolian gerbils of 100 mg/kg body weight/day for five days.)

I therefore conclude that this study by Lorke, et al. presents inconclusive evidence that sodium cyclamate does not induce chromosomal aberrations in the spermatogonia of Chinese hamsters. I would have more confidence in these negative results if the positive control values for cyclophosphamide had been in the normal range and if Dr. Lorke had used more animals per group.

(4) *Other Matters:* For the record, I note that the sodium cyclamate portion of the study just discussed as A–811, App. 19 is also contained in the record as A–827, App. 13.

6. *Deficient Studies.* I have classified the following 15 studies as "deficient" because they fail to meet the minimum criteria set forth in Subsection C.2. and 3. above. Accordingly, these studies are not entitled to any weight.

I note that all but the first of these studies (Collin, G–27) were classified by Abbott as being negative, including the two studies (Dick, A–177 and Ford, A–297) claimed to be exact replicates of Legator's study (G–9). thus, the weakness of Abbott's position on the mutagenicity issue is due in large part to the high number of deficient studies on which it relies.

For organizational purposes, I have arranged these studies according to the type of cells analyzed: bone marrow, blood, or sperm cells.

a. *Bone Marrow Studies:* (1) *Collin* (G–27).

(a) *Study Design:* This was a rat feeding study using sodium cyclamate as the test compound. The test population consisted of four rats. The dose size was stated in terms of being 5% of the feed. The length of exposure ranged from two to six months. The number of cells analyzed was not specified. There is also no mention of a negative control.

(b) *Study Results:* The investigator reported chromosomal damage, including breaks, but no data were presented and no analysis of statistical significance was reported.

The ALJ characterized this study as "positive" (*id.* at 35), based upon results which included "chromosome breaks, the absence of satellites on chromosomes and numerous achromatic areas" (*id.* at 26).

(c) *Analysis:* Abbott takes exception to the ALJ's statement that Collin found "numerous achromatic areas," claiming that this finding came from the *in vitro,* not *in vivo,* portion of the study (Abbott's Exceptions at 57). Abbott also criticizes the study's design for using too few animals and analyzing too few cells (*id.*). The Bureau does not discuss this study in any detail, but merely lists it as one of a group that "produced clear positive results" (Bureau's Brief at 75).

Unlike the other studies with positive findings discussed above, Collin's work is deficient and warrants no weight at all. The primary deficiency in this study is that it contains no meaningful presentation of data, and therefore insufficient information exists to evaluate it properly. I also note that the study contains no comparison, in terms of statistical significance, between the findings of the treated group and a negative control group. Given this conclusion, I need not reach more specific criticisms raised by Abbott.

(2) *Dick, et al.* (A–177). (a) *Study Design:* This study was conducted on Holtzman rats using CHA as the test compound. 14 rats were given daily injections of CHA base at a dose of 50 mg/kg body weight for five days. 17 rats were given equivalent doses of CHA–HCL. A negative control group of 12 rats were injected with water. In addition, two positive control groups were dosed for two days. The first group of 10 rats was injected with triethylenemelamine (TEM) at a dose of 0.5 mg/kg body weight. The second group of 8 rats received injections of tris-(2-methly-1-aziredinyl) phosphine oxide (METEPA) at a dose of 20 mg/kg body weight. Cells from the bone marrow were analyzed. The cells numbered 700, 850, 600, 470 and 400, respectively, for the five groups.

(b) *Study Results:* Cells were scored for two categories of chromosomal aberrations: (i) gaps and breaks (combined); and (ii) reunion figures and fragmented metaphases (combined). The investigators reported *fewer* gaps and breaks in the treated groups than in the negative control. No reunion figures or fragmented metaphases were found in any of these three groups. For the positive controls, there was a statistically significant increase for METEPA over negative controls in both categories. For TEM, however, there was a statistically significant increase only for the second category. For gaps and breaks, the findings for TEM were virtually the same as for the negative control (A–177 at Table III).

The ALJ found this study to be negative despite "appear[ing] to have used the exact protocol used by Dr. Legator" (*id.* at 25; see also *id.* at 35). Thus, the ALJ concluded that Dr. Dick's findings helped place Dr. Legator's in doubt (*id.*).

(3) *Analysis:* Abbott agrees with the ALJ (Abbott's Exceptions at 72; see also Abbott's Brief at 56–57). The Bureau criticizes the Dick study on two related grounds. First, the Bureau challenges the validity of Dick's grouping of "reunion figures" and "fragmented metaphases" together, arguing that the former are the best indicators of heritable genetic damage while the latter are the least reliable (Bureau's Brief at 89). Assuming this to be true, the Bureau attempts to eliminate the fragmented metaphases from the incidence found for the positive control, TEM, and then asserts that the remaining incidence for reunion figures for TEM is far too low (*id.* at 89–90).

I agree with the ALJ to the extent that Dr. Dick used a very similar protocol as Legator, et al. (G–9) (bone marrow portion). Both investigators tested CHA on male Holtzman rats using five daily injections. Although Legator used five dose levels, Dick matched his highest dose level. Thus, were Dick's findings credible, they would indeed place Legator's findings in some doubt.

However, as explained above in my discussion of the bone marrow portion of the Legator study, the Dick study has a fatal flaw. According to Table III of A–177, the incidence of breaks and gaps for the positive control, TEM is virtually identical to that of the negative control (water). This means that some unknown factor severely compromised the experiment's ability to detect breaks and gaps. Thus, since Dick was not even able to detect an increased incidence of breaks and gaps where they should have been, *a fortiori* no conclusion can be

drawn from Dick's failure to observe breaks and gaps in the test compound, CHA. (See general discussion of positive controls in Subsection C.3.c. above.)

Due to this finding, I need not reach the Bureau's argument regarding the insufficiency of Dick's positive control values.

I therefore conclude that this study is entitled to no weight at all because Dick's inability to detect a statistically significant increase in breaks and gaps in the positive control, TEM, invalidates any negative findings.

(3) *Ford, et al.* (A–297). (a) *Study Design:* This study was also conducted on Holtzman rats using CHA as the test compound. Three treatment groups with three animals per group were used. Each received daily injections for five days. The first group received 50 mg/kg body weight of CHA base, and the other two groups were given an equivalent amount of CHA–HCL (obtained from two different suppliers). A negative control group was given water. Two positive control groups, using TEM and METEPA, were each dosed for two days. Cells from the bone marrow were analyzed, numbering 150 per group (250 for TEM group).

(b) *Study Results:* The investigators reported their findings as simply "negative" for the CHA treated groups and "positive" for the positive controls (See Table IV). No statistical analysis was mentioned in the text or presented in table form.

As with the Dick study, the ALJ found that Dr. Ford "appear[ed] to have used the exact protocol used by Dr. Legator (Ex. No. G–9) but ha[s] failed to replicate his results" (*id.* at 25–26; see also *id.* at 35).

(c) *Analysis:* Abbott agrees with the ALJ (Abbott's Exceptions at 72; 56; see also Abbott's Brief at 56–57). The Bureau has not made specific comments on this study.

I find this study to be deficient because the investigators used only three animals per group. As explained in Subsection C.3.a. above, this test is simply too intensitive for any confidence to be placed in its negative results. Additionally, I find that Dr. Ford did not present statistical information demonstrating that the treated groups were statistically negative and the positive controls statistically positive. Along this line, I note that in at least one category, "Average percent breaks," the incidence for the CHA base group (1.3) was virtually the same as for the positive control, TEM (1.2) (A–297 at Table 1). This study, therefore, is entitled to no weight at all.

(4) *Friedman, et al.* (A–195). (a) *Study Design:* This was a cyclamate feeding study conducted on Holtzman rats. One group of ten rats received 1% cyclamate as part of their feed for an unspecified period of time. A negative control group consisted of six rats. No positive control was used. An unspecified number of bone marrow cells were analyzed.

(b) *Study Results:* The investigators reported that the "range of values [found for breaks] is considered to be well within the expected 'background' range of values for normal untreated males of this strain and age" (A–195 at 754). No data of any consequence was presented.

The ALJ found these results to be "negative" (*id.* at 26).

(c) *Analysis:* Abbott agrees with the ALJ (Abbott's Exceptions at 56; Abbott's Brief at 56). The Bureau makes no specific comments on this study.

I find this study to be deficient because there are virtually no data which would enable me to evaluate it adequately. Neither is any kind of statistical analysis presented, nor a positive control used. (See discussion in Subsections C.2 and C.3 above.) This mutagenicity research was clearly peripheral to the more fully presented carcinogenity experiment. Indeed, the authors themselves characterized the mutagenicity portion as "limited" (A–195 at 752). I therefore conclude that this study is entitled to no weight.

(5) *Khera, et al.* (A–222). (a) *Study Design:* The cells analyzed in this study were taken from female Wistar rats used in a reproduction study. For cytogenetic purposes, two groups of five rats each were given cyclohexylamine sulfate (CHS) as part of their feed for an extended period ranging from 52 to 67 days. The dose, stated as a percentage of the feed, ranged from 5.56% to 11.12%. A negative control was given distilled water. No positive control was used. 100 bone barrow cells from each rat were analyzed.

(b) *Study Results:* The investigators reported "no abnormality in distribution of chromosome number or incidence of structural aberrations" (A–222 at 267). No additional commentary or data was presented.

The ALJ found this study to be "negative," but noted that it had been criticized by the Bureau (*id.* at 26).

(c) *Analysis:* The Bureau's principal objections were that: (i) the test population was too small; and (ii) no details are given in terms of data, background rate, or how the cells were scored (Bureau's Brief at 85). Abbott admits that "(t)here are reasons for giving less weight to this study," but contends that the investigators did present adequate data (Abbott's Exceptions at 58).

I find this study to be deficient for several reasons. First, the test population of only five rats per group is too small. Second, no data are presented (G–124 at 19). Third, no positive control was used. Therefore, for the reasons explained in Subsection C.3. above, I attribute no weight at all to this study.

(6) *Oser, et al.* (A–274). (a) *Study Design:* This was a multigeneration feeding study on Wistar rats using CHA as the test compound. The F_0 generation rats were fed doses of 50 or 150 mg/kg body weight as part of their diet for periods of 6, 12 or 18 months. Group sizes ranged from three to five rats. Negative control groups of the same size were used for each dose size and period, but no positive controls. Bone marrow cells were analyzed, averaging about 250 per group. (See Tables 49–50.)

Cells were also analyzed from the offspring. For the F_1 and F_2 generations, fetal tissue was taken at Caesarian section. For the F_3 generation, bone marrow was taken from weanlings. Negative controls were used for each group, but no positive controls. Group and cell populations were as follows: F_1, 10 rats and 250 cells; F_2, 4 rats and 100 cells; F_3, 6 rats and 300 cells. (See A–274 at 25a and Table 51).

(b) *Study Results:* Cells were scored only for the number and percent of abnormalities. The only abnormalities found were in cells "exhibiting a subnormal number of chromosomes" but these "occurred in no greater proportion in the test groups than in the controls" (A–274 at 25c). The investigators explicitly stated that "[n]o abnormalities in chromatin morphology [*e.g.*, breaks, gaps, exchange figures] were observed" in either treated or control groups (A–274 at Tables 49–50 and 51).

The ALJ found this study to be "negative" (*id.* at 26).

(c) *Analysis:* Abbott agrees with the ALJ's finding (Abbott's Exceptions at 56; Abbott's Brief at 56–57). The Bureau makes no comments on this study.

Although this study has some interesting aspects in its design (*i.e.*, multi-generation analysis, long duration of exposure), the study has two fatal weaknesses which render it deficient. First, with the exception of the F_1 generation, all test groups had six or fewer animals. Second, the fact that no chromosome abnormalities (such as breaks, gaps or exchange figures) were found in any of the groups, treated or control, raises a serious question about the study's sensitivity, especially considering the long duration of exposure. This is in contradiction to virtually all the other credible studies and is a prime example of where concurrent positive controls are

essential. (See general discussion in Subsection C.3.c. above.) I therefore conclude that this study is entitled to no weight at all.

b. *Blood cell studies (animals).* 1. *Mostardi, et al.* (A–264).

(a) *Study Design:* This study was conducted on Wistar rats using CHA as the test compound. There were two treatment groups of three rats each. The first group received CHA injections at a dose of 20 mg/kg body weight; the second group at a dose of 50 mg/kg body weight. Injections were given daily for five consecutive days during each of seven weeks. Blood was drawn 24 hours after the fifth injection of each week. A negative control group of three animals was also used. A positive control was not. 50 metaphase spreads were analyzed for each group.

(b) *Study Results:* The investigators reported "no discernible differences" between the treated and controls in terms of both "the number of abnormal spreads and percent of cells with abnormal chromosomes" (A–264 at 316). However, no statistical analysis was presented. Nor did the investigators explain (beyond the characterization "abnormal") how the cells were scored.

The ALJ found this study to be "negative" (*id.* at 26).

(c) *Analysis:* Abbott agrees with the ALJ's finding (Abbott's Exceptions at 59; Abbott's Brief at 57; 59). The Bureau criticizes this study on the grounds that: (i) the test population was too small; (ii) too few cells were analyzed; (iii) the investigators did not specify what chromosome aberrations were scored for; and (iv) there were "enormous" variations in the negative controls from week to week (Bureau's Brief at 84–85).

I find this study to be deficient because the number of rats per group (3) is too small, the presentation of data is inadequate in that the investigators do not state for which chromosome aberrations they scored the cells, and no statistical analysis is presented. (See G–124 at 18–19 and general discussion in Subsections C.2. and 3. above.) I also note that these findings are not confirmed by positive controls. I therefore conclude that no weight at all should be attributed to this study.

(2) *Lisker and Cobo* (A–241). Although the ALJ found that this study "failed to show any positive effects" (*id.* at 27), it appears in the record only in a Spanish version. This is not an acceptable form for my evaluation, especially since Abbott has presented no expert testimony favorably interpreting it. Moreover, according to the Bureau's Brief at 85, this study employed only two animals (rabbits) per group. I therefore conclude that no weight at all should be attributed to it.

c. *Blood cell studies (humans).* (1) *Dick, et al.* (A–177).

(a) *Study Design:* In this experiment, four persons (two men and two women) were given sodium cyclamate capsules at a dose of 5 g per day for the men and 4 g per day for the women, for a total of four days. These persons had previously been tested and found to be able to convert cyclamate to CHA. In addition, a similar group of non-converters were placed on the same dosing regimen. Urine analyses were conducted throughout the experiment to verify whether the "converters" and "non-converters" maintained that status. A negative control group was also established. Blood samples were obtained and at least 100 metaphases (cells) were examined for each sample. The cell slides were coded so that the person analyzing them did not know whether they came from a treated or control group.

(b) *Study Results:* The investigators reported no increased incidence of chromosomal abnormalities in either the converters or non-converters. What abnormalities were found were predominantly gaps, with a few breaks. No exchange figures were observed. However, one of the "converters" acted as a "non-converter" during the course of the experiment.

The ALJ found this study to be "negative" (Id at 27).

(c) *Analysis:* Abbott agrees with the ALJ's finding (Abbott's Exceptions at 60–61; Abbott's Brief at 58–59). The Bureau criticizes this study on three principal grounds: (i) small test population; (ii) small cumulative dose when compared to Bauchinger (J–1); and (iii) inadequate presentation of raw data (Bureau's Brief at 88, relying upon G–124 at 21; Tr. at 971–72). In its exceptions, Abbott defends the adequacy of the data as presented in Tables II and III of the study (Abbott's Exceptions at 64).

I find this study to be deficient because of the small population size which consisted of only three subjects that were demonstrated converters (see G–124 at 21 and Tr. at 971–72; see general discussion in Subsection C.3.a. above). I need not reach the issue of dose size because that would go to the weight of the study had it met the minimum criteria. I also do not reach the issue of the adequacy of the data in Table II of A–177 (Table III contains data on the rat portion of the experiment). I therefore conclude that no weight at all should be attributed to this study.

(2) *Coulson* (A–703). (a) *Study Design:* This study was conducted using prisoners as test subjects. Sodium cyclamate capsules were administered orally for either eight or thirteen weeks. For the eight week period, group and dose sizes were as follows: 5 subjects, 5 g/day; 5 subjects, 10 g/day; 2 subjects, 3 g/day (after having 16 g/day for 6 days); and 6 subjects, placebo. For the thirteen week period, these were: 2 subjects, 5 g/day; 3 subjects, 10 g/day; 3 subjects, 3 g/day (after having 16 g/day for 6 days); and 4 subjects, placebo. Five of the subjects were used for both time periods. Approximately 10 blood cells from each sample were examined.

(b) *Study Results:* The authors reported simply, "No chromosomal abnormalities were observed" (A–703 at final page (unnumbered)). No mutagenicity data was presented.

The ALJ found this study to be "negative" (*id.* at 27).

(c) *Analysis:* Abbott agrees with the ALJ's finding (Abbott's Exceptions at 60–61; Abbott's Brief at 58), emphasizing that Coulson used dosage levels comparable or exceeding that of Bauchinger (J–1) and well above the use level proposed by Abbott in its food additive petition (Abbott's Brief at 59–60). The Bureau criticizes this study on four grounds: (i) no data was presented; (ii) too few cells (10) were analyzed per subject; (iii) the investigators did not specify how the cells were scored; and (iv) the study is unpublished and therefore has never been subject to peer review (Bureau's Brief at 90).

I find this study to be deficient because there are insufficient data presented for evaluation. (See general discussion in Subsection C.3.b. above.) In fact, I have reviewed this study in detail and have found no data at all relating to mutagenicity. I do note that one expert stated that he thought some data were presented (Tr. at 526–27). That conclusion, however, was based on an admittedly cursory review of the study conducted that same day (*id.*). A careful review disclosed that the numerous tables containing blood analyses data related to concentrations of various compounds in the blood (*e.g.*, protein-bound iodine, thyroxiniodine, free thyroxin, thyroxin-binding globulin, and plasma cortisol) rather than findings of chromosome abnormalities. The abundance of this irrelevant data strongly suggests that the chromosome analysis was a peripheral part of this study. This may explain why chromosome data were not presented.

Moreover, I find the fact that the investigators found no chromosome abnormalities at all raises a serious question about the study's sensitivity, especially considering the long duration of exposure. As noted above in

connection with the bone marrow portion of the Oser study (A–274), the absence of any chromosome abnormalities contradicts the findings of virtually all the credible studies presented in this record and therefore presents a second, independent basis for classifying the study as deficient.

Due to these two major flaws, I need not reach the other objections raised by the Bureau. I conclude that no weight at all should be attributed to this study.

d. *Sperm cell studies.* (1) *Ford (A–297), Friedman (A–195), and Oser (A–274) studies.* These three studies have already been discussed in connection with the deficient bone marrow studies, Subsection F.6.a. above. In addition, each investigator also analyzed sperm cells from male rats. The ALJ found the sperm cell portions of these studies to be "negative" (*id.* at 28). Abbott agrees, emphasizing that these and other studies rebut Legator's (G–9) positive sperm cell findings (Abbott's Exceptions at 65–66; Abbott's Brief at 61).

The Bureau offers no additional comments. I find that, except for the difference in cells analyzed, the design and reporting of the sperm cell portions of these studies are identical to that of the bone marrow portions. I therefore adopt and incorporate here my previous discussion of these three studies and conclude that, for the same reasons stated in Subsection F.6.a. above, each is deficient and thus should be accorded no weight at all.

(2) *Kaziwara, et al.* (A–217). (a) *Study Design:* This study was conducted on adult male mice using C.H.A. as the test compound. An unspecified number of mice were injected with a single dose of CHA, either at 40 mg/kg body weight or 80 mg/kg body weight. No mention was made of either a negative or positive control. Cells analyzed were spermatogonia and primary and secondary spermatocytes. Ten cells were analyzed per group.

(b) *Study Results:* The investigators reported only that "[n]o chromosome aberrations were observed in male reproductive cells of mice treated with either 40 or 80 mg/kg of CHA " (A–217 at 6). No data was presented. No statistical analysis was presented. No explanation was given as to how the cells were scored.

The ALJ found this study to be "negative" (*id.* at 28).

(c) *Analysis:* Abbott agrees with the ALJ's finding (Abbott's Exceptions at 66; Abbott's Brief at 61). The Bureau has not commented on this study.

I find this study to be deficient for several reasons. First, no data are presented to enable an adequate evaluation of the study. Along this line, I note that the cytogenetic portion of this study was clearly peripheral to the larger teratology portion, which may account for this shortcoming. Second, no statistical comparison between the treated and negative controls was presented. In fact, there is no indication that a negative control was even used. Third, the test population, although unspecified, appears to be grossly inadequate. Only ten cells were analyzed per dose level which suggests that only one or two mice were used per group. Moreover, this number of cells is "totally unacceptable" (Tr. at 527). I also note that no positive control was used. (See general discussion in Subsection C.2. and 3. above.) I therefore conclude that no weight at all should be given to this study.

(3) *Leonard and Linden* (A–240). (a) *Study Design:* This was a sodium cyclamate feeding study conducted on mice. The cyclamate was added to the drinking water at concentrations of 2.667 g/liter, 5.334 g/liter, or 10.668 g/liter and given to the mice for periods of 30, 60 or 150 days. One mouse was used for each dose level and time period. Negative controls consisting of one mouse per time period were also established. No positive control was reported. The investigators examined 200 dividing spermatocytes for each mouse.

(b) *Study Results:* The investigators "detected no evidence of chromosome anomaly. The rate of univalents was practically the same (±5%) in the different groups" (A–240 at 1–2). No data relating to the chromosome analysis was presented.

The ALJ found this study to be "negative" (*id.* at 28).

(c) *Analysis:* Abbott agrees with the ALJ's finding (Abbott's Exception at 66; Abbott's Brief at 61). The Bureau makes no specific comments on this study.

I find this study to be deficient for the following reasons. First, the size of the test population (one animal per dose level per time period) is totally inadequate. Second, no data relating to the chromosomal analysis is presented so as to allow me to make a proper evaluation. Finally, I note that: (i) it is unclear whether a proper statistical analysis was performed; and (ii) no positive controls were used to validate the findings. (See general discussion in Subsection C.2. and 3. above.) I therefore conclude that no weight at all should be given to this study.

7. *Additional Support: In Vitro Cytogenetic Studies:* a. *Summary:* The parties also submitted *in vitro* cytogenetic experiments performed by 13 different investigators. Like the *in vivo* studies of this class, *in vitro* cytogenetic experiments are designed to measure a test compound's effects upon chromosomes (*i.e.*, breaks, gaps, and exchange figures). The principal difference between *in vitro* and *in vivo* studies are that *in vitro* experiments are performed using cells in test tubes rather than live animals (A–800 at 6; G–124 at 10).

The parties agree that the information derived from *in vitro* studies is of limited value. The major limitation of *in vitro* cytogenetics is that cells in culture media represent an artificial setting which cannot imitate a live animal's metabolism (G–124 at 10 and 30; see A–800 at 7). Because of this limitation, *in vitro* studies serve merely as initial screens to determine if a compound is "active" or "inactive" from a mutaggenicity standpoint (G–124 at 25 and 30; Abbott's Brief at 44). Positive findings in *in vitro* cytogenetic studies are therefore insufficient, by themselves, to declare a compound a mutagen. Such findings can, however, buttress more definitive *in vivo* findings, if present. That is precisely the situation here.

A review of the *in vitro* cytogenetic studies has shown that the evidence, taken as a whole, strongly suggests that cyclamate is "active" from a mutagenicity standpoint. Several studies found a statistically significant increase of breaks (G–10, G–11, G–25 (CHA portion), G–33, G–35, G–39, G–46, and A–722), and one study found such an increase in exchange figures (G–35). Moreover, three of these studies found a dose response (G–25, G–33 and A–722). These findings outweigh the negative ones found in the studies relied upon by Abbott (A–143, A–205 (calcium cyclamate portion), A–259, and A–300). I therefore conclude that the *in vitro* cytogenetic studies provide additional support for my conclusion that Abbott has not shown that there is a reasonable certainty that cyclamate does not cause heritable genetic damage.

b. *The Studies' Findings:* Because of the limited utility of the *in vitro* results, I will discuss these studies only briefly.

(1) *Suggestive Studies:* The record contains 7 studies which found a statistically significant increase in chromosome damage which may reasonably be attributed to cyclamate or its metabolites.

(a) Stone, et al. (G–10) found that calcium and sodium cyclamate caused a statistically significant increase in breaks in human blood cells at concentrations of 250–500 mcg/ml.

(b) Stoltz, et al. (G–11) found that cyclamate, CHA and N-OHCHA each caused a statistically significant increase in chromosome aberrations (primarily breaks and gaps) in human blood cells at concentrations of 10^{-3},

10^{-4} and 10^{-5} (for cyclamate, equivalent to approximately 179, 17.9 and 1.79 mcg/ml, respectively; for CHA, 99, 9.9 and 0.99 mcg/ml; and for N-OHCHA, 115, 11.5, and 1.15 mcg/ml).

(c) Green, et al. (G–25) found that CHA caused a statistically significant increase in breaks in rat-kangaroo cells at concentrations of 50, 100 and 500 mcg/ml. A dose response trend was also observed (see Table 2).

(d) Kristoffersson (G–33) found that cyclamate caused a statistically significant increase in breaks and gaps in Chinese hamster cells at concentrations between 100 and 1,000 mcg/ml. A dose response trend was also observed (G–33 at 278).

(e) Tokumitsu (G–35) found that sodium cyclamate caused a statistically significant increase in breaks and exchange figures in human blood cells at a concentration of 0.01 M (approximately 2000 mcg/ml).

(f) Perez-Requejo (G–37; A–722) found that sodium cyclamate caused a statistically significant increase in chromosome aberrations (primarily breaks and gaps) in human blood cells at concentrations of 4.5 and 9.0 mg/ml (equal to 4500 and 9000 mcg/ml). A dose response was also found (A–722 at 5).

(g) Ebenezer (G–39) found that sodium cyclamate caused a statistically significant increase in chromosome aberrations (breaks, gaps and fragments, grouped together) in human blood cells at concentrations of .02 and .04 mg/ml (equal to 20 and 40 mcg/ml).

(2) *Negative Studies:* The record also contains four studies which found no statistically significant increase in chromosome damage which may reasonably be attributed to cyclamate or its metabolites.

(a) Brewen, et al. (A–143) found no CHA or N-OHCHA induced increase in chromosome aberrations in human blood cells at concentrations of 20, 100 and 500 mcg/ml CHA or 25, 50, 100, 200 and 250 mcg/ml N-OHCHA.

(b) Green, et al. (G–25) found no calcium cyclamate induced increase in chromosome breaks in rat-kangaroo cells at concentrations up to 200 mcg/ml.

(c) Shamberger, et al. (A–300) found no significant increase in breaks in human blood cells treated with sodium cyclamate treated in concentrations of 100 mcM (approximately 20 mcg/ml).

(d) Meisner, et al. (A–259), found no statistically significant increase in breaks in human fibroblasts after exposure to cyclamate in a concentration of 500 mcg/ml.

(3) *Deficient Studies:* Three studies, Schoeller et al. (G–18), Dixon (G–34) and Lederer, et al. (G–46; A–235) are deficient because they do not present sufficient data for a full evaluation (see Subsection C.3.b. above). Accordingly, they have been eliminated from consideration.

(c) *Analysis:* The ALJ made the following conclusion with respect to the *in vitro* cytogenetic evidence:

> Most of the *in vitro* cytogenetic studies, including the tests on human leukocytes, human lymphocytes and kangaroo rat cells produced significant positive results of serious chromosomal aberrations (Ex. Nos. G–11, G–17, G–25, F–33, G–34, G–35, G–39). In addition, a statistically significant increase in chromosome breaks and gaps and dose-dependent results were found in the studies on Chinese hamster cells, Chinese hamster fibroblasts and human fibroblasts (Ex. Nos. G–17, G–33, G–34). Even if the incidence of breaks and gaps does not represent serious genetic damage a contention with which many scientists disagree—the presence of a statistically significant effect cannot be disregarded.

(*id.* at 34; see also *id.* at 21–23).

Abbott raises three types of exceptions to the ALJ's findings. First, Abbott challenges the ALJ's characterization of several of the studies' findings (Abbott's Exceptions at 43–45 and 68–69). As is evident from my description of these studies' findings, I agree with Abbott that only Tokumitsu (G–35) found a statistically significant increase in exchange figures. Moreover, my finding that the Dixon study (G–34) is deficient and that the Meisner study (A–259) is negative dismisses any concerns that Abbott may have with the ALJ's characterization of those results. I disagree, however, with Abbott's attempt to dismiss the results of the Stoltz study (G–11) (increased incidences of breaks) due to "cytotoxicity." Cytotoxicity means cell death. As explained above, it is true that chromosomes with breaks will sometimes die rather than repair themselves or join with other broken chromosomes to form exchange figures (see Subsection F.2.a. above). However, findings of breaks are nevertheless biologically significant for the reasons set forth in detail in Subsection F.2.c. above. Abbott's exception that this study is insignificant due to observed cytotoxicity is therefore without merit.

Second, Abbott asserts more directly that breaks do not constitute serious genetic damage (Abbott's Exceptions at 68–69). Again, I have already addressed this issue extensively in subsection F.2.c. above and need not repeat it here.

Finally, Abbott claims that the positive findings were achieved only through the use of massive doses which are not relevant to human experience (Abbott's Exceptions at 46; Abbott's Brief at 51; A–800 at 8). In response, the Bureau asserts that findings from *in vitro* studies are relevant only to the question of whether the compound is "active" or "inactive" (Bureau's Brief at 100; G–124 at 30).

I agree with the Bureau on this point. As Dr. Legator explained:

> I know of no way in which one can with any degree of validity determine dosages from *in vitro* tests and apply them to *in vivo* studies. When we talk about *in vitro* testing we, of course, have a very artificial situation that does not occur in the animal system. The only conclusions that one can make on the basis of *in vitro* studies is that the compound is active or inactive. To try to read anything further into the results, for example, to try to make quantitative extrapolations, is probably to push the method far beyond its possible usefulness.

(G–124 at 30).

I therefore conclude that the suggestive *in vitro* cytogenetic studies of record are relevant and provide strong evidence that cyclamate and its metabolites are "active" in terms of mutagenicity. Accordingly, these studies provide additional support for my conclusion that cyclamate has not been shown to a reasonable certainty not to cause heritable genetic damage.

G. Other Studies Insufficient to Outweigh Suggestive Evidence

1. *Summary:* In addition to the *in vivo* cytogenetic studies discussed above, the record contains three other types of *in vivo* mutagenicity studies: (a) host-mediated assay; (b) dominant lethal assay; and (c) drosophila. Several additional *in vitro* tests were also performed. The studies from each of these groups produced predominantly negative results. These findings, however, are insufficient to outweigh the strongly suggestive *in vivo* cytogenetic findings discussed above because known mutagens have been found to show mutagenic effects in some *in vivo* test methods but not in others (G–124 at 9–10 and 31; Tr. at 933–34; Tr. at 498–501; Tr. at 717–18 and 734; see discussion in Subsection D.1. above), and because the last group of *in vitro* studies are by their very nature insufficient to outweigh suggestive *in vivo* findings (see Subsection A.3 above).

2. *Host-Mediated Assay:* The ALJ made the following findings with regard to the four host-mediated assay studies:

> The *host-mediated assay* is a mutagenicity test which involves placing a known indicator organism into the interperitoneal cavity of a treated animal, considered a host. The host animal is then treated with the test compound, in this case, sodium or calcium cyclamate or CHA. Upon conclusion of the testing, the indicator is removed and examined for mutations. The primary

advantage of the host-mediated assay is that it provides the sensitivity of *in vitro* tests combined with exposure to a metabolic process as in *in vivo* tests. Mice were used as the host animal and either Salmonella typhimirium or Serratia morescens was used as an indicator. Chinese hamsters were used in one experiment using human leukocytes as an indicator. The results of testing both cyclamate and CHA were negative in the four studies conducted (Ex. Nos. A–143, A–268, A–325, A–375).

(*Id.* at 24). Abbott takes no substantive exceptions to this portion of the ALJ's opinion. Abbott does note, however, that all the sentences but the final one "appear without cites." Abbott therefore suggests that these constitute "non-substantive statements" rather than "finding[s]" (Abbott's Exceptions at 50). The Bureau makes no comments on these studies.

I adopt the above-quoted statement of the ALJ and agree that these four studies are negative. I also find that the ALJ's description of the host-mediated assay method is amply supported by the texts of the studies themselves, and that therefore the ALJ's citations are adequate. Thus, these statements constitute substantive findings. As noted above, however, negative studies using the host-mediated assay are insufficient to outweigh the suggestive cytogenetic experiments (see discussion in Subsection D.1. above).

3. *Dominant Lethal Assay.* a. *Description of Test Methods:* A dominant lethal assay is a study designed to detect geneticall caused deaths in the next (F_1) generation. The study is conducted in three principal steps: (1) dosing the animals; (2) allowing the animals to mate; and (3) examining each female's uterus for evidence of fetal deaths (G–124 at 11).

In the dosing stage, usually only the males are treated with the test compound (*e.g.*, G–29 and A–151), although sometimes both sexes (*e.g.*, A–827. App. 17) or only the females (*e.g.*, A–827, App. 11) are dosed. Dosing may be in single dose (*e.g.*, A–827, App. 11), several doses over a few days (*e.g.*, G–29), or many doses over several weeks (*e.g.*, A–827, App. 17).

After mating is completed, each female's uterus is examined when the animal reaches mid-pregnancy. The most important factor to be looked for is called "post-implantation loss." This means that an embryo has died *after* the egg has implanted itself into the uterus. Embryotic death may be observable either as a dark spot, called a "resorption" (or "deciduomata"), or as a recognizable embryo which is no longer viable (A–827, App. 9 at 8; A–827, App. 15 at 7). For example, if a subject female has ten implanted embryos, three of which later died, the subsequent examination of the uterus will reveal seven live embryos and a total of three dead embryos or resorptions. The mutagenic significance of post-implantation loss is that it is caused by chromosome damage, such as exchange figures (G–29 at 128; A–151 at 472).

A second, less significant factor to be looked for is called "pre-implantation loss." This means that an embryo has died *before* it has implanted itself in the uterus. This figure is obtained by subtracting the number of implant sites in the uterus (both viable and non-viable) from the number of "corpora lutea" in the ovaries (*i.e.*, sites from where eggs were shed) (A–827, App. 9 at 9; A–827, App. 15 at 4). Pre-implantation loss is less significant from a mutagenicity standpoint because, given the state of scientific knowledge, it is not certain that such losses are necessarily due to genetic damage (A–827, App. 15 at 8; G–121 at 14).

The dominant lethal assay, in one respect, is an excellent mutagenicity test method because it enables one to examine the mutagenic effects of a test compound on progeny (G–124 at 11). In another respect, however, this method is quite limited because it only measures "lethal" effects; thus, non-lethal mutagenic effects, which may still be serious, go undetected (*id.* at 11–12).

b. *The Studies' Findings:* Findings from the 15 dominant lethal studies are described below. One of these studies produced findings suggestive of mutagenicity (G–29), and nine studies produced negative findings. Also described briefly below are the five studies found to be deficient.

(1) *Suggestive Studies:* (a) Peterson, et al. (G–29) found a statistically significant increase ($P=.05$) of post-implantation loss [27] for inbred C57B1/Fe mice. The males had been treated with a total of 500 mg/kg CHA over five days and then mated with untreated females of the same strain for three weeks. Both positive and negative controls were used. These results (Table II) confirmed earlier, similar findings by the same authors in a pilot study (Table III).

(2) *Negative Studies:* (a) *Mouse Studies:* (i) Cattanach, et al. (A–151) found no statistically significant increase in pre-implantation or post-implantation loss after mating hybrid male mice (dosed with a total of 250 or 500 mg/kg CHA over 5 days) with untreated females.

(ii) Lorke (A–827, App. 15) found no statistically signficant increase in pre- or post-implantation loss after mating NMRI/BOM strain male mice (dosed with a total 50 g/kg of cyclamate over 5 days) with untreated females.

(iii) Lorke, et al. (A–827, App. 9) found no statistically significant increase in pre- or post-implantation loss after mating NMRI male mice (dosed with a total of 750 mg/kg of CHS over 5 days) with untreated females.

(iv) Lorke, et al. (A–827, App. 17) found no statistically significant increase in pre- or post-implantation loss after mating NMRI strain male and female mice. Both sexes were treated for ten weeks prior to mating. Doses were either 2,000 mg/kg/day of sodium cyclamate (1% of feed) or 200 mg/kg/day of CHA (0.11% of feed).

(v) Lorke, et al. (A–827, App. 11) found no statistically significant increase in pre- or post-implantation loss after mating NMRI strain treated female mice (single dose of 10 g of sodium cyclamate) with untreated males. The published version of this study (A–811, App. 18) shows that a positive control (cyclophosphamide) was used and that positive dominant lethal results were obtained.

(vi) Ford, et al. (A–297) found no statistically significant increase in post-implantation loss after mating male Cox Swiss albino mice (given a single injection of 50/mg CHA) with untreated females. These negative findings were confirmed by statistically significant positive findings in several positive control groups (see Table II).

(vii) Epstein, et al. (A–182) found no statistically significant ($P<.05$) increase in post-implantation loss after mating treated male ICR/Ha Swiss Mice with untreated females. The dosing regimen for calcium cyclamate was either a single injection of 132 or 660 mg/kg, or five doses totalling 500 or 1000 mg/kg (A–182 at 305). The dosing regimen for CHA ranged from a single dose of 5 mg/kg to three doses totalling 75 mg/kg (A–182 at 314). This study was part of a massive experiment in which 174 compounds were tested for dominant lethality. Numerous compounds produced statistically significant positive results (see Table 5).

(b) *Rat Studies:* (i) Green, et al. (A–206) found no statistically significant increase in post-implantation loss after mating Holtzman strain albino male rats (dosed with a total of 100 or 300 mg/kg CHA) with untreated females of the same strain. These negative findings were validated by a positive control group, dosed with triethylenemelamine (TEM), in which a statistically significant ($P<.05$) increase in post-

[27] The ALJ mistakenly called this "pre-"implantation loss. However, the parties agree, as do I, that the actual findings were of "post"-implantation loss (Abbott's Exceptions at 52; Bureau's Brief at 81).

implantation loss was found (see Table 2). The only positive findings ($P<.05$) in CHA-treated animals was in terms of pre-implantation loss, but the authors concluded that this result was not of genetic origin (A–206 at 33).

(ii) Kennedy, et al. (A–220) (rat portion) found no statistically significant increase in post-implantation loss after mating Charles River albino male rats (dosed with 1.5 or 15.0 mg/kg day of CHS for 60 days) with females of the same strain (similarly dosed, but only for the 14 days immediately prior to mating).

(3) *Deficient Studies:* For the following reasons, five studies which reported negative findings are deficient and therefore entitled to no weight. Friedman, et al. (A–195) fed male Holtzman rats a diet of 2% calcium cyclamate and then mated them with untreated females. The presentation of data in this study, however, is inadequate because there is no comparison shown between the number of dead implants per female and the total number of implants per females (see A–195 at 754). Khera, et al. (A–221) also did not provide adequate data to support their finding that CHS did not cause a statistically significant increase in post-implantation loss. The only data presented are in a rough graph (Figure B) which lacks the necessary precision to permit an adequate evaluation. Finally, the two Oser studies (A–273 and A–274) and the rabbit portion of the Kennedy study (A–220) were actually teratology studies rather than dominant lethal experiments. A teratology study is where the females are treated with the test compound *during pregnancy* to determine if any effect is produced on the growing fetus. Because dosing takes place after conception, this type of study can not possibly detect mutagenic effects on germ cells prior to conception (as is the purpose of a dominant lethal study). The fact that these are indeed teratology studies is reflected in the descriptions of test methods (see A–273 at 9–10, A–274 at 6–7 and A–220 at 6–7). Thus, these studies are not entitled to any weight in the evaluation of the potential dominant lethality of cyclamate or its metabolites.

c. *Analysis:* The studies in dispute are Peterson, et al. (G–29), the four studies by Lorke, et al. (A–827, App. 15; A–827, App. 9; A–827, App. 17; and A–827, App. 11), and Epstein, et al. (CHA portion) (A–182).

(1) *The Peterson Study (G–29).* This suggestive study, as noted above, was the only dominant lethal study to observe statistically significant ($P<.05$) positive findings of post-implantation loss. Abbott criticizes this study on several grounds: (1) small number of animals; (2) a typically low number of implanted and live embryoes in untreated controls; (3) findings not replicated by any other researcher; and (4) the ALJ wrongly said that the dominant lethality observed by Peterson increased over time (Abbott's Exceptions at 52–53). Abbott's first two points are based on a brief evaluation of this study by Lorke, et al. contained in one of their dominant lethal studies described above (A–827, App. 9 at 10). The Bureau defends the Peterson study by stating that: (1) a small animal population is adequate where the findings are positive; (2) the raw number of live and dead implants are less important than the ratio between the two; and (3) Peterson was able to replicate his own results, even if other researchers were not (Bureau's Brief at 81–82).

With respect to the small number of animals used, I agree with the Bureau that where a study's findings are positive, a small animal population does not negate the validity of the study (see Subsection C.3.a. above). However, I also agree with Abbott to the extent that a small animal population detracts somewhat from the weight to be given to that study (*id.*). I therefore consider the Peterson study to be facially valid but entitled to slightly less weight than would similar results from a larger animal population.

The second issue regarding the total number of implants being atypically small is also a question of weight rather than validity. As Dr. Green explained, the threshold issue in a dominant lethal study is the ratio between the living and dead implants rather than their total number (G–123 at 5). The Bureau does admit, however, that the total number of implants per female was low (Bureau's Brief at 81); this has the effect of reducing the "test population" (*id.* at 81, n. 20) for purposes of statistical analysis, and hence reduces the study's sensitivity. Indeed, as Dr. Legator, stated: ". . . one of the serious shortcomings in the Peterson study was the low number of implants per female" (Tr. at 948–49). Thus, again, I consider this study to be facially valid but entitled to somewhat less weight than would similar results from a study with more implants per female.

The third issue regarding replicability requires only brief discussion. I agree with the Bureau that Peterson did replicate his results with similar statistically significant ($P<.05$) findings of post-implantation loss (G–29 at Tables II and III). This replicability adds credence to Peterson's findings. The fact that these results were not replicated by other investigators using other strains and species is a separate issue to be discussed below.

Abbott's final criticism results from a misinterpretation of a statement made by the ALJ with respect to this study. The statement in question is as follows:

> However, the effect seen in the CHA treated animals was significantly higher than that of the saline control which increased over the three or six weeks [sic] period.

(*Id.* at 25). Abbott suggests that this statement wrongly implies that Peterson found an increase in dominant lethality over time. I agree with Abbott that Peterson did not find such a time-related increase. However, I do not interpret that ALJ's statement to convey this fact. Rather, I interpret the ALJ's statement to mean that Peterson found an increase over *both* the three and six week periods, not that the findings in the sixth week were greater than those in the third week. This is consistent with the facts and should satisfy Abbott's exception.

I therefore conclude that the Peterson study is strongly suggestive of mutagenicity, especially since his findings were replicated. I would have more confidence in these results, however, if Peterson had used more animals, if the total number of implants had been greater, and if his findings had been replicated by an independent investigator.

(2) *The Lorke Studies (A–827, Apps. 9, 11, 15 and 17).* Abbott places great reliance upon four studies conducted by Lorke, et al. As described above, Lorke used several different procedures, including not only the traditional method of mating treated males with untreated females, but also the less common modes of mating untreated males with treated females and of treating both sexes before mating. All four studies produced negative findings for post-implantation loss.

The Bureau attacks the validity of these studies on two grounds: (1) that Lorke failed to perform preliminary experiments necessary to determine the "maximum tolerated dose" to be used in the dominant lethal studies on cyclamate (Bureau's Brief at 93–94) but rather used only mathematical extrapolations (Bureau's Reply at 26); and (2) the alleged failure to use positive controls in the experiment where the females were treated rather than the males (Bureau's Brief at 94). Abbott defends these studies by asserting that Lorke did properly ascertain the maximum tolerated dose (Abbott's Brief at 69; Abbott's Exceptions at 54) and that a positive control was used in the

study using treated females. (Abbott's Brief at 70). A review of the record shows that Abbott is correct on both points.

The maximum tolerated dose ("MTD") is "the dose just below [the one in] which one sees obvious toxicity" (Tr. at 847). The parties agree that in a dominant lethal study it is important to use the MTD in order to maximize the chances of detecting a positive effect. The parties interpret differently, however, the following statement in one of the Lorke studies which describes how the dose level used was ascertained:

> This dose was chosen because preliminary experiments had demonstrated that it is well-tolerated by the animals. The administration of higher doses would have created considerable difficulties due to the large quantity of substance involved.

(A–827, App. 8 at 4). This statement makes clear, first of all, that Lorke used "preliminary studies" and not "Mathematical extrapolations" to ascertain the proper dose. Second, I interpret Lorke's statement, when read as a whole, to mean that the "maximum" dose arrived at was maximum in terms of potential toxicity. I therefore reject the Bureau's criticism that Lorke did not properly determine the MTD. Moreover, I note that the total dose used by Lorke in the CHA study using treated males (A–827, App. 9) (approximately 510 mg/kg) [28] was comparable to that used by Cattanach, et al. (A–151) (500 mg/kg). I therefore find that Lorke's dose levels were adequate.

The second issue regarding postitive controls may be disposed of easily. The Bureau complains that one specific study (A–827, App. 11) in which untreated males were mated with treated females lacked a necessary positive control. Although it is true that no positive control information is reported in the unpublished version of this study (A–827, App. 11), the published version (A–811, App. 18) shows that a positive control (cyclophosphamide) was used and that positive results were obtained. (A comparison of the data in Table 1 of A–827, App. 11 with that in Table II of A–811, App. 18 shows that they are indeed the same study.)

I therefore conclude that the four studies by Lorke are all negative in terms of post-implantation loss and are entitled to considerable weight.

[28] This study actually employed CHS rather than CHA. The total dose of CHS was 750 mg/kg. The authors stated, however, that 150 mg of CHS equals approximately 102 mg of CHA base (A–827, App. 9 at 4). The total dose, when converted to CHA, is therefore approximately 510 mg/kg.

(3) *The Epstein Study (CHA portion).* The parties agree that the findings of post-implantation loss in this study are not statistically significant at the $P \leq .05$ level. Based upon the testimony of Dr. Epstein, however, the Bureau maintains that borderline findings of *pre-*implantation loss make the study "suggestive" rather than "negative" (Tr. at 865–66). I disagree. As noted above in the description of dominant lethal assay test methods, findings of pre-implantation loss, even if statistically significant at the $P \leq .05$ level, are not necessarily tied to mutagenicity. Even Dr. Epstein admits this (G–121 at 14; Tr. at 866), as does Dr. Green in his dominant lethal study which did find a statistically significant ($P \leq .05$) increase in pre-implantation loss (A–206 at 29). Indeed, in Dr. Green's study, he concluded that the pre-implantation loss was not of genetic origin (A–206 at 33). This does not mean that findings of pre-implantation loss would never be considered biologically significant, but corroborating evidence would be needed (such as statistically significant findings of post-implantation loss in the same study). Accordingly, I have attributed no weight to the findings in this record of pre-implantation loss.

(d) *The Evidence As a Whole.* The ALJ found that the positive findings in the Peterson study (G–29) are not completely rebutted by the negative mouse studies because of the difference in mouse strain tested:

> In evaluating the results of various tests, it must be remembered that various strains react with various degrees of sensitivity to chemical mutagens.

(ID at 25). The ALJ therefore concluded that "the results cannot be disregarded" (ID at 35). Abbott takes exception to this finding of the ALJ and maintains that the dominant lethal studies, when viewed in the aggregate, are negative (Abbott's Exceptions at 71; 53).

As is evident from the above description of the dominant lethal studies, most of the evidence in this test method are negative, and these studies encompass several mouse and rat strains. Nevertheless, none of these negative studies used the same strain of (C57B1/Fe) mice as did Peterson. The record is clear that differences in strains are important. As Dr. Green explained:

> Therefore, when one considers the fact that the strain of mouse utilized by Peterson was not employed by the other investigators, one has to consider the possibility that the effect observed by Peterson *et al.* was genuine. This study raises the possibility that cyclohexylamine can produce dominant lethality in animals possessing certain genetic constitutions.

(G–123 at 5). I therefore conclude that, although most of the dominant lethal evidence is negative, some question still remains about the mutagenic potential of cyclamate and CHA in at least one strain of mouse. Thus, the evidence is not conclusive. However, even were the evidence conclusively negative in the dominant lethal studies, such findings would be insufficient to outweigh the suggestive cytogenetic experiments (see Subsection D.1. above). Indeed, the dominant lethal assay technique has been known to report negative findings for compounds that are proven mutagens in other test methods (G–124 at 9–10; Tr. at 498–500).

4. *Drosophila.* The final type of *in vivo* mutagencity testing performed on cyclamate or its metabolites was conducted using Drosophila (fruit flies). The specific type of Drosophila test which was conducted is called a "sex-linked recessive lethal" test. The ALJ described this test as follows:

> A recessive lethal mutation present on a male's only X chromosome will cause the male to die. If the compond being tested induces a recessive lethal, and the affected gene is carried in the X chromosome of the sperm, the mating with untreated females will produce offspring (F1), which when mated together produce males (F2), half of which have X chromosomes from the original treated males. If this group is absent [in the F2 generation] recessive lethals were produced.

(*Id.* at 23; see G–122 at 8–10 for more detailed description). The sex-linked recessive lethal test is "the most efficient and informative procedure in Drosophila testing" (G–122 at 8). It will detect a wide range of genetic damage, principally in the gene mutation category (*id.* at 11).

The Drosophila evidence in this record consists of two negative studies (A–712 and A–728) and five deficient ones (G–24, G–122 at 20, A–263, A–289 and A–305). The two negative studies require some discussion because the Bureau has questioned how much weight should be attributed to them.

Vogel, et al. (A–728) and Knapp, et al. (A–712) each conducted sex-linked recessive lethal tests as described above. Vogel, et al. conducted adult feeding tests using sodium cyclamate and CHA as the test compounds. Knapp, et al. conducted adult injection and larvae feeding tests using the metabolites CHA and N-OHCHA as the test compounds. The ALJ found both of these studies to be negative (*Id.* at 23), and I agree.

The Bureau maintains, however, that neither of these studies has a large enough test population to establish safety for this test system (Bureau's

Brief at 91–92; Bureau's Reply at 23). The Bureau based this position on the testimony of Dr. Zimmering, who explained that a population size of 12,000 X chromosomes (F1 flies) in both the treated and control groups would be necessary to achieve a test sensitivity capable of detecting a doubling over the control rate (G–122 at 13; Tr. at 485–86). Dr. Zimmering considered this degree of sensitivity to be necessary to establish safety because Vogel, et al. found the frequency of recessive lethals in the treated group of Brood 3 (0.68%) to be roughly double the frequency of recessive lethals in the controls (0.36%) (G–122 at 13). Although this difference was not statistically significant for Vogel's population size (approximately 1700–3200 for all Broods combined), a test with a population of 12,000 would detect positive findings *if* the relative frequencies between the treated and control groups remained the same (*id.*). Thus, Dr. Zimmering would require the larger experiment to test his hypothesis (*id.* at 13–14). He noted that Drosophila tests of this size are "carried out routinely in most laboratories" (*Id.* at 13). Dr. Zimmering make a similar analysis with respect to the Knapp study (*id.* at 16).

I agree with the Bureau on this point, but I emphasize that the issue goes to the issue goes to the weight to be attributed to these studies, not their validity. The studies as carried out and reported are valid negative studies. Dr. Zimmering's point, with which I agree, is simply that given the frequencies of recessive lethals found in these experiments, much larger tests would be necessary to establish safety in this test system.

The Bureau also contends that the Drosophila evidence is incomplete in that no experiment tested cyclamate (as opposed to the metabolites) using the adult injection method (Bureau's Brief at 92). Again, the Bureau relies upon the testimony of Dr. Zimmering (G–122 at 14–15). Although Dr. Zimmering's testimony is quite persuasive as to why each different route of administration must be used, he does not explain why the parent compound (*i.e.*, cyclamate) must be tested using each such route where, as here, the metabolites have already been so tested, and where cyclamate itself has been tested in an adult feeding study. I therefore reject this criticism raised by the Bureau.

The record also contains five Drosophila studies which I have found to be deficient, all due to an inadequate presentation of data (see Subsection C.3.b. above). These studies are Stith, et al, (A–305), Majundar, et al. (G–24), Moon, et al. (A–263), Browning (discussed in G–122 at 20), and Rotter, et al. (A–289). All of these studies were available only as abstracts without the data necessary for a full evaluation (G–122 at 21; see Tr. at 484). Moreover, I note that for the two abstracts which reported positive findings (A–305 and G–24), Commissioner Kennedy asked in his Remand Order that the parties supply more information (44 FR 47623). The parties have since stipulated that the requested data is unavailable (Stipulation dated September 17, 1979 at 5–7). I therefore am attributing no weight to these five studies.

In summary, the available Drosophila evidence is negative, but the sensitivity of these studies is such that they do not establish the safety of cyclamate and its metabolites in this test system. Even were those studies to establish safety in this test system, however, evidence in Drosophila would be insufficient to outweigh the cytogenetic findings (see Subsection D.1. above).

5. *Additional In Vitro Testing.* The final category of mutagenicity evidence contained in the record involves additional *in vitro* testing performed on cyclamate or its metabolites. As noted above, however, *in vitro* studies by their very nature are useful only as preliminary screens and cannot outweigh positive or suggestive *in vivo* findings (see Subsection A.3 above).

The ALJ made the following findings with respect to these *in vitro* test:

> *In Vitro Tests.* The *Ames* test has previously been described. Additional results using the Ames test were introduced concerning the mutagenicity issue. Both cyclamate and CHA were tested by several scientists using Salmonella typhimurium. All the results were negative (Ex. Nos. A–736, A–808, G–124). However, positive results were found using CHA in Saccharomyces cerevisia (Ex. No. A–268).
>
> In addition, Chinese hamster cells were cultured with CHA or N-hydroxychlohexylamine (N–OH–CHA) added in a study to examine gene mutation. A forward mutation change was seen with N–OH–CHA (Ex. No. G–47). Upon prolonged treatment with N–OH–CHA the cell survival rate was reduced to 20% and a significant increase in mutations was seen over the controls (Ex. No. G–47).
>
> Plant cell studies were also performed using onion seeds or Haworthia (Ex. Nos. A–250, A–251, A–295). The results of the plant cell studies on sodium cyclamate were negative.

(*Id.* at 21). Abbott's exceptions to this portion of the ALJ's opinion primarily involve clarifications rather than disagreements. For example, Abbott correctly notes that in the first paragraph discussing the Ames test, the ALJ failed to cite fidings by Dr. Legator (A–268) involving negative Ames test results on calcuim cyclamate, CHA and N–OHCGA (Abbott's Exceptions at 40–41). Similarly, Abbott correctly observes that in the second paragraph discussing the Chu study (G–47), the ALJ failed to expressly state that the CHA portion of the study was negative (*id.* at 41) However, Abbott incorrectly suggests that the ALJ omitted to cite negative findings from studies A–736 and A–808 (*id.* at 42), for citations of these studies are contained in the ALJ's first paragraph quoted above. I find these minor omissions by the ALJ to be inconsequential.

Finally, Abbott contends that the ALJ was wrong in one instance. With respect to the Chu study (G–47) (second paragraph), the ALJ stated that the cell survival rate was reduced after "prolonged treatment," while Abbott contends that the cell survival rate was reduced at "increased concentrations" (Abbott's Exceptions at 42). A review of G–47 shows that Abbott is correct on this point. All this means, however, is that the increase in mutations was seen at increased concentrations rather than after prolonged treatment. The study, therefore, still reports positive findings.[39]

In summary, these *in vitro* studies were predominantly negative, although two investigators did find positive results in studies that were not directly rebutted. Thus, the evidence is not conclusive. Even were these studies conclusively negative, however, such finding would be insufficient to outweigh the suggestive *in vitro* experiments (see Subsection A.3 above).

H. Miscellaneous Mutagenicity Issues

1. *The Relationship Between Mutagenicity and Cancer.* The ALJ found that "[m]utagens in somatic cells can lead to cancer" (*Id.* at 35), thereby suggesting a causal link between mutagenicity and carcinogenicity (see also *Id.* at 21). The Bureau agrees (Bureau's Brief at 71–72 and Bureau's Reply at 21). Abbott, however, takes strong exception to this finding by the ALJ both as a matter of general scientific principle and as applied to the evidence on cyclamate (Abbott's Exceptions at 39–40 and 74).

A review of the record in this proceeding shows that adquate expert testimony was not elicited as to any of the issues concerning the relationship between mutagenicity and carcinogenicity—*e.g.*, what, if any, types

[39] As noted in Subsection F.2.c. (2) above, however, since this study is reported only as abstract, both its positive findings with N–OHCHA and negative findings with CHA are entitled to little, if any, weight.

of genetic damage cause cancer; what, if any, types of mutagenicity study results would serve as an indicator that the test compound may cause cancer; and finally, the applicability of these issues, if any, to the evidence of mutagenicity. Given the inadequacies of the record in this respect, I make no findings concerning what, if any, relationship exists between mutagenicity and cancer.

2. Findings of the Temporary Committee. The Temporary Committee, in its Review of Data on the Carcinogenicity of Cyclamate, make several findings with respect to the mutagenicity studies (G–41 at 32–36). These findings provide additional support for my conclusion that the evidence in this record, particularly the cytogentic studies, strongly suggests that cyclamate or its metabolites may cause heritable genetic damage. Indeed, the Temporary Committee concluded as follows with respect to the mutagenicity studies:

> . . . the fact that several laboratories have shown that cyclamate and cyclohexylamine can produce chromosome damage in both rodents and humans following *in vivo* administration of doses approximating human usage raises the possibility that these compounds may adversely affect genetic activity.

(*Id.* at 36.) [30]

VI. Acceptable Daily Intake and Safe Conditions for Use

Two additional issues were addressed by the parties during the hearing phase of this proceeding. The ALJ described these issues as follows:

> [1.] Apart from the [carcinogenicity and mutagenicity] issues . . ., what does the evidentiary record show is an acceptable daily intake level for cyclamate?
> [2.] Whether apart from the [carcinogenicity and mutagenicity] issues . . ., because of probable consumption patterns, safe conditions of use can be prescribed.

(*Id.* at 4). As explained in more detail below, I find it is unnecessary to decide either of these two issues since they are mooted by the conclusions I have already reached with respect to carcinogenicity and mutagenicity.

The two issues of acceptable daily intake and safe conditions for use are interrelated. The acceptable daily intake level is the level (expressed in mg/kg body weight/day) immediately below the lowest level which produces significant adverse or toxic effects. For cyclamate, the parties introduced evidence concerning testicular atrophy and reproductive effects. Once the acceptable daily intake level is determined, the probable consumption patterns of cyclamate must be calculated to determine whether, if cyclamate is added to the food supply as Abbott proposes, actual consumption would exceed the acceptable daily intake level. This latter calculation is the safe conditions for use issue. For the purposes of this discussion, it is not necessary to state precisely how these calculations are made.

The ALJ found that the administrative record would support a finding "that the acceptable daily intake is five mg cyclamate/kg body weight/day or less" (*Id.* at 38). This is consistent with the Bureau's position, although Abbott advocates a higher level. On the second issue, the ALJ found that since each party either overestimated or underestimated the probable consumption figures to support its respective position, "neither can be relied on to give an accurate picture of the probable consumption" of cyclamate (*Id.* at 37). Accordingly, the ALJ found that the safe conditions for use issue was not resolvable on this record (*Id.* at 38).

It is clear that the questions involving acceptable daily intake and safe conditions for use are only important if Abbott prevails on both the carcinogenicity and mutagenicity issues. This is because under the act, as explained in Section II. above, the agency must deny approval of Abbott's food additive petition if Abbott fails to prove either that cyclamate is not carcinogenic or that cyclamate is not mutagenic. Since Abbott has failed to make either of these two showings, I find it unnecessary to decide the acceptable daily intake and safe conditions for use issues.

VII. Miscellaneous Matters

A. Allegations Concerning 21 CFR 12.120(b)

Abbott made several general objections which relate primarily to the form of the Initial Decision. Abbott contends that the Initial Decision fails to comply with 21 CFR 12.120(b) in that it does not contain (1) sections entitled "findings of fact" and "conclusions of law;" (2) a full articulation of the reasons for the findings and conclusions that are made, and (3) full citations to the record (Abbott's Exceptions at 3–6).

A careful review of these exceptions leads me to conclude that they go primarily to form rather than substance. I therefore find that there is no merit in the argument that the Initial Decision does not comply with 21 CFR 12.120(b). It is not necessary for the Initial Decision to contain a detailed discussion of every item of evidence in order to have evaluated it adequately; nor is it necessary to provide a record citation for every factual statement in the decision so long as the decision is supported by the record. Although the main text of the Initial Decision is brief in its explanation of the reasons for resolution of the scientific issues, I find that the Initial Decision's discussion of the issues and citations to the record are sufficient both to support the ultimate findings and conclusions made, and to adequately inform Abbott of the reasons for those findings and conclusions. Moreover, it is clear from the ALJ's detailed description of the studies submitted that the ALJ examined the record in detail. Accordingly, the Initial Decision complies with 12.120(b).

B. Alleged Failure To Comply With 21 U.S.C. 348

Abbott contends that the Initial Decision is not a "fair evaluation of the record" in that the ALJ unfairly evaluated the evidence. Thus, Abbott asserts that the ALJ failed to comply with section 409 of the act, 21 U.S.C. 348 (Abbott's Exceptions at 3).

I find that this exception is also without merit. In most respects, this exception faults the Initial Decision simply because the decision did not accept the arguments offered by Abbott (*e.g.*, finding a study "suggestive" even though the results of the study are not significant at the .05 level). Abbott's specific arguments concerning the ALJ's unfair evaluation of the evidence are discussed in detail in the body of this decision. I find that the ALJ did carefully consider Abbott's arguments. See, *e.g.*, ID at 5, 6, 7, 8, 9, 13, 14, 15, 16, 17, 19, 22, 25, 26–27, 28, 31, 32, 36, and 37. With minor exceptions discussed in the body of this decision, the ALJ correctly evaluated the evidence. Although the Initial Decision does contain some errors, virtually all of these are inconsequential. They clearly do not reflect any prejudice or unfairness in evaluating the evidence, but rather an impartial and conscientious effort to resolve the issues.

[30] The only significant finding made by the Temporary Committee that is at variance with my mutagenicity findings relates to the dominant lethal assay evidence. The Temporary Committee reported that "there is no evidence that either cyclamate or cyclohexylamine possess dominant lethal effects" (G–41 at 33). I, however, found that one study by Peterson, et al. (G–29) contains statistically significant (P=.05) findings of post-implantation loss (see Subsections G.3.b.(1)(a) and G.3.c.(1) above). It is quite possible, however, that the Temporary Committee never reviewed this study since the Temporary Committee does not specify any dominant lethal study by author, and since the Temporary Committee reviewed only 11 dominant lethal studies whereas the hearing record contains 15. I therefore conclude that this finding by the Temporary Committee does not necessarily contradict the findings made in this decision.

C. Allegations That the Initial Decision Is a Repudiation of Science

Abbott further contends that the Initial Decision is a repudiation of science in that it reflects a lack of understanding of the scientific evidence and rejects fundamental principles such as statistical significance, replicability, presence of uncontrolled variables, and scientific peer review (Abbott's Exceptions at 6–12). Almost all of these general exceptions are discussed by Abbott in connection with Abbott's criticism of the ALJ's evaluation of specific studies. I have therefore discussed those significant exceptions in detail in connection with my evaluation of each specific study. In general, Abbott's contention that the Initial Decision is a repudiation of science is without merit. Although there were some minor errors in the Initial Decision and in some instances its phrasing could be improved, when evaluated on an overall basis, the limitations of the Initial Decision do not undercut the validity of its basic finding that cyclamate has not been shown to be safe. That finding is supported by a large body of scientific studies and expert testimony contained in the record.

D. Documents Relating to the Internal Deliberative Process

Abbott moves to admit into evidence two sets of documents which reflect the decisionmaking process that led to the agency's decision in 1976 to deny approval of the food additive petition for cyclamate. For purposes of identification, the first set of documents is attached to Abbott's Exceptions to the Initial Remand Decision, dated February 25, 1980; the second set is attached to a Motion to Include Documents, dated April 17, 1980. Both sets were obtained by Abbott through civil discovery ordered in *Abbott Laboratories* v. *Harris,* 481 F. Supp. 74 (N.D. Ill. 1979). Abbott argues that the Bureau should have disclosed these documents to Abbott in accordance with 21 CFR 12.85(a)(2).

I have reviewed these documents in their entirety and find that they do not fall within the purview of 21 CFR 12.85(a)(2). That section provides that, prior to the issuance of a notice of hearing, the director of the responsible bureau shall disclose:

> All documents in the director's files containing factual information, whether favorable or unfavorable to the director's position, which relate to the issues involved in the hearing. . . .

The documents at issue, however, contain internal, predecisional opinions and recommendations which clearly do not fall within the category of "factual information." As the preamble to Subpart B of FDA's Administrative Practice and Procedure Regulations explains:

> . . . The Commissioner advises that the requirement of this section [21 CFR 12.85] does not extend to documents reflecting the agency's internal deliberative process, e.g., documents expressing the point of view of agency employees who reviewed an NDA, even though such documents are contained in and administrative file relating to a matter that is the subject of the hearing.

(41 FR 51714, November 23, 1976). Indeed, section 12.85(a)(2) has since been amended to express clearly this longstanding agency interpretation: "Internal memoranda reflecting the deliberative process * * * are not required to be submitted" (44 FR 22344, April 13, 1979).

I therefore find that the Bureau did not act improperly in withholding the documents at issue. In any event, the court in *Abbott Laboratories* v. *Harris,* Civil No. 79–C–3732 (N.D. Ill., decided July 12, 1980) flatly rejected Abbott's claim that these documents show that its food additive petition would have been approved had not then Commissioner Schmidt included improper considerations in making his decision:

> The picture which merges from the record is one of good faith uncertainty caused by the limitations of prior testing and differing interpretations of the results. Many in the scientific community believed and believe limited use of cyclamates to be safe to a reasonable certainty; others have not been able to so conclude. Virtually no one is of the opinion that the limited use of cyclamates is demonstrably unsafe.
>
> Given those circumstances, it is not surprising that there were differences of opinion among advisors and that, in collegial discussion, views changed. The initial denial was viewed, as stated by one witness, as a very close call, a very difficult judgment. Possibly the views of some advisors were influenced by their perception of the Commissioner's tentative judgment. Possibly the Commissioner, despite his recognition of the proper legal standard, was himself somewhat influenced by his own perceptions of the need or lack of need, for cyclamates in the marketplace. Without doubt public focus upon and plaintiff's interest in the question caused the decisionmaking process to be somewhat more cautious and ponderous than it otherwise might have been.
>
> The record does not, however, support the conclusion that defendants and their predecessors acted in bad faith or for improper reasons. Rather, the deposition of the Commissioner making the initial decision reveals a somewhat acerbic gentlemen with strong views and a willingness to express them who, after considering numerous opinions, made a technical judgment he was authorized to make.

I therefore reject all of Abbott's exceptions concerning FDA's internal deliberative process (Abbott's Remand Ex. at 4–17).

E. Separation of Functions

FDA regulations governing the conduct of agency officials in administrative hearings, such as the cyclamate proceeding, provide that:

> * * * Representatives of the bureau shall not participate or advise in any decision except as witness or counsel in public proceedings. There is to be no other communication between representatives of the bureau and representatives of the Commissioner concerning the matter [involved in the hearing] before the decision of the Commissioner.

21 CFR 10.55(b)(2)(i). Abbott complains that this regulation was violated because Dr. Vasillios Frankos and Dr. Constantine Zervos, who presently work in the Commissioner's Office of Health Affairs, have served as advisers to the Bureau of Foods in this proceeding. Dr. Frankos also served as a witness for the Bureau. Abbott contends that Dr. Frankos and Dr. Zervos might "taint" other scientists who are responsible for advising me (Abbott's Remand Brief at 25–26). Abbott further contends that Dr. Zervos and a Bureau attorney may have contacted other scientists in the Office of Health Affairs and "tainted" them by asking them to advise the Bureau (*id.*).

The requirement of separation of functions is designed to ensure that the same persons do not serve as both advocate and judge in the same proceeding. Thus, in the context of the cyclamate hearing, representatives from the bureau of Foods (the "advocate") are forbidden from having certian communications with representatives from the office of the Commissioner (the "judge"). This restriction is intended to "avoid even the appearance of unfairness" (40 FR 40691; September 3, 1975). At the same time, the restriction on communications is limited to "the matter" which is involved in the hearing. Here, that matter is the substantive issue of whether cyclamate has been shown to be safe.

The mere fact that two former representatives of or advisors to the Bureau (Dr. Frankos and Dr. Zervos) now work in the office of the Commissioner is insufficient by itself to constitute a violation of separation of functions. The regulations prohibit certain communiations, not mere proximity of offices. Abbott has not presented any credible evidence that either of these two scientists has had substantive communications regarding

the safety of cyclamate with any person advising me on this issue.

Neither has Abbott demonstrated that either Dr. Zervos or a Bureau attorney contacted other scientists on my staff to ask them to advise the Bureau. Even if such communications were made, however, I do not consider them to violate separation of functions because the communications would not have involved any substantive discussion on the safety of cyclamate.

Finally, Abbott complains that the separation of functions regulation prohibits Bureau attorneys from representing the Commissioner in a lawsuit filed by Abbott which sought a declaratory judgment that Abbott's food additive petition be approved (Abbott's Remand Brief at 27). That lawsuit has not involved an evaluation of the evidence on cyclamate's safety, but rather allegations by Abbott concerning whether improper considerations played a role in the agency's prior decision to deny approval of Abbott's petition. Since separation of functions does not apply to the latter subject, it is appropriate for Bureau of Food's attorneys to participate in that lawsuit.

I therefore reject all of Abbott's contentions regarding separation of functions.

F. Admissibility of the IRLG Report

The ALJ refused to admit into evidence exhibit G–142, the report of the Interagency Regulatory Liaison Group ("IRLG"), on the ground that the document was not in final form and was thus subject to further alteration by the FDA (IRD at 5). As the ALJ explained, the report contains a discussion of scientific concepts and methods concerning the evaluation of substances that may pose a risk of cancer in humans (*id.*). The document in its current form is a government proposal subject to public notice and comment procedures (*id.*). The Bureau takes exception to this ruling, contending that lack of finality is not a proper basis for the exclusion of evidence (Bureau's Remand Ex. at 5–6). Abbott urges me to uphold this ruling (Abbott's Remand Reply at 7).

I agree with the Bureau that the report's lack of finality goes to weight rather than admissibility. The IRLG report is therefore admitted into evidence because its purported subject matter is relevant. However, because the document contains preliminary views only, and because, in any event, the Bureau has not adequately shown exactly how these views should be applied to this record, the document has not been given any weight in determining whether cyclamate has been shown to be safe.

VIII. Conclusion

Based on the foregoing findings, conclusions, and discussion, I affirm the Initial Decisions and conclude that:

1. Section 409(c)(3)(A) of the act, 21 U.S.C. 348(c)(3)(A), requires FDA to deny approval of a food additive petition if a fair evaluation of the data presented fails to establish that the food additive will be safe under its proposed use. See Section II.

2. "Safe" means a reasonable certainty of no harm. See Section II.

3. The act places the burden of proving safety on the company seeking approval of the food additive petition. See Section II.

4. For Abbott to obtain approval of its food additive petition, it must prove that the data in the record establish that there is a reasonable certainty of no harm from the proposed use of cyclamate. See Section II.

5. The data in the record do not establish that there is a reasonable certainty that cyclamate does not cause cancer. See Sections III and IV.

6. The data in the record also do not establish that there is a reasonable certainty that cyclamate does not cause heritable genetic damage. See Section V.

7. Abbott has failed to meet its burden of proving that cyclamate is safe under its proposed use. See Sections III, IV and V.

8. In light of these findings and conclusions, the issues involving acceptable daily intake and safe conditions for use need not be decided. See Section VI.

The foregoing decision in its entirety constitutes my findings of fact and conclusions of law.

IX. Order

In accordance with subsections (c)(3)(A), (f)(1) and (f)(2) of section 409 of the act (21 U.S.C. 348(c)(3)(A), (f)(1) and (f)(2)) and 21 CFR 12.130, and under the authority delegated to the Commissioner (21 CFR 5.1), the food additive petition (FAP 4A 2975) for approval of cyclamate for use as a sweetening agent in food and for technological purposes in food is denied. The Initial Decisions are affirmed, as modified and supplemented herein.

In accordance with section 409(f)(3) of the act (21 U.S.C. 348(f)(3)), the effective date of this order is December 15, 1980.

Dated: September 4, 1980.

Jere E. Goyan,

Commissioner of Food and Drugs.

[FR Doc. 80–27590 Filed 9–4–80; 1:54 pm]

BILLING CODE 4110–03–M