

D R A F T

Toxicological Principles

for the Safety Assessment of

Direct Food Additives and

Color Additives Used in Food

"Redbook II"

EXTRA copy

US Food and Drug Administration

Center for Food Safety and Applied Nutrition

1993

52454 1795

Preface

File - Redbook
on
FDA Redbook

This document is the first revision of the Food and Drug Administration's 1982 "Redbook I" (*Toxicological Principles for the Safety Assessment of Direct Food Additives and Color Additives Used in Food*). The revised "Redbook II" is intended 1) to provide guidance regarding criteria used for safety assessment of direct food additives and color additives used in food and 2) to assist petitioners in developing and submitting for Agency review data for the safety assessment sections of petitions for these food additives under Section 21 of the Code of Federal Regulations (CFR). While the guidelines in this document do not preclude the petitioner from demonstrating safety by using other types of data, a submission conforming to the recommended scheme of toxicity testing would normally provide sufficient scientific information to assess safety.

In 1982, FDA and the Center for Food Safety and Applied Nutrition (CFSAN, then the Bureau of Foods) first published the "Redbook I" to describe the criteria the Agency employed for assessing the safety of direct food additives and color additives used in food. In revising "Redbook I" the Agency is taking into account developments in toxicity testing since 1982 and comments received from the scientific community and public concerning the 1982 "Redbook I." As with the 1982 "Redbook I," the tiered system for determining concern levels and minimum sets of toxicity tests for compounds assigned to each concern level are discussed in this document. In addition to conventional types of toxicity tests, new or significantly expanded sections include metabolism and pharmacokinetics, immunotoxicity, neurobehavioral toxicity, alternatives to whole animal testing, emerging issues in toxicity testing, pathology and statistical considerations, human studies, epidemiological studies, and carcinogenic risk assessment.

A major objective of the 1982 "Redbook I" was to make public the Agency's policy of cyclic review of the safety of additives in food. Since that time, the concept of cyclic review was abandoned and a program entitled "Priority-Based Assessment of Food Additives (PAFA)" was established. The PAFA program maintains a database of administrative, chemical and toxicological information on "Everything Added to the Food in the United States" (EAFUS), including the "Generally Recognized as Safe" (GRAS) compounds and all CFR regulated direct food additives and color additives used in food. It is beyond the scope of this document to provide a comprehensive list of all types of information in PAFA, or to provide a complete description of the procedures now used to evaluate data prior to inclusion in the database. This information will be available to the public by requesting a supplemental document to "Redbook II" entitled "FDA/CFSAN's Priority-Based Assessment of Food Additives Database".

Redbook II should provide useful guidelines to the petitioner in developing the toxicological safety data and documentation section in petition submissions for direct food additives and color additives used in food. A petitioner may follow the guidelines and protocols in "Redbook II," or may choose to use alternative procedures. If a petitioner chooses to use alternative procedures, however, he/she should discuss the procedures informally with the Agency to prevent expenditure of money and effort on activities that may later be determined to be unacceptable to the FDA.

* A notice of availability of the document entitled "FDA/CFSAN's Priority-Based Assessment of Food Additives Database" will be published, and information concerning the document may be requested by contacting the Division of Health Effects Evaluation (HFS-225), CFSAN, FDA, 200 C Street S.W., Washington, D.C.

Acknowledgements

The success of this revision of the Food and Drug Administration's Toxicological Principles for the Safety Assessment of Direct Food Additives and Color Additives Used in Food (Bureau of Foods, 1982) is due to the hard work of many scientists in FDA's Center for Food Safety and Applied Nutrition. Without the conscientious effort of these scientists, and the patience of their supervisors, the publication of this document would not have been possible. We thank all those listed below. If we omitted anyone's name, we extend our sincerest apologies for the oversight.

Michael A. Adams, Deborah Aub, Marilyn F. Balmer, R. Daniel Benz, Garfield N. Biddle, Marvin J. Bleiberg, Herbert Blumenthal, Ivan Boyer, June A. Bradlaw, Mona S. Calvo, Thomas F. X. Collins, Susan Conrath, Gregory M. Cramer, Michael J. DiNovi, Prem N. Dua, Virginia C. Dunkel, Karen B. Ekelman, Lawrence Fishbein, W. Gary Flamm, Charles V. Frazier, Frederick S. Fry, Jr., Chandrakon P. Giri, Walter H. Glinzmann, Stuart L. Graham, Sidney Green, James C. Griffiths, Rosa M. Gryder, David G. Hattan, Sarah H. Henry, Walter Hill, Fred A. Hines, Dennis M. Hinton, Hiltje Irausquin, Benjamin A. Jackson, Carl B. Johnson, Charles J. Kokoski, Chiu Lin, Francis S. Lin, Peng Tu Liu, Ronald J. Lorentzen, Carol A. Mapes, James H. Maryanski, Edwin J. Matthews, Krishna P. Misra, Ronald W. Moch, John Modderman, Michael W. O'Donnell, Jr., Michael J. Prival, William L. Roth, Alan M. Rulis, Robert J. Scheuplein, Samuel I. Shibko, Susan Shoaf, Thomas J. Sobotka, Janet A. Springer, Jean M. Taylor, Linda R. Tollefson, D. Jesse Wagstaff, and John J. Welsh.

Table of Contents

Chapter	Title	Page
Preface	i
Acknowledgements	ii
Table of Contents	iii
List of Figures and Tables	xiv

Chapter I. Introduction

I A.	Major Changes in the Revised Guidelines	2
1.	Introduction	2
2.	Changes in Determining Concern Levels and Recommended Toxicity Tests for Food Additives and Color Additives Used in Food	3
3.	Changes in Toxicity Testing Guidelines	3
a.	General Recommendations for Toxicity Tests	3
b.	Short-Term Tests for Genetic Toxicity	4
c.	Acute Toxicity Tests	4
d.	Short-Term Tests with Rodents and Non-Rodents	4
e.	Subchronic Toxicity Tests with Rodents and Non-Rodents	4
f.	Carcinogenicity Studies with Rodents	4
g.	Combined Chronic Toxicity/Carcinogenicity Studies with Rodents	4
h.	Reproduction and Developmental Toxicity Studies	5
4.	Other Changes	5
a.	Special Toxicity Studies	5
b.	Human Clinical Studies	5
c.	Emerging Issues in Assessment of Safety of Direct Food Additives and Color Additives Used in Food	5
d.	Glossary	5
I B.	Flexibility and Consistency in Guidelines for Toxicity Testing	6
I C.	Applicability of These Guidelines to the Safety Evaluation of Indirect Food Additives	7

Table of Contents Continued

Chapter	Title	Page
Chapter II. Agency Review of Toxicology Information in Petitions for Direct Food Additives and Color Additives Used in Food		
II A.	Introduction	8
II B.	Expediting Review of Toxicology Information	10
II C.	Evaluating Toxicology Information	11
1.	Introduction	11
2.	No-Observed-Effect Level (NOEL)	11
3.	Safety Factors	11
4.	Acceptable Daily Intake (ADI)	12
5.	Carcinogenic Risk Assessment	12
a.	CFSAN's Cancer Assessment Committee (CAC)	12
	Figure 1: Flow Chart Depicting the Various Groups Involved in the Assessment of Cancer Risk at the Center for Food Safety and Applied Nutrition (CFSAN) of the Food and Drug Administration	13
	Figure 2: Four Steps in the Risk Assessment of Additives in Food at the Food and Drug Administration	14
b.	CFSAN's Quantitative Risk Assessment Committee (QRAC)	15

Chapter III. Concern Levels and Recommended Toxicity Tests

III A.	Introduction	18
III B.	Concern Levels	19
1.	Determining Concern Levels	19
	Figure 3: Concern Levels as Related to Chemical Structure and Exposure	21
2.	Structure Category Assignment	22
a.	Introduction	22
b.	Structure Category Assignment of Additives	23
i	Structure Category A Chemicals	24
ii	Structure Category B Chemicals	26
iii	Structure Category C Chemicals	29

Table of Contents Continued

Chapter	Title	Page
IIIB3.	Estimation of Human Exposure to Direct Food Additives and Food Ingredients	43
a.	Parameters for the Exposure Estimate	43
	Table 1: Summary of Characteristics of Major Databases Used for Estimating Intake of Substances	44
b.	Estimates of Food Intake	46
i.	Annual Poundage Information (Disappearance Data)	46
ii	Dietary Surveys of Food Intake	46
iii	Substance Concentration Data (Petitioner Supplied Data)	47
c.	Preparing Exposure Estimates	47
i.	Estimating Exposure for Pre-Market Evaluations	47
ii	Exposure Estimates for Substances Currently in the Market Place	49
d.	Conversion Factors	49
e.	Summary	50
III C.	Recommended Toxicity Tests	51
1.	Recommended Minimum Set of Toxicity Tests	51
a.	Minimum Set of Toxicity Tests for Concern Level III Substances	51
	Figure 4: Summary of the Toxicity Tests Recommended for Substances in Different Levels of Concern	52
b.	Minimum Set of Toxicity Tests for Concern Level II Substances	53
c.	Minimum Set of Toxicity Tests for Concern Level I Substances	53
2.	Selecting Additional Toxicity Tests	54
a.	Acute Toxicity Tests	54
b.	Short-term Toxicity Tests with Rodents and Non-Rodents	54
c.	Subchronic Toxicity Tests with Rodents	54
d.	One-Year Toxicity Tests with Non-Rodents	55
e.	Carcinogenicity Studies with Rodents	55
f.	Two-Generation Reproduction Studies with a Teratology Phase	55
g.	Gavage Administration of the Test Compound in Teratology Studies	55
h.	Metabolism and Pharmacokinetic Studies	55
i.	Neurotoxicity Studies	56
j.	Immunotoxicity Studies	56

Chapter IV. Guidelines for Toxicity Tests

IV A.	Introduction	57
IV B.	General Recommendations for Toxicity Studies	58
1.	General Guidelines for Toxicity Studies	58
a.	Good Laboratory Practice	58

Table of Contents Continued

Chapter	Title	Page
IVB1 b.	Test Animals	58
i.	Care, Maintenance and Housing	58
ii.	Selection of Rodent Species, Strains and Sex	58
iii.	Infected Animals	58
iv.	Animal Identification	59
v.	Caging	59
vi.	Feed and Water	59
vii.	Diet	59
viii.	Assignment of Control and Experimental Animals	60
ix.	Mortality	60
x.	Autolysis	60
xi.	Necropsy	60
c.	Test Substances	61
i.	Composition	61
ii.	Route of Administration	61
iii.	Controls	62
iv.	Selection of Treatment Doses	62
d.	Observations and Clinical Tests	62
i.	Observation of Test Animals	62
ii.	Body Weight and Food Intake Data	63
iii.	Clinical Testing	63
e.	Necropsy and Microscopic Examination	65
i.	Gross Necropsy	65
ii.	Organ Weight	65
iii.	Preparation of Tissues for Microscopic Examination	65
iv.	Microscopic Evaluation	67
v.	Histopathology	67
2.	Summary Guidelines for Reporting the Results of Toxicity Studies	68
a.	Identification	68
b.	Good Laboratory Practice for Nonclinical Laboratory Studies	68
c.	Body of the Report	68
i.	Protocol and Amendments	68
ii.	Summary and Conclusion	69
iii.	Materials	69
iv.	Methods	70
v.	Results	70
vi.	References	72
3.	Pathology Considerations in Toxicity Studies	73
a.	Description of the Process for Review of Pathology Data	73
b.	Common Problems Encountered During Review of Pathology Data	75
i.	Lack of Morphologic Descriptions of Lesions	75
ii.	Inconsistency in Applying Diagnostic Terminology	75
iii.	Incomplete Descriptions of the Results of Gross Pathology Examinations	75
iv.	Inaccurate Summaries of Data	76
v.	Failure to Adequately Discuss the Results of Pathology Examinations	76

Table of Contents Continued

Chapter	Title	Page
IVB3 c.	General Recommendations for Reporting Pathology Data	76
i.	Arranging Tabular Data and Morphological Observations	76
ii.	Summary Tables	77
iii.	Cross Reference Table	77
iv.	Animal Disposition Table	77
v.	Pathologist's Narrative	77
4.	Statistical Considerations in Toxicity Studies	78
a.	Specific Statistical Issues	79
i	Study Protocol and Design	79
ii	Presentation of Collected Data	80
iii	Presentation and Interpretation of Analytical Results	81
iv	Support from CFSAN Statistical Reviewers	83
5.	Diets for Toxicity Studies	84
a.	Types of Diets	86
i.	Natural Ingredient Diets.....	86
ii.	Purified Diets	86
b.	Issues to Consider when Selecting Diets for Animals in Toxicity Studies	86
IV C.	Guidelines for Recommended Toxicity Studies	88
1.	Short-Term Tests for Genetic Toxicity	88
a.	Definition of Genetic Toxicity Testing	88
b.	Rationale for Selection of Specific Genetic Toxicity Endpoints	88
c.	Test Battery for Genetic Toxicity Testing	89
i.	Gene Mutation in <i>Salmonella typhimurium</i>	89
ii.	Gene Mutation in Mammalian Cells <i>In Vitro</i>	90
iii.	Cytogenetic Damage in Mammalian Cells	92
d.	Additional Scientifically Justified Genetic Toxicity Tests	93
i.	<i>In Vitro</i> Mammalian Cell Transformation Assay	93
ii.	DNA Damage in Mammalian Cells (Unscheduled DNA Synthesis)...	94
e.	Standards for Acceptability of Test Results	95
i.	Experimental Parameters	95
ii.	Test Substance Parameters	96
iii.	Additional Concerns	97
f.	Data Collection and Evaluation	98
g.	Reporting Requirements	99
2.	Acute Oral Toxicity Tests	100
a.	Limit Tests	101
b.	Dose-Probing Tests	101
c.	Up-and-Down Tests	102
d.	Pyramiding Tests	102
3.	Short-Term Toxicity Tests with Rodents and Non-Rodents	103
a.	Experimental Animals	103
i.	Species and Age	103
ii.	Number and Sex	103
b.	Administration of the Test Substance	103
i.	Duration of Testing	103
ii.	Number of Doses	103

Table of Contents Continued

Chapter	Title	Page
IVC3	c. Observations and Clinical Tests	104
	i. Observation of Test Animals	104
	ii. Neurotoxicity Screening	104
	iii. Clinical Testing	104
	d. Necropsy and Histopathology Examination	105
4.	Subchronic Toxicity Tests with Rodents and Non-rodents	106
	a. Experimental Animals	106
	i. Species and Age	106
	ii. Number and Sex	106
	b. Administration of the Test Substances	106
	i. Duration of Testing	106
	ii. Dosed Groups	106
	c. Observations and Clinical Tests	107
	i. Observations of Test Animals	107
	ii. Neurotoxicity Screening	107
	iii. Clinical Testing	107
	d. Necropsy and Histopathology Examination	108
5.	One-Year Toxicity Tests with Non-Rodents	109
	a. Experimental Animals	109
	i. Species and Age	109
	ii. Number and Sex	109
	b. Administration of the Test Substance	109
	i. Duration of Testing	109
	ii. Dosed Groups	109
	c. Observations and Clinical Tests	110
	i. Observations of Test Animals	110
	ii. Clinical Testing	110
	d. Necropsy and histopathology examination	110
6.	Carcinogenicity Studies with Rodents	111
	a. Experimental Animals	111
	i. Age	111
	ii. Species and Strain	111
	iii. Number and Sex	111
	b. Administration of the Test Substance	113
	i. Duration of Testing	113
	ii. Dosed Groups	113
	c. Observations and Clinical Tests	115
	i. Observations of Test Animals	115
	ii. Clinical Testing	115
	d. Necropsy and Histopathology Examination	115
7.	Combined Chronic Toxicity/Carcinogenicity Studies with Rodents.....	116
	a. Experimental Animals	116
	i. Age	116
	ii. Species and Strain	116
	iii. Number and Sex	117
	b. Administration of the Test Substance	117
	i. Duration of Testing	117
	ii. Dosed Groups	117

Table of Contents Continued

Chapter	Title	Page
IVC7 c.	Observations and Clinical Tests	118
i.	Observations of Test Animals	118
ii.	Clinical Testing	118
d.	Necropsy and Histopathology Examination	119
8.	<i>In Utero</i> Exposure Phase for Addition to Carcinogenicity Studies with Rodents ...	120
a.	Experimental Animals	120
i.	Species and Strain Selection	120
ii.	Age	120
iii.	Number and Sex	120
iv.	Caging and Animal Maintenance	120
b.	Dose Selection, Treatment Period, and Method of Dosing	121
i.	Duration of Treatment	121
ii.	Route of Administration	121
iii.	Selection of Dose Levels	121
iv.	Mating Procedures	121
v.	Standardizing the Number of Pups per Litter	121
vi.	Selection of F ₁ Animals	122
c.	Clinical Observations	122
i.	Parental Animals	122
ii.	F ₁ Animals	122
d.	Other Recommendations	122
i.	Termination of P and F ₁ Animals not Selected for the Post-Weaning Phase	122
ii.	Data Reporting	122
9.	Reproduction and Developmental Toxicity Studies	123
a.	Guideline for Reproduction Studies	123
i.	Experimental Animals	123
	Figure 5: Schematic for a 2-Generation Reproduction Study with a Teratology Phase	124
ii.	Dose Selection, Treatment Period, and Method of Dosing ...	125
	Figure 6: Schematic for a 3-Generation Reproduction and Teratology Study ...	127
iii.	Clinical Observations and Histopathology Examination	128
	Figure 7: Continuous Breeding Protocol	129
iv.	Endpoints of Reproductive Toxicity	131
v.	Reporting the Results of Reproduction Studies	133
b.	Guideline for Developmental Toxicity Studies	133
i.	Experimental Animals	134
ii.	Dose Selection, Treatment Period, and Method of Dosing ...	134

Table of Contents Continued

Chapter	Title	Page
Chapter V. Additional Recommended Studies		
V A.	Introduction	138
1.	Metabolism and Pharmacokinetics	138
2.	Neurotoxicity	138
3.	Immunotoxicity	139
V B.	Metabolism and Pharmacokinetic Studies	140
1.	Considerations in the Design of, Analysis of, and Use of Data from Metabolic and Pharmacokinetic Studies	140
a.	Design and Analysis of Metabolic and Pharmacokinetic Studies	140
i.	Test Compound	141
ii.	Animals	141
iii.	Route of Administration	142
iv.	Dosage Regimen	142
v.	Sampling	143
vi.	<i>In Vitro</i> Studies	143
vii.	Analysis of Data	144
b.	Use of Data from Metabolism and Pharmacokinetic Studies	146
2.	Recommended Metabolism and Pharmacokinetic Studies	147
3.	Additional Metabolism and Pharmacokinetic Studies	148
V C.	Neurotoxicity Studies	149
1.	Evaluating Neurotoxicity	151
a.	Screening	152
b.	Characterization	154
c.	Dose-Response	155
2.	Recommended Criteria for Designing a Neurotoxicity Screen	155
V D.	Immunotoxicity Studies	157
1.	Immunity: A Brief Review	157
a.	Humoral Immunity	157
b.	Cell-Mediated Immunity	158
c.	Non-Specific Immunity	158
2.	Key Concepts in Immunotoxicity Testing	159
3.	Indicators of Possible Immune Toxicity	160
a.	Indicators from Short-Term and Subchronic Toxicity Studies	160
b.	Indicators from Developmental Toxicity Studies	161
4.	Expanded Type 1 Immunotoxicity Tests	162
5.	Type 2 Immunotoxicity Tests	164
6.	Relevance of Primary Indicators of Immune Toxicity to Health	166
a.	Hematological Indicators	166
b.	Serum Protein Indicators	166
c.	Histopathology Indicators	167
d.	Body and Organ Weights	167
7.	Adequacy and Reliability of Primary Indicators of Immune Toxicity	167

Table of Contents Continued

Chapter	Title	Page
VD 8.	Recommendations for Further Immunotoxicity Testing when Primary Indicators are Positive	169
a.	Immunotoxicity Studies for Compounds that have been Assigned to Concern Level III	169
b.	Immunotoxicity Studies for Compounds that have been Assigned to Concern Level II	172
c.	Immunotoxicity Studies for Compounds that have been Assigned to Concern Level I	172
9.	Animal Models for Immunotoxicity Tests	173
a.	Rodent Models	173
b.	Non-rodent Models	173
10.	Recommended Strategy for Assessing the Immunotoxic Potential of Food Additives and Color Additives Used in Food	174
11.	Conclusion	174
	Table 2: Summary of Immunotoxicity Testing Requirements for Direct Food Additives and Color Additives Used in Food	175

Chapter VI. Human Studies

VI A.	Clinical Evaluation of Foods and Food Additives	177
1.	General Considerations for Clinical Studies of Foods and Food Additives ...	177
2.	Specific Considerations for Clinical Studies of Foods and Food Additives ...	179
a.	Protocol Design	179
b.	The Study Population	180
c.	Statistical Analyses	182
3.	Sequence of Clinical Studies for Foods and Food Additives	183
a.	Early Clinical Studies	183
b.	Further Clinical Studies	184
4.	Submitting Reports of Clinical Studies on Foods and Food Additives	185
5.	Appendix A--Principles of Institutional Review and Informed Consent	186
a.	Principles of Institutional Review	186
b.	Principles of Informed Consent	186
VI B.	Epidemiology Studies	187
1.	Descriptive Epidemiology Studies	187
a.	Correlational Studies	187
b.	Case Reports	188
2.	Analytic Epidemiology Studies	188
a.	Cross-Sectional Studies	188
b.	Prospective Studies	188
c.	Retrospective Studies	189
d.	Meta-Analyses	190

Table of Contents Continued

Chapter	Title	Page
Chapter VII. Emerging Issues in Safety Assessment of Food Additives and Color Additives Used in Food		
VII A.	Introduction	191
VII B.	Macro-Additives	191
1.	Nutritional Concerns in Animal Toxicity Tests	192
2.	Absorption, Metabolism, Distribution and Elimination Studies	192
3.	Impurities and By-products	193
4.	Clinical Studies	193
VII C.	Safety of Foods and Food Additives Derived from New Plant Varieties by Biotechnology	194
	Figure 8: Summary of Safety Assessment of New Varieties	195
VII D.	Enzymes	197
VII E.	Microbially Derived Food Ingredients	198
VII F.	Advances in the Development of Alternatives to Whole Animal (Vertebrate) Testing	199
1.	Reasons for Developing Alternative Tests	199
2.	Possible Applications of Alternative Tests	200
3.	Limitations of Alternative Tests	201
4.	Current Use of <i>In vitro</i> Tests	201
VII G.	Heritable and Somatic Genetic Toxicity	203
1.	Rationale for Testing for Heritable and Somatic Genetic Toxicity	203
2.	Rationale for Selecting a Specific Test Battery	203

Chapter VIII. Glossary

VIII	Glossary: Acronyms and Definitions	205
------	--	-----

Table of Contents Continued

Chapter	Title	Page
Chapter VII. Emerging Issues in Safety Assessment of Food Additives and Color Additives Used in Food		
VII A.	Introduction	191
VII B.	Macro-Additives	191
1.	Nutritional Concerns in Animal Toxicity Tests	192
2.	Absorption, Metabolism, Distribution and Elimination Studies	192
3.	Impurities and By-products	193
4.	Clinical Studies	193
VII C.	Safety of Foods and Food Additives Developed by Biotechnology	194
	Figure 8: Summary of Safety Assessment of New Varieties	195
VII D.	Enzymes	197
VII E.	Microbially Derived Food Ingredients	198
VII F.	Advances in the Development of Alternatives to Whole Animal (Vertebrate) Testing	199
1.	Reasons for Developing Alternative Tests	199
2.	Possible Applications of Alternative Tests	200
3.	Limitations of Alternative Tests	201
4.	Current Use of <i>In vitro</i> Tests	201
VII G.	Heritable and Somatic Genetic Toxicity	203
1.	Rationale for Testing for Heritable and Somatic Genetic Toxicity	203
2.	Rationale for Selecting a Specific Test Battery	203

Chapter VIII. Glossary

VIII	Glossary: Acronyms and Definitions	205
------	--	-----

Table of Contents Continued

Chapter	Title	Page
Chapter IX. References		
I	Introduction	207
I A (1)	Major Changes in the Revised Guidelines (Introduction).....	207
I B	Flexibility and Consistency in Guidelines for Toxicity Testing	208
II A	Introduction	208
II C (5)	Evaluating Toxicological Information for Food Additives and Color Additives Used in Food (Carcinogenic Risk Assessment)	208
III B (2)	Concern Levels (Structure Category Assignment)	209
III B (3)	Concern Levels (Estimation of Human Exposure to Direct Food Additives and Food Ingredients)	211
IV A	Introduction	212
IV B (1 d)	General Recommendations for Toxicity Studies (General Guidelines for Toxicity Studies).....	212
IV B (3)	General Recommendations for Toxicity Studies (Pathology Considerations in Toxicity Studies)	212
IV B (4)	General Recommendations for Toxicity Studies (Statistical Considerations in Toxicity Studies)	212
IV B (5)	General Recommendations for Toxicity Studies (Diets for Toxicity Studies) ...	212
IV C (1)	Guidelines for Oral Toxicity Studies (Short-Term Tests for Genetic Toxicity) ...	215
IV C (2)	Guidelines for Oral Toxicity Studies (Acute Toxicity Studies)	219
IV C (6)	Guidelines for Oral Toxicity Studies (Carcinogenicity Studies with Rodents) ...	220
IV C (9)	Guidelines for Oral Toxicity Studies (Reproductive and Developmental Toxicity Studies).....	221
V B	Metabolism and Pharmacokinetic Studies	222
V C	Neurotoxicity Studies	224
V D	Immunotoxicity Studies	227
VI A	Guidelines for the Clinical Evaluation of Foods and Food Additives	233
VI B	Epidemiology Studies	233
VII C	Advances in the Development of Alternatives for Whole Animal Testing	234
VII D	Safety of Foods and Food Additives Developed by Biotechnology	236
VII E	Short-Term Tests for Heritable and Somatic Genetic Toxicity.....	236

List of Figures and Tables

Number	Title	Page
--------	-------	------

Figures

1	Flow Chart Depicting the Various Groups Involved in the Assessment of Cancer Risk at the Center for Food Safety and Applied Nutrition (CFSAN) of the Food and Drug Administration	13
2	Five Steps in the Risk Assessment of Additives in Food at the Food and Drug Administration	14
3	Concern Levels as Related to Chemical Structure and Exposure	21
4	Summary of the Toxicity Tests Recommended for Different Levels of Concern	52
5	Schematic for a 2-Generation Reproduction Study with a Teratology Phase ...	124
6	Schematic for a 3-Generation Reproduction and Teratology Study	127
7	Continuous Breeding Protocol	129
8	Summary of Safety Assessment of New Varieties	195

Tables

1	Summary of Characteristics of Major Databases Used for Estimating Intake of Substances	44
2	Summary of Immunotoxicity Testing Requirements for Direct Food Additives and Color Additives Used in Food	175

Chapter I

Introduction

One of the responsibilities of the Food and Drug Administration (FDA) and its Center for Food Safety and Applied Nutrition (CFSAN) is to evaluate the safety of food additives and color additives used in food. Although the Food, Drug, and Cosmetic Act (the Act)^a defines food additives generally, the Agency has further divided the universe of food additives into direct food additives (which are of interest here) and indirect food additives (see Chapter I C). Direct food additives are substances deliberately added to food to achieve a specific technical effect, such as emulsification or calorie reduction. The "safety" of these additives is defined in sections 70.3 and 170.3 of Title 21 of the Code of Federal Regulations (CFR) as a reasonable certainty that a substance is not harmful under the intended conditions of use.^b

Under the Food, Drug, and Cosmetic Act, the safety of food additives and color additives used in food must be established prior to marketing by evaluating the probable exposure to the substance and appropriate toxicological and other scientific information. Thus, approval of any new food additive or color additive used in food depends in part upon the outcome of toxicity tests that are performed and evaluated prior to marketing.

FDA consistently has taken the position that various types of scientifically valid information can support a finding that the proposed use of an additive will not cause harm to the consumer. Thus, the Agency continues to adjust testing recommendations for direct food additives and color additives used in food as necessary to reflect both the steady progress of science and current information about population exposure to additives.

In 1982, FDA's Bureau of Foods published its guidelines: *Toxicological Principles for the Safety Assessment of Direct Food Additives and Color Additives Used in Food*.^c The guidelines set out a system of tiered information recommendations for additives in food. They describe how FDA incorporates information about expected human exposure and chemical structure/activity relationships into initial Concern Levels for food and color additives used in food. The Concern Levels provide guidance on how much toxicity testing should be done for different levels of estimated human exposure. The 1982 guidelines also set forth the toxicological safety evaluation criteria that FDA uses in judging the safety of additives.

This document is the Agency's first published revision of the 1982 guidelines. "A submission conforming to these recommendations would normally provide sufficient scientific information to evaluate safety. However, these guidelines are not intended as rigid rules and they do not preclude the petitioner from demonstrating safety by using other types of toxicological data and information. The flexibility of FDA recommendations contained in this document is discussed in Chapter I B.

^a Committee of Labor and Public Welfare (1988)¹

^b Committee of Labor and Public Welfare (1988);¹ Code of Federal Regulations (1992);² Code of Federal Regulations (1992)³

^c U.S. FDA report (1982)⁴

I A. Major Changes in the Revised Guidelines

1. Introduction

This section summarizes major changes in this revision of the 1982 *Toxicological Principles for the Safety Assessment of Direct Food Additives and Color Additives Used in Food*.^a In general, these changes derive from three major sources: 1) Changes in the purpose of the publication; 2) comments received on the 1982 publication; and 3) increased scientific knowledge and technological advances since 1982. Attempts also have been made to achieve consistency with guidelines published by other agencies, countries, and international organizations, when such consistency does not compromise FDA's ability to ensure the safety of direct food additives and color additives used in food.

A major objective of the 1982 publication was to make public the principles of the Agency's priority-based assessment of food additives (PAFA). For example, the 1982 publication described in addition to the "current" guidelines, "core standards" for toxicity studies. Core standards define standards to be used in determining whether previously conducted toxicity studies provide information that would be a useful addition to the PAFA database. There has been some confusion about whether core standards represent minimally acceptable protocols for conducting toxicity studies to support the safety of newly petitioned food and color additives used in food; in general, they do not. While FDA will continue to make information about PAFA available to the public upon request, it will not be presented in this publication. A separate document is available containing information on the PAFA database.^b

Other changes in this revision are aimed at clarifying how toxicology review fits into the overall petition review process for direct food additives and color additives used in food. Thus, guidelines on how to submit machine-readable data for review by FDA (see Chapter II B), and information about how the Agency assesses the safety of food and color additives used in food (see Chapter II C) have been incorporated into the revised guidelines.

After publication of *Toxicological Principles for the Safety Assessment of Direct Food Additives and Color Additives Used in Food* in 1982, the Agency received thoughtful comments about its recommended guidelines from scientists, consumer interest groups, health agencies in other countries, companies in the food industry, and manufacturers associations. These comments concerned such things as the appropriate balance between exposure and structure/activity information in assigning chemicals to Concern Levels, assessing the safety of food additives that are expected to be consumed in large quantities, and the recommended duration of rodent carcinogenicity bioassays. Some changes in this publication resulted, in part, from consideration of these comments.

Finally, changes in this publication derive from increased knowledge about toxicological processes and outcomes, from technological advances in the food industry, and from changes in public opinion that focus on the need to pay attention to the humane and economically efficient use of laboratory animals in scientific research.

^a FDA (1982).¹

^b A notice of availability of the document entitled "FDA/CFSAN's Priority-Based Assessment of Food Additives Database" will be published, and information concerning the document may be requested by contacting the Division of Health Effects Evaluation (HFS-225), CFSAN, FDA, 200 C Street S.W., Washington, D.C.

I A Major Changes in the Revised Guidelines Continued

2. Changes in Determining Concern Levels and Recommended Toxicity Tests for Food Additives and Color Additives Used in Food

■ Estimation of Human Exposure to Food Additives and Food Ingredients: Information about how the Agency estimates human pre-market exposure to direct food additives and food ingredients is provided (see Chapter III B 3) in this document.

■ Structure Category Assignments: Several changes in structure category assignments have been made. In general, these changes derive from scientific information available since 1982. Some changes also were designed to enhance the reader's understanding of how additives are assigned to Structure Categories A, B, and C (see Chapter III B 2).

■ Minimum Sets of Toxicity Tests: Changes have been made to the recommended minimum set of toxicity tests for additives assigned to each Concern Level (see Figure 3 in Chapter III B 1); these changes are listed below:

i) Concern Level I: Screens for neurotoxicity and immunotoxicity have been added to the recommended short-term toxicity test with rodents.

ii) Concern Level II: Metabolism and pharmacokinetic studies now are recommended for these substances. Screens for neurotoxicity and immunotoxicity have been added to the recommended subchronic toxicity tests with rodents and non-rodents and the reproduction study with a teratology phase. The recommended reproduction study now consists of two generations, with one litter per generation.

iii) Concern Level III: Metabolism and pharmacokinetic studies now are recommended for these substances. Screens for neurotoxicity and immunotoxicity have been added to the recommended subchronic toxicity test with rodents and the reproduction study with a teratology phase. The recommended reproduction study now consists of two generations, with one litter per generation.

■ Subchronic Toxicity Test with Rodents: For Concern Level III substances, FDA now recommends that a subchronic feeding study with rodents be completed before carcinogenicity bioassays are begun.

3. Changes in Toxicity Testing Guidelines

a. General Recommendations for Toxicity Tests

General recommendations for toxicity tests are discussed in Chapter IV B. These include guidelines for test animals and test substances (see Chapter IV B 1) and for reporting the results of toxicity studies (see Chapter IV B 2); recommendations for pathology and statistical considerations in toxicity tests (see Chapters IV B 3 and 4, respectively); and recommendations concerning the use of various types of animal diets for toxicity studies (see Chapter IV B 5).

I A Major Changes in the Revised Guidelines Continued

b. Short-Term Tests for Genetic Toxicity

This guideline recommends a modified battery of short-term tests for genetic toxicity that includes: 1) *Salmonella typhimurium* reverse mutation assay, 2) *in vitro* mutagenicity assay in mammalian cells, and 3) *in vivo* cytogenetics assays (chromosomal aberrations in mouse or rat bone marrow and the mouse micronucleus test) (see Chapter IV C 1 c). Additional, scientifically justified genetic toxicity tests are also discussed in the chapter (see Chapter IV C 1 d).

c. Acute Toxicity Tests

Guidelines in Chapter IV C 2 stress that acute toxicity data are not required for making the final decision on the safety of direct food additives and color additives used in food. If petitioners decide to conduct acute toxicity studies for new materials that may be added directly to food, this guideline recommends alternatives to the classic LD₅₀ test.

d. Short-Term Toxicity Tests with Rodents and Non-Rodents

The guideline for this test has been modified to include screens for neurotoxicity and immunotoxicity (see Chapter IV C 3). In addition, FDA recommends that rodents be single-caged (instead of gang-caged) and that a complete histopathology evaluation be performed for all animals in the study (see Chapter IV B 1).

e. Subchronic Toxicity Tests with Rodents and Non-Rodents

The guideline for this test has been modified to include screens for neurotoxicity and immunotoxicity (see Chapter IV C 4). In addition, FDA recommends that rodents be single-caged (instead of gang-caged) and that a complete histopathology evaluation be performed for all animals in the study (see Chapter IV B 1).

f. Carcinogenicity Studies with Rodents

Important changes in the guideline for this study include recommendations that rodents be single-caged (instead of gang-caged), that bioassays begin with at least 50 animals of each sex per experimental and control groups, that rodent bioassays be terminated after 104 weeks of exposure to the test substance, and that microscopic examination of recommended tissues and organs be performed on all animals in the study (see Chapter IV C 6).

g. Combined Chronic Toxicity/Carcinogenicity Studies with Rodents

Changes in the guideline for this study are similar to changes in the guideline for carcinogenicity bioassays with rodents, and include recommendations that rodents be single-caged (instead of gang-caged), that bioassays begin with at least 50 animals of each sex per experimental and control groups, that the carcinogenicity segment of the study be terminated after 104 weeks of exposure to the test substance, and that microscopic examination of recommended tissues and organs be performed on all animals in the study (see Chapter IV C 7).

I A Major Changes in the Revised Guidelines Continued

h. Reproduction and Developmental Toxicity Studies

Two generations, with one litter per generation, are recommended as the minimum reproduction study (see Chapter IV C 9). If results from the minimum reproduction study or other toxicity tests indicate that a test compound may be associated with reproductive toxicity, the minimum reproduction study should be expanded. For example, the guideline includes optional procedures for inclusion of additional litters per generation, additional generations, a test for teratogenic effects, and reproductive assessment by continuous breeding. Guidelines for reproduction and developmental toxicity studies have been modified to include an expanded assessment of the effects of the test compound on males and to provide a screen for neurotoxicity and immunotoxicity.

4. Other Changes

a. Special Toxicity Studies

FDA now recognizes that information about metabolism and pharmacokinetics, neurotoxicity and immunotoxicity are significant endpoints in assessing the safety of direct food additives and color additives used in food. Recommended strategies for assessing these endpoints are described in Chapters V B, C and D, respectively.

b. Human Clinical Studies

FDA does not require petitioners to conduct human clinical studies to support the safety of direct food additives and color additives used in food. However, when petitioners elect to perform such studies, the Agency recommends that the studies conform to the guidelines presented in Chapter VI B.

c. Emerging Issues in the Assessment of the Safety of Direct Food Additives and Color Additives Used in Food

Chapter VII discusses special tests or approaches to testing that may be useful in assessing the safety of additives intended for use at high levels of exposure (macro-additives), bioengineered additives, additives that are enzymes, and microbiologically-derived additives. In addition, this chapter discusses alternatives to the use of whole animals in assessing the safety of food and color additives and the Agency's acknowledgement that tests for heritable and somatic genetic toxicity have been developed and may be useful in evaluating the safety of food and color additives used in food in the future.

d. Glossary

A glossary (see Chapter VIII) has been provided in this document.

I B. Flexibility and Consistency in Guidelines for Toxicity Testing

Although many different agencies regulate the same chemicals (for example benzene may be regulated for different uses by FDA, the Environmental Protection Agency and the Occupational Safety and Health Administration), the toxicity testing guidelines developed separately by the various health regulatory agencies are not always uniform. Differences among guidelines can result in unnecessary duplication of effort and inefficient use of scarce testing resources. When possible, the guidelines presented in this section are consistent with guidelines of other agencies and organizations. However, it must be emphasized that food additives can present special needs for testing and the guidelines presented in this section continue to reflect such needs. Thus, we have retained the recommendation that *in utero* exposure be added to one of two recommended carcinogenicity bioassays (see Chapters IV C 6,7, and 8).

Changes occurring in the global economy are now having, and will continue to have, effects on the food chemical regulatory work of FDA as well as on the industry it regulates. The European Economic Community is expected to unite under new legislation that promises to reduce trade barriers between the member European nations;⁴ in December 1986 Canada and the United States signed a Free Trade Agreement; in 1992 Mexico, Canada and the United States signed the North American Free Trade Agreement (NAFTA). A goal of these agreements is to harmonize regulatory requirements and, where possible, to reduce or eliminate trade barriers between the signatory nations. Food and food chemicals clearly constitute an important area of trade likely to be affected by these agreements.

Much work needs to be done to harmonize international food chemical regulation. Nations have different regulatory schemes and often different permitted substances in food. Several European nations, for example, regard flavor chemicals differently, compared with the United States or the United Kingdom.⁵ Canada and the United States regulate packaging materials differently.

FDA's guidelines for toxicity tests for direct food additives and color additives used in food continue to emphasize that there is no substitute for sound scientific judgement. These guidelines are recommendations--not hard and fast rules. If an investigator believes that he/she can provide the Agency with useful toxicological information by modifying a recommended study protocol, and is able to support the modification with sound scientific arguments, then the investigator should propose the modified protocol to toxicologists at CFSAN. As always, we urge petitioners to consult with the Agency about protocols for toxicity tests before the studies begin.

⁴ Elkes (1989)¹

⁵ Grignolo (1989);² Schneebaum (1989)³

I C. Applicability of These Guidelines to the Safety Evaluation of Indirect Food Additives

As with the 1982 edition of the guidelines, the tiered system of determining concern levels outlined in this document for safety assessment applies to direct food additives and to color additives used in food. It does not apply to indirect food additives. Indirect food additives are not intentionally added to food; they are substances used as articles or components of articles that are intended for use in packaging, transporting, or holding food. As such, indirect food additives are not intended to become components of the food itself; their potential presence in food may be a result of migration or inadvertent extraction from the food contact surface.

The indirect additives comprise a wide diversity of food-contact situations -- long-term contact with food, as in a final consumer package; intermediate contact, as in a holding container in a food processing plant; short-term, incidental contact, as from a moving belt on a feed line in a food manufacturing operation. The indirect additives also involve a wide range of different chemical structure classes -- from reactive chemical agents used as components of food packaging material or biocides, to inert polymers used for food containers. Thus, indirect food additives present problems for estimating consumer exposure which are different from those associated with substances added directly to food.

FDA traditionally has applied a separate system of tiered information recommendations for indirect food additives that differ somewhat in scope and substance from those for direct additives. The outline for toxicity testing of indirect additives will be provided upon request to the FDA. However, when it is determined that one or several toxicity studies will be required to demonstrate safety of an indirect food additive, the guidelines outlined in this document for conduct of these studies will be applicable.

Chapter II

Agency Review of Toxicology Information in Petitions for Direct Food Additives and Color Additives Used in Food

A. Introduction

The food additive petition review process came into existence in 1958 when Congress enacted the Food Additives Amendment^a to the Federal Food, Drug, and Cosmetic Act (the Act).^b This Amendment creates a pre-market safety evaluation process for new substances added to food, "food additives." A similar statute, the Color Additive Amendments of 1960,^c created analogous requirements for color additives used in foods, drugs, cosmetics, or medical devices. "Color additive" used in food is defined in section 201(t) of the Act; "food additive" is defined in section 201(s) of the Act.

Since 1958, before a food additive may be used, an authorizing regulation must be in effect. Approval of a petition for an additive and issuance of an authorizing regulation require that the Agency conclude that the additive is safe for its intended conditions of use. This safety requirement, embodied in section 409(c)(3)(A), is often referred to as the general safety clause for food additives. When the proponent of the proposed use of the additive has shown that the additive is safe for its intended use, the Agency publishes a regulation in the *Federal Register* establishing permitted conditions for the use of the additive.

When a petition for a direct food additive or color additive used in food is submitted to the Agency, or when the petitioner first contacts FDA, a Consumer Safety Officer (CSO) generally is assigned to the petition. One of the CSO's tasks is to coordinate FDA's review of the petition. When appropriate, the CSO can arrange for the petitioner to meet with other individuals in the Agency to discuss specific issues or problems that arise during review of the petition. All communication with the Agency concerning the status or review of the petition should be made through the assigned CSO. General information about the petition review process has been published;^d specific questions should be addressed to the CSO assigned to the petition.

^a Federal Food, Drug and Cosmetic Act amendment (1958)¹

^b Federal Food, Drug, and Cosmetic Act (1958);² Committee of Labor and Public Welfare (1988)³

^c Color Additive Amendment (1960)⁴

^d Rulis (1990)⁵

II A Introduction *Continued*

The Act and the Code of Federal Regulations* specify the basic elements that a petition must contain. One of these elements is safety data on the additive, which is usually provided in the form of toxicity studies. Toxicologists, pathologists, and mathematicians evaluate any toxicity studies included in the petition. If appropriate, toxicologists can recommend that carcinogenicity studies be evaluated by special CFSAN committees: the Cancer Assessment Committee (CAC) and the Quantitative Risk Assessment Committee (QRAC); for more information on these committees, see Chapter II C 5 i and ii.

Review of toxicity studies and other toxicology information results in an estimate of the acceptable daily intake (ADI) for the direct food additive or color additive used in food. The ADI is typically based on the dose level of the additive in animal studies that was shown to cause no adverse effect, multiplied by an appropriate safety factor (often 1/100; see 21 CFR 170.22). Chronic ingestion of the additive at the ADI is considered consistent with a reasonable certainty of no harm.

FDA urges individuals or corporations preparing to submit petitions for direct food additives or color additives used in food to consult with the Agency early in the planning stages. For example, before the petition is submitted, petitioners can submit toxicity study protocols to FDA for review by Agency scientists. This can help the petitioner perform toxicity studies and prepare data in a form that will expedite the Agency's review of the information in the petition (for more information on expediting review, see Chapter II B).

This document delineates the toxicology information deemed appropriate for assessing the safety of direct food additives and color additives used in food. However, guidelines contained in this document are only one possible approach among many to providing the toxicological basis for an assessment of safety. We urge petitioners to discuss alternative approaches and toxicity test protocols with the Agency before toxicity tests are begun.

* USCFR (1992)*

II B. Expediting Review of Toxicology Information

The Agency recommends that petitioners use the following approaches to minimize requests for additional data and to expedite review of direct food additive and color additive petitions:

- Make sure that petitions are formatted properly and contain complete and adequate information before submitting them for review. Guidelines and recommendations contained in this publication should be consulted before the petition is submitted.
- Initiate interactions between petitioner's representatives and Agency CSOs and scientists before the petition is submitted. Such interactions can involve Agency review of toxicity study protocols and Agency recommendations about the extent of toxicity testing that may be recommended to adequately assess the safety of the food additive or color additive used in food.
- Submit toxicology data in machine-readable form. During review of the safety of a food additive or color additive used in food, it may be necessary for scientific reviewers to re-analyze some of the data in a submission. A large proportion of the work in such a re-analysis is computer entry and verification of data. Therefore, much time would be saved if data are submitted in a machine-readable form (magnetic tape for the IBM mainframe standard or floppy disks for IBM personal computers. Please note that the Agency no longer has the capability to read punched cards). General guidelines for submitting machine-readable data follow, but petitioners are urged to contact the Agency before submitting machine-readable data to discuss modifications to these guidelines.
- Enclosed with the machine-readable data should be:
 - i) the name of a contact person;
 - ii) a printout of the first 100 to 200 records; and
 - iii) the layout of the data. This would include the location of each variable in the record, the type of variable (e.g. character, integer), the permissible range of values, and information about how missing data are stored.
- Magnetic tape format needs to be 9-track, with 6250 bpi preferred (although 800 and 1600 bpi are also readable). Data should be recorded in IBM-EBCDIC or ASCII, or should be in IBM-TSO or statistical package datasets; please consult with the Agency statisticians about appropriate datasets. Interior labels should be IBM standard with volume number and dataset names. Unlabeled tapes should be accompanied by the record format, record length, blocking factor, and the name of the program that created the tape.
- Floppy disks should be submitted in duplicate; these should be copy-protected because accidental erasure and destruction of disks can occur. The data should be submitted in a form readable by software programs to which the Agency has access; please consult with Agency statisticians about acceptable software.

II C. Evaluating Toxicology Information

1. Introduction

Toxicity testing requirements for assessing the safety of food and color additives used in food have evolved over the past years as knowledge in the field of toxicology has expanded. While short-term or acute studies were considered adequate even for major food additives several decades ago, *today's recommendations generally include comprehensive, long-term toxicity studies.* CFSAN toxicologists exercise their best scientific judgement in determining what toxicity studies are needed for the Agency to adequately assess the safety of a direct food additive or color additive used in food. In making these decisions, the toxicologists take into account what is already known about the properties of a compound, its intended conditions of use, and current standards for toxicity testing.

From data submitted by the petitioner in support of the safety of a direct food additive or color additive used in food, Agency toxicologists determine the no-observed-effect level (NOEL), select an appropriate safety factor, and calculate the acceptable daily intake (ADI) for the substance. These steps are briefly summarized below.

2. No-Observed-Effect Level (NOEL)

Non-treatment-related variations in the incidence of toxic endpoints occur and may depend on a number of factors, including the source of the animals, sex, genetic variations, diet, age at death, environmental conditions and the histological criteria used by pathologists.

However, Agency scientists determine the most sensitive treatment-related toxic endpoint (adverse effect) from the data submitted in support of the petition. This endpoint is the adverse or toxic effect that occurs in test animals at the lowest exposure to the test substance. The highest exposure that does not produce this adverse effect is called the no-observed-effect level (NOEL) or the no-observed-adverse-effect level (NOAEL).

3. Safety Factors

Use of safety factors is based on the observation that toxic substances usually have thresholds below which toxic effects cannot be detected. The safety factor attempts to account for differences between animals and humans and differences in sensitivity among humans. Use of the safety factor is intended to provide an adequate margin of safety for consumers.

For non-cancer endpoints, the NOEL is divided by a safety factor to obtain an estimate of the maximum acceptable daily intake (ADI) of the additive for humans. The selection of a safety factor is based on the biological significance of the endpoint, uncertainties inherent in extrapolating information about adverse effects from toxicity studies in animals to human populations, and other judgmental factors. The food additive procedural regulations (21 CFR 170.22) state that a safety factor of 100 will be used as a general rule in applying animal test data to man. However, exceptions to a safety factor of 100 are permitted in accordance with the nature and extent of data available and the circumstances of use of the food additive. For example, safety factors may be modified because of potentially sensitive sub-populations such as children, geriatrics, individuals with deficiency states, and lack of developed enzyme metabolic systems.

II C Evaluating Toxicology Information Continued

II C 4. Acceptable Daily Intake (ADI)

The acceptable daily intake (ADI) is generally estimated by dividing the no-observed-effect level (NOEL) of a test substance by the safety factor. The NOEL may be expressed as mg test substance per kg body weight of the test animal or as percent or ppm (parts per million) of the test diet for the animal. The ADI is usually expressed in mg additive per kg body weight of humans. A food additive generally is considered safe for its intended use if the estimated daily intake (EDI) of the additive is less than, or approximates, the ADI. Because the ADI is calculated to protect against the most sensitive adverse effect, it also protects against other adverse effects occurring at higher exposures to the ingredient.

5. Carcinogenic Risk Assessment

FDA has found risk assessment to be useful for estimating the risk from carcinogenic contaminants of food or color additives used in food, for helping the Agency to set priorities, and for determining the urgency of a regulatory action.*

Under the general safety clause of the Act, FDA has used risk assessment procedures to determine the upper limit of risk to the consumer from the presence of a carcinogenic contaminant or constituent chemical. For example, FDA approved for permanent listing D&C Green No. 6, which had not been shown to be a carcinogen in appropriate tests, even though it contains the carcinogenic impurity, *para*-toluidine. In this decision, FDA stated its belief that the lifetime upper limit of risk could adequately be estimated from animal data and extrapolated to humans. Although FDA continues to be concerned about carcinogenic contaminants in the food supply, the Agency believes that this approach can be used, where appropriate, without compromising FDA's mandate to protect the public health.

a. CFSAN's Cancer Assessment Committee (CAC)

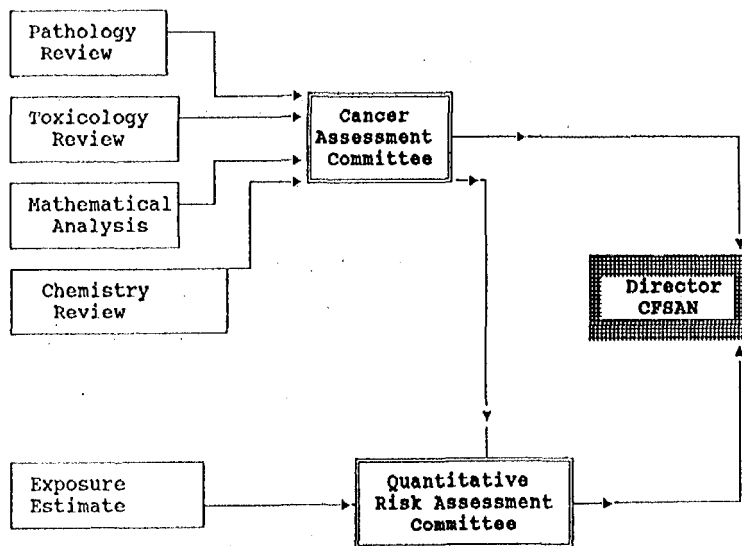
The Cancer Assessment Committee (CAC) is comprised of CFSAN experts in such fields as pathology, toxicology, mathematics, food chemistry and technology, epidemiology, and nutrition. These experts are charged with ensuring a uniform and consistent scientific approach for dealing with diverse problems of carcinogenicity throughout the broad regulatory purview of CFSAN. The CAC reviews all lifetime feeding studies submitted to the Agency in support of the safety of direct food additives and color additives used in food. The risk assessment process also can be triggered when a newly petitioned or previously regulated food or color additive presents a question of possible carcinogenicity. If the CAC determines that a substance is a carcinogen, and if it is believed that a quantitative risk assessment may have impact on the regulation of the substance, the CAC informs the Quantitative Risk Assessment Committee (QRAC, see Chapter II C 5 b) of this decision.

Figure 1 is a flow chart depicting in schematic fashion the groups involved in the risk assessment process at CFSAN. Figure 2 identifies the steps involved in risk assessment at CFSAN; each of the steps in Figure 2 is associated with a particular group or set of groups in Figure 1.

* Lorentzen (1984)¹

Figure 1

Flow Chart Depicting the Various Groups Involved in
the Assessment of Cancer Risk at the
Center for Food Safety and Applied Nutrition (CFSAN)
of the Food and Drug Administration



52454 1823

Figure 2

**Four Steps in the Risk Assessment of Additives in Food
at FDA's Center for Food Safety and Applied Nutrition**

1. Toxicological Evaluation

- **Input Obtained from Internal Experts**
 - Toxicologists
 - Chemists
 - Other Experts
 - Pathologists
 - Biostatisticians
 - Epidemiologists
- **Input Obtained from External Experts (where need is indicated)**

2. Cancer Assessment Committee (CAC) Evaluation

- **CAC Reviews input from Internal and External Experts**
 - Is the Substance a Likely Carcinogen?
If Yes: CAC recommends the studies, tissue sites, species, and sex suitable for quantitative risk evaluation if risk assessment is allowed under the statute
If No: No further consideration by CAC or QRAC is needed

3. Quantitative Risk Assessment Committee (QRAC) Evaluation

- **QRAC Reviews Data and Exposure Potential**
- **QRAC Chooses Risk Assessment Model and Procedure**
- **QRAC Estimates Magnitude of Potential Human Risk**
 - Calculate the Upper Bound Lifetime Risk

4. Action Taken by Director of CFSAN, FDA

- **Makes Risk Management and Policy Recommendations to the Commissioner**

II C Evaluating Toxicology Information Continued

II C 5. Carcinogenic Risk Assessment Continued

As indicated in Figure 1, the CAC plays a central role in the risk assessment process at CFSAN. This standing committee, which was established in 1978, is made up of 10 CFSAN individuals with expertise in the various scientific disciplines related to chemical carcinogenesis: pathology, toxicology, mathematics, and food chemistry and epidemiology. The decisions of the CAC with respect to issues of science are authoritative and invariably form the basis for CFSAN's recommendations to the Commissioner.

In addition to reviewing information presented by the disciplines indicated in Figure 1, the CAC may request additional information from internal and external experts, such as a review of available epidemiological data or a special review of mutagenicity data. The CAC may choose to postpone a final decision on the carcinogenicity of a compound pending the outcome of ongoing or anticipated animal or analytical experiments. In some cases, the CAC may request that CFSAN pathologists review microscope slides from an animal bioassay. External scientific peer review is sometimes requested by the CAC when a particularly difficult or controversial scientific issue is involved.

In general, FDA and CFSAN follow the National Research Council guidelines for risk assessment, described in *Risk Assessment in the Federal Government: Managing the Process*.^a FDA and CFSAN also follow the set of principles for risk assessment contained in the 1985 Office of Science and Technology Policy document, *Chemical Carcinogens: A Review of the Science and its Associated Principles*.^b

There are no universally agreed upon ways of evaluating carcinogenicity data. It is necessary that there be interaction between pathologist, toxicologist and statistician. The role of the pathologist is to decide whether an observed lesion is cancerous or noncancerous.^c The role of the toxicologist is to determine whether the lesion is related to the treatment. The statistician's role is to analyze the mathematical probability of occurrence of the tumors by chance or as a result of treatment.

Some suggested approaches to the assessment of the evidence of carcinogenicity of a substance are discussed in the following sections.

- i) Evaluation of the Adequacy of the Design and Conduct of the Bioassay: The first step in the analysis is a general review of the adequacy of design and conduct of the bioassay to decide whether it is acceptable for evaluation and for deriving conclusions about safety. For example: Was the test chemical properly identified and characterized? Were an adequate number of animals of each sex used per group? Was the test chemical administered for the major part of the life span of the animals? Did sufficient numbers of animals in each group survive long enough for possible late-developing tumors to be manifested? Were there unforeseen events, such as an outbreak of infectious disease, that might invalidate the bioassay? Did the bioassay utilize adequate matched control animals for statistical comparison? Were detailed pathological examinations performed for every tissue?

^a National Research Council (1983)²

^b U.S. Office of Science and Technology Policy report (1985);³ Anonymous (1986)⁴

^c Dua and Jackson (1988)⁴

II C Evaluating Toxicology Information Continued

II C 5. Carcinogenic Risk Assessment Continued

ii) Evaluation of the Possible Increase in Tumor Incidence: Since it is generally believed that cancers arise independently in various parts of the body, it has become customary to treat each potential target site (e.g., brain, lung, liver, kidney, urinary bladder) separately for evaluation. One general exception is the evaluation of types of tumors that may be multicentric in origin, including leukemia and, possibly, tumors originating in blood vessels or nerves, such as hemangioendotheliomas or neurofibrosarcomas. In general, tumor incidence is defined as the number of tumor-bearing animals having tumors at a specific organ site divided by the total number of animals with that organ examined histopathologically.

Judgment of an experienced pathologist is important for proper diagnosing and grouping of lesions for statistical analysis to determine whether or not observed increases in tumor incidence implicate a compound as a carcinogen. The grouping of tumors for statistical evaluation should be based on commonality of histogenic origin. Because it is frequently a matter of arbitrary definition and expert pathologists may disagree about how to designate tumors on the borderline of the continuum between benign and malignant, and because of practical difficulties in categorizing certain tumors as benign or malignant, it is usually necessary to combine the incidence of certain benign tumors with that of malignant tumors occurring in the same tissue and organ for statistical evaluation.

Having recorded the tumors present for each animal, the statistical analysis can be undertaken to evaluate the internal consistency of the data, the reproducibility of the test results, the level of statistical significance, the increase in tumor incidence, the evidence for dose-response relationship or shortened latency period, etc. Methods of statistical analysis for carcinogenicity are available.*

iii) Evaluation of the Extent of Evidence for Carcinogenicity: Because the power of carcinogenesis bioassays that use groups of a few dozen animals is relatively weak for determining carcinogenic activity, it is not surprising that evidence of carcinogenicity is sometimes difficult to establish from a single bioassay. This is so for several reasons, including problems of histological diagnosis, sensitivity of the bioassay, and variability of the background tumor incidence. For these reasons, other correlative information may be necessary to add to the weight of evidence of carcinogenicity of a chemical. In general, the extent of the evidence for carcinogenicity can be determined by considering the following information:

- the number of species or strains with an increased tumor incidence;
- the number of positive studies (with different routes of administration and/or doses), if tested in more than one bioassay;
- the degrees of tumor response (incidence, site, type, multiplicity, etc.);
- evidence of structure-activity relationship;
- prevalence of dose-response relationship;
- the results of short-term tests for genetic toxicity;
- the presence of preneoplastic lesions; and

* Peto *et al.* (1980);⁷ Bickis and Krewski (1985);⁸ McKnight (1988)⁹

II C Evaluating Toxicology Information Continued

II C 5. Carcinogenic Risk Assessment Continued

- a reduced latency for tumor development or increase in the severity (malignancy of the neoplasia).

Other information, such as whether there was a shortened survival due to the toxicity of the test substance or whether the chemical is tested at or near the MTD, can also add weight to or confound the evidence of carcinogenicity. Information on dose-dependent or nonlinear kinetics from metabolic and pharmacokinetic studies in experimental animals and humans can supplement the assessment of the potential carcinogenic hazard of the additive to humans.

It should be noted that, although general approaches to animal carcinogenesis bioassays are well accepted by the scientific community, opinions about the design, conduct, and interpretation of such test results are not always in agreement and are often the source of scientific debate. This may be due, in large degree, to our lack of knowledge about the mechanisms of cancer induction and progression. Because the Act prohibits the use of carcinogenic food and color additives, the interpretation of carcinogenicity test results has enormous potential societal and economic impact. Consequently, proper assessment of carcinogenicity data has become an extremely critical function of CFSAN.

b. CFSAN's Quantitative Risk Assessment Committee (QRAC)

The QRAC was formed in 1983. Although quantitative risk assessments were performed under the auspices of the CAC prior to this, the QRAC was formed because of the need for an increasing number of quantitative risk assessments related to food and color additive petitions. Based on its evaluation of all relevant data on a substance, the CAC recommends to the QRAC the bioassays and epidemiological studies most appropriate for low-dose extrapolation. The CAC also recommends to the QRAC the tissue site(s), species, and sex most suitable for quantitative evaluation.

The QRAC then performs a quantitative risk assessment. This portion of the risk assessment process is often controversial, even among experts. Currently, the QRAC uses a linear-at-low-dose approach, similar to that described by Gaylor and Kodell.* The QRAC cannot determine the most probable expected human risk for almost any case because of the uncertainties and sources of error inherent in quantitative risk assessment using high-dose animal data. However, the QRAC believes that, in cases where dose-response data are suitable, it can predict a lifetime upper limit of risk with some degree of confidence.

* Gaylor and Kodell (1980)⁵

Chapter III

Concern Levels and Recommended Toxicity Tests

A. Introduction

This chapter describes how FDA determines which toxicity tests are recommended to assess the safety of food additives (direct food additives and color additives used in food) that are proposed for new or expanded use. Chapter III B explains how these additives are assigned to levels of concern (see Chapter III B 1) that indicate their potential for posing significant health risks to humans, if approved. A substance is assigned to a Concern Level based on available toxicology information or on the substance's structural similarity to known toxicants (see Chapter III B 2) and on the estimated human exposure to the substance from its proposed use (see Chapter III B 3). As in the previous edition of these guidelines (1982), exposure is weighted more heavily than structure in assigning substances to Concern Levels (see Figure 3).

Chapter III C describes the toxicity tests recommended for assessing the safety of additives (direct food additives and color additives used in food) assigned to each Concern Level. Different minimum testing levels are recommended for compounds assigned to Concern Levels I, II and III (see Figure 4 in Chapter III C 1). Because Concern Level III substances may present more significant health risks than substances assigned to Concern Levels I and II, more rigorous and longer-term toxicity testing is recommended to assess the safety of Concern Level III substances. (Note that some tests in the minimum set of toxicity tests recommended for compounds assigned to Concern Levels I, II and III have been changed from those recommended in the 1982 publication; these changes are summarized in Chapter I A 2.) Chapter III C 2 explains how the Agency develops additional testing recommendations for assessing the safety of direct food additives and color additives used in food proposed for new or expanded use. These tests augment the minimum set of toxicity tests, as appropriate; examples are provided.

Detailed guidelines for specific toxicity tests are not included in this chapter. However, guidelines for the conduct of short-term tests for genetic toxicity, acute toxicity tests, short-term toxicity tests with rodents and non-rodents, subchronic toxicity tests with rodents and non-rodents, one-year toxicity tests with non-rodents, carcinogenicity studies with rodents, combined chronic toxicity/carcinogenicity studies with rodents, and reproduction and developmental toxicity studies, can be found in Chapter IV C. Guidelines to assist the petitioner in developing strategies for assessing the metabolism and pharmacokinetics, immunotoxicity and neurotoxicity of food additives and color additives used in food can be found in Chapter V and recommended strategies for conducting human clinical trials with direct food additives and color additives used in food can be found in Chapter VI B.

III B. Concern Levels

1. Determining Concern Levels

In 1982, FDA introduced the concept of tiered testing requirements for obtaining information about the safety of direct food additives and color additives used in food. This concept is based on the assumption that the degree of effort expended to reduce uncertainty about the safety of a direct food additive or color additive used in food should relate in some logical way to the likelihood that the additive poses a health risk to the public.

In evaluating the toxicological safety of direct food additives and color additives used in food, two factors are of primary importance: the extent of human exposure (dose) and the toxicological effects on various biological systems (nature of effect, target, magnitude of response per unit dose, etc.). These factors determine the extent of the Agency's initial concern about the safe use of an additive. The greater the Concern Level, the greater the potential for toxicity.

In the absence of toxicological information about a compound, potential toxicity can be evaluated on the basis of structural similarity to known toxicants (see Chapter III B 2). Information about a compound's potential toxicity and estimated human exposure from a designated use (see Chapter III B 3) are sufficiently useful to permit semi-quantitative categorization of direct food additives and color additives used in food into three broad initial Concern Levels. For example, high toxic potential and high exposure would result in a compound being assigned a high initial Concern Level (*i.e.* Concern Level III), and low toxic potential and low exposure would result in a compound being assigned a low initial Concern Level (*i.e.* Concern Level I). Thus, Concern Levels are relative measures of the degree to which the use of an additive may present a hazard to human health.

Available toxicology information can, of course, change the Concern Level to which an additive has been assigned and alter the recommended set of toxicity tests for the additive. Subsequent and final Concern Levels, therefore, may be different from the initial Concern Level, and will be based on estimated human exposure and actual information about the metabolism and toxicology of the compound. For example, an additive may be transformed by metabolic activity into a substance of greater potential toxicity, or a potentially toxic additive may be distributed or metabolized in a manner that protects the target tissue or organ from the toxic effects of the chemical (blood-brain barrier; placental barrier; metabolic deactivation).

The minimum set of recommended toxicity tests for each additive (*i.e.* direct food additives and color additives used in food) is determined by the initial Concern Level to which it is assigned (see Chapter III C 1). Recommended toxicity tests are designed to reduce uncertainty about the safety of direct food additives and color additives used in food that have been proposed for new or expanded use. In addition, these testing recommendations allow more resources to be concentrated on additives that present the highest probable risk to human health (*i.e.* Concern Level III substances); fewer resources per additive can be expended on additives where use levels and potential toxicity are minimal (*i.e.* Concern Level I substances). Such a system for development of toxicology information is expected to be more cost-effective than one in which all additives are made to undergo the same regimen of testing irrespective of any other considerations.

III B 1 Determining Concern Levels Continued

In general, the procedure for determining the initial Concern Level for a direct food additive or color additive used in food is as follows:

- On the basis of information about its molecular structure, an additive will be placed in one of three broad categories: Category C is for additives whose toxicological potential is considered to be high; Category A is for additives whose toxicological potential is considered to be low; and Category B is for additives whose toxicological potential is likely to be intermediate between Categories A and C (see Chapter III B 2).
- Human exposure to each additive will be estimated (see Chapter III B 3).
- Within each structure category (A, B, and C), estimated human exposure will determine the initial Concern Level to which each additive is assigned (see Figure 3 below).

The choice of three broad Concern Levels reflects the traditional division of toxicity studies into three broad classes based on duration of exposure to the test compound: Short-term, subchronic, and chronic. As the duration of the exposure increases, the lowest-effect dose and the types of effects observed usually are determined with increasing sensitivity. Similarly, as the Concern Level to which an additive has been assigned increases from I to III, the recommended duration of toxicity studies and exposure to the test compound also increase (see Chapter III C 1 and Figure 4). As data from the minimum set of toxicity tests are obtained, the results can be used to refine or adjust the type, sensitivity, and rigor of subsequent tests, and therefore the precision of the estimate of an additive's toxicity.

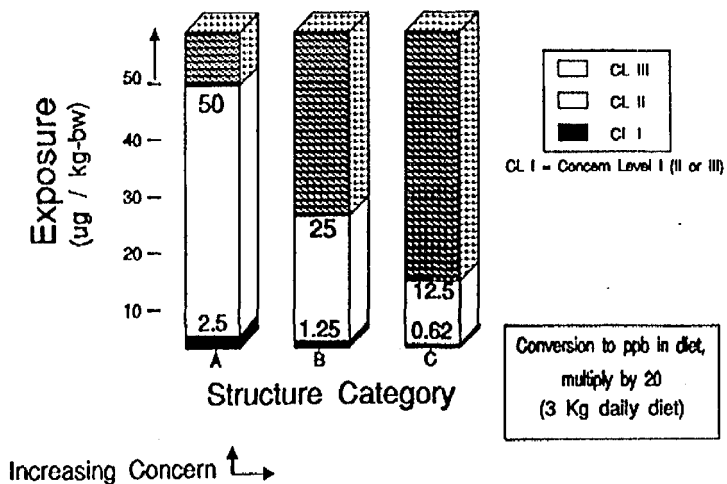
Levels of exposure that define which substances in each Structure Category are assigned to Concern Levels I, II, and III (see Figure 3) were selected in 1982 on the basis of recommendations by experienced toxicology experts within CFSAN. While exposures may range over approximately 6 orders of magnitude, the structure category of the substance has the effect of only having the breakpoints for determining Concern Level assignments between structure categories A or B or between structure categories B or C. Structure category is allowed only this limited influence in determining minimum testing levels partly because of the considerable uncertainty still surrounding the use of chemical structure to estimate potential toxicity.*

As noted previously, a food or color additive is considered safe if there is a reasonable certainty that no harm will result from its use (see Chapter II C). The level of exposure for which there is a reasonable certainty of no harm usually can be extrapolated from data obtained from toxicity studies. Thus, for each toxic effect associated with a food or color additive, the degree of concern can be defined as the extent to which actual exposure is expected to exceed the acceptable daily intake (acceptable exposure) determined from toxicological information. Because the degree of concern is also a function of the nature of the toxic effect, information that indicates a more severe toxic effect (for example, irreversible and life-threatening effects) may increase the Concern Level of a substance, regardless of exposure.

* Rulis *et al.* (1984)¹

Figure 3

Concern Levels as Related to Chemical Structure and Exposure



52454 1831

III B 2. Structure Category Assignment

a. Introduction

The toxic action of a compound is a consequence of the physical and chemical interaction of the compound or its active form with a critical molecular target--receptor, enzyme, DNA or other cellular constituent--within the living organism. Thus, it is reasonable to expect that the structure and associated physicochemical properties of a compound play an important role in its toxicity. This relationship between toxicity and chemical structure forms the basis for various systematic schemes and approaches developed over the years in attempts to estimate the toxic potential of untested chemicals or to prioritize chemicals for toxicity testing*.

In recent years, a number of computer-assisted, structure-activity relationship (SAR) models have been developed for predicting or estimating the toxicity of untested compounds. A general approach to developing such a model is to derive a correlation equation that relates structural features and physicochemical parameters of compounds to the toxicological endpoint of interest. The correlation equation is based upon a database assembled from a series of structurally related compounds or a set of heterogeneous structures.* The parameters (or variables) commonly used in SAR modelling fall into four major categories: topological, geometric, electronic and physicochemical,* as illustrated below:

- Topological parameters: counts of atoms and bonds, molecular weight, counts of rings and ring atoms, presence or absence of selected functional groups and substructural fragments;
- Geometric parameters: molecular size and shape;
- Electronic parameters: partial charge, dipole moments and bond strength; and
- Physicochemical parameters: partition coefficient.

Using this general approach, Enslein and co-workers,* Jurs and co-workers,* and Klopman and Rosenkranz* have constructed SAR models for a number of toxicological endpoints, such as acute toxicity in rodents (i.e. LD₅₀), carcinogenicity, mutagenicity in *Salmonella typhimurium*, and teratogenicity.

* Dehn and Helmes (1974);¹ Gori (1977);² Cramer *et al.* (1978);³ Lutz (1984);⁴ Woo *et al.* (1985);⁵

* Enslein (1984)⁶

* Ramiller (1984)⁷

⁴ Enslein and Craig (1978);⁸ Enslein and Craig (1982);⁹ Enslein *et al.* (1983);¹⁰ Enslein *et al.* (1983);¹¹ Enslein *et al.* (1987);¹² Enslein *et al.* (1987)¹³

* Chou and Jurs (1979);¹⁴ Jurs *et al.* (1979);¹⁵ Yuan and Jurs (1980);¹⁶ Tinker (1981);¹⁷ Yuta and Jurs (1981);¹⁸ Stouch and Jurs (1985);¹⁹ Rohrbaugh *et al.* (1988);²⁰ Randic and Jurs (1989)²¹

* Klopman (1984);²² Klopman and Rosenkranz (1984);²³ Rosenkranz *et al.* (1984);²⁴ Klopman *et al.* (1985a);²⁵ Klopman *et al.* (1985b);²⁶ Rosenkranz *et al.* (1985);²⁷ Rosenkranz and Klopman (1990)²⁸

III B 2 Structure Category Assignment Continued

The Agency recognizes that certain chemical structures bear some relationship to biological activity. While these SAR models hold great promise for specific applications in the future, they are subject to several major limitations at this time. Because of these limitations, the Agency believes that information about such relationships should be used only to guide recommendations about the acquisition of toxicological data, and not as a substitute for such data. Acting on this premise, the Agency continues to incorporate information on chemical structures into its recommendations about the initial level of testing recommended to demonstrate the toxicological safety of a direct food additive or a color additive used in food.

The purpose of this chapter is to provide a general guideline whereby a chemical that has been proposed as a direct additive or a color additive for use in food can be assigned to a Chemical Structure Category based on substructural features—specifically, the presence or absence of chemical groups that have been associated with certain types of toxicity. This information will be integrated with data on predicted human exposure to determine the potential for toxicity, and thus the recommended initial level of toxicity testing for the proposed additive.

The guidelines provided in this Chapter are not intended to be comprehensive or as a rigid set of rules. Substructural fragments and functional groups are illustrative of those groups identified in the Structure Categories Section below. The initial, usually temporary, Concern Level to which a substance has been assigned is based on its structure category assignment (see Chapter III B 2 b below) and the estimated human exposure to the substance from its petitioned use (see Chapter III B 3). This initial Concern Level will be modified during the review process based on chemical or biological information, such as: 1) the functional groups of known or predicted metabolites of the additive are judged to be of more or less concern than the structure of the additive; 2) there is evidence of potential bioaccumulation of the additive or its metabolites; 3) there is unequivocal evidence that the additive is poorly or not absorbed; or 4) qualitative or quantitative information is available on secondary component(s) or contaminants.

b. Structure Category Assignment of Additives

The initial step in assigning a proposed direct food or color additive to its correct Structure Category is to identify its complete chemical structure(s) and functional group(s). A direct food additive or color additive used in food may be a single chemical (arbitrarily defined as a chemical that is $\geq 90\%$ pure), or a compound that is a mixture of two or more chemicals. Each chemical component in an additive is evaluated for the presence of one or more functional groups. Based on this information, the additive under consideration can be placed in the appropriate Chemical Structure Category. Structure Categories are divided into three classes of potential for toxicological significance (e.g. Categories A, B, and C), with Category A having the least potential for toxicity and Category C having the highest potential for toxicity.

This Chapter is an updated version of the "Chemical Structure Category Section" in the 1982 Agency guidelines. While the major groups of chemical structure categories presented in 1982 Agency guidelines have remained unchanged, the majority of these categories have been subdivided into smaller groups of chemicals that share common functional groups. For those petitioners who would like additional information on the assignment of chemicals to different structural categories, the Agency has a supplemental document entitled "Structure Category Assignments of Chemicals in the Priority-Based Assessment of Food Additive Database" available upon request.

III B 2 Structure Category Assignment Continued

Structure Category A

i. Structure Category A Chemicals

In general, Structure Category A includes compounds with chemical structures (substructural fragments and functional groups) believed to be of low toxic potential. It includes substances that are normal cellular constituents (e.g. certain fats and carbohydrates), but it excludes amino acids, proteins and certain intermediates of lipid and carbohydrate metabolism. The aliphatic organic chemicals in this category have relatively simple structures that are saturated.* Inorganic chemicals in this category are certain endogenous salts of alkali metals (e.g. sodium and potassium) and alkaline-earth metals (e.g. calcium and magnesium).

Chemicals in Structure Category A can be divided into three general groups, including: 1) aliphatic hydrocarbons^b (saturated + un-functionalized^c or mono-functional^d), 2) fats and carbohydrates, and 3) inorganic chemicals.

Aliphatic Hydrocarbons

- **Aliphatic hydrocarbons: un-functionalized and non-cyclic^e (C=2 to 30):** Group includes saturated straight- and branched-chain alkanes.



* Hydrocarbons with only single bonded carbon atoms are referred to as saturated; conversely, hydrocarbons with one or more double bonded carbon atoms are referred to as unsaturated.

^b Chemicals that contain only carbon (C) and hydrogen (H) are known as hydrocarbons. Hydrocarbons are further divided into two main classes, aliphatic and aromatic. Aliphatic chemicals include both open chain and closed ring structures with single, double and triple bonded carbon atoms. In contrast, aromatic chemicals include benzene and chemicals with substituted benzene ring structures.

^c Hydrocarbons with exclusively single bonded carbon atoms are referred to as aliphatic "un-functionalized" hydrocarbons (i.e. hydrocarbons with no functional groups).

^d Mono-functional aliphatic hydrocarbons refers to hydrocarbons containing one type of functional group in addition to the single bonded carbon atoms (e.g. a carbon/carbon double bond or a carbon oxygen double bond). In addition, chemicals containing two or more of the same functional group are arbitrarily referred to as mono-functional. In contrast, chemicals with two or more different functional groups are referred to as "multi-functional".

^e Non-cyclic chemicals have open carbon chains. In contrast, cyclic chemicals have closed carbon ring structures, and these rings may be saturated or unsaturated.

III B 2 Structure Category Assignment Continued

Structure Category A

- Aliphatic hydrocarbons: un-functionalized, saturated and mono-cyclic (C=6 to 20):

e.g. cyclohexane



- Aliphatic hydrocarbons: mono-functional, saturated and non-cyclic (C=2 to 30):
Group includes mono-functional:

aliphatic acids ($R^1\text{-COOH}^a$) and alcohols ($R^1\text{-OH}$),
aliphatic aldehydes ($R^1\text{-CH=O}$) and esters ($R^1\text{-COO-R}^2$),
aliphatic ethers ($R^1\text{-CH}_2\text{-O-CH}_2\text{-R}^2$) and ketones ($R^1\text{-CO-R}^2$), and
aliphatic mercaptans ($R^1\text{-SH}$).

- Aliphatic hydrocarbons: mono-functional, saturated and mono-cyclic (C=6 to 20):
Group includes mono-functional, mono-cyclic acids; mono-functional, mono-cyclic alcohols; mono-functional, mono-cyclic aldehydes; mono-functional, mono-cyclic esters; mono-functional, mono-cyclic ethers; mono-functional, mono-cyclic ketones; and mono-functional, mono-cyclic mercaptans.

Fats and Carbohydrates

- Fats, fatty acids, fatty acyl esters and their salts: Group includes: fats, unsaturated and saturated fatty acids and fatty acyl esters.

fats (e.g. butter esters, coconut and peanut oil),
unsaturated fatty acid (e.g. oleic acid: $\text{CH}_3\text{-(CH}_2)_7\text{-CH=CH-(CH}_2)_7\text{-COOH}$),
saturated fatty acid (e.g. caprylic acid: $\text{CH}_3\text{-(CH}_2)_6\text{-COOH}$), and
fatty acyl esters ($R^1\text{-COO-R}^2$).

- Intermediates and products of carbohydrate and lipid metabolism in humans:

intermediates of carbohydrate metabolism (e.g. citric acid) and
intermediates of lipid metabolism (e.g. lecithin)

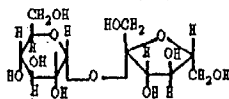
* Certain letters in the chemical structures presented in this chapter represent codes for functional groups that could be placed within the chemical structure, including:
 R^1 includes aliphatic (open chain hydrocarbons) and alkyl functional groups used in Categories A and B, as well as the aryl (e.g. benzyl) functional groups used in Category C; and
 R^2 includes only the aliphatic alkyl [e.g. $\text{-CH}_2\text{-}$].

III B 2 Structure Category Assignment Continued

Structure Category A

- Simple and complex carbohydrates: Group includes carbohydrates which are components in the human diet, including: saccharides, oligosaccharides, and polysaccharides.

simple carbohydrates (e.g. gluconic acid: $\text{HOOC}-(\text{CHOH})_4-\text{CH}_2\text{OH}$ and sucrose)



complex carbohydrate (e.g. starch)

Inorganic Chemicals

- Endogenous (normal cellular constituents) inorganic salts: Group includes alkali metals (Na^+ , K^+), alkaline-earth metals (Mg^{2+} , Ca^{2+}), simple ammonium salts (NH_4^+), hydrochloric acid, sodium hydroxide and anions (Cl^- , CO_3^{2-} , NO_3^- , PO_4^{3-} , and SO_4^{2-}).

e.g. sodium chloride NaCl

- Inert gases: Group includes certain inert gases (e.g. argon (Ar), helium (He) and nitrogen (N_2)). It also includes carbon dioxide (CO_2) and elemental carbon.

ii. Structure Category B Chemicals

Structure Category B includes compounds with chemical structures that have been associated with adverse effects other than mutagenicity and carcinogenicity in animals or humans. Structure Category B also includes indeterminate structures and structures believed to have a potential for toxicity that is intermediate between structures in Structure Categories A and C. Chemicals in Structure Category B can be divided into four general groups, including: 1) aliphatic hydrocarbons (certain *mono-functional* and *saturated*, as well as *mono-functional* or *multi-functional*,^a *unsaturated* and *non-conjugated* chemicals^b); 2) amino acids, proteins and certain nitrogenous chemicals; 3) inorganic chemicals; and 4) mixtures of defined chemicals (*with only Category A or B chemicals*).

^a Multi-functional hydrocarbons are classified as chemicals that have two different functional groups in addition to single bonded carbon atoms (e.g. a carbon/carbon double bond and a carbon/oxygen double bond). In contrast, a chemical with two of the same functional groups is referred to as mono-functional.

^b Conjugated double bonds are hydrocarbons with alternating double and single carbon bonds (e.g. $\text{R}^1-\text{CH}=\text{CH}-\text{CH}=\text{CH}-\text{R}^2$); conversely, non-conjugated hydrocarbons have more than one single carbon bond between the double bonded carbon atoms (e.g. $\text{R}^1-\text{CH}=\text{CH}-\text{CH}_2-\text{CH}_2-\text{CH}=\text{CH}-\text{R}^2$).

III B 2 Structure Category Assignment Continued

Structure Category B

Aliphatic Hydrocarbons

- Aliphatic hydrocarbons: mono-functional hydrocarbons not listed in Categories A & C:
Group includes:

mono-functional aliphatic acetals $[(R'O)_2CH-R']$;

mono-functional glycol ethers

(e.g. ethylene glycol monomethyl ether $HO-CH_2-CH_2-O-CH_3$); and

methyl alcohol (CH_3-OH) and methyl esters ($R^A-COO-CH_3$).

- Aliphatic hydrocarbons: mono-functional and mono-unsaturated: Group includes both cyclic and non-cyclic mono-functional and mono-unsaturated hydrocarbons.

mono-functional and mono-unsaturated, non-cyclic hydrocarbons

(e.g. 2-hexene: $CH_3-CH=CH-(CH_2)_3-CH_3$)

mono-functional and mono-unsaturated, cyclic hydrocarbons

(e.g. cyclohexene)



- Aliphatic hydrocarbons: mono-functional and poly-unsaturated (& non-conjugated) (C=6 to 30): Group includes both cyclic and non-cyclic mono-functional and polyunsaturated (non-conjugated) hydrocarbons.

mono-functional and poly-unsaturated (non-conjugated), non-cyclic hydrocarbons

(e.g. 1,4-pentadiene: $HC_2=CH-CH_2-CH=CH_2$)

mono-functional and polyunsaturated (non-conjugated), cyclic hydrocarbons

(e.g. 1,5-cyclononene)



- Aliphatic hydrocarbons: multi-functional and saturated, mono-unsaturated or polyunsaturated (non-conjugated): Group includes: multi-functional, saturated, non-cyclic hydrocarbons; multi-functional, saturated, cyclic hydrocarbons; multi-functional, mono-unsaturated (or polyunsaturated & non-conjugated), non-cyclic hydrocarbons; multi-functional, mono-unsaturated (or polyunsaturated & non-conjugated) cyclic hydrocarbons. Examples of multi-functional chemicals included in this group are unsaturated carboxylic ethers and anhydrides, polyaldehydes and polyols.

III B 2 Structure Category Assignment Continued

Structure Category B

multi-functional, saturated, non-cyclic hydrocarbons

e.g. 2-hydroxypropionaldehyde: $\text{CH}_3\text{-CH(OH)-CHO}$

multi-functional, saturated, cyclic hydrocarbons

e.g. 4-hydroxycyclohexanoic acid

multi-functional, mono-unsaturated (or polyunsaturated and non-conjugated)
non-cyclic hydrocarbons (e.g. 3-hexenol: $\text{CH}_3\text{-CH}_2\text{-CH=CH-CH}_2\text{-CH}_2\text{-OH}$)multi-functional, mono-unsaturated (e.g. 3-cyclohexen-1-ol, *center*) or
poly-unsaturated & non-conjugated cyclic hydrocarbons
(e.g. cyclonona-3,7-dien-1-ol, *right*)

Amino Acids, Proteins and Certain Nitrogenous Chemicals

- **Amino acids:** Group includes amino acids, unless they contain functional groups listed in Category C.

e.g. alanine: $\text{CH}_3\text{-CH(NH}_2\text{)-COOH}$

- **Proteins and polypeptides**
proteins (e.g. yeast protein extract) and
polypeptides (e.g. protein hydrolysate)
- **Certain nitrogenous chemicals:** Group includes quaternary ammonium salts, alkylated ammonium compounds, and urea.

e.g. quaternary ammonium salts $[(\text{R})_4\text{N}^+ \text{X}^-]$ and urea $[\text{NH}_2\text{-CO-NH}_2]$

III B 2 Structure Category Assignment Continued

Structure Category B

Inorganic Chemicals

- **Inorganic salts of Fe, Cu, Mn, Zn, and Sn:** Category also includes simple iodide salts (e.g. sodium iodide), sulfur dioxide (SO_2) and silicates (e.g. NaSiO_3). Furthermore, this category includes organic salts of the same inorganic chemicals, so long as the metal is not covalently bonded to the organic substance (and the organic substance is not included in Category C).

e.g. ferric sulfate: $\text{Fe}_2(\text{SO}_4)_3$

Mixtures

- **Mixtures of chemicals:** Group includes only mixtures of chemicals of defined composition, and all of the chemicals in the mixture must be assigned to Category A, Category B, or Categories A and B.

iii. Structure Category C Chemicals

In contrast to the two previous categories, Structure Category C contains compounds or metabolites that are structurally related to a reported mutagen or carcinogen, or chemicals that are structurally related to compounds demonstrated to produce carcinogenicity in humans or laboratory animals. A total of 55 individual subgroups of chemicals have been pooled into six major groups based upon the presence or absence of specific types of chemical functional groups, including:

- aliphatic (*multi-functional & conjugated*) alkene and alkyne hydrocarbons (*with and without C and O functional groups*);
- aromatic (*mono- and polycyclic*) hydrocarbons (*mono- and multi-functional*);
- aliphatic and aromatic (*mono- and multi-functional*) hydrocarbons with functional groups containing N, P and S atoms;
- heterocyclic chemicals^{*};
- inorganic and organometallic chemicals; and
- mixtures of chemicals (*with Category C and unknown chemicals*).

Structure category assignment of chemicals in Category C is relatively straightforward when the additive has only one functional group. When the additive has more than one specified functional group, a conservative approach is used by the FDA. Chemicals with more than one Category C functional group are assigned to all of the appropriate Category C functional groups.

^{*} Heterocyclic chemicals have a closed ring structure that contains one or more atoms within the ring that differ from carbon (e.g. nitrogen, oxygen and sulfur).

III B 2 Structure Category Assignment Continued

Structure Category C

Aliphatic Alkenes & Alkynes

This group of chemical structure categories includes chemicals with relatively simple aliphatic and aromatic structures that are devoid of nitrogen, sulfur and phosphorus functional groups

- **Aliphatic hydrocarbons: unsaturated (& conjugated) and non-aromatic:** Group includes conjugated (non-cyclic and cyclic, but not aromatic) alkenes, aldehydes, and ketones. It also contains α,β -unsaturated (non-cyclic or cyclic, but not aromatic) carbonyl^a acids and esters. In addition, all conjugated non-cyclic and cyclic chemicals with an allyl fragment (e.g. $\text{CH}_2=\text{CH}-\text{CH}_2-$) are included in this group.

conjugated alkenes [$\text{R}'-\text{CH}=\text{CH}-\text{CH}=\text{CH}-\text{R}''$];
 conjugated aldehydes [$\text{R}'-\text{CH}=\text{CH}-\text{CH}=\text{O}$] and ketones [$\text{R}'-\text{CH}=\text{CH}-\text{CO}-\text{R}''$];
 α,β -unsaturated carbonyl acids [$\text{R}'-\text{CH}=\text{CH}-\text{COOH}$] and esters [$\text{R}'-\text{CH}=\text{CH}-\text{CO}-\text{O}-\text{R}''$]; and
 chemicals with an allyl fragment ($\text{CH}_2=\text{CH}-\text{CH}_2-$).

- **Alkynes:**

alkynes: $\text{R}'-\text{C}\equiv\text{C}-\text{R}''$

Aromatic Hydrocarbons

- **Aromatic hydrocarbons:** Group includes mono-aromatic hydrocarbons, including: mono-aromatic chemicals with or without alkyl functional groups; mono-aromatic chemicals with conjugated alkenes (including the allyl functional group); mono-aromatic α,β -unsaturated carbonyl acids and esters; mono-aromatic, conjugated aldehydes and ketones; mono-aromatic chemicals with the oxy functional group (e.g. methoxy, ethoxy, etc.); and mono-aromatic chemicals with one or more hydroxy ($-\text{OH}$) functional groups.

mono-aromatic benzene + alkyl functional groups (e.g. benzene)



mono-aromatic conjugated alkene (e.g. 2-phenyl-2-butene)



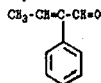
^a The carbonyl group has an *alpha* (α), *beta* (β) unsaturated double bonded carbon and oxygen (e.g. $\text{R}''-\text{CH}=\text{CH}-\text{C}(\text{O})-\text{R}'$).

III B 2 Structure Category Assignment Continued

Structure Category C

mono-aromatic α,β -unsaturated carbonyl (e.g. benzoic acid)

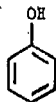
mono-aromatic, conjugated aldehydes and ketones (e.g. 2-phenyl-2-butenal)



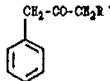
mono-aromatics with the oxy functional group (e.g. anisole)



mono-aromatics with hydroxyl functional group (e.g. phenol)



- **Benzylic hydrocarbons:** Group includes aromatic hydrocarbons with the benzylic functional group, including: benzylic acids, alcohols, aldehydes, and esters.

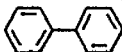
benzylic acid (*left*), benzyl alcohol (*center*), and benzyl ethers (*right*)phenylacetaldehyde (*left*), benzyl ketones (*center*) and benzylic esters (*right*)

III B 2 Structure Category Assignment Continued

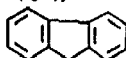
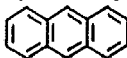
Structure Category C

- Polycyclic aromatic hydrocarbons: Group includes:

biphenyl, aromatic hydrocarbons (e.g. biphenyl) and



polycyclicaromatic hydrocarbons [e.g. anthracene (*center*) and fluorene (*right*)]



Aliphatic and Aromatic Chemicals Containing Halogen, Nitrogen, Phosphorus, and Sulfur Functional Groups

This group of chemical structure assignment categories includes chemicals that contain halogen, nitrogen, phosphorus, and sulfur chemical functional groups.

Halogenated Chemicals

- Halogenated chemicals: Group includes:

aliphatic [$R^1-CH(X)-R^2$] and aromatic halides (e.g. 1,2-dichlorobenzene);



halocarbonyl acids ($R^1-CH(X)-COOH$) and aldehydes ($R^1-CH(X)-CHO$);
halocarbonyl amides ($R^1-CH(X)-CO-NH_2$) and esters ($R^1-CH(X)-COO-R^2$);
haloethers (e.g. α -alkyl haloether: $R^1-CH(X)-O-CH_2-R^2$); and
halohydrins ($R^1-CH(X)-CH_2-OH$).

* Certain letters in the chemical structures presented in this chapter represent codes for functional groups that could be placed within the chemical structure. The halogen functional groups are coded "X," and they include bromine (Br), chlorine (Cl), fluorine (F) and iodine (I).

III B 2 Structure Category Assignment Continued

Structure Category C

Nitrogen Functional Groups

Chemicals containing nitrogen functional groups include hydrazides, hydroxylamides, imines, un-substituted amides, and lactams; aliphatic and aromatic amines; nitro and nitroso groups; N-nitroso group; nitriles; azo and di-azo chemicals; azoxy chemicals; azide and triazene chemicals; hydrazines; carbamic acid esters; urea derivatives; guanidines; isocyanates; isothiocyanates; carbodiimides; and organic nitrates and nitrites. In contrast, heterocyclic chemicals that contain nitrogen within the ring structure are presented later in the section entitled "Nitrogen Heterocyclic Chemicals" along with other heterocyclic chemicals.

- Hydrazides; hydroxylamides and hydroxylamines; imines and hydroxylimines; and un-substituted amides:

hydrazides $[R^1-CO-NH-NH_2]$;
hydroxylamides $(R^1-CO-NH-OH)$ and hydroxylamines $(R^1-NH-OH)$;
imines $(R^1-CH=NR^2)$ and hydroxylimines $(R^1-CH=N-OH)$; and
unsubstituted amides (e.g. primary-amide $R^1-CO-NH_2$).

- Aliphatic and aromatic amines: Group includes:

1°-amines (R^1-NH_2) , 2°-amines (R^1_2-NH) and 3°-amines (R^1_3-N)

- Nitro and nitroso groups:

nitro (R^1-NO_2) and nitroso (R^1-NO)

- N-nitroso group:

N-nitroso: e.g. $R^1-NH-NO$

- Nitriles:

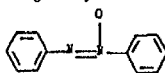
nitriles: $R^1-C\equiv N$

- Azo and poly-azo chemicals:

mono-azo $(R^1-N=N-R^2)$ and di-azo $(R^1-CH=N^+=N^-)$

- Azoxy group:

e.g. azoxybenzene



- Azides and triazenes:

azides $(R^1-N=N^+=N^-)$ or R^1-N_3 and triazenes $(R^1-N=N-NH-R^2)$

- Hydrazines:

hydrazines: $R^1-NH-NH-R^2$

III B 2 Structure Category Assignment Continued

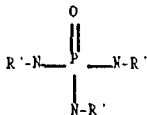
Structure Category C

- **Carbamic Acid Esters:** Group includes:
carbamic acid ester ($R^1-NH-C(O)-OR^2$),
halogenated carbamic acid esters ($R^1-CH(X)-NH-C(O)-OR^2$), and
thiocarbamic acid esters ($R^1-NH-C(S)-OR^2$).
- **Substituted ureas:**
substituted ureas: $R^1-NH-CO-NH-R^2$
- **Guanidines:**
guanidines: $NH_2-C(=NH)-NH-R^1$
- **Isocyanates and cyanates:**
isocyanates ($R^1-N=C=O$) and cyanates ($R^1-O-C\equiv N$)
- **Isothiocyanates:**
isothiocyanates: $R^1-N=C=S$
- **Carbodiimides:**
carbodiimides: $R^1-N=C=N-R^2$
- **Organic nitrates and nitrites:**
organic nitrates (R^1-O-NO_2) and organic nitrites R^1-O-NO

Phosphorus Functional Groups

Chemicals containing phosphorus functional groups include phosphoramides; phosphates ($-PO_4$) and phosphites ($-PO_3$); and phosphonate esters and phosphonium functional groups. Chemicals containing both sulfur and phosphate functional groups include the mono- and dithio-phosphate esters.

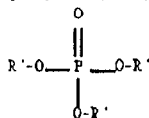
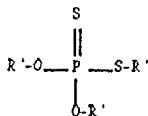
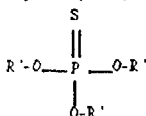
- **Phosphoramides**



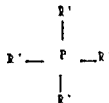
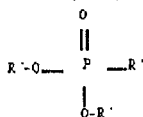
III B 2 Structure Category Assignment Continued

Structure Category C

■ Phosphates and thiophosphates:

phosphates (*center*)thiophosphates (*center*) and dithiophosphates (*right*)

■ Phosphonate esters and phosphonium salts:

phosphonate esters (*center*) and phosphonium ion (*right*)

Sulfur Functional Groups

Chemicals which contain sulfur functional groups include: thioamides; substituted thioureas; thioethers; sulfamates; sulfate ($-\text{SO}_4$) and sulfite ($-\text{SO}_3$) esters; sulfonate and sulfinyl esters; and dithiols and aromatic thiols. Chemicals containing both sulfur and nitrogen functional groups include thiocarbamates and isothiocyanates. In contrast, heterocyclic chemicals with sulfur atoms in the ring structure are included in the section entitled "Sulfur Heterocyclic Chemicals" along with other heterocyclic chemicals.

■ Thiocarbamic acids:

thiocarbamic acid esters: $\text{R}'\text{-NH-C(S)-OR}'$

■ Isothiocyanates:

isothiocyanates: $\text{R}'\text{-N=C=S}$

■ Thioamides:

thioamides: $\text{R}'\text{-CS-NH-R}'$

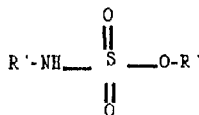
■ Substituted thioureas:

substituted thioureas: $\text{R}'\text{-NH-CS-NH-R}'$

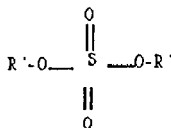
III B 2 Structure Category Assignment Continued

Structure Category C

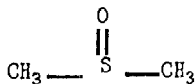
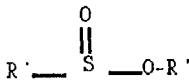
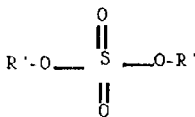
- **Thioethers:** Group includes thioethers, disulfides and trisulfides
e.g. thioethers: $R'-S-R'$
- **Sulfamates:**



- **Sulfate and sulfite esters:**
sulfate ester (*center*) and sulfite esters (*right*)



- **Sulfonate esters, sulfinyl esters and sulfoxides:**
sulfonate esters (*left*), sulfinyl esters (*center*)
sulfoxides (*right*, e.g. dimethyl sulfoxide) and sulfones.



- **Aromatic thiols and dithiols:** Group includes both aromatic (and other cyclic) thiols and dithiols (cyclic and non-cyclic).
aromatic thiols (e.g. benzenethiol)



dithiols (e.g. 1,2-propanedithiol: $CH_3-CH(SH)-CH_2-SH$)

III B 2 Structure Category Assignment Continued

Structure Category C

Heterocyclic Chemical Structure Categories

Heterocyclic chemicals include chemicals that contain within the ring structure a nitrogen, oxygen, or sulfur atom. In addition, some heterocyclic chemicals contain ring structures with both nitrogen and oxygen atoms; nitrogen and sulfur atoms; oxygen and sulfur atoms; and all three, nitrogen, oxygen and sulfur atoms.

Nitrogen Heterocyclic Chemicals

Heterocyclic chemicals containing a nitrogen atom within the aromatic ring include: acridines; aziridines; carbazoles; imidazoles, triazoles, and benzotriazoles; indoles; lactams; piperidines; pteridines; purines; pyrazoles and pyrazolones; pyridines and pyrazines; pyrimidines and pyrimidinetriones; pyrroles; pyrrolidines; quinolines, isoquinolines and benzoquinolines; and triazines and benzotriazines.

■ Aziridines:

■ Lactams: Category includes lactams (*center*) and lactims (*right*)

■ Purines:



■ Pyrimidines, pyrimidinetriones, triazines and benzotriazines:

pyrimidines (*center*) and pyrimidinetriones (*right*)



benzotriazines (*center*) and triazines (*right*)



■ Pyrroles:



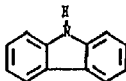
III B 2 Structure Category Assignment Continued

Structure Category C

- Pyrazoles and pyrazolones:
pyrazoles (*center*) and pyrazolones (*right*)



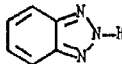
- Carbazoles:



- Indoles:



- Imidazoles, triazoles and benzotriazoles:
imidazoles (*left*), triazoles (*center*) and benzotriazoles (*right*)

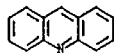


- Pyrrolidines

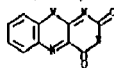


- Additional nitrogen heterocyclic chemicals [e.g. acridines, alloxazines, benzoquinolines, naphthyridines, phthalimides, piperazines, piperidines, pteridines, pyrazines, pyridines, and quinolines]:

acridines (e.g. *acridine*, *center*) and naphthyridines (e.g. *naphthyridine*, *right*)



alloxazines (*center*) and phthalimides (*right*)

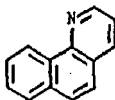


III B 2 Structure Category Assignment Continued

Structure Category C

piperidines (*center*) and piperazines (*right*)pyridines (*center*) and pyrazines (*right*)

pteridines

benzoquinolines (*center*) and quinolines (*right*)

Oxygen Heterocyclic Chemicals

The heterocyclic chemicals containing an oxygen atom within the ring structure include: alkene/phenoxy chemicals; dioxanes; epoxides; furans and benzofurans; oxetanes; pyrans and benzopyrans; saturated lactones, and α,β -unsaturated lactones. In addition, certain oxygen substituted heterocyclic chemicals have been included in this section, including: anthraquinones, benzoquinones, quinones, and thioxanthones.

- **Epoxides:** Group contains three membered mono- and poly-functional epoxides. This category also contains peroxides which are not heterocyclic.

mono-epoxides (*e.g.* ethylene oxide)peroxides (*e.g.* hydrogen peroxide: H_2O_2)

- **Saturated lactones:**



III B 2 Structure Category Assignment Continued

Structure Category C

- **α,β -Unsaturated lactones:**

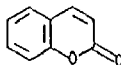


- **Dioxanes:**

e.g. 1,4-dioxane

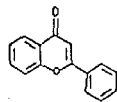
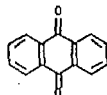


- **Furans, benzofurans and coumarins:**
furans (*left*), benzofurans (*center*), and coumarins (*right*)

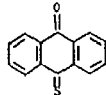


- **Antraquinones, benzoquinones, flavones, pyrones and thioxanthenes:**

e.g. anthraquinone (*left*), benzoquinone (*center*), and flavones (*right*)

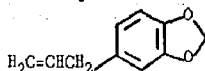


e.g. pyrone (*center*) and thioxanthone (*right*)



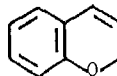
- **Aromatic ethers with alkene functional groups:** Group contains chemicals with saffrole-like structures (*i.e.* mono-aromatic ethers with a conjugated alkene functional group).

e.g. saffrole



- **Oxygen heterocyclic chemicals [e.g. oxetanes, pyrans, and benzopyrans]:**

oxetanes (*left*), pyrans (*center*) and benzopyrans (*right*)



III B 2 Structure Category Assignment Continued

Structure Category C

Sulfur Heterocyclic Chemicals

- Heterocyclic chemicals containing sulfur [e.g. sulfones, trithianes, thienes, thiones, and thiophenes]:

sulfones (center) and trithianes (right)



thiones (center) and thiophene (right)



Nitrogen and Oxygen Heterocyclic Chemicals

- Heterocyclic chemicals containing nitrogen and oxygen [e.g. morpholines and oxazoles]:

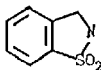
morpholines (center) oxazoles (right)



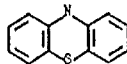
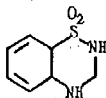
Nitrogen and Sulfur Heterocyclic Chemicals

- Heterocyclic chemicals containing nitrogen and sulfur [e.g. sulfinimides, thiazadiazoles, thiazides, thiazines and thiazoles]:

sulfinimides (left), thiazoles (center) and thiadiazoles (right)



thiazides (center) and thiazines (e.g. phenothiazine, right)



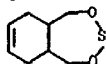
III B 2 Structure Category Assignment Continued

Structure Category C

Oxygen and Sulfur Heterocyclic Chemicals

- Heterocyclic chemicals containing oxygen and sulfur [e.g. oxythiepins]:

oxythiepins (e.g. 2,4,3-benzodioxathiepin)



Nitrogen, Oxygen and Sulfur Heterocyclic Chemicals

- Heterocyclic chemicals containing nitrogen, oxygen and sulfur [e.g. oxythiazins]:



Inorganic and Organometallic Chemicals

- Inorganic salts: Group contains inorganic salts that are not included in Categories A and B (e.g. aluminum). It also contains non-covalent complexes of these inorganic chemicals with organic chemicals.

e.g. aluminum ammonium sulfate: $\text{AlNH}_4(\text{SO}_4)_2$

- Organometallic chemicals:

e.g. vitamin B_{12}

Mixtures

- Mixtures: Group contains compounds that are mixtures of chemicals, and two types of mixtures are distinguished. The first type of mixture includes defined mixtures containing one or more substances which possess functional groups listed in Category C. The second type of mixture includes mixtures of undefined composition, including compounds in which all known components contain functional groups listed in Categories A and B. This conservative approach was taken, because it is possible that a minor, undefined constituent of a compound mixture could have a chemical with a functional group listed in Category C.

III B 3. Estimation of Human Exposure to Direct Food Additives and Food Ingredients

A key factor in the safety evaluation of a food additive or food ingredient is the relationship of its probable human exposure to the level at which adverse effects are observed in animal and/or clinical studies. Estimates of probable human exposures require knowledge of the specific uses and use levels of a substance under consideration and quantitative information on intakes of the foods in which the substance is used. Individuals' food intakes are distributed over a range determined by lifestyles and localized patterns of food availability and can be expected to change in response to changes in economic circumstances, education, health, the media, and the availability of products in the food supply.^a Because of the many factors affecting food intakes and the uncertainties in the eventual marketing of a petitioned food ingredient/additive, the estimation of probable exposures is a complex exercise. The Agency's assumptions concerning intake patterns, market penetration, and substance concentrations result in a conservative estimate of exposure^b. These assumptions are used because detailed information that can replace these assumptions is usually unavailable.

CFSAN's estimates of probable human exposure are based on food intake or food availability data obtained over relatively short time frames (one day to one year) and are used to represent chronic or "lifetime" exposure. We typically use the 90th percentile to represent probable exposure for a "heavy" consumer of a substance.

a. Parameters for the Exposure Estimate

In the broadest sense, two factors are required for making an estimate of exposure to a substance in the food supply. The first is the daily intake of the food in which the substance is used or can be found. The second is the concentration or use level of the substance in the food. Simple multiplication of these two factors gives an estimate of exposure to the substance from consumption of the food.^c

These two factors can be derived from a number of sources. For pre-market approval of new substances, information on the expected use level (or in the case of processing aids, expected residue concentration) in food is generally supplied by the petitioner. For substances already in the food supply, for which a cumulative exposure estimate incorporating proposed new uses is needed, use levels in food may be obtained from additional sources, such as Agency records, users of the substance, or by chemical analyses of the foods in which the substance is known to be used.

The daily intake of foods can be derived from a variety of data bases. The three most commonly used sources of food-intake data are: per-capita data derived from annual poundage surveys of producers or distributors; survey data on the frequency of consumption of foods ("food-frequency" surveys); and food-intake survey data. These three data base types will be described in more detail below. A number of data bases currently available for determining food intakes for estimating exposure to substances in the diet are shown in Table 1 below.^d

^a FASEB report (1988)¹

^b When comparing estimated daily intakes (EDI's), a more conservative EDI is higher.

^c It should be noted that this discussion does not pertain to the estimation of exposure to indirect food additives.

^d Abridged from footnote a, page 1.

Table 1

Summary of Characteristics of Major Databases Used for Estimating Intake of Substances

	NFCS ¹	NFCS Foods Commonly Consumed ²	NFCS CSFII ³	NHANES II ⁴	NRCA Menu Census ⁵	WPD National Eating Trends ⁶	Seafood Consumption ⁷	USDA ERS ⁸	FOA Total Diet Study ⁹
Date of data collection	1977-78 1987-88 ¹⁰	1977-78	annual since 1985	1976-80	every 5 years since 1957	annual	1973-74	annual since 1909	annual since 1961
Population surveyed	civilian; free- living; all ages	see NFCS	core & poverty groups of women 19-50 yr & their children 1-5 yr	civilian; free-living; 6 mo to 74 yr	participants in consumer survey; all ages	participants in consumer survey; all ages, 11 age-sex groups reported	participants in consumer survey; all ages, single person households excluded	net commodities in US commerce	since 1982 nationally representative diets for 8 age/sex groups
Type of instrument for intake	1-d recall & 2-d record	see NFCS	self; proxy for child	self (private interview); proxy for child	homemaker reports for all	one respondent for household	one respondent for household	NA	NA
Estimates possible	single day; usual	esters only intake/d (mean and percentiles)	single day; usual	single day	single day; usual	single day; usual	single day; usual	per capita availability	usual
Form of Estimate	mean frequency & distribution for raw data total population and esters only	mean and distribution of esters	mean frequency & distribution for raw data total population and esters only	mean frequency & distribution for raw data total population and esters only	mean frequency & distribution for raw data total population and esters only	mean & distribution for total population and esters only	mean number of servings; % of individuals using product	population mean	since 1982, mean + sd

See next page for footnotes.

Footnotes for Table 1

1. U.S. Department of Agriculture. Nutrient Intakes: individuals in 48 states, year 1977-78. Nationwide Food Consumption Survey (NFCS) Report No. 1-2. Available from: U.S. Government Printing Office, Washington DC.
2. Pao, E.M., Fleming, K.N., Guenther, P.M., and Mickle, S.J. (1982) Foods Commonly Eaten by Individuals: Amount per Day and per Eating Occasion. Home Economics Research Report No. 44. Available from the U.S. Government Printing Office, Washington, DC.
3. U.S. Department of Agriculture (1985) Nationwide Food Consumption Survey: Continuing Survey of Intake by Individuals, women 19-50 years and their children 1-5 years, 1 day. CSFII Report No. 85-1. Available from the U.S. Government Printing Office, Washington, DC.
4. McDowell, A.D., Engel, A., Massey, J.T., and Maurer, K. (1981) Plan and Operation of the Second National Health and Nutrition Examination Survey, 1976-1980. Vital and Health Statistics. Series 1, No. 15. DHHS Publication No. (PHS)81-1317. Available from the U.S. Government Printing Office, Washington, DC.
5. Abrams, I.J. (1981) Mail diary method for collecting food purchasing and food usage information from consumer panels. In, Assessing changing food consumption patterns. Washington, DC: National Academy Press, pp. 119-134.
6. The NPD Group of Marketing and Research Services (1988) National Eating Trends. A research study on in-home food and beverage consumption. Available from The NPD Group of Marketing and Research Services, Park Ridge, IL.
7. Hu, T.-W. (1985) Analysis of Seafood Consumption in the US: 1970, 1974, 1978, 1981. Report prepared for the National Marine Fisheries Service, Washington, DC, under grant no. NA82AA-M-00053 by the Institute for Policy Research and Evaluation, the Pennsylvania State University, University Park, PA. p. 94. Available from The National Marine Fisheries Service, Washington, DC.
8. Bunch, K.L. (1985) Food consumption, prices, and expenditures 1985. Statistical Bulletin No. 749. Available from: U.S. Government Printing Office, Washington, DC.
9. Pennington, J.A.T. and Gunderson, E.L. (1987) A history of the Food and Drug Administration's Total Diet Study, 1961-1987. J.Assoc. Off.Anal.Chem. 70:772-782.
10. This survey has been criticized in a General Accounting Office Report entitled "NUTRITION MONITORING: Mismanagement of Nutrition Survey Has Resulted in Questionable Data", GAO/RCED-91-117, Washington DC. Care should be exercised before employing data taken from this survey.

III B 3 Estimation of Human Exposure to Direct Food Additives and Food Ingredients Continued

b. Estimates of Food Intake

i. Annual Poundage Information (Disappearance Data)

Information on the poundages of commodities entering commerce is usually available from government and industry sources on an annual basis. These data are referred to as "disappearance" data. It is generally not possible to separate out the fraction actually consumed as food from that remaining in inventory and from non-food expenditures from inventory (wasted, exported, used in pet food or animal feed, etc.) at the end of the reporting period. Annual disappearance figures can be divided by the national population and by 365 days to obtain a "per capita" daily intake of the commodity.

Annual poundages of some substances produced and used solely for addition to food have been compiled as a part of the National Academy of Sciences Survey of Industry on the Use of Food Additives (National Research Council, 1977, 1982, and 1987). Industry responses to these surveys are voluntary, and the reliability of these data depends heavily on the completeness of the industry response for a given substance. In order to correct for under-reporting in such surveys, a correction factor is generally employed. This factor is related to the percentage of users of a substance that submitted information to the survey.

ii. Dietary Surveys of Food Intake

Food-intake surveys provide data that are commonly used to estimate exposure to a food additive or ingredient. Two different types of surveys exist: daily-consumption surveys and "food-frequency" surveys, i.e. surveys of the frequency of consumption or number of eating occasions of a given food on a given day. Daily-consumption surveys require participants to record or recall the amounts and types of each food eaten during the day. Food-frequency surveys require participants to record only the number of times each food is consumed during each day of the survey period; these frequencies need to be multiplied by a portion size to obtain the daily food-intake information.

These survey methods have the advantage of providing several different kinds of information about food consumption. That is, food intakes of various sub-groups (e.g. sex, age, eaters-only, total sample etc.) can be obtained for either the total diet or for specific foods. Eaters-only intake data are useful for determining intake of a food used by a small percentage of the population and by individuals selecting for a particular product. In such cases, use of the information derived from the responses of the total sample will generally yield intake figures that are much lower than actual intake. Daily-consumption surveys and food-frequency surveys can also provide valuable information for short-term intake (i.e. days to years).

Food-intake survey data are essential when information is required concerning the potential for very high use among consumers. Because these data are compiled from information obtained from individual consumers, it is possible to determine a distribution of intakes. The 90th percentile intake estimate (that intake which is equal to or higher than 90 percent of the intakes for all individuals surveyed) is used to represent the intake of heavy consumers of a substance.

For a substance not expected to be consumed frequently, the use of surveys of short duration (1-3 days) often leads to an overestimate of chronic intake of the food containing the substance. This is due to a variety of circumstances beyond the scope of this discussion (see the FASEB report noted previously on page 43). Additionally, the use of longer-term surveys (e.g. 14 days) is generally preferred for estimating exposure that is more likely to reflect chronic intake.

III B 3 Estimation of Human Exposure to Direct Food Additives and Food Ingredients Continued

iii. Substance Concentration Data (Petitioner-Supplied Data)

When seeking pre-market approval or approval for a new use of a regulated substance, the petitioner is required to supply information concerning the intended use levels of the substance in food. This information is often supplied as a maximum use level. If demonstrable, a technologically self-limiting concentration can be supplied and used in the estimation of probable intake. Usually, the petitioner will supply a "typical" or "recommended" use level, based on in-house experimentation, which can be used in the exposure estimation process. The Agency can determine which type of information is most pertinent on a case-by-case basis, usually using the information that yields a conservative, yet reasonable exposure estimate (see previous discussion, Chapter II C).

c. Preparing Exposure Estimates

i. Estimating Exposure for Pre-Market Evaluations

Pre-market estimations are intended to represent conservative yet reasonable estimates of exposure to a new substance used in food. Information concerning potential use levels is supplied by the petitioner. Food-intake data are obtained by Agency reviewers from the above mentioned data bases and other appropriate sources, including the petitioner. One basic assumption for making an exposure estimate is that all food ingested by a consumer that may contain the additive or ingredient, does contain it at the recommended or maximum level of use.

A major issue in the pre-market estimation of exposure is the choice of the data base used to determine representative food intake. While broad generalizations can be made, each case requires an examination of the suitability of various data bases with respect to the availability of, or necessity for, information concerning age, sex or other sub-groupings, the extent of consumer awareness of the substance, and the potential ubiquity of the substance in the diet.

a) Per-capita Estimates

In the absence of food intake data, per-capita disappearance data for commodities may be used to make a pre-market exposure estimate for a food additive or ingredient that is expected to have no market appeal (one that will not be sought out by consumers) or that is expected to be ubiquitous in the diet. For these purposes, "per-capita" generally refers to the number of people in the United States. For example, the annual poundage of the commodity first is converted to grams per day (the calculation can be modified if information about non-food uses of the commodity is known). This daily intake figure is then multiplied by concentration of the substance in the commodity (based on the intended use) to yield the probable exposure. One example of this procedure would be the estimation of exposure to an anti-dust spray used in grain silos. The petitioner would have determined the optimal amount of substance sprayed onto a given amount of grain. This concentration would be multiplied by the daily per-capita disappearance of the grain to determine the per-capita exposure to the anti-dust agent. Other information could be factored into this calculation, if available; for example, information about the loss of treated grain during storage and the effects of processing on the quantity of the anti-dust agent that would ultimately be ingested by a consumer.

Conservatism in a per-capita estimate arises from the inability to determine how much of the commodity is lost in storage, waste, or processing, remains in inventory, is exported, or is used in non-food applications. This type of estimate cannot directly produce an upper percentile intake estimate for a substance (see discussion below for making such estimates).

III B 3 Estimation of Human Exposure to Direct Food Additives and Food Ingredients Continued

Per-capita estimates of exposures to substances that have been surveyed by the NAS (food use only) can be made by dividing the reported annual poundages by the current population and converting to daily usage. Conservatism in this type of estimate again arises from the inability to separate wastage, inventory, and loss of the substance in processing from the total actually consumed via the food supply.

Per-capita estimates usually are inappropriate in cases where the use of the additive or ingredient is highly limited, when only a limited number of consumers are eaters of foods containing the added substance, or when a consumer can select for the substance in food.

b) Survey-based Estimates

Because dietary intake or food-frequency surveys contain the most detailed information about the subjects' eating habits, they are the preferred source of food-consumption data for use in estimating exposure. In the simplest case, the average daily intake or the intake for a given percentile of a food containing the substance of interest is multiplied by the concentration of the substance in that food to yield the exposure estimate. For a substance expected to be used in several foods the problem is more complex. The Agency currently has access to food-intake information that is aggregated across groups of individuals and food categories, (i.e. data bases containing consumption information based on general food categories and sub-categories, for example, baked goods, milk and milk products). Using the aggregated food-intake data, the Agency can estimate the mean total exposure (i.e. total sample basis) to a petitioned substance by adding mean exposures to the substance calculated for the individual food categories.

When considering eaters-only intakes, additional considerations should be included. Simple addition of mean, eaters-only intakes may lead to an exaggeration of the mean intake. For example, if a substance is to be used in both regular and diet soft drinks, an overestimation of intake is likely to result from addition of the potential exposures to the substance from each type of soft drink. Consumers usually drink one or the other of these beverages, but not both. Typically, the higher exposure estimate (for, in this case, diet or regular soft drinks) would be used in place of the summed value. The same would be true for potential exposure to an additive from different types of snack foods, such as pretzels and potato chips. Therefore, in calculating the mean eaters-only exposure to additives, caution must be exercised in determining whether exposures derived from aggregated food categories should be added.

A specific percentage intake of an additive that may be used in different food groups can be estimated using different methodologies. We have noted that 90th percentile intakes are typically 2 to 3 times the mean intake. The intake of a heavy consumer of an additive can be approximated by multiplying the derived mean intake by a factor of two to three.^a Also, computer-based modeling, such as a Monte Carlo simulation, can be used to statistically derive distributions of intake for a total sample or eaters-only population. Monte Carlo modeling methods have been described in the literature.^b

^aSurvey intake data from the individual foods provide 90th percentile/mean ratios, which can be averaged to determine the factor used.

^b Thompson *et al.* (1992);² Burmaster and von Stackelberg (1991);³ McKone (1990);⁴ Finkel, (1990);⁵ Rubenstein (1981)⁶

III B 3 Estimation of Human Exposure to Direct Food Additives and Food Ingredients Continued

An alternate approach for estimating eaters-only intakes involves the estimation of exposure derived from dietary survey analyses based on the food consumption of each surveyed individual. Such exposure estimates can be made only by accessing the raw data from the survey. Although the Agency does not have ready access to such data, petitioners have in some instances contracted with owners of raw survey data to provide exposure estimates based on specific information about food uses and use levels of petitioned additives or ingredients. The Agency uses its judgement in considering the manner in which intake estimates for individual food categories should be combined to estimate total exposure. The ability to manipulate the raw survey data permits the actual intake of each food for each surveyed individual to be combined with the proposed use level for the additive or ingredient in the specific foods eaten by that individual. This allows the construction of a distribution curve based on total additive or ingredient intake for each individual; from this curve the desired percentile information on exposure may be obtained. Given that the conservatism inherent in the use of aggregate data has been removed, exposure analyses based on intakes of individuals that have been submitted by a petitioner are carefully evaluated for their appropriateness for predicting probable chronic intake of the substance.

Finally, special cases may arise, particularly in the area of substances that could become macro-ingredients in the diet, for which food-consumption and use-level information necessary for estimation of exposure are inadequate or unavailable. For example, difficulties can arise in estimating intake when current eating habits cannot be reliably extrapolated to include the new substance. In such cases, new approaches to the pre-market estimate will have to be devised. The use of substitutes for added fats illustrates this point. In this example, the diet as a whole, especially the amount of energy needed to maintain normal function, needs to be considered if such a substitute would be marketed to consumers with no restrictions on its use.

ii. Exposure Estimates for Substances Currently in the Market Place

Updated exposure estimates are needed for substances on the market when a new use of an approved substance is petitioned or when intake is believed to have changed appreciably from the time of the original estimate. The approaches available for making this type of estimate are similar to those for pre-market approvals, with the advantage that more information is generally available on the substance, including, but not limited to, actual levels in foods.

Detailed intake estimates can be made using dietary survey information and actual substance use levels. For a new use of an existing substance, a cumulative estimate can be made by combining the appropriate use level and food-intake data for the new use, and adding this estimate to the more accurate estimate available for the existing uses. Alternatively, new data based on an analysis of intakes by individuals and covering both regulated and proposed uses may be submitted by the petitioner. As discussed above for pre-market approvals, estimates for desired sub-groups (age, sex, 90th percentile eaters, eaters-only) can be obtained using these dietary survey data.

d. Conversion Factors

Exposure estimates are commonly presented in grams per person per day (g/p/d), milligrams per kilogram body weight (mg/kg bw), or parts per million of the daily diet. To convert among these unit types, we typically use the following factors: a 60 kg "typical adult," and a total daily diet of 3000 g food and water (1500 g solid food, 1500 g liquid food). For those cases where information concerning children is needed, we have used a body weight of 15 kg for a 2-5 year old child.

III B 3 Estimation of Human Exposure to Direct Food Additives and Food Ingredients Continued

c. Summary

The Federal Food, Drug, and Cosmetic Act (sec. 409 (c)(5)(A)) requires that probable consumption of an additive and of any substance formed in or on food because of its use be considered in determining whether the proposed use is safe. FDA's estimates of probable consumption are generally made using existing commodity disappearance data and food-intake and food-frequency data bases, occasionally supplemented with ad hoc approaches and reasoned judgments. Reasonable exposure estimates for chemicals used in food are critical to the maintenance of a safe food supply. Additional information concerning the preparation of estimates of exposure to food additives is available from the FDA's Center for Food Safety and Applied Nutrition.*

* DiNovi, M.J. (1992)⁷

III C. Recommended Toxicity Tests

1. Recommended Minimum Set of Toxicity Tests

The extent and type of toxicity testing recommended for direct food additives or color additives used in food will depend on the initial Concern Level to which that additive has been assigned and available information about the metabolism, chemical composition, and toxicity of the additive. Recommendations for minimum testing are associated with each Concern Level, and these recommendations reflect the Agency's consensus that extensive toxicity testing should be reserved for additives with high exposures and potentially reactive structures and for additives that induce adverse toxic effects at low doses or after short exposures (see Figure 4 below).

The final extent and type of toxicity testing recommended for a food or color additive will be determined by estimated exposure and potential toxic effects (dose, onset, duration, type, extent, etc.) observed in the minimum set of tests recommended for the additive.

a. Minimum Set of Toxicity Tests for Concern Level III Substances

The recommended tests for Concern Level III substances are sensitive enough to detect nearly all types of observable toxicity, including malignant and benign tumors, pre-neoplastic lesions, and most other signs of chronic toxicity. They include:

- short-term tests for genetic toxicity;
- metabolism and pharmacokinetic studies;
- a subchronic feeding study (at least 90 days in duration) in a rodent species, which includes an evaluation of the potential neurotoxicity and immunotoxicity of the test substance;
- a multi-generation reproduction study (two generations, one litter per generation) with a teratology phase (developmental toxicity study) in a rodent, which includes an evaluation of the potential developmental neurotoxicity and immunotoxicity of the test substance;
- a long-term (at least one year in duration) feeding study in a non-rodent species; and
- carcinogenicity studies on two rodent species. At least one of these studies should be a combined chronic toxicity/carcinogenicity study with an *in utero* exposure phase.

The results of short-term tests for genetic toxicity may be used to determine priority for the conduct of lifetime carcinogenicity bioassays, and may assist in evaluating the results of bioassays. Results of metabolism and pharmacokinetic studies can be used to help set appropriate dose levels in toxicity studies and evaluate the results of those studies; information from metabolism and pharmacokinetic studies also may be used to modify the set of toxicity studies recommended for a particular additive (for example, concern about an additive may be reduced if the additive is shown to be largely unabsorbed by humans). Results from the reproduction study with teratology phase may indicate the need for expanded reproduction and/or developmental toxicity tests. Results of immunotoxicity and neurotoxicity screens in subchronic studies and developmental toxicity studies may indicate the need for further testing in these areas.

Figure 4

Summary of the Toxicity Tests Recommended for Different Levels of Concern

Toxicity Tests ^a	Concern Levels		
	I	II	III
Short-term Tests for Genetic Toxicity	X	X	X
Metabolism and Pharmacokinetic Studies		X	X
Short-term Toxicity Tests with Rodents	X ^b		
Subchronic Toxicity Tests with Rodents		X ^b	X ^b
Subchronic Toxicity Tests with Non-Rodents		X ^b	
Reproduction Study with Teratology Phase		X ^b	X ^b
One-year Toxicity Tests with Non-Rodents			X
Carcinogenicity Study with Rodents			X ^c
Chronic Toxicity/Carcinogenicity Study with Rodents			X ^{c,d}

^a Not including dose range-finding studies, if appropriate^b Including neurotoxicity and immunotoxicity screens^c An *in utero* phase is recommended for one of the two recommended carcinogenicity studies with rodents, preferably the study with rats^d Combined study may be performed as separate studies

III C Recommended Toxicity Tests Continued

b. Minimum Set of Toxicity Tests for Concern Level II Substances

The tests recommended for Concern Level II substances are sensitive enough to detect most toxic phenomena other than late-developing histopathological changes in tissues and organs. Tests recommended for food and color additives used in food assigned to Concern Level II are:

- short-term tests for genetic toxicity;
- metabolism and pharmacokinetic studies;
- a subchronic feeding study (at least 90 days in duration) in a rodent species, which includes an evaluation of the potential neurotoxicity and immunotoxicity of the test substance;
- a subchronic feeding study (at least 90 days in duration) in a non-rodent species, which includes an evaluation of the potential neurotoxicity and immunotoxicity of the test substance; and
- a multi-generation reproduction study (two generations, one litter per generation) with a teratology phase (developmental toxicity study) in a rodent. This study includes an evaluation of the potential developmental neurotoxicity and immunotoxicity of the test substance.

The results of short-term tests for genetic toxicity may be used to identify compounds assigned to Concern Level II for which carcinogenicity testing may be recommended. Results of metabolism and pharmacokinetic studies can be used to help set appropriate dose levels in toxicity studies and evaluate the results of these studies; information from metabolism and pharmacokinetic studies also may be used to modify the set of toxicity studies recommended for a particular additive (for example, concern about an additive may be reduced if the additive is shown to be largely unabsorbed by humans). Results from the reproduction study with teratology phase may be used to indicate the need for expanded reproduction and/or developmental toxicity testing. Results of immunotoxicity and neurotoxicity screens in subchronic studies and developmental toxicity studies may indicate the need for further testing in these areas.

c. Minimum Set of Toxicity Tests for Concern Level I Substances

Recommended tests for Concern Level I substances include:

- short-term tests for genetic toxicity and
- a short-term feeding study (at least 28 days duration) in a rodent species, which includes an evaluation of the potential neurotoxicity and immunotoxicity of the test substance.

The results of short-term tests for genetic toxicity may suggest the need for information about the additive that can be obtained from chronic toxicity or carcinogenicity tests. The short-term feeding study is sensitive enough to detect any acute, life-threatening toxicity and to provide an indication of target organs and doses for toxicity tests of longer duration, if such tests are recommended. Results of immunotoxicity and neurotoxicity screens in the short-term feeding study may indicate the need for further testing in these areas.

III C Recommended Toxicity Tests Continued

III C 2. Selecting Additional Toxicity Tests

Deciding how much information is sufficient to assess the safety of an additive is a problem that has long been recognized both by the Agency and industry. Results from the initial set of recommended toxicity tests for direct food additives and color additives used in food may indicate a need for additional or specialized testing to assess the safety of the additive. Additional recommended tests will depend, in large part, on effects observed in the initial set of recommended toxicity tests. The purpose of this section is to provide examples of how FDA decides what additional toxicological information needs to be developed for a direct food additive or color additive used in food, based on evaluation of data obtained from studies submitted by the petitioner in support of the safety of an additive. The examples are not intended to be comprehensive. Decisions about the need for additional toxicology information on food and color additives used in food will be made on a case-by-case basis, will always include a significant element of expert scientific judgement, and thus may differ from examples presented below.

a. Acute Toxicity Tests

Acute toxicity tests (usually single-dose tests in which animals are observed for 7-14 days following administration of the test substance) may be recommended for compounds when there is no other information that can be used to select appropriate dose levels for short-term or subchronic toxicity tests.

b. Short-Term Toxicity Tests with Rodents and Non-Rodents

Short-term feeding tests with rodents or non-rodents (usually studies in which animals are exposed to continuous oral doses of the test substance for one month or less) may be recommended for compounds when there is no other information that can be used to select appropriate dose levels for subchronic or chronic studies.

c. Subchronic Toxicity Tests with Rodents

Subchronic toxicity tests (usually studies in which animals are exposed to continuous oral doses of the test substance for 90 days to 12 months) may be recommended for Concern Level I compounds with a lowest observed effect level (LOEL) from a shorter-term study which is less than 2000 times the estimated human consumption of the compound.*

Subchronic toxicity studies may be recommended for compounds when there is no other information that can be used to select appropriate dose levels for longer-term toxicity studies.

* Differences in recommendations for additional studies in rodents (subchronic toxicity tests) and non-rodents (one-year toxicity tests) are due, in part, to the greater certainty that can be obtained from tests in which 10-20 rodents per sex per group are used compared to tests in which 2-4 dogs per sex per group are used.

III C Recommended Toxicity Tests Continued

d. One-Year Toxicity Tests with Non-Rodents

One-year toxicity tests in non-rodents may be recommended for Concern Level II compounds when the lowest observed effect level (LOEL) from a shorter-term, non-rodent study is less than 1000 times the estimated human consumption of the compound, particularly if the non-rodent species is the species most sensitive to the effect and is appropriate for extrapolation to man.

One-year toxicity tests in non-rodents may be recommended for Concern Level II compounds when available toxicology information suggests the probability that the compound bioaccumulates and/or is associated with late-occurring toxicity in rodents; such late-occurring toxicity may not be observed or may be poorly quantified in subchronic studies.

e. Carcinogenicity Studies with Rodents

Carcinogenicity bioassays in two rodent species may be recommended for Concern Level I and II compounds when data from other studies indicate treatment-related hyperplasia, metaplasia, or other proliferative lesions, or when data from other studies indicate progressive and irreversible lesions, such as treatment-related necrosis. Carcinogenicity bioassays also may be recommended for Concern Level I or II compounds that have demonstrated significant carcinogenic potential, based on the results of short-term tests for genetic toxicity.

f. Two-Generation Reproduction Studies with a Teratology Phase

Two-generation reproduction studies with a teratology phase may be recommended for Concern Level I compounds when results from other toxicity studies indicate that the compound may be associated with reproductive organ toxicity.

Two-generation reproduction studies with a teratology phase may be recommended for Concern Level I compounds that have demonstrated significant carcinogenic potential, based on the results of short-term tests for genetic toxicity.

g. Gavage Administration of the Test Compound in Teratology Studies

Gavage administration of the test compound in teratology studies may be recommended when the estimated human exposure exceeds 0.625 mg/kg/day in the diet.

Gavage administration of Concern Level III test compounds in teratology studies may be recommended when the compound is expected to be added to beverages that may be consumed by pregnant women.

Gavage administration may be recommended for compounds with adverse reproductive effects that suggest possible teratogenicity.

h. Metabolism and Pharmacokinetic Studies

Additional metabolism and pharmacokinetic studies may be recommended for any compound when results of the recommended set of studies do not resolve important metabolic information, such as whether or not the food additive is absorbed in significant amounts from the gastrointestinal tract.

III C Recommended Toxicity Tests Continued

i. Neurotoxicity Studies

Neurotoxicity studies may be recommended for any compound when results from the neurotoxicity screen or other information suggests that the compound may be associated with neurotoxic effects.

j. Immunotoxicity Studies

Immunotoxicity studies may be recommended for any compound when results from the immunotoxicity screen or other information suggests that the compound may be associated with immune system toxicity.

Chapter IV

Guidelines for Toxicity Tests

A. Introduction

A major difficulty in preparing a safety profile for an additive is the lack of common, consistent, and clearly defined testing guidelines for the design and conduct of toxicity studies. To help eliminate this difficulty, the Agency published its first set of guidelines for toxicity tests for food and color additives used in food in 1982.* The revised guidelines reflect the most up-to-date scientific knowledge relevant to safety evaluation of direct food additives and color additives used in food.

Chapter IV B presents 1) general recommendations for conducting and reporting the results of toxicity studies (see Chapters IV B 1 and IV B 2, respectively) and pathology examinations (see Chapter IV B 3), 2) suggests appropriate statistical analyses of data from toxicity studies (see Chapter IV B 4), and 3) discusses the suitability of purified diets for toxicity studies (see Chapter IV B 5).

Chapter IV C presents additional guidelines for the conduct of specific toxicity tests, including 1) short-term tests for genetic toxicity (see Chapter IV C 1), 2) acute toxicity tests (see Chapter IV C 2), 3) short-term toxicity tests with rodents and non-rodents (see Chapter IV C 3), 4) subchronic toxicity tests with rodents and non-rodents (see Chapter IV C 4), 5) one-year toxicity tests with non-rodents (see Chapter IV C 5), 6) carcinogenicity studies with rodents (see Chapter IV C 6), 7) combined chronic toxicity/carcinogenicity studies with rodents (see Chapter IV C 7), 8) *in utero* exposure phase for addition to carcinogenicity studies (see Chapter IV C 8), and 9) reproduction and developmental toxicity studies (see Chapter IV C 9).

Important changes in the current guidelines from those published in 1982 are described in Chapter I A.

* FDA (1982)¹

IV B. General Recommendations for Toxicity Studies

1. General Guidelines for Toxicity Studies

Guidelines that are common to several or all toxicity studies are described in this section. Guidelines specific to recommended toxicity studies are found in the Chapter IV C, including: genetic toxicity studies (Chapter IV C 1), acute toxicity tests (Chapter IV C 2), short-term toxicity tests with rodents and non-rodents (Chapter IV C 3), subchronic toxicity tests with rodents and non-rodents (Chapter IV C 4), one-year toxicity tests with non-rodents (Chapter IV C 5), carcinogenicity studies with rodents (Chapter IV C 6), combined chronic toxicity/carcinogenicity studies with rodents (Chapter IV C 7), *in utero* exposure phase for addition to rodent toxicity studies (Chapter IV C 8), and reproduction and teratology studies (Chapter IV C 9).

a. Good Laboratory Practice

Toxicity studies should be conducted according to good laboratory practice regulations, issued under Part 58, Title 21, *Code of Federal Regulations*. Title 21, *Code of Federal Regulations*, may be obtained from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 20402, (202) 783-3238.

b. Test Animals

i. Care, Maintenance and Housing:

Recommendations about the care, maintenance, and housing of animals contained in DHEW publication no. 74-23 (NIH publication no. 85-23), "*Guide for the Care and Use of Laboratory Animals*," should be followed unless they conflict with specific recommendations in these guidelines.

ii. Selection of Rodent Species, Strains and Sex:

Healthy animals that have not been subjected to previous experimental procedures should be used. Both male and female test animals should be used in toxicity studies.

In selecting rodent species and strains for toxicity studies, it is important to consider the test animals' general sensitivity and the responsiveness of particular organs and tissues of test animals to toxic chemicals. At this time there is no scientific basis for guiding the petitioner in selecting among the use of inbred, out-bred, or hybrid rodent strains for toxicity tests; instead, the important consideration is that test animals come from well-characterized and healthy colonies. Because recent information suggests survivability problems exist for some strains of rats (see Chapter IV C 6 a), test animals should be selected that are likely to achieve the recommended duration of the study. Additional information on selecting species and strains for particular toxicity tests is presented in guidelines to the specific tests. FDA encourages petitioners to consult with Agency scientists before toxicity testing is begun if they have questions about the appropriateness of particular species and strains.

iii. Infected Animals:

Generally, it is not possible to treat animals for infection during the course of a study without risking interaction between the drug used for treatment and the test substance.

IV B 1 General Guidelines for Toxicity Studies Continued

iv. Animal Identification:

Test animals should be characterized by reference to their species, strain, sex, and weight or age. Each animal must be assigned a unique identification number.

v. Caging:

Animals should be housed one per cage or run (single-caged) except during mating and lactation and for acute toxicity studies. This recommendation reflects two principal concerns:

- organs and tissues from single-caged animals that die during the study should not be lost due to cannibalism; and
- changes in body weight (and body weight gain) of animals consuming diets containing test substances may be indicators of toxicity or ill health, or they may be responses to changes in the palatability of test diets. Thus, in oral toxicity studies, it is important to be able to determine if body weight changes are accompanied by changes in feed consumption. Without such information, it may not be possible to distinguish body weight changes due to palatability problems with the test diet from changes due to toxicity of the test substance. Because the amount of feed consumed by each animal in the study cannot be determined when more than one animal is housed in each cage, these guidelines recommend that test animals be single-caged.

vi. Feed and Water:

In general, feed and water should be provided *ad libitum*. However, gavage administration of the test diet or pair-feeding may be recommended when feed rejection or unexplained depression in growth occurs that may be due to problems with the palatability of the experimental diet. For more information on diets for toxicity studies, see Chapter IV B 5.

vii. Diet:

The test animals' diet should meet all of the nutritional requirements of the species. If more than 5% of the diet is being replaced by the test substance, special care should be taken to ensure that the diet for the control group is equivalent in nutritional value to the diets of dosed groups. For example:

- When the test substance is itself a nutrient, incorporation of such materials at relatively high concentration in the diet (for example, 10%) may interfere with normal nutrition. Under these circumstances, a nutritional deficit is operationally defined as a greater than 10% loss of body weight in the group tested with the test substance, compared with the untreated (basal diet) control group. If the percentage of the test substance added to the diet is small, a single type of control (untreated) may be sufficient. However, if the percentage of the test substance added to the diet is large, the types of control groups needed will be similar to those described below for non-nutritive substances.

IV B 1 General Guidelines for Toxicity Studies Continued

- When the test substance is non-nutritive and comprises a substantial amount of the diet (for example 10%), a control group with comparable dilution of the diet is recommended for long-term studies. For example, one group of control animals should be fed the usual (undiluted, basal) diet; a second group of control animals should be fed the basal diet supplemented with another, preferably inert, non-nutritive substance *in lieu* of the test substance, at a percentage equal to the highest percentage of the test substance in the diet.
- Paired feeding may be recommended when administration of materials in a diet produces inanition in animals from an interference with their feed consumption. Such materials may have an unpleasant taste, poor texture, or other property which makes the new diet radically different from the animal's accustomed diet. In a paired feeding experiment, pairs of litter-mate weanling rats of the same sex and approximate size are used for the control and experimental diets. One member of each pair is placed on the experimental diet and the other is placed on the control diet. The experimental dosage level selected for the paired-feeding study should be one that definitely produced inanition or other toxic manifestation in a shorter-term toxicity study. In the paired-feeding study, food consumption is determined daily, and the control animal is fed an amount of food equal to that which the paired experimental animal ate on the preceding day. If the test substance is non-nutritive, the control animal should be fed an amount of food that is nutritionally equivalent to that which the paired experimental animal ate on the preceding day. An *ad libitum*-fed control group of animals also should be included in the study.

viii. Assignment of Control and Experimental Animals:

Animals should be assigned to control and experimental groups in a stratified random manner; this will help minimize bias and assure comparability of pertinent variables across experimental and control groups (for example, mean body weights and body weight ranges) for statistical purposes.

ix. Mortality:

Excessive mortality due to poor animal management is unacceptable and may be cause to repeat the study. For example, under normal circumstances, mortality in the control group should not exceed 10% in short and intermediate length (not lifetime) toxicity studies.

x. Autolysis:

Adequate animal husbandry practices should result in considerably less than 10% of animals and tissues or organs lost to a study because of autolysis. Autolysis in excess of this standard may be cause to repeat the study.

xi. Necropsy:

Necropsy should be performed soon after an animal is killed or found dead, so that loss of tissues due to autolysis is minimized. When necropsy cannot be performed immediately, the animal should be refrigerated at a temperature that is low enough to prevent autolysis but not so low as to cause cell damage. If histopathological examination is to be conducted, tissue specimens should be taken from the animals and placed in appropriate fixatives when the necropsy is performed.

IV B 1 General Guidelines for Toxicity Studies Continued

c. Test Substance

i. Composition

The specific substance or mixture of substances to be tested should be determined in consultation with the Agency. The composition of the test substance should be known: Information should include the name and quantities of all major components, known contaminants and impurities, and the percentage of unidentifiable materials. The test substance in toxicity studies should be the same substance that the petitioner intends to market.

A single lot of test substance should be used throughout the study, when possible. Alternatively, lots that are similar in purity should be used. The test sample should be stored under conditions that maintain its stability, quality, and purity from its production until the studies are complete.

ii. Route of Administration

The route of administration of the test substance should approximate that of normal human exposure, if possible. The same method of administration should be used for all test animals throughout the study. Animals in all experimental groups should be placed on study on the same day; if this is not possible because of the large number of animals in a study, animals may be placed on study over several days. If the latter recommendation is followed, control and experimental animals should be placed on the study each day in order to maintain concurrence.

Food additives and color additives used in food should be administered in one of the following ways:

- In the diet, if human exposure to the test substance is likely to be through consumption of solid foods or a combination of solid and liquid foods. If the test substance is added to the diet, animals should not be able to selectively consume either basal diet or test substance in the diet on the basis of color, smell, or particle size. If the compound is mixed with ground food and pelleted, nothing in the pelleting process should affect the test substance (for example, heat-labile substances may be destroyed during pellet production by a steam process). When the test substance is administered in the diet, dietary levels should be expressed as mg of the test substance per kg of food.

- Dissolved in the drinking water, if the test substance is likely to be ingested in liquid form (for example, in soft drinks or beer), or if administration in the diet is inappropriate for other reasons. Amount of test substance administered in drinking water should be expressed as mg of test substance per ml of water.

- By encapsulation or oral intubation (gavage), if the two previous methods are unsatisfactory or if human exposure is expected to be through daily ingestion of single, large doses instead of continual ingestion of small doses. If the test substance is administered by gavage, it should be given at approximately the same time each day. The maximum volume of solution that can be given by gavage in one dose depends on the test animal's size; for rodents, the volume should not exceed 1 ml/100 g body weight (if gavage vehicle is oil, see Chapter IV B 5 b). If the test substance must be given in divided doses, all doses should be administered within a 6-hour period. Doses of test substance administered by gavage should be expressed as mg of test substance per ml of gavage vehicle. Finally, the petitioner should provide information, such as metabolism data, that can allow the reviewer to conclude that administration of the

IV B 1 General Guidelines for Toxicity Studies Continued

test compound by encapsulation or gavage is equivalent in all toxicologically important respects to administration in the diet or drinking water.

iii. Controls

A concurrent control group of test animals is required. For dietary studies, the control group should be fed the basal diet; exceptions to this are discussed above and in Chapter VII B. Pair-feeding may be recommended when an unexplained food rejection and/or growth depression occurs in animals receiving diets containing the test substance during short-term and/or subchronic toxicity studies.

A carrier or vehicle for the test substance should be given to control animals at a volume equal to the maximum volume of carrier or vehicle given to any dosed group of animals. Sufficient toxicology information should be available on the carrier or vehicle to ensure that its use will not compromise the results of the study. If there is insufficient information about the toxic properties of the vehicle used to administer the test substance, an additional control group that is not exposed to the carrier or vehicle should be included. In all other respects, animals in the control group should be treated the same as animals in dosed groups.

iv. Selection of Treatment Doses

It generally is not acceptable to select doses for toxicity studies based on information unrelated to the toxicity of the test substance. For example, the highest dose in a bioassay should not be selected so as to provide a pre-determined margin of safety over the maximum expected human exposure to the test substance, assuming that the results of testing at that dose will be negative.

In addition to a concurrent control group, at least three dose levels of the test substance should be used in oral toxicity studies. For all oral toxicity studies except carcinogenicity studies (see Chapters IV C 6 and IV C 7) and reproduction and developmental toxicity studies (see Chapter IV C 9): 1) the high dose should be sufficiently high to induce toxic responses in test animals; 2) the low dose should not induce toxic responses in test animals; and 3) the intermediate dose should be sufficiently high to elicit minimal toxic effects in test animals (such as alterations in enzyme levels or slight decreases in body weight gains). No dose used should cause an incidence of fatalities that prevents meaningful evaluation of the data. Administration of the test substance to dosed groups should be done concurrently (additional information can be found in Chapter IV B 1 b viii).

d. Observations and Clinical Testsi. Observations of Test Animals

Individual records should be maintained for each animal. Animals should be observed at least twice a day throughout the study; the usual interval between observations should be at least 6 hours. Observations should be sufficiently frequent to detect all toxic and pharmacologic effects; the onset and progression of these effects should be recorded. Tumor development, particularly in long-term studies, should be followed: the time of onset, location, dimensions, appearance and progression of each grossly visible or palpable tumors should be recorded. During the course of a study, toxic and pharmacologic signs may suggest the need for additional clinical tests or expanded post-mortem examinations.

IV B 1 General Guidelines for Toxicity Studies Continued

For short-term and subchronic toxicity studies in rodents and non-rodents and developmental toxicity studies in rodents, an expanded set of cage-side tests should be performed to detect neurological disorders, behavioral changes, autonomic dysfunctions, and other signs of nervous system toxicity. Specific information about these tests is contained in Chapter V C.

ii. Body Weight and Food Intake Data

Recommendations are described in guidelines for specific toxicity tests (see Chapter IV C).

iii. Clinical Testing

Ophthalmological examination, hematology profiles, clinical chemistry tests, and urinalyses should be performed as described in the following sections:

a) Ophthalmological Examination: This examination should be performed by a qualified individual on all animals before the study begins and on control and high-dose animals during and at the end of the study. The recommended time intervals for ophthalmological examination are found in individual toxicity guidelines (see Chapter IV C). If the results of examinations subsequent to the initial examination indicate that changes in the eyes may be associated with administration of the test substance, ophthalmological examinations should be performed on all animals in the study for that time interval and for all subsequent time intervals.

b) Hematology: Rodents sampled during and at the end of a study should be the same animals that were sampled before dosing. For dogs, two pre-dose samples should be drawn approximately one week apart. Blood samples should be analyzed individually, and not pooled. If animals are sampled on more than one day during a study, blood should be drawn at approximately the same time each sampling day. The recommended number of animals and time intervals for hematology assessment are found in individual toxicity guidelines (see Chapter IV C).

The following determinations are recommended: hematocrit, hemoglobin, erythrocyte count, total and differential leukocyte counts, and a measure of clotting potential (such as clotting time, prothrombin time, thromboplastin time, or platelet count).

c) Clinical Chemistry: Rodents sampled during and at the end of the study should be the same animals that were sampled before dosing. For dogs, two pre-dose samples should be drawn approximately one week apart. Before blood is drawn for clinical chemistry tests, dogs should be fasted overnight and blood samples should be drawn before feeding. Blood samples should be analyzed individually, and not pooled. If animals are sampled on more than one day during a study, blood should be drawn at approximately the same time each sampling day. The recommended number of animals and time intervals for clinical chemistry assessment are found in individual toxicity study guidelines (see Chapter IV C).

Clinical chemistry tests that are appropriate for all test substances include measurements of electrolyte balance, carbohydrate metabolism, and liver and kidney function. Specific determinations should include:

alanine aminotransferase
albumin
alkaline phosphatase

IV B 1 General Guidelines for Toxicity Studies Continued

aspartate aminotransferase
bilirubin (total)
calcium
chloride
creatinine
 γ -glutamyl transpeptidase
glucose (in fasted animals)
ornithine carbamyl transferase
phosphorous
potassium
protein (total)
sodium
urea nitrogen.

However, when adequate volumes of blood cannot be obtained from test animals, the following determinations should be given priority:

alkaline aminotransferase
alkaline phosphatase
chloride
creatinine
 γ -glutamyl transferase
glucose (in fasted animals)
ornithine carbamyl transpeptidase
potassium
protein (total)
sodium
urea nitrogen.

Additional clinical chemistry tests may be recommended to extend the search for toxic effects attributable to a test substance. The selection of specific tests will be influenced by observations on the mechanism of action of the test substance. Clinical chemistry determinations that may be recommended to ensure adequate toxicological evaluation of the test substance include analyses of acid/base balance, cholinesterases, hormones, lipids, methemoglobin, and proteins.

In spite of standard operating procedures and equipment calibration, it is not unusual to observe considerable variation in the results of clinical chemistry analyses from day to day.^a Therefore, clinical chemistry analyses for all dose groups should be completed during one day. If this is not possible, samples from each control and dose group should be analyzed on each day, with the process repeated on subsequent days.

d) Urinalyses: Microscopic evaluation of urine sediment and determination of specific gravity of urine samples are recommended for animals and time intervals described in the individual toxicity study guidelines (see Chapter IV C).

e) Immunotoxicity: For short-term, subchronic and developmental toxicity studies, results of clinical tests that are included in the list of primary indicators for immune toxicity (see Chapter V D) should also be evaluated as part of an immunotoxicity screen.

^a Gaylor *et al.* (1987)¹

IV B 1 General Guidelines for Toxicity Studies Continued

e. Necropsy and Microscopic Examination

i. Gross Necropsy

All test animals should be subjected to complete gross necropsy, including examination of external surfaces, orifices, cranial cavity, carcass, and all organs. The gross necropsy should be performed by, or under the direct supervision of, a qualified pathologist, preferably the pathologist who will later perform the histopathological examination (see below).

ii. Organ Weight

Organs that should be weighed include the adrenals, brain, kidneys, liver, spleen, testes, thyroid/parathyroid, and thymus. Before being weighed, organs should be carefully dissected and trimmed to remove fat and other contiguous tissue. Organs should be weighed immediately after dissection to minimize the effects of drying on organ weight.

iii. Preparation of Tissues for Microscopic Examination

For 28-day, subchronic, and long-term toxicity studies with rodents, the following tissues should be fixed in 10% buffered formalin (or another generally recognized fixative) and stained with hematoxylin and eosin (or another appropriate stain) for preparation of microscope slides:

- adrenals
- aorta
- bone (femur)
- bone marrow (sternum)
- brain (at least 3 different levels)
- cecum
- colon
- corpus and cervix uteri
- duodenum
- epididymis
- esophagus
- eyes
- gall bladder (if present)
- Harderian gland
- heart
- ileum
- jejunum
- kidneys
- liver
- lung (with main-stem bronchi)
- lymph nodes (representative)
- mammary glands
- nasal turbinates
- ovaries and fallopian tubes
- pancreas
- pituitary
- prostate
- salivary gland
- sciatic nerve

IV B 1 General Guidelines for Toxicity Studies Continued

seminal vesicle
skeletal muscle
skin
spinal cord (at least 2 different locations)
spleen
stomach
testes
thymus (or thymic region)
thyroid/parathyroid
trachea
urinary bladder
vagina
Zymbal's gland
all tissues showing abnormality

For subchronic and one-year toxicity studies in dogs and other non-rodents, the following tissues should be fixed in 10% buffered formalin (or other generally recognized fixative) and stained with hematoxylin and eosin (or other appropriate stain) for preparation of microscopic slides:

adrenals
aorta
bone
bone marrow
brain (at least 3 different levels)
cecum
colon
corpus and cervix uteri
duodenum
epididymis
esophagus
eyes
gall bladder
heart
ileum
jejunum
kidneys
liver
lung (with main-stem bronchi)
lymph nodes (representative)
mammary glands
ovaries and fallopian tubes
pancreas
pituitary
prostate
rectum
salivary gland
sciatic nerve
seminal vesicle
skeletal muscle
skin
spinal cord (at least 2 different locations)
spleen
stomach

IV B 1 General Guidelines for Toxicity Studies Continued

testes
thymus (or thymic region)
thyroid/parathyroid
trachea
urinary bladder
vagina
all tissues showing abnormality

iv. Microscopic Evaluation

All gross lesions should be examined microscopically. In addition, organs and tissues listed above from all animals in the study should be examined microscopically.

v. Histopathology

Histopathology evaluation of the lymphoid organs should be performed as described in the section on immunotoxicity testing (see Chapter V D) for all animals in short-term and subchronic toxicity studies and developmental toxicity studies.

IV B 2. Summary Guidelines for Reporting the Results of Toxicity Studies

Summary guidelines for reporting the results of toxicity studies are contained in this section. More complete information can be found in Chapters IV B 3 c and IV B 4 a and in sections describing guidelines for the conduct of specific toxicity studies (see Chapter IV C). Guidelines for submitting machine-readable data are presented in Chapter II B.

a. Identification

Each test report should be signed by the persons responsible for the test, and should identify:

- the laboratory where the test was performed by name and address;
- the inclusive dates of the test; and
- each person primarily responsible for separate components of the test, including: 1) conduct of the test, 2) pathology, 3) analysis of the data, 4) writing the report, and 5) any other information contained in the report.

b. Good Laboratory Practice for Nonclinical Laboratory Studies

The Good Laboratory Practice Regulations (GLP's) were designed to establish basic standards for conduct and reporting of nonclinical safety testing and are intended to assure the quality and integrity of safety data submitted to the FDA. Each nonclinical toxicity study submitted to the Agency should include either a statement that the study was conducted in compliance with Good Laboratory Practice Regulations, as specified in Section 21 of the CFR 58, or, if the study was not conducted in compliance with GLP's, a statement of the reason for the noncompliance. In the latter case, the petitioner should list the specific areas of noncompliance. Each study report should also include a record of periodic inspections conducted by the Quality Assurance Unit (QAU) showing the date of the inspection, the phase or segment of the study inspected, the date the findings were reported to management and properly signed by the appropriate individuals within the QAU.

c. Body of Report

The test report should include all information necessary to provide a complete and accurate description and evaluation of the test procedures and results. The following sections should be included:

i. Protocol and Amendments

A written protocol that clearly indicates the objectives and methods for the conduct of the study. The protocol should fulfill all the requirements set forth in Section 21 of CFR § 58.120, including the inclusion of all changes in or revisions of the protocol and the reasons for those changes.

IV B 2 Summary Guidelines for Reporting the Results of Toxicity Studies Continued

ii. Summary and conclusions

This section of the test report should contain a brief description of the methods, summary of the data, analysis of the data, and a statement of the conclusions drawn from the analysis.

The summary should highlight all positive data or observations and any deviations from control data which may indicate toxic effects of the test substance.

Neurotoxicity and immunotoxicity screens should be performed on rodents and non-rodents in short-term and subchronic toxicity studies and on the offspring in reproductive toxicity studies (see Chapters V C and D). Reports of these studies should contain a summary statement about the neurotoxic and immunotoxic potential of the test substance.

The summary should include a description of all circumstances that may have affected the quality or integrity of the data.

iii. Materials

The materials section of the test report should include, but not be limited to, the following information:

■ Identification of the test substance:

- i) Chemical name, Chemical Abstracts Service (CAS) registry number (or code number), molecular structure, and a qualitative and quantitative determination of its chemical composition, including names and quantities of known contaminants and impurities and the percentage of unidentifiable materials;
- ii) Manufacturer and lot number of the substance tested, and such information as physical state, pH, stability, and purity; and
- iii) Exact identification of diluents, suspending agents, emulsifiers, excipients, or other materials used in administering the test substance.

■ Animal data:

- i) Species and strain used and, particularly if a strain other than a common laboratory strain is used in the study, rationale for selection of the strain;
- ii) Source of supply of the animals, diet (including lot number, composition, etc.) and water;
- iii) Description of any pre-test conditioning (such as quarantine procedures);
- iv) Description of the method used to randomize animals to test and control groups; and
- v) Numbers, age, and condition of animals of each sex in all test and control groups at the beginning of the study.

IV B 2 Summary Guidelines for Reporting the Results of Toxicity Studies Continued

vi) Data on and fate of, each individual animal in the study.

■ Data on experimental facilities:

i) Descriptions of the caging condition, diet, bedding material, ambient temperature, humidity, and lighting conditions.

iv. Methods

The methods section of the test report should include, but not be limited to, the following information:

■ Deviation from guidelines: This section should indicate all ways in which the test procedure deviates from these guidelines and should state the rationale for each deviation.

■ Specification of test methods: This section should include a full description of the experimental design and procedures, the length of the study, and the dates on which the study began and ended.

■ Statistical analyses: All statistical methods used should be fully described or identified by reference. For a complete discussion of the information that should be contained in this section of the study report, see Chapter IV B 4.

■ Data on dosage administration:

i) All dose levels administered, expressed as mg/kg body weight;

ii) Method, frequency, and time of day of administration; and

iii) Total volume of dose plus vehicle administered to each animal, if the test substance is administered by gavage.

■ Data on observation methods:

i) Duration; and

ii) Method and frequency of observation of the test animals.

v. Results

Presentation of individual results and tabulation of data must accompany each report in sufficient detail to permit independent evaluation of the results.

The following information should be included for each test animal:

■ Time of first observation of each abnormal sign and its subsequent course; [Toxic response data should be organized, when appropriate, by litter.]

■ Time of death during the study for each test animal; [For those animals that are not sacrificed on schedule, cause of death should be determined and reported, if possible.]

IV B 2 Summary Guidelines for Reporting the Results of Toxicity Studies Continued

- Food consumption data (and water consumption data, if the test compound is administered in the drinking water) for each animal;
- Body weights and body weight changes;
- Hematology, clinical chemistry, and other clinical findings;
- Results of neurotoxicity and immunotoxicity studies, as appropriate;
- Gross necropsy findings; and
- Microscopic findings.

Each test animal placed on the study should be accounted for; for animals found dead or moribund during the study, the cause of death should be indicated.

Data also should be summarized in tabular form, organized by sex and dose group; when appropriate, data also should be organized by litter. When numerical averages are presented, they should be accompanied by an appropriate measure of variability, such as the standard error. For each summarized parameter, the following information should be included:

- The number of animals at the beginning of the study;
- The number of animals evaluated for each parameter;
- The time when animals were evaluated for each parameter; and
- The number and percentage of animals positive and negative for each parameter.

All numerical results should be evaluated by an appropriate statistical method; for detailed guidelines about statistical considerations in toxicity studies, see Chapter IV B 4.

- Evaluation of the results should include:

- i) Statements about the relationship, if any, between exposure to the test substance and the incidence and severity of all general and specific toxic effects (such as lesions and tumors, organ weight effects, and mortality effects).
- ii) Statements about the relationship between clinical observations made during the course of a study and post mortem findings.
- iii) An indication of the dosage level at which no toxic effects attributable to the test substance were observed.

IV B 2 Summary Guidelines for Reporting the Results
of Toxicity Studies Continuedvi. References

This section of the study report should include the following information:

- Availability of original data, specimens and samples of the test substance; location of all original data, specimens, and samples of the test substance.
- Literature or references, including, when appropriate, references for: test procedures; statistical and other methods used to analyze the data; compilation and evaluation of results; and the basis upon which conclusions were reached.

IV B 3. Pathology Considerations in Toxicity Studies

Pathology data make up an essential part of the toxicology information submitted to FDA in support of the safe use of food additives, color additives used in food, and other products regulated by FDA. The interpretation of pathology data and other safety data forms the basis for judgement about the safety of a product.

Specific recommendations concerning necropsy of test animals and microscopic examination of organs and tissues for short-term toxicity tests with rodents and non-rodents, subchronic toxicity tests with rodents and non-rodents, one-year toxicity tests with non-rodents, carcinogenicity studies with rodents, combined chronic toxicity/carcinogenicity studies with rodents, reproduction studies, and developmental toxicity studies can be found in Chapter IV B 1 e. In general, these guidelines recommend that all animals in the studies be subjected to complete gross necropsy, all gross lesions should be examined microscopically, and tissues and organs (see Chapter IV B 1 e iii) from all animals in the study should be examined microscopically.

This section on pathology considerations in toxicity studies describes the review process for pathology data, identifies common problems reviewers encounter in reviewing such data, and presents general guidelines for reporting pathology data. Although not addressed in this chapter, CFSAN pathologists also review and provide advice to petitioners on protocols for proposed toxicity studies; requests for such review should be directed to the CSO assigned to the petition (see Chapter II A).

a. Description of the Process for Review of Pathology Data

Review of pathology data may begin with a request for pathology evaluation from regulatory review scientists or from the CAC. This happens when questions about the interpretation of pathology data arise during the scientific review of the toxicology information submitted in support of the safety of food additives and color additives used in food. Requests for review are generally limited to specific interpretative questions, directing the reviewing pathologist's attention to findings in a particular organ or tissue. Occasionally, a reviewing pathologist is asked to examine all the pathology findings in a study.

The pathology portion of the study report usually contains mean and individual organ weights, clinical chemistry results, hematological measurements, summary incidences of observed pathological changes, and gross and microscopic pathology observations for individual animals. An evaluation memorandum from the regulatory review scientist may accompany the material; the memorandum contains summaries of toxicology information, including the results of previous toxicity studies and information from relevant scientific literature.

The pathologist usually begins his/her review by examining the experimental design and methods. He/she carefully reviews general indices of toxicity in test animals (for example, body weight gain, food consumption, clinical or hematologic findings, and organ weight changes); particular attention is paid to the survival of the animals and the number of animals alive at termination. This knowledge helps the reviewing pathologist evaluate the relation of observed pathology changes to treatment.

IV B 3 Pathology Considerations in Toxicity Studies Continued

Although the pathology review is not performed in any strict order, the elements listed below are considered in all reviews. The reviewing pathologist:

- Determines how the percentage of animals with lesions in summary incidence tables has been calculated; for example, was the denominator the total number of animals in the study, or was it the number of animals for which a particular tissue or organ was examined microscopically.
- Compares gross and microscopic findings to ensure that all gross observations are accounted for by microscopic findings or by suitable explanations;
- Examines the diagnostic terminology applied to lesions to determine whether it is contemporary and conventional;
- Checks to see that individual animal data provide adequate information on the location, size, and distribution of reported lesions;
- Considers the qualitative characteristics, severity of lesions, and the incidence figures in evaluating treatment-related differences among groups of experimental animals;
- Carefully evaluates control data before interpreting findings;
- Evaluates the discussion of significant pathological findings prepared by the study pathologist; and
- Tries to correlate pathology findings with other observations about treatment-related effects on test animals during the study.

When the pathology review is completed, a formal written report to the collaborating regulatory review scientist is submitted. The report discusses the pathological findings based on review of submitted material and the relationship of pathological findings to treatment. If questions about the pathology data remain, the report may contain recommendations for requesting additional, clarifying material.

A follow-up pathology review requires additional data. The additional information most often requested by the Agency is clarification of the diagnostic criteria used and historical control data on a specific lesion. The Agency may ask to review the existing microscope slides; in some cases, the petitioner will be asked to prepare new slides from paraffin blocks or wet tissue for FDA review. The Agency's review of slides from a toxicity study provides an independent characterization of the lesions and enables the incidence of lesions to be verified.

In reviewing the microscope slides, the pathologist initially examines the slides without knowledge of the group or dose levels of the compound administered. Having earlier reviewed the pathology data, it is not possible for the pathologist to be completely unaware of the type and incidence of the lesions of concern. However, the pathologist reviews the slides without referring to previous observations, i.e., diagnoses initially are made without specific knowledge of earlier findings.

When microscope slides and other materials are requested by the Agency for a follow-up review, the Agency provides instructions for their submission. Usually, microscope slides from an organ or tissue site should be arranged by treatment group, sex, and in the order of pathology accession numbers. If microscope slides are submitted according to the Agency's directions, and the reviewing pathologist does not have to rearrange the slides before his/her review, the follow-up review will be expedited.

IV B 3 Pathology Considerations in Toxicity Studies Continued

b. Common Problems Encountered during Review of Pathology Data

The timely review of pathology data is sometimes hindered by missing, inaccurate, or incomplete information. These problems are often encountered in submissions to the Agency; a general discussion of problems resulting from information deficiencies is presented below. A more detailed discussion of this subject is available in a recent publication.*

i. Lack of Morphologic Descriptions of Lesions

One of the most common problems causing delay in the review of pathology data is the lack of adequate morphologic descriptions of lesions. Usually, only the diagnoses and numeric incidence data are available for initial review by FDA's toxicologists and pathologists. It is difficult to assess the significance of reported lesions without information on their diagnostic criteria, distribution, and severity. This is a particularly important problem when the terminology for lesions is controversial.

ii. Inconsistency in Applying Diagnostic Terminology

The use of multiple diagnostic terms without explanation for describing a single type of lesion can present problems for the reviewing pathologist. Further clarification is needed to indicate whether two or more terms are being used interchangeably or the results of the study have been evaluated by more than one pathologist, each using different terms for the same diagnosis. For example, in one study the terms "hepatocellular carcinoma" and "hepatoma, malignant" were used in the same set of diagnoses. In another report, four different terms--"c-cell," "clear cell," "light cell," and "parafollicular cell"--were used to describe rat thyroid lesions. In both reports, reasons for using multiple terms for the same diagnosis were not provided.

Differences in the use of diagnostic terms have been encountered when more than one pathologist has examined slides: for example, a study was submitted in which tissues from about one-third of the animals were reviewed by the study pathologist and the remainder were reviewed by a consulting pathologist. The diagnostic terminology was not consistent between pathologists and no attempt was made to explain the inconsistencies in the study report. Although the data appeared to show treatment-related effects, these were subsequently attributed to the way different categories of lesions were summarized.

iii. Incomplete Descriptions of the Results of Gross Pathology Examinations

Incomplete gross descriptions have made it difficult to correlate gross pathology findings with microscopic diagnoses. When microscopic findings do not correlate with gross descriptions, the reviewer must attempt to determine if important information is missing. The report should describe steps taken to resolve discrepancies between gross findings and microscopic diagnoses (for example, recuts of paraffin blocks or additional samples taken from wet tissues).

* Dua and Jackson (1988)¹

IV B 3 Pathology Considerations in Toxicity Studies Continued

iv. Inaccurate Summaries of Data

Inaccurate summary numbers resulting from incorrect counts or calculations have caused difficulty in reviewing pathology data. When pathology data are summarized, all experimental animals should be accounted for and incidence figures should be based on the numbers of animals, organs, and tissues actually examined.

v. Failure to Adequately Discuss the Results of Pathology Examinations

Often, submissions fail to adequately discuss the significance of the results of pathology evaluations. Some reports summarize conclusions but do not explain how the conclusions were deduced from the available pathology data. Some reports base conclusions solely on the results of statistical analyses of data, ignoring broader conclusions that may be discerned from considering all relevant biological information from a study.

c. General Recommendations for Reporting Pathology Data

The pathology section of the report of a toxicity study generally includes an introductory statement and sections on materials and methods, results and discussion, and summary and conclusions.

When pathology data are reported separately from the toxicity study, some brief information about the experimental design and methodology of the toxicity study should be included. This information should include the species and strain of the experimental animals, details about the administration of the test compound, number of experimental and control groups, number of animals in each group, type and frequency of in-life observations including clinical chemistry measurements and hematological examinations, and the scope of gross and microscopic evaluation of tissues. In general, information provided should be sufficient to enable a reviewer to evaluate the quality of the pathology data and to identify its strengths and weaknesses.

Deviations from the original protocol should be explained. For example, if tissues from low- and mid-dose groups were not scheduled for microscopic examination but were examined, the reason for this deviation should be given.

i. Arranging Tabular Data and Morphological Observations

The arrangement of tabular information in an easily comprehensible format is especially important for facilitating review. Table titles and row and column headings should be brief but informative. In the tables showing the individual animal findings, descriptive diagnostic categories should be informative. Redundancy of categories of lesions should be avoided. Morphologic diagnoses should reflect currently accepted criteria. Whenever multiple categories of lesions are grouped under a common "diagnosis," the rationale for grouping should be provided. When multiple diagnoses are not grouped under a common diagnosis, it will be assumed that morphologic differences preclude grouping. Severity grades as well as information on the distribution of a lesion within an organ or tissue should be provided; these observations are particularly important when progression of lesions and effects of different dosages are being studied. In paired organs such as adrenal glands, gonads, and kidneys, lesions should be listed as unilateral or bilateral. All gross lesions should be accounted for by microscopic findings or a written explanation.

IV B 3 Pathology Considerations in Toxicity Studies Continued

ii. Summary Tables

Summary tables in the results section of a report should clearly indicate the number of animals, organs, and tissues actually examined. Unless the number of tissues examined in animals of each group is known and indicated, the incidence figures or mean values indicating effects are subject to question. Summary tables should be free of double counting. In determining incidence, the denominators should reflect actual numbers of animals whose tissues were examined, not just the starting number of animals in each group. The figure for the number of tissues examined should clearly show any adjustments that reflect loss, autolysis, or missing tissue: For example, the accurate incidence of lesions involving the adrenal medulla should be based upon how many adrenal sections (for an animal) from both adrenal glands contained sufficient medullary tissue for microscopic examination. In summarizing lesions that are disseminated, e.g., tumors of the lymphoreticular tissue, the incidence figures should reflect the number of animals with these lesions, not just the presence of the disease in individual organs.

iii. Cross-Reference Table

A cross-reference table that lists individual lesions on the vertical axis and individual animal numbers along the horizontal axis should be included, if possible. This is convenient both for reviewing lesions within an animal and for comparing lesions across animals in a group or among different groups. With the increasing use of computer programs for manipulating pathology data, cross-reference tables can be generated easily.

iv. Animal Disposition Table

The report should generally contain an animal disposition table that provides the pathology accession number, sex, group designation, number of days on the study, and fate of the animals (for example, interim sacrifice, moribund sacrifice, found dead, or terminal sacrifice). This serves as a ready reference for the Agency's scientific reviewers and eliminates the need to develop this information from individual animal data.

v. Pathologist's Narrative

Finally, the report should include a section that specifically discusses the pathology data. This pathology narrative should provide an overview of the pathology findings from the study pathologist's perspective. A discussion that includes qualitative description of lesions and that highlights differences among treated and control groups is an essential part of the interpretation and evaluation of pathology data. The description of morphologic characteristics of lesions is particularly important where terminology may be controversial or misunderstood. Remarks about possible pathogenesis, strengthened by references to the scientific literature, could be an important part of the pathologist's narrative. Significant events, such as a disease outbreak during the study, and the impact of such events on the study outcome should be discussed. If microscope slides have been evaluated by more than one pathologist, any differing diagnoses in the report should be addressed in the narrative. Differences in the incidence of key histopathologic findings among groups should be discussed; if observed differences are not regarded as treatment-related, then the basis for this conclusion should be provided.

IV B 4. Statistical Considerations in Toxicity Studies

The regulations governing approval for marketing new color additives used in food (21 CFR Part 71) and food additives (21 CFR Part 171) imply that petitions should contain both statistical analyses of toxicology data presented in the petition and documentation of the analyses. The purpose of this section is to guide the petitioner in documenting statistical aspects of toxicity studies contained in direct food additive and color additive petitions so that CFSAN reviewers can evaluate these studies efficiently. Additional advice in the form of Standard Operational Procedures (SOPs) prepared by the Division of Mathematics of CFSAN's Office of Toxicological Sciences is available upon request from the CSO assigned to the petition (see Chapter II A).

To ensure the validity of safety assessments of a food or color additive obtained from well-conducted toxicity studies, statistical expertise should be used routinely in the planning, design, execution, analysis, and interpretation of results. This guideline highlights factors that are of primary importance in assessing the validity of evidence from toxicity studies. These factors are 1) study protocol and design, 2) presentation of collected data (individual animal data), 3) presentation and interpretation of analytical results (including tables of summary data), and 4) other considerations.

FDA emphasizes that communication between statisticians and the scientists conducting a particular toxicity study can help ensure that the statistics used are relevant to the biology of the toxicity test. For example, statistical outliers are not always biological outliers, and a "significant" statistical test ($p \leq 0.05$) does not always indicate biological significance. FDA encourages petitioners to consult with Agency statisticians during the design and conduct of the study and the interpretation of data from the study, as appropriate.

The following recommendations offer general guidance to the petitioner in organizing and documenting the results of toxicity studies:^a

- Data should be submitted in a form that will enable FDA reviewers to easily verify the results by duplicating the analysis or, if necessary, performing an alternative analysis. The best way to accomplish this is to submit the data in tabular form in the petition and, at the same time, in a machine-readable form (see Chapter II B for additional information about submission of machine-readable data).
- Summary tables of the data also should be submitted.
- The submission should be organized and documented so as to enable Agency reviewers to move easily between the data and the summary tables.

For example, if the report of a bioassay involving 50 rats in a dose group includes a summary table indicating that the incidence of a given tumor is 3/40, there should be auxiliary tables showing which three rats had the tumor, which 37 rats were examined but did not have the tumor, and which ten rats were not examined for the tumor.

- When outliers are removed for statistical reasons, the statistical test upon which the decision to remove them was based should be specified.

^a Dubey (1985)¹

IV B 4 Statistical Considerations in Toxicity Studies Continued

- The description of a statistical inference should include a statement about the model used, summary data appropriate for the model, analysis of the data with estimates of treatment effects, and reasonable statistical checks on the adequacy of the model.
- In presenting tables of summary data that reference statistical tests of hypotheses, a statement should identify the null and alternative hypotheses, the statistical test, the sampling distribution of the test statistic under the null hypothesis, the value of the test statistic, the degrees of freedom of the test statistic (when appropriate), and the p-value, and whether the test is one or two tailed.
- Statistical analyses should be directly linked to specific questions regarding the safety of the additive (i.e. comparing results for treated groups with results for a control group and evaluating the effects of various animal characteristics [sex, species, age, etc.] on the results of an experiment).
- Results of the statistical analyses of all toxicity studies (e.g., p-values, confidence intervals) should be tabulated. Additionally, an effort should be made to explain how these results contribute to resolving questions about the safety of the direct food additive or color additive.
- The submission should cross-reference related information (e.g. data tabulations, statistical hypotheses tested, models used, etc.) that will facilitate FDA's statistical review of the study.

a. Specific Statistical Issues

i. Study Protocol and Design

The submitted petition should contain the original protocol and a complete account of protocol modifications made during the course of the study. The protocol is a critical document in the evaluation of a bioassay, shaping both the conduct of the study and the ultimate analyses. It sets forth the objectives of the study and relates these objectives to the statistical hypotheses that are tested. It describes critical features of the study's design and execution, such as the purpose of the study, experimental design (subchronic, short-term, multi-generation), selection of species, selection of parameters to be assessed, planned interim analyses of data, planned interim and final sacrifices, events that would trigger early termination of the study, roles and responsibilities of data monitoring boards or quality assurance boards, and proposed statistical methods. By designating in advance the treatment groups and the variables that will be considered to be primary endpoints for statistical analyses, the protocol appropriately defines and limits the hypotheses that the study is able to test.

A well-designed experimental protocol will normally contain, as a minimum, the following items:

- Statement of objectives: In addition to the primary objective(s), secondary objective(s) should be stated explicitly. The precise hypotheses that the study is attempting to prove or disprove also should be stated explicitly.
- Source of test animals: A clear statement about the species, strain, sex and source of the test animals in the study and how animals are screened from the study (i.e. will "runs" be eliminated; why?).

IV B 4 Statistical Considerations in Toxicity Studies Continued

- Experimental design: This should include information about initial baseline periods (if any), the study configuration (short-term, lifetime, etc.), the treatment levels, the control group(s), the number of animals in each group (sample size), and the criteria for terminating the study.
- Randomization procedures: A description of the randomization procedure(s) used to assign animals to experimental groups. Generally, a computer-driven procedure using a random number generator is better than a table of random numbers.
- Route of administration: A statement about the route of administration and frequency of administration of the test compound.
- Diets: A complete description of any diets used in the study.
- Experimental parameters minimized: A statement about how the effects of confounding response variables interest (i.e. caging effects) were minimized.
- Experimental parameters measured: A description of the parameters that will be measured and a statement about how frequently they will be measured.
- Power analysis: A power analysis or a statement about the differences in study parameters between compared groups that the study should detect.
- Quality control: A description of the steps taken to ensure accurate, consistent, and reliable data (e.g. training sessions, standard operating procedures, instruction manuals, data verification, cross-checking or audits).
- Data analysis: A description of planned interim analyses of the data, including monitoring procedures, variables to be analyzed, statistical analyses to be used (including the choice of significance level for each interim analysis), and frequency of analysis.
- Statistical Methods: A description of the statistical methods to be applied to the data. Here, specific questions that the statistical analyses will address in support of the study objectives are identified. For example, a description of the methodology that would be used to detect outliers may be important. The major end-points for analysis should be identified. If multiple comparisons are to be made, they should be pre-planned.

ii. Presentation of Collected Data

Information on every animal in the study should be presented. Data should be organized so that the reviewer can easily find all information about any animal used in the study. For example, data should be organized so that the reviewer can view all study parameters for a single animal and a single parameter for all animals. Individual animal records can be presented or data can be tabulated, depending on the study and the type of data collected. The liberal use of data tables and submission of machine-readable data is strongly encouraged (see Chapter II B). Steps taken to assure the numerical accuracy of the collected data should be documented in detail sufficient to permit the reviewer to judge their accuracy.

IV B 4 Statistical Considerations in Toxicity Studies Continued

As described previously, the identifying number, age upon entry into the study, dose level, sex, initial body weight, and cage identification should be presented for each animal in the study. There also should be a table showing how animals were randomized into their respective dose groups. Other information should include:

- For each animal, length of time in the study, date of death, type of death (e.g. scheduled sacrifice, moribund sacrifice, animal found dead, etc.), and reason for early withdrawal from the study, if this occurred (e.g. escaped from cage).
- Food, water, and test compound consumption at each interval specified in the protocol.
- All measured values for defined parameters and the times at which these measurements were taken. If deviations from standard operating procedures occurred in taking the measurements, the nature of the deviation, the reason for the deviation, and its impact on the study should be discussed.
- For all lesions: Identification of the type of lesion, the organ where the lesion occurred, and whether the lesion was metastatic; the time the lesion was observed; and the severity of the lesion (e.g. mild, moderate, severe).

iii. Presentation and Interpretation of Analytical Results

Presentation of results of statistical analysis should include a description of, and rationale for, all statistical methods used. Unless the method is well-known (e.g. analysis of variance), references should be provided. A thorough discussion of the statistical analysis, including reasons for the use of a particular analysis, assumptions, conduct of the analysis, and validity of the conclusions, will guide FDA in deciding whether re-analysis of the data is needed. For each analysis of a relevant variable that is submitted, the following information should be provided:

- Specific variables and analysis of variance: A statement identifying the specific variable; if not obvious, a discussion of its relevance to the objectives of the study should be included.
- Statistical model: The statistical model underlying the analysis; references should be provided, if necessary.
- Hypothesis: A statement of the hypothesis being tested and of the alternative hypothesis.
- Power calculation: A power calculation for tests that failed to reject the null hypothesis, particularly to justify the adequacy of the sample size.
- Confidence intervals: The statistical methods used to estimate effects, construct confidence intervals, etc.; literature references should be supplied when appropriate.
- Outliers: The methods used to detect outlying data points (outliers) and the reasons why particular methods were selected. Identified outliers should be studied in an attempt to determine the reason for their deviation from other data in the set.
- Assumptions underlying the statistical methods: It should be shown that, insofar as is statistically reasonable, the data satisfy crucial assumptions, especially when such assumptions are necessary to confirm the validity of an inference. For example, in deciding whether to use parametric or non-parametric methods, tests for normality and for equality of variances should be conducted.

IV B 4 Statistical Considerations in Toxicity Studies Continued

- Survival analyses: Such analyses will address the question of whether treated animals died earlier than control animals and will help determine if treated animals lived long enough to enable treatment-related tumors to be detected. Animals that were killed or died accidentally too soon for the animal to have been at risk for a tumor should not be included in the survival analyses.
- Analysis of tumors: Analysis of tumors (benign and malignant) and other lesions for each group of test animals. Whether the tumor is an incidental finding upon death or a cause of death should dictate the method of analysis used. The major theoretical difference between these analyses is the manner in which the number of animals at risk in each time interval is defined. This needs to be taken into account in performing tests such as the standard Cox Life Table test.
- Trend test: A trend test, when appropriate. This includes not only a test for linearity, but a test for lack of fit as well.
- Plots or graphs of summary data: Care should be taken to generate plots that will convey the most information: For example, in studies with many animals in each dose group, it may be better to plot the mean and confidence limits or ± 1 standard deviation than to attempt to plot individual data.

The following points are also important in the presentation of collected data:

- Transformation of data: Unnecessary data transformations should be avoided. If data transformation has been performed, a rationale for the transformation and an interpretation of the estimates of treatment effects based on transformed data should be provided.
- Parametric and non-parametric analyses of data: Parametric and non-parametric analyses of the same parameter at different time periods should be avoided. For example, if equality of variances in a parameter measured over time is tested, and some tests turn out significant and others do not, the statistician should arrive at a consensus (i.e. does the preponderance of evidence point to equality of variances or not). We recommend that this be done by converting p-values obtained to standard normal deviates (z-scores) and obtaining the p-value for the average score times the square root of the number of p-values.
- Litter and caging effects: Litter and caging effects should be taken into account in determining the statistical model. If this is not possible, that fact should be stated along with the reason for the inability to account for these effects and its possible impact on the study.
- Repetitive measurements: For parameters that are measured across time, a repeated measures analysis should be considered.
- Dependent experimental parameters: If a given parameter depends biologically on another parameter (i.e. organ weight depends on body weight), then the dependent parameter should be adjusted, as in analysis of co-variance.
- Time of death: Time of death should be reported as days from the start of the study. For example, if a study begins on January 1, 1990 and the animal dies on January 1, 1992, then the animal died on Day 730.
- Reproduction studies: In reproduction studies, if a dam continues in the study after all pups have died, the number of pups in her litter should be counted as 0.

IV B 4 Statistical Considerations in Toxicity Studies Continued

- Statistical comparisons: When statistical comparisons of data were not pre-planned, a statement on how bias was avoided in choosing the particular analysis should be included.
- Statistic: The statistic, the sampling distribution of the test statistic under the null hypothesis, the value of the test statistic, the significance level (i.e. p-value), a statement of whether the test used was one or two tailed, and intermediate summary data should be presented in a format that will enable the reviewer to verify the results of the analysis quickly and easily. In most cases, a copy of the computer output will provide the necessary information. For example, documentation of a two-sample t-test should include the two sample sizes, the mean and variance for each of the samples, the pooled estimate of variance, the value of the t-statistic, the associated degrees of freedom, and the p-value.
- Computer programs: When possible, commonly available computer programs should be used; please consult with FDA statisticians about appropriate programs. If it is necessary to use a program written by the petitioner itself, the program should be fully documented, including:
 - i) the source code;
 - ii) test runs against "known" results; that is, textbook examples, examples worked by hand, or examples run with packaged programs. These test runs should cover every case that could arise in connection with the data in the petition. Test cases should be run both before and after the program is used for the submitted data.

iv. Support from CFSAN Statistical Reviewers

In the case of a complex toxicity test or carcinogenicity bioassay, the petitioner is encouraged to consult with CFSAN before submitting the petition to discuss relevant statistical considerations. Requests for comments by statistical reviewers on protocols for proposed toxicity studies can be sent to the CSO assigned to the petition (see Chapter II A).

If unusual concerns arise during the conduct of a study, the petitioner may submit preliminary tabulations of the data and materials pertaining to the statistical analysis to CFSAN for advice and guidance.

IV B 5. Diets for Toxicity Studies

The effects of diet composition on the responses of experimental animals to xenobiotics have been reviewed.^a Some of the most important effects include:

- Diet composition may influence experimental results through effects on background rates of toxicology parameters, such as tumor incidence.^b
- Unrecognized or inadequately controlled nutritional and other dietary variables may alter the outcome and reproducibility of long-term toxicity studies.^c
- A number of nutrients and non-nutritive dietary components have been shown to enhance or inhibit carcinogenesis; these include calories or energy, fat, protein, fiber, vitamins C and E, selenium, and lipotropes (methionine, choline, folacin, and vitamin B₁₂). Dietary fibers have been shown to reduce, enhance, or have no effect on the toxicity and carcinogenicity of chemicals.^d Detailed reviews of the interactions of nutrients and carcinogens have been reported.^e

a. Types of Diets

i. Natural Ingredient Diets

Natural ingredient diets are the most widely used diets in toxicology research. They are prepared from unrefined plant and animal materials such as wheat, corn, oats, fish meal, soybean meal, or wheat bran and are characterized as open formula or closed formula diets. The percentages of ingredients in open formula diets are known, but the composition of closed formula diets is proprietary information.^f Natural ingredient diets support growth and reproduction and are economical, commercially available, and satisfactory for studies involving additives that will not affect nutrient balances.

^a Belinsky *et al.* (1987);¹ Clayson (1975);² Conner and Newberne (1984);³ Meydani (1987);⁴ National Research Council Report (1982);⁵ Park and Ioannides (1981)⁶

^b Mitchell *et al.* (1987);⁷ Rader (1989)⁸

^c Rader (1989)⁸

^d Kritchesky *et al.* (1986);⁹ Omaye (1986)¹⁰

^e Conner and Newberne (1984);³ Ip (1987);¹¹ Kritchesky *et al.* (1985);¹² Newberne and Rogers (1986);¹³ Rao (1988);¹⁴ Reddy *et al.* (1980)¹⁵

^f American Institute of Nutrition Report (1977)¹⁶

IV B 5 Diets for Toxicity Studies Continued

Limitations of natural ingredient diets for toxicity studies include:

- Variations in types and quantities of nutrients and other dietary components are due to several factors; for example, the composition of fibers may vary with their sources,^a the mineral content of natural ingredient diets can vary significantly among production batches, and specifications for essential dietary elements are not always met.^b
- Diet composition cannot be altered to study the effects of varying a particular nutrient, which makes natural ingredient diets poor choices for research protocols in which nutrition may influence outcome.
- Nutrient excesses well beyond their requirements, and the presence of other non-nutrients substances in natural ingredient diets support rapid weight gain, pregnancy, and lactation in experimental animals and decrease the effects of many xenobiotics.
- Finally, common contaminants of natural ingredient diets that can alter the response of laboratory animals to experimental treatment include pesticides and mycotoxins.^c

ii. Purified Diets

The use of purified diets has been recommended to avoid some of the limitations associated with the use of natural ingredient diets.^d Purified diets usually contain refined proteins, carbohydrates, and fat. Vitamin and mineral mixtures including highly purified vitamins and inorganic salts also are added to purified diets.^e AIN-76A, the most commonly used purified diet,^f was formulated to provide a diet of known composition that was intended to meet the known nutrient requirements of rodents; it supports growth, reproduction (generally, one or two generations), and lactation in a manner similar to natural ingredient diets.^g

Advantages of using purified diets for toxicity studies include:

- Ability to reproduce nutrient concentrations from batch to batch, to maintain the nutrient composition of a diet within a narrow range, and to alter the type and composition of dietary components.^h

^a Wise and Gilburt (1980)¹⁷

^b Rader *et al.* (1984)¹⁸

^c Newberne and Rogers (1986);¹³ Fox *et al.* (1976)¹⁹

^d Newberne and Rogers (1986);¹³ National Research Council Report (1978);²⁰ Ross *et al.* (1980)²¹

^e American Institute of Nutrition Report (1977)¹⁴

^f Rao (1988)¹⁴

^g American Institute of Nutrition Report (1977)¹⁴

^h National Research Council Report (1978)²⁰

IV B 5 Diets for Toxicity Studies Continued

- Use of purified diets usually decreases dietary intake of contaminants such as pesticide residues, heavy metals, enzyme inducers and other agents that may alter the responses of test animals to experimental treatment.^a

Disadvantages of using purified diets for toxicity studies include:

- Difficulty in assessing the impact of purified diets on animal survival and toxicology endpoints because adequate historical data regarding the use of such diets is lacking;
- Lack of information about the suitability of purified diets for long-term studies, although some researchers have used purified diets successfully for up to 56 weeks;^b
- Errors that may occur in the preparation of purified diets may be more critical than similar errors in the preparation of natural ingredient diets because, in purified diets, each ingredient may be the sole dietary source of an essential nutrient.^c In general, practical experiences with purified diets in long-term studies have not been satisfactory.^d

b. Issues to Consider when Selecting and Preparing Diets for Animals in Toxicity Studies

The following are important issues to consider when selecting diets for animals in toxicity studies:

- Protein requirements for maintenance and growth of laboratory animals are well characterized,^e but this is not true for most nutrients. Nutrient needs^f and metabolism of xenobiotics^g change with age. Hence, the general practice of feeding a single diet throughout the life cycle of experimental animals may be inappropriate--nutritional deficiencies may occur during phases of rapid growth and development in young animals and nutrient excess may occur in older animals.
- Individual ingredients in purified diets may cause problems in long-term studies. For example, purified diets high in ingredients such as casein and sucrose may stick to the hair of rodents and cause excessive grooming. Purified sugars as the sole source of carbohydrates in diets that are low in dietary fiber may cause diarrhea, resulting in problems of digestion and absorption of other nutrients.

^a Rao (1988);¹⁴ National Research Council Report (1978);²⁰ Ross *et al.* (1980)²¹

^b Cruse *et al.* (1978);²² Cruse *et al.* (1978)²³

^c National Research Council Report (1978)²⁰

^d Mitchell *et al.* (1987);² Hamm *et al.* (1982);²⁴ Nguyen and Woodard (1980);²⁵ Harwood (1982);²⁶ Medinsky *et al.* (1982)²⁷

^e National Research Council Report (1978)²⁰

^f Munro (1985)²⁸

^g Garattini (1985);²⁹ Welling (1985)³⁰

IV B 5 Diets for Toxicity Studies Continued

- For reasons that are incompletely understood, animals may not reproduce well when fed purified diets. The components in natural ingredient diets that are required to support reproduction have not been defined.
- Toxic chemicals in the diet and induced nutrient deficiencies can lead to decreased food intake by experimental animals and reduced rates of growth and development. When such an effect is expected to occur in a long-term study, pair-feeding can be used to eliminate differences in food intake among experimental groups; this is the preferred method for ensuring that differences in energy or nutrient intake have not caused the observed experimental results or complicated their interpretation. For example, a moderate restriction of energy intake may increase the life-span, decrease the background cancer rates, and decrease the potency of carcinogens in rodents, thereby potentially modulating the action of a chemical carcinogen. When pair-feeding studies are recommended to eliminate differences in food intake among experimental groups, animals should be single-caged and food consumption should be carefully and accurately determined for each animal in the study.
- When the test substance is added to the diet, accurate records of food consumption must be maintained to determine the administered dose and food intake must be equalized across control and experimental groups of animals. When the test substance is a carbohydrate, protein, or fiber that will be added to the diet in large quantities, it must replace a dietary ingredient or the nutrient and energy contents of the diet will be significantly diluted (see Chapter VII B 1). The nutrient and energy contents of control diets also must be adjusted to match those of experimental diets. One recommended strategy is to make the control and test diets isocaloric. If food consumption among groups of experimental animals has been equalized, then equal densities of metabolizable energy in the diets will equalize nutrient intake across the groups.^a
- When oil is used as the gavage vehicle for fat-soluble test substances, the necessity of including a vehicle-control group in the study may introduce some problems.^b If the quantity of oil administered daily by gavage contributes significantly to the total dietary energy of the animals, results for experimental and vehicle-control groups may be significantly different than results for the untreated control group. If a decision is made to administer a test substance by gavage, the volume of oil given as a vehicle should be limited to 0.3 to 0.4 ml/100 g of body weight and the use of a low-fat diet should be considered.
- Related issues are discussed in the following chapters: 1) control diets for test animals in Chapter IV B 1 b-c; 2) survivorship and recommendations concerning the duration of carcinogenicity bioassays in Chapter IV C 6 a; and 3) nutritional concerns for food substitutes (macro-additives) in Chapter VII B.

^a Leveille and Cloutier (1987)³¹

^b Nutrition Foundation Report (1983)³²

IV C. Guidelines for Recommended Toxicity Studies

1. Short-Term Tests for Genetic Toxicity

FDA recommends the use of a variety of short-term genetic toxicity tests for all chemicals that are direct food additives or color additives used in foods, including chemicals associated with Concern Levels I, II, and III (see Figure 3 in Chapter III B 1). The Agency uses the data from genetic toxicity assays to assist in the evaluation of animal carcinogenicity data. It is also recognized that genetic toxicity assays can be used for determination of heritable effects of chemicals (refer to Chapter VII G entitled "Short-term Tests for Heritable and Somatic Genetic Toxicity").

a. Definition of Genetic Toxicity Testing

Genetic toxicity tests are used to determine the ability of chemicals to cause molecular changes in the DNA or structural or numerical changes in chromosomes of cells. These tests are performed for two distinct reasons: 1) to test chemicals for potential carcinogenicity or 2) to assess whether or not a chemical may induce heritable genetic damage.

Tests used to evaluate genetic toxicity are diverse and include *in vitro* tests using microorganisms and cells from multi-cellular animals, as well as *in vivo* tests using insects, plants, and mammals. Both *in vitro* and *in vivo* tests can be further characterized and grouped on the basis of the endpoint detected. Presently, genetic toxicity assays can be divided into three major groups: 1) forward and reverse mutations [e.g. point mutations, deletion mutations, etc.]; 2) clastogenicity assays detecting structural and numerical changes in chromosomes [e.g. chromosome aberrations, micronuclei, etc.]; and 3) assays that identify DNA damage [e.g. DNA strand breaks, unscheduled DNA synthesis, etc.].

b. Rationale for Selection of Specific Genetic Toxicity Endpoints

Increasing evidence indicates that, although mechanisms that do not directly involve changes in the DNA are also possible, multiple genetic events including suppressor gene loss or inactivation and oncogene activation can contribute to the neoplastic transformation of cells. Studies in several rodent models imply that oncogenes are activated by chemical carcinogens and that this activation process can be a significant early step in tumor induction. Although the mechanism of carcinogen-induced oncogene activation is not understood, activation of members of the *ras* family of oncogenes has been shown to involve a single point mutation. Other genetic alterations that result in oncogene activation include chromosomal rearrangements and gene amplification. Which of these changes is important in neoplastic transformation is not known, but it is now generally accepted that multiple events must occur in a cell before it becomes malignant. Taken together, these studies provide support for the use of genetic toxicity tests, and they corroborate the significant association between the carcinogenicity and mutagenicity of many chemicals.

Over the past 15 years, considerable effort has been directed to the development of genetic toxicity tests and to evaluating their ability to identify chemical carcinogens. Although recent analyses have shown that overall correlations between carcinogens and mutagens are imperfect,^a promising alternative approaches are being developed. Because of the complexity of chemical carcinogenesis, which involves activation, detoxification and other complex interactions within the host, as well as the stages of initiation, promotion and progression through which oncogenesis is generally agreed to proceed, there will probably never be complete agreement between the results of

^a Tennant *et al.* (1987);¹ Ashby and Tennant (1988);² and Ashby and Tennant (1991)³

IV C 1 : Short-Term Tests for Genetic Toxicity. Continued

in vivo carcinogenicity tests and those obtained in genetic toxicity tests. Since some chemical carcinogens do not induce all types of genetic toxicity endpoints, and many others do not interact directly with cellular DNA, genetic toxicity tests will result in some "false negative" results. This indicates that the usefulness of such tests is limited to detection of those carcinogenic agents that are directly active at the genetic level. Because of this, the particular battery of tests used always should be chosen knowledgeably.

c. Test Battery for Genetic Toxicity Testing

Multiple tests are recommended by FDA to provide an adequate perspective on the genetic toxicity activity of a chemical, unless information to the contrary is available. Several tests are needed because it is important to have parallel evaluation for different molecular mechanisms, i.e., gene mutations and chromosomal aberrations. Additionally, no single test can detect the activity of all chemicals, and it is known that certain substances that are not responsive in prokaryotic systems induce responses in eukaryotic cells.

In the evaluation of the genetic toxicity of any substance, FDA considers assays with endpoints for point mutation and chromosomal aberrations to be particularly useful. These endpoints reflect different underlying molecular events, and certain chemicals may cause one or more of these effects. In the absence of information that would indicate that these tests are inappropriate, or not useful for a particular test substance, the Agency recommends the use of a routine battery of three types of genetic toxicity tests:

- gene mutation in *Salmonella typhimurium*;
- gene mutation in mammalian cells *in vitro*^a; and
- cytogenetic damage *in vivo*.^b

i. Gene Mutation in *Salmonella typhimurium*

a) Endpoint Description: The *Salmonella typhimurium* mutagenicity assay measures reversion from histidine dependence (*his*⁻) to histidine independence (*his*⁺) in several strains of bacteria. The changes are induced by agents that cause base-pair substitutions or frameshifts in genes of the histidine operon.

^a FDA recommends the use of either the L5178Y mouse lymphoma cells (thymidine kinase locus [TK⁺]) or another cell line with an autosomal locus that has a documented high sensitivity to mutagenic chemicals (e.g. Chinese hamster ovary [CHO] ASS2 cells).

^b CFSAN highly recommends the concurrent detection of micronuclei and chromosome aberrations in the mouse bone marrow; however, it will consider data from a mouse micronucleus test alone, or a chromosome aberration test using mouse (or rat) bone marrow.

IV C 1 Short-Term Tests for Genetic Toxicity Continued

- b) Test Procedure: The recommended procedure to be used for the *Salmonella typhimurium* reverse mutation assay is the plate incorporation method described by Ames and co-workers;^a however, data from the pre-incubation and suspension method will also be considered.^b

The recommended *Salmonella typhimurium* tester strains for the mutation assay are TA98, TA100, TA1535 and TA1537. TA98 and TA100 should contain the R-factor plasmid (pKM-101) which enhances sensitivity to some mutagens, presumably by modifying an endogenous bacterial DNA repair polymerase complex involved with mismatch-repair processes. In addition, strains TA98, TA100, TA1535, and TA1537 should contain the *rfa* and *uvrB* mutations which enhance their sensitivity to mutagens. The *rfa* mutation results in the loss of one of the enzymes responsible for the synthesis of part of the liposaccharide barrier of the bacterial cell wall which in turn results in increased permeability to certain classes of chemicals. The *uvrB* mutation results in a deficient DNA excision repair system. The use of strain TA1538 is not considered to be generally necessary if strain TA98 is used. The use of other strains (i.e. TA97, TA97a and TA102) should be justified, and experiments with these strains are subject to the same controls and considerations of the recommended four strains in the test battery.

The *Salmonella typhimurium* mutation assay can be performed as a direct plate incorporation assay, a preincubation assay, or a suspension assay. The direct plate incorporation *Salmonella typhimurium* mutation assay involves mixing the test chemical dose, bacteria and molten agar (± S9) and overlaying the mixture on a basal agar layer followed by an incubation of the cultures for 48 hours at 37°C. In contrast, in a suspension assay the bacteria are exposed to the test chemical (± S9) in a liquid suspension, washed free of the test chemical, and plated on selective medium; bacteria from the same suspension are diluted and plated on a similar medium containing biotin and histidine to determine viable counts.^c

In the preincubation assay the bacteria and test chemical are mixed in a tube (± S9) and incubated at 30-37°C for 20-30 minutes. The test chemical is added after the bacteria (± S9) to ensure that the bacteria are not subjected to excessively high, and possibly toxic, concentrations of the chemical. To ensure the integrity of the S9, the reaction tube should be kept on ice one minute or less prior to addition of the test chemical. The top agar with the test chemical is added to the tubes after the preincubation, the contents are mixed, and the mixture is poured immediately into the Petri dish containing the base agar.

ii. Gene Mutation in Mammalian Cells In Vitro:

- a) Endpoint Description: FDA currently recommends the use of an *in vitro* mutation assay that employs a cell line capable of measuring single gene point mutations, frame-shift mutations, and chromosomal mutations (i.e. mutations that affect or involve more than one gene or multiple loci). In contrast, the Agency does not recommend mutation assays that measure the recovery of mutants at one specific gene, but do not permit recovery of chromosomal mutations which may include neighboring essential genes (i.e. systems that select for mutations at the *hprt* locus which is located on a sex chromosome).

^a Ames et al. (1975);^d Maron and Ames (1983)^e

^b Refer to Code of Federal Regulations at 40 CFR §798.5265 (1990).

^c Mitchell (1978)^f

IV C 1 Short-Term Tests for Genetic Toxicity. Continued

To satisfy these criteria, FDA recommends the use of the mouse lymphoma (ML) mutation assay which measures the conversion of L5178Y cells from thymidine kinase independence (*tk*⁻) to thymidine kinase dependence (*tk*⁺). The marker generally used in L5178Y cells to detect the mutagenic event is resistance to trifluorothymidine (TFT) which results from a loss of thymidine kinase (TK) activity. The ML assay was selected over other assays because it detects both specific gene and chromosome mutations, it has a large database^a and it had been chosen by the National Toxicology Program (NTP) testing program to evaluate the mutagenic activity of chemicals that had been tested in rodent bioassays.^b

Nevertheless, FDA will consider data from other *in vitro* mammalian cell mutation tests which detect both site specific and chromosome mutations. For example, the Agency believes that the mutation assays using CHO-AS52 cells (*gpt* locus)^c and the parental and transgenic human lymphoblastoid cell line AHH-1 (*tk*⁺ locus)^d, or any of its transgenic sublines containing human cytochrome p450 cDNA have been sufficiently developed and validated. If a petitioner uses the *hprt* locus in Chinese hamster ovary (CHO)^e or V79 cells^f, then he/she should also submit data from a second system measuring mutations at an autosomal locus (i.e. CHO-AS52 cells). Likewise, if the petitioner wishes to submit data from another system entirely, this test system should have a sensitivity for detecting mutagens comparable to the ML system, the system should have a large and validated database, and the data should be accompanied by a scientific justification for use of the alternative test procedure.

b) **Test Procedure:** General guidelines for detection of gene mutations in somatic cells in tissue culture have been reported.^g In the ML assay, exponentially growing cells in suspension are exposed to the test substance both in the presence and absence of an exogenous metabolic activation system. [If a transgenic cell line transfected with a specific P450 cDNA is used, justification for testing in the absence of exogenous metabolic activation should be supplied by the petitioner.] After removal of the test substance, cytotoxicity is determined by measuring growth rate or cloning efficiency. The remaining treated cells are cultured for sufficient time, depending on the selective marker, to allow for phenotypic expression of induced mutants. Cells are then seeded into both selective and non-selective medium to determine the mutant frequency per surviving cell. General procedures for testing are described for the L5178Y system by Clive and coworkers^h

^a Refer to Dr. W. Caspary, NTP Chemtrack System, Division of Toxicology Research and Testing, National Institute of Environmental Health Science (NIEHS)

^b McGregor *et al.* (1987);⁷ McGregor *et al.* (1988);⁸ McGregor *et al.* (1988);⁹ Mitchell *et al.* (1988);¹⁰ Myhr and Caspary (1988);¹¹ Myhr *et al.* (1990)¹²

^c Tindall *et al.* (1984);¹³ Stankowski and Tindall (1987);¹⁴ Tindall and Stankowski (1987)¹⁵

^d Crespi and Tilly (1984);¹⁶ Crespi *et al.* (1989);¹⁷ Liber and Tilly (1982)¹⁸

^e Hsieh *et al.* (1981);¹⁹ Li *et al.* (1990)²⁰

^f Bradley *et al.* (1981)²¹

^g Refer to Code of Federal Regulations at 40 CFR §798.5300.

^h Clive and Spector (1975);²² Clive *et al.* (1979);²³ Clive *et al.* (1990)²⁴

IV C 1 Short-Term Tests for Genetic Toxicity Continued

iii. Cytogenetic Damage in Mammalian Cells:

a) Endpoint Descriptions: Cytogenetic damage can be evaluated *in vivo* using several different endpoints, including: chromosome aberrations, micronuclei, sister chromatid exchanges, and non-disjunction events (i.e. aneuploidy/polyploidy). FDA recommends the use of assays which detect micronuclei or chromosome aberrations for the assessment of cytogenetic damage. Structural chromosomal aberrations include a variety of cytogenetic damage such as breaks, terminal and interstitial deletions, rings, translocations and dicentrics.

To meet these criteria, FDA recommends the concurrent detection of micronuclei of circulating erythrocytes and chromosome aberrations in marrow cells of the mouse. Nevertheless, the Agency will consider data from only one of these two assays, or a test for chromosome aberrations using rat marrow. While chromosome aberrations can be detected in many mammalian species, and detection of micronuclei is not limited to the mouse, the above mentioned assays were selected because they have large databases and standardized protocols have been developed.

b) Test Procedures: In *in vivo* cytogenetic assays for chromosome aberrations,^a the animals are treated with the test substance and, prior to harvesting, the marrow cells are treated with a spindle inhibitor to arrest the cells in metaphase. Chromosome preparations are made, stained, and analyzed for chromosome aberrations.^b In the micronucleus assay,^c DNA-containing micronuclei are detected in polychromatic erythrocytes through special staining techniques.^d The implicit advantage of both of these *in vivo* assays over *in vitro* assays is that cells within a whole living animal are given the opportunity to metabolize the test substance under natural conditions.

The Agency also considered recommending the use of the sister chromatid exchange (SCE) assay, but decided against this based on the results and analysis of data from the NTP study of 73 compounds. The results indicated that responses in the *in vitro* SCE assay using CHO cells were essentially independent of exogenous metabolic activation requirements.^e

^a Preston *et al.* (1981);²⁵ Preston *et al.* (1987)²⁶

^b The recommended procedure for testing for the induction of chromosome aberrations is described in the Code of Federal Regulations at 40 CFR §798.5385.

^c Heddle *et al.* (1983);²⁷ McGregor *et al.* (1987)²⁸

^d The recommended procedure for the mouse micronucleus assay can be found in the Code of Federal Regulations at 40 CFR §798.5395.

^e Tennant *et al.* (1987)¹

IV C 1 Short-Term Tests for Genetic Toxicity Continued

d. Additional Scientifically Justified Genetic Toxicity Tests

FDA recognizes the necessity of taking into account structure/activity information before selecting specific genetic toxicity tests,^a and acknowledges the existence of non-genotoxic carcinogens which should be dealt with separately.^b Furthermore, the Agency acknowledges that additional, scientifically justifiable, short-term *in vitro* tests may be needed to fully evaluate the genetic toxicity of a test substance. Thus, FDA may recommend expanding the recommended test battery, on a case-by-case basis, to include either variations of tests described above, or different tests to evaluate the genetic toxicity of test chemicals.

There are a number of genetic toxicity tests being developed that can provide information about the potential carcinogenicity/mutagenicity of a substance. These assays are considered to be useful and data from them can supplement the information obtained from the recommended battery of tests. Two such tests are described briefly below.

i. In Vitro Mammalian Cell Transformation Assay

a) Endpoint Description: A morphological change (i.e. transformation) is observed in a colony of wild type cells within a contact-inhibited monolayer of normal cells. The morphological change is characterized by piling of the cells in an irregular, criss-cross pattern that represents a loss of normal growth inhibition and cell-cell orientation.

b) Test Procedure: The methodology for BALB/c-3T3 cell transformation assay (-S9) was first described by Kakunaga,^c and revisions of that procedure were recommended by an IARC/NCI/EPA Working Group^d and a recent NTP program.^e In the BALB/c-3T3 transformation assay, rapidly growing cells are seeded and grown as a monolayer tissue culture. These cultures are exposed to the test substance for 48-hours (days 2-4 after seeding), washed to remove the test chemical, and refed bi-weekly for a total culture period of 28-days. Cytotoxicity of the test chemical is measured in a co-culture clonal survival assay.^f

FDA currently recommends use of the BALB/c-3T3 cell transformation assay over other transformation assays because this assay has been performed on >200 chemicals tested under identical experimental conditions. In addition, this assay has been shown to be capable of selectively detecting non-mutagenic carcinogens.^g Finally, this is the only transformation assay for which a

^a Refer to Chapter II C in the 1992 Agency Guidelines; Ashby and Tennant (1991);² Klopman *et al.* (1990);²⁹ Rosenkranz and Ennever (1988);³⁰ Rosenkranz and Klopman (1990);³¹ Rosenkranz and Klopman (1990);³² Rosenkranz *et al.* (1990)³³

^b Ashby and Tennant (1988);² Ashby and Tennant (1991)³

^c Kakunaga (1973)³⁴

^d IARC/NCI/EPA Working Group (1985)³⁵

^e Matthews (1986);³⁶ Matthews *et al.* (1993)³⁷

^f Matthews (1993)³⁸

^g Matthews *et al.* (1993)³⁹ Matthews *et al.* (1993)⁴⁰

IV C 1 Short-Term Tests for Genetic Toxicity Continued

structure/activity relationship model has been developed to interpret the transformation responses of genotoxic and non-genotoxic chemicals.^a

The Agency acknowledges that a number of different cell transformation assays are available for measuring chemically-induced morphological transformation of cells; however, the Agency believes that these systems have not been sufficiently developed at this time. Such additional cell transformation assays use continuous cell lines (e.g. the C3H10T_{1/2} assay), 2) primary or early passage cells (e.g. the Syrian hamster embryo [SHE] colony assay) and 3) virus-infected cells (e.g. the SHE infected with Simian adenovirus SA7 assay).

ii. DNA Damage in Mammalian Cells (Unscheduled DNA Synthesis)

a) Endpoint Description Unscheduled DNA synthesis (UDS) occurs during the repair of DNA damage induced by a variety of agents in non-S-phase cells. It involves excision of DNA adducts followed by strand polymerization and ligation to restore the original DNA structure.

b) Test Procedure In the standard hepatocyte/DNA repair assay,^b primary rat hepatocytes are exposed to the test substance in medium containing ³H-thymidine. At the end of the treatment period, the cells are fixed and exposed to autoradiographic emulsion to determine the amount of labeled thymidine incorporated into the DNA. At the end of the exposure period, the slides are developed and the cells are stained. Nuclear and cytoplasmic grains are then counted and used to calculate net nuclear grains.

The Agency recognizes that the standard *in vitro* UDS assay is insensitive to some hepatocarcinogens and other species- and organ-specific carcinogens, and has a high false-negative rate, as determined in an evaluation by the NTP.^c More recent studies, however, indicate that the sensitivity of this assay can be enhanced by pre-treatment of the animals with mixed-function oxidase (MFO) inducers.^d Accordingly, if a chemical gives a negative response with hepatocytes from un-induced animals, the test should be repeated with hepatocytes from induced animals.

An *in vivo/in vitro* variation of the UDS assay also has been developed and can be used as an alternative test.^e In this procedure, young animals are pre-treated with the test substance and, after an appropriate period of time, the liver is perfused and the liver cells are placed into culture. Uptake of ³H-thymidine is determined by autoradiography, as in the standard *in vitro* procedure.

^a Matthews *et al.* (1993)⁴⁰

^b Williams, G.M. *et al.* (1982);⁴¹ Butterworth, B.E. *et al.* (1987);⁴² Butterworth, B.E. *et al.* (1987)⁴³

^c Tennant, R.W. *et al.* (1987)¹

^d Shaddock, J.G. *et al.* (1989);⁴⁴ Shaddock, J.G. *et al.* (1990)⁴⁵

^e Butterworth, B.E. *et al.* (1987b)⁴⁶

IV C 1 Short-Term Tests for Genetic Toxicity Continued

e. Standards for Acceptability of Test Results

FDA has minimum standards of acceptability for short-term genetic toxicity tests. Tests which do not meet these standards may be unacceptable for evaluating genetic toxicity of the test substance.

i. Experimental Parameters

- a) Replicate Experiments: For *in vitro* tests, the test substance should be tested in independent (i.e. different times), replicate experimental trials; the repeat experiment should use the same method but with fresh reagents and cells. In addition, FDA recommends that the replicate experiment use doses of the test chemical that are adjusted from the first experiment to optimize detection of genetic toxicity. The Agency recommends that the test substance be evaluated in replicate experiments whether the substance is considered to be active, inactive, or equivocal in the first experiment.
- b) Replicates of Treatment Doses: The Agency recommends that substances evaluated in the *in vitro* assays use the current recommended number of replicates of each dose (e.g. triplicate doses for the *Salmonella typhimurium* assay and triplicate doses for the ML assay). In the *in vivo* chromosome aberration assay, it is recommended that the study evaluate at least 100 cells in mitosis; in the mouse micronucleus test, it is recommended that the study evaluate at least 1000 erythrocytes.
- c) Negative Control: Acceptable concurrent negative control data should be submitted for each test. Negative control data should be obtained from animals or cultures treated with the solvent used to solubilize the test compound. The solvent should be used at the maximum concentration used in experimental groups; this concentration should have no effect on animals, or cell growth, cell survival, or mutagenic response. To ensure that a non-aqueous solvent is not having an adverse effect on mammalian cells in culture, an aqueous medium control should also be included.
- d) Positive Control: Acceptable, concurrent positive control data should be submitted for each test. Dose levels for positive control chemicals should be selected so that they are high enough to elicit a significant response, but low enough to fall on the rising portion of the dose response curve. In the absence of an exogenous activation system, a positive control chemical must either directly induce genetic toxicity, or the target animal or cell must be capable of using the chemical to induce genetic toxicity. In contrast, the positive control in an experiment which uses an exogenous activation system should employ a chemical which is inactivate when the exogenous activation system is not included in the experiment and active only when the exogenous activation system is present.
- e) Treatment Duration: The Agency recommends that duration of treatment with the test chemical in genetic toxicity assays be consistent with current optimal treatment time (e.g. 48-hours for the direct plate incorporation method of the *Salmonella typhimurium* reverse mutation assay).
- f) Exogenous Activation System: In *in vitro* assays detecting chemical-induced mutations in prokaryotic and eukaryotic cells, the test substance should be tested in the presence of an exogenous activation system. The most common exogenous activation system is Aroclor 1254-

IV C 1 Short-Term Tests for Genetic Toxicity Continued

induced rat liver.* While other types of S-9 may be acceptable, their use should be justified in the report. When the chemical being tested is known to be activated more effectively by a different type of S-9, then that S-9 should be used in addition to the Aroclor 1254-induced rat liver S-9. For chemicals likely to be metabolized extensively by pathways that do not occur in the liver S-9 activation system, additional modifications may be necessary.

Filtration of the S-9 or S-9 mix may lead to loss of enzyme activity.⁵ Therefore, liver S-9 should be prepared using aseptic techniques so that subsequent filter-sterilization is not required. Each batch of S-9, whether produced by the testing laboratory or obtained commercially, should be tested for sterility and discarded if contaminated.

The composition of the S-9 mix used should be described completely in the report rather than exclusively by reference to published literature. It is particularly important to specify the amount of S-9 used in the S-9 mix since this value is not specified by Ames *et al.*⁴ or by Maron and Ames⁵. In general, the amount of S-9 in the S-9 mix should be in the range of 4% to 10%, corresponding to 20-50 μ l per plate. Amounts outside of this range should be justified in the report, for example, by documenting that chemicals of a particular class are more readily detected as mutagens at another S-9 concentration. Also, any deviations from cofactor mix components or concentrations specified by Maron and Ames⁵ should be justified.

ii. Test Substance Parameters

a) Selection of the Solvent Control: Whenever possible, the solvent vehicle control should be an aqueous buffered solution. When a chemical is insufficiently soluble in aqueous solvents, then an appropriate solvent vehicle should be used to maximize solubility of the test compound in culture medium. The solvent vehicle could be an organic solvent such as acetone, dimethyl sulfoxide, or ethanol or a non-ionic surfactant such as pluronic F68. In some cases, experiments with and without the solvent are necessary to document that the solvent itself has no mutagenic effect. In addition, positive control chemicals should be dissolved in alternative solvents other than a buffered aqueous solution to show that the solvent does not affect results obtained from positive control chemicals.¹⁰

b) Range-finding Experiment: Preliminary toxicity tests should be performed to assist in selecting the highest dose used in mutagenicity and cytogenetic assays. Such tests should be executed using precisely the same protocol that will be used for the standard assay; however, the recommended number of replicate doses (or *Salmonella typhimurium* tester strains) may be reduced. In the *in vitro* assay, cytotoxicity is usually manifested by a significant decrease in the number of spontaneous colonies of cells (or revertants) per plate. In the *in vivo* cytogenetics assay, toxicity is measured in terms of a demonstrable effect on rodent marrow (e.g. cell cycle delay).

* The Agency is aware of the current safety and disposal considerations with aroclor; thus, other inducing agents are under consideration and may become more widely accepted in the near future. Critical to the acceptance of other inducing agents will be the existence of appropriate validation studies and sufficient data to compare their effectiveness with the aroclor S9. Please consult with CFSAN scientists before selecting an alternate to aroclor.

⁵ Maron and Ames (1983)⁵

IV C 1 Short-Term Tests for Genetic Toxicity Continued

- c) Selection of Treatment Doses: Test substances should be tested using five or more treatment doses in *in vivo* mutation assays using procaryotic or eukaryotic cells. Furthermore, when possible, one or more dose levels should be clearly cytotoxic to the target organism.

In the *in vivo* cytogenetics assays, the highest treatment dose (HTD) should be either the HTD to a maximum of 25 mM, or a dose producing some indication of cytotoxicity (e.g. partial inhibition of mitosis). Toxic chemicals should be tested in at least one experiment with three or more doses in which a dose-related change in cytotoxicity induced by the test substance can be detected. In contrast, a single dose utilizing an acceptable number of animals may be used in either a preliminary study or in a study with a non-toxic chemical tested at the HTD.

In all of the recommended genetic toxicity tests, as well as additional scientifically justified tests, test chemicals should be tested using a dilution scheme which includes more than two doses per 10-fold dilution of the test chemical (i.e. 1:2-fold dilutions of 1000, 500, 250 $\mu\text{g/ml}$, etc.). In general, doses should be approximately evenly (geometrically) spaced. Testing at 10-fold intervals between doses, or using other dilution schemes, incurs the risk of missing a crucial intermediate dose.

- d) Highest Treatment Dose: Test substances which are relatively non-cytotoxic to the target cell should be tested at the highest, scientifically justified, treatment dose (HTD). Since test substances have a wide range of molecular weights, we recommend a HTD of 25 mM, in the absence of solubility problems. Thus, a solid, non-cytotoxic chemical with a molecular weight of 200 and no solubility problems would have a HTD of 5 mg/ml *in vitro* (or 5 mg/kg bw *in vivo*). This dose can be lowered if the test substance elicits one or more of the physicochemical problems listed below. This HTD rule is based upon the Agency's concern that some chemicals may be tested at concentrations that could significantly affect the osmolarity of the culture medium. If the highest dose used is lower than the HTD, then the highest dose used should be clearly toxic to the cells in each test (as shown by decreased colony counts) or it should be at, or close to, the limit of solubility of the chemical in the solvent used.

When a chemical mixture is being tested, the composition of the mixture should be stated as completely as possible. The mixture should be tested at doses such that 25 mM of the principal or active ingredient in the mixture is added to each plate, unless this is not possible because of toxicity or limited solubility.

iii. Additional Concerns

- a) Criteria for a Valid Experiment: The study should describe the criteria that were used to determine whether the experiment is valid. For example, the *Salmonella typhimurium* mutation assay usually has six or more criteria for a valid experiment:

- agar culture vessels were tested and shown to be sterile;
- tester strains with the *rfa* cell wall mutation were sensitive to crystal violet;
- TA98 and TA100 were tested and found resistant to ampicillin;
- adequate titers of tester strains were used in the study (e.g. 5×10^4 cells);
- all positive controls induced ≥ 3 -fold increase in revertants/plate; and
- a minimum of 3 non-toxic dose levels were used.

- b) Mycoplasma and Microbial Contamination: All continuous cell lines used for genetic toxicity tests should be checked routinely for mycoplasma contamination. Documentation

IV C 1 Short-Term Tests for Genetic Toxicity Continued

should be provided to show that cells used in the assays for which data are submitted have been checked for mycoplasma contamination.

c) Experimental Losses due to Contamination: The loss of experimental samples because of microbial contamination and other accidental causes should be minimal. Assays having too many doses or replicates missing because of contamination or other technical errors may not be considered acceptable. In general such assays should be repeated.

d) Additional Protocol Deviations: The experimental methods actually used in genetic toxicity tests may vary somewhat from laboratory to laboratory and from the recommended guidelines. Such variations may be acceptable as long as they are described fully in the study. For example, alternative tester strains have been developed for use in the *Salmonella typhimurium* reverse mutation assay. Likewise, the medium in which cells were grown (including the manufacturer of the medium) should be declared. An other example is that Maron and Ames¹ recommend oxoid nutrient broth No. 2 (CM 67) as the growth medium for cells; if another growth medium was chosen for the study, its use should be scientifically justified.

f. Data Collection and Evaluation

Genetic toxicity data can be obtained by hand or automatic colony counters, but the method used should be specified in the report. If an automatic counter is used, the type of counter should be specified. The report should state whether the reported data are uncorrected counts taken directly from the automatic colony counter or whether some method of calibration was used. If the colony counter was calibrated to correct for the decrease in apparent counts as the number of colonies per plate increased, then the method used for the calibration should be explained in the report. A calibration curve (hand counts vs. uncorrected automatic colony counts) should be included in the report if corrected counts are submitted.

Test substance cytotoxicity sometimes results in the appearance of relatively small colonies of revertant cells in the *Salmonella typhimurium* and ML mutagenesis assays. Since the biological meaning of these small colonies is ambiguous in both assays, the Agency recommends that studies using these procedures carefully describe the criteria used to accept or reject small colonies as part of the test substance activity. If unusually small colonies are considered to be significant observations, then a number of the small colonies should be isolated and analyzed for stability of this phenotypic change. Only true revertants should grow under the standard selection conditions.

There is no single, generally accepted method for distinguishing a positive from a negative result in the *Salmonella typhimurium* or ML mutation assays. Fortunately, results are usually clearly negative or a chemical induces a clear dose-related increase in activity. However, some chemicals induce weak or marginal responses that are not completely reproducible. The most widely used criteria for determining whether or not a result is positive include: 1) an increase of at least 2-fold over the spontaneous level at two or more consecutive doses or at the highest non-toxic dose tested and 2) a reproducible dose-dependent response. The so-called "2-fold rule" is often modified to a "3-fold rule" when the spontaneous count is low, for example 10^3 or less.

These criteria are only general guidelines. When only a single dose appears to give a positive response, it is important for repeat tests to be performed at smaller dose intervals to see if the response is reproducible and if a dose-dependent response can be seen. When marginal or not

¹ Maron and Ames (1983)⁴

IV C 1 Short-Term Tests for Genetic Toxicity Continued

completely reproducible results are obtained, it may be useful to vary the protocol somewhat, for example by altering the amount or type of S-9 or by using a preincubation protocol. In some cases it is not possible to decide, without reservation, whether or not a chemical induced genetic toxicity; in such cases it is necessary to conclude simply that the activity is weak, marginal, equivocal, or not completely reproducible.

g. Reporting Requirements

Reports of genetic toxicity tests submitted to the Agency should be as complete and detailed as possible so that FDA reviewers can be assured that the assays were performed appropriately (see Chapter IV B 2). The report should include:

- a detailed protocol that contains the information required by Good Laboratory Practice regulations for Non-clinical Laboratory Studies;
- scientific justification(s) for deviations from recommended guidelines;
- an adequate description of the test system [For example, information should be included on the source of the bacterial tester strains and mammalian cell lines, as well as the methods used for their storage and for the preparation of cultures for testing];
- all raw data (individual counts), in addition to the mean counts, for any transformed data submitted;
- historical negative and positive control data for a recent sequence of experiments, in addition to concurrent negative and positive control data [These data will be used to determine the acceptability of the concurrent solvent and positive controls; the solvent control values should be within the historical range established by a particular laboratory and should be consistent with published values for each particular system];
- a statement by the petitioner as to why he/she feels that the tests that were done are capable of detecting genetic toxicity in the specific chemical(s) tested; and
- a description of physicochemical properties of any test substance that could cause technical difficulties in testing the compound, as well as any problems incurred in the test experiments.

Physicochemical properties that can cause technical difficulties may include, but are not limited to:

- volatility (for liquid test substances);
- acidity and alkalinity (these compounds could alter the normal physiologic pH of the culture medium);
- solubility in culture medium, organic solvents, and/or non-ionic surfactants (e.g. pluronic F68); and
- reactivity (Reactivity problems may include reaction of the test substance with the plastic culture vessel, with functional groups on biochemicals in the medium (or cells), light and temperature sensitivity, sensitivity to air (i.e. oxygen), and other problems that could affect the activity of the test substance in the genetic toxicity assay.).

IV C 2. Acute Oral Toxicity Tests

Acute toxicity tests can provide preliminary information on the toxic nature of a material for which no other toxicology information is available. Such information can be used to:

- deal with cases of accidental ingestion of a large amount of the material (e.g., for poison control information);
- determine possible target organs that should be scrutinized and/or special tests that should be conducted in repeated-dose toxicity tests;¹ and
- select doses for short-term and subchronic toxicity tests when no other toxicology information is available.

In most acute toxicity tests, each test animal is administered a single (relatively high) dose of the test substance, observed for 1 or 2 weeks for signs of treatment-related effects, then necropsied. Some acute toxicity tests (such as the "classical" LD₅₀ test) are designed to determine the mean lethal dose of the test substance. The median lethal dose (or LD₅₀) is defined as the dose of a test substance that is lethal for 50% of the animals in a dose group. LD₅₀ values have been used to compare relative acute hazards of industrial chemicals, especially when no other toxicology data are available for the chemicals. However, many important observations of toxicity are not represented by LD₅₀ values or by slopes of dose-response curves for lethality. For example, information about morbidity and pathogenesis may have more toxicological significance than mortality, and these endpoints also should be evaluated in short term toxicity tests.

The Agency does not recommend that petitioners determine the median lethal dose (or LD₅₀) for direct food additives or color additives used in food. However, if a petitioner decides to conduct an acute oral toxicity test, alternative test protocols can provide useful information about the acute toxicity of a substance.² These protocols generally use fewer animals, and are thus more cost efficient, than tests designed to determine LD₅₀s.³ The following guidelines should help the petitioner design acute oral toxicity tests when the petitioner has decided that such information is useful:

- The main focus of the acute toxicity test should be on observing the symptoms and recovery of the test animals, rather than on determining the median lethal dose (LD₅₀) of the substance.
- The rat often is used as the animal model in acute toxicity tests, but other species also may be used.
- Often only one sex is studied in an acute toxicity test; generally, the female is assumed to be more sensitive to the acute toxic effects of chemicals than the male.⁴

¹ Gad and Chengelis (1988)^{1a}

² Litchfield and Wilcoxon (1949)³

³ FDA LD₅₀ test policy (1988)²

⁴ Gad and Chengelis (1988)^{1b}

IV C-2 Acute Oral Toxicity Tests Continued

- Before deciding on the dose of a test compound that will be used in studying its acute toxicity, the compound's chemical and physical characteristics (including molecular weight, partition coefficient, and the toxicity of related chemicals) should be considered; otherwise, oral toxicity—including lethality—caused by relatively large doses of a chemical may have no biological relevance to the chemical's effects at lower doses.^a

The following brief descriptions of oral toxicity tests may help the petitioner choose a test that meets his needs; detailed information about each type of test is available in the referenced material.

a. Limit Tests

To determine the acute toxicity of a new food additive that is not expected to be particularly toxic, 5 gm (or ml) of the compound/kg body weight of the test animal should be administered orally by gavage to several (perhaps 5) animals that have been fasted (overnight for rats, 4 hours for mice). Test animals should be observed closely for up to 14 days; symptoms of toxicity and recovery should be noted. Gross and histopathological examination of the test animals at the end of the study may help identify toxic effects on target organs. If no animals die as a result of this dose, there is no need to test higher dosages. The acute toxicity of the compound can then be expressed as being greater than 5 gm (or ml)/kg body weight of the test animal. This method is called the "limit test." In general, 5 gm or 5 ml of the test substance/kg body weight is the practical upper limit for the amount of test material that can be administered in one oral gavage dose to a rodent.

If there are deaths following administration of an acute dose of 5 gm/kg body weight, then a lower dose should be administered to several (perhaps 5) animals and the results evaluated as discussed above. For compounds expected to be acutely toxic at 5 gm/kg body weight, it would be wise to select a lower initial "limit" dose.

b. Dose-Probing Tests

Dose-probing acute toxicity protocols may have value when the petitioner has no preliminary information about the test substance that would help him select appropriate doses for toxicity studies. In a dose-probing acute toxicity test, one animal per each of 3 widely spaced dosages should be used and a sufficient observation period should follow administration of the doses. Subsequent toxicity studies may be based on the results of the dose-probing study.^b Variations of dose-probing acute toxicity studies are described in the literature.^c Other methods of determining appropriate doses for longer-term toxicity studies include a simple test wherein 3 or 4 doses are each administered to 1 or 2 test animals and the animals are observed for up to 14 days. If some of the animals die, one can estimate an approximate median lethal dose, termed ALD.^d

^a Gad and Chengelis (1988);^{1c} Zbinden and Flury-Roversi (1981)⁵

^b Gad and Chengelis (1988)^{1c}

^c Lorke (1983);⁴ Schutz and Guchs (1982)⁷

^d Deichmann and LeBlanc (1943)⁶

IV C 2 Acute Oral Toxicity Tests Continued

c. Up-and-Down Tests

The "up-and-down" procedure involves dosing animals one at a time: First one animal at one dose, then another animal one or two days later at a higher dose (if the first animal survives) or a lower dose (if the first animal dies). This process continues until the approximate LD_{50} has been determined. One disadvantage to this test is the length of the study. Each animal should be observed for at least seven days after dosing so that delayed deaths can be recorded. However, this method usually requires only six or eight test animals as compared with the 40 to 50 test animals that may be used in the "classical" LD_{50} test.*

d. Pyramiding Tests

Pyramiding studies involve a minimum number of animals: Two animals are given successively increasing doses of the test substance on alternate days until an acutely toxic dose or some practical upward limit is reached. This test does not yield a lethality curve and often is used to assess acute toxicity in non-rodents. This test, although more like a short-term, repeated dose toxicity study than a true acute toxicity study, can provide useful preliminary information on the toxic nature of a new material for which no other toxicology information is available.

* Bruce (1985);⁹ Gad *et al.* (1984);¹⁰ Muller and Kley (1982)¹¹

IV C 3. Short-Term Toxicity Tests with Rodents and Non-Rodents

Short-term toxicity tests with rodents are recommended for substances in Concern Level I. Short-term toxicity tests with rodents and non-rodents are generally conducted for 14 or 28 days (one month). Results of these tests (1) can help predict appropriate doses of the test substance for future subchronic or chronic toxicity tests, (2) can be used to determine NOELs for some toxicology endpoints, and (3) allow future tests in rodents and non-rodents to be designed with special emphasis on identified target organs.

Unless specific exceptions are noted below, general recommendations for toxicity studies (see Chapter IV B 1) and for reporting the results of toxicity studies (see Chapter IV B 2) apply to short-term toxicity studies with rodents and non-rodents.

a. Experimental Animals

i. Species and Age

This guideline is for use with rodents (usually rats) and non-rodents (usually dogs); if other species are used, modification of the guideline may be necessary.

Testing should be performed on young and healthy laboratory animals. Dosing of rodents should begin as soon as possible after weaning and acclimation and before the rodents are 6 weeks old. If dogs are used, dosing should begin at 4 to 6 months of age.

ii. Number and Sex

Equal numbers of males and females of each species and strain should be used for the test. For short-term toxicity studies of 30 days duration or less, experimental and control groups should have at least 10 rodents per sex and at least 4 dogs per sex. If the study will be used to determine appropriate doses for longer-term dog studies, but will not be used to determine a NOEL for the test substance, experimental and control groups may have 2 dogs per sex. The number of animals that survive until the end of the study must be sufficient to permit a meaningful evaluation of toxicological effects.

b. Administration of the Test Substance

i. Duration of Testing

Animals should be exposed to the test substance 7 days per week for the duration of the study (from 2 to 4 consecutive weeks).

ii. Number of doses

At least three dose levels of the test substance should be used per sex (one dose level per group); ideally, 4 or 5 dose levels of the test substance should be used. A concurrent control group should be included. Information from acute toxicity studies can help determine appropriate doses for sub-chronic toxicity studies.

IV C 3 Short-Term Toxicity Tests with Rodents and Non-Rodents Continued

c. Observations and Clinical Tests

i. Observations of Test Animals

Food consumption (or water consumption if the test substance is administered in the drinking water) should be measured every week during the short-term toxicity test. Petitioners should also attempt to quantify spillage of food by test animals, and to determine if spillage is greater with test diets than with control diets. Test animals should be weighed at least once a week. Petitioners should use this information to calculate intake of the test substance during each week of the study.

ii. Neurotoxicity Screening

Screening for neurotoxic effects should be routinely carried out in all short-term toxicity studies with rodents (preferably rats) and non-rodents (preferably dogs or miniature swine). The neurotoxicity screen should include: (1) a specific histopathological examination of tissue samples representative of major areas of the brain, spinal cord, and peripheral nervous system (see organs and tissues listed in Chapter IV B 1 e) and (2) a functional battery of quantifiable observations and manipulative tests selected to detect signs of neurological, behavioral, and physiological dysfunctions. References to published literature that can guide the petitioner in selecting an appropriate battery of observations and tests for the neurotoxicity screen are included in Chapter V C.

Reports of short-term toxicity tests should include an assessment of the potential for the test substance to adversely affect the structural or functional integrity of the nervous system. This assessment should evaluate data from the neurotoxicity screen and other toxicity data from the study, as appropriate. Based on this assessment, the petitioner should make an explicit statement about whether or not the test substance presents a potential neurotoxic hazard which requires further neurotoxicity testing. Additional neurotoxicity tests are discussed in Chapter V C but should not be undertaken without first consulting with the Agency.

iii. Clinical Testing

Ophthalmological Examination: This examination should be performed on designated animals before and at the end of the study.

Hematology: For rodents, hematologic tests should be performed on 10 animals of each sex per group before dosing and at the end of the study. For dogs, hematological tests should be performed on all animals in the study before dosing and at the end of the study.

Clinical Chemistry: For rodents, clinical chemistry tests should be performed on 10 animals of each sex in each group before dosing and at the end of the study. For dogs, clinical chemistry tests should be performed on all animals in the study before dosing and at the end of the study.

Urinalyses: Microscopic evaluation of urine sediment and determination of specific gravity of urine samples are recommended before dosing and at the end of the study. For rodents, these tests should be performed on 10 animals of each sex in each group; for dogs, the tests should be performed on all animals in the study.

IV C 3. Short-Term Toxicity Tests with Rodents and
Non-Rodents Continued

Immunotoxicity Screening: Results from tests that are included in the list of primary indicators of immune toxicity (see Chapter V D) should also be evaluated as an immunotoxicity screen.

Reports of short-term toxicity tests should include an assessment of the potential for the test substance to adversely affect the immune system. This assessment should evaluate data from the list of primary indicators included in the immunotoxicity screen and other toxicity data from the study, as appropriate. Based on this assessment, the petitioner should make an explicit statement about whether or not the test substance presents a potential immunotoxic hazard which requires further immunotoxicity testing. Additional immunotoxicity tests are discussed in Chapter V D but should not be undertaken without first consulting with the Agency.

d. Necropsy and Histopathology Examination

See Chapter IV B 1 e for appropriate tissues and organs.

IV C 4. Subchronic Toxicity Tests with Rodents and Non-Rodents

Subchronic toxicity tests with rodents are recommended for substances in Concern Levels II and III; subchronic toxicity tests with non-rodents are recommended for substances in Concern Level II. These tests are generally conducted for 90 days (3 months), but they may be conducted for up to 12 months. Results of subchronic toxicity tests (1) can help predict appropriate doses of the test substance for future chronic toxicity tests, (2) can be used to determine NOELs for some toxicology endpoints, and (3) allow future long-term toxicity tests in rodents and non-rodents to be designed with special emphasis on identified target organs. However, subchronic toxicity tests usually cannot determine the carcinogenic potential of a test substance.

Unless specific exceptions are noted below, general recommendations for toxicity studies (see Chapter IV B 1) and for reporting the results of toxicity studies (see Chapter IV B 2) apply to subchronic toxicity tests with rodents and non-rodents.

a. Experimental Animals

i. Species and Age

The guideline is for use with rodents (usually rats) and non-rodents (usually dogs); if other species are used, modification of the guideline may be necessary.

Testing should be performed on young laboratory animals. Dosing of rodents should begin as soon as possible after weaning and acclimation, and before they are 6 weeks old. If dogs are used, dosing should begin at 4 to 6 months of age.

ii. Number and Sex

Equal numbers of males and females of each species and strain should be used for the test. At the beginning of the test, experimental and control groups should have at least 4 dogs per sex and at least 20 rodents per sex. These recommendations will help ensure that the study can provide a meaningful evaluation of toxicological effects.

If interim necropsies are planned, the number of animals per sex per group should be increased by the number scheduled to be killed before completion of the study; for rodents, at least 10 animals per sex per group should be available for interim necropsy.

b. Administration of the Test Substance

i. Duration of Testing

Animals should be exposed to the test substance 7 days per week for at least 90 consecutive days (3 months).

ii. Dosed Groups

At least three dose levels of the test substance should be used (one dose level per group per sex). Information from acute (see Chapter IV C 2) and short-term (see Chapter IV C 3) toxicity studies can help determine appropriate doses for subchronic studies.

IV C 4 Subchronic Toxicity Tests with Rodents and Non-rodents Continued

c. Observations and Clinical Tests

i. Observations of Test Animals

Food consumption (or water consumption if the test substance is administered in the drinking water) should be measured every week during the subchronic toxicity test. Petitioners should also attempt to quantify spillage of food by test animals, and to determine if spillage of test diets is greater than spillage of control diets. Test animals should be weighed at least once a week. Petitioners should use this information to calculate intake of the test substance during each week of the study.

ii. Neurotoxicity Screening

Screening for neurotoxic effects should be routinely carried out in all subchronic toxicity studies with rodents (preferably rats) and non-rodents if appropriate tests are available. The neurotoxicity screen should include: (1) a specific histopathological examination of tissue samples representative of major areas of the brain, spinal cord, and peripheral nervous system (see organs and tissues listed in Chapter IV B 1 e) and (2) a functional battery of quantifiable observations and manipulative tests selected to detect signs of neurological, behavioral, and physiological dysfunctions. References to published literature that can guide the petitioner in selecting an appropriate battery of observations and tests for the neurotoxicity screen are included in Chapter V C.

Reports of subchronic toxicity tests should include an assessment of the potential for the test substance to adversely affect the structural or functional integrity of the nervous system. This assessment should evaluate data from the neurotoxicity screen and other toxicity data from the study, as appropriate. Based on this assessment, the petitioner should make an explicit statement about whether or not the test substance presents a potential neurotoxic hazard which requires further neurotoxicity testing. Additional neurotoxicity tests are discussed in Chapter V C but should not be undertaken without first consulting with the Agency.

iii. Clinical Testing

Ophthalmological Examination: This examination should be performed on designated animals before and at the end of the study.

Hematology: For rodents, hematology tests should be performed on at least 10 animals per sex in each group before dosing, on days 30 and 60, and at the end of the study. For dogs, hematology determinations should be made on all animals in the study before dosing, on days 30 and 60, and at the end of the study.

Clinical Chemistry: For rodents, clinical chemistry tests should be performed on at least 10 animals per sex in each group before dosing, on days 30 and 60, and at the end of the study. For dogs, clinical chemistry tests should be performed on all animals in the study before dosing, on days 30 and 60, and at the end of the study.

Urinalyses: Microscopic evaluation of urine sediment and determination of specific gravity are recommended before dosing, at 30 and 60 days, and at the end of the study. For rodents, these tests should be performed on at least 10 animals of each sex in each group; for dogs, the tests should be performed on all animals in the study.

IV C 4 Subchronic Toxicity Tests with Rodents
and Non-rodents Continued

Immunotoxicity Screening: Results from tests that are included in the list of primary indicators of immune toxicity (see Chapter V D) should also be evaluated as an immunotoxicity screen.

Reports of subchronic toxicity tests should include an assessment of the potential for the test substance to adversely affect the immune system. This assessment should evaluate data from the list of primary indicators included in the immunotoxicity screen and other toxicity data from the study, as appropriate. Based on this assessment, the petitioner should make an explicit statement about whether or not the test substance presents a potential immunotoxic hazard which requires further immunotoxicity testing. Additional immunotoxicity tests are discussed in Chapter V D but should not be undertaken without first consulting with the Agency.

d. Necropsy and Histopathology Examination

See Chapter IV B 1 e for appropriate tissues and organs.

IV C 5. One-Year Toxicity Tests with Non-Rodents

Long-term one-year toxicity tests with non-rodents are recommended for substances in Concern Level III and should be conducted for a minimum of 12 months (one year). Results of these tests can be used to (1) characterize the toxicity of the test substance in non-rodents and (2) determine the dose of the test substance that produces no observed adverse effects (NOEL or NOAEL). One-year toxicity tests are not conducted for the purpose of assessing carcinogenicity, although data from these tests may reveal information about the carcinogenicity of the test substance.

The following guideline is written for dogs; if other non-rodents are used, modifications to the guideline may be necessary. Unless specific exceptions are noted below, general recommendations for toxicity studies (see Chapter IV B 1) and for reporting the results of toxicity studies (see Chapter IV B 2) apply to one-year toxicity tests with non-rodents.

a. Experimental Animals

i. Age

Dosing of dogs should begin at 4 to 6 months of age, at which time they should have received the appropriate vaccinations.

ii. Number and Sex

Equal numbers of males and females should be used for one-year toxicity studies; at the beginning of the study, experimental and control groups should have at least 4 dogs per sex. If interim necropsies are planned, the total number of dogs of each sex per group should be increased by the number scheduled to be killed before completion of the study. The number of animals that survive until the end of the study should be sufficient to permit a meaningful evaluation of the toxicological effects of the test substance.

b. Administration of the Test Substance

i. Duration of Testing

Animals should be exposed to the test substance 7 days per week for at least 52 weeks (one year).

ii. Dosed Groups

At least three dose levels should be used (one dose level per group per sex). Information from 90-day toxicity studies in non-rodents can help determine appropriate doses for the one-year toxicity study in non-rodents (see Chapter IV C 4).

IV C 5 One-Year Toxicity Tests with Non-Rodents Continued

c. Observations and Clinical Tests

i. Observation of Test Animals

Food consumption (or water consumption if the test substance is administered in the drinking water) should be measured every week during the one-year toxicity test. Petitioners should also attempt to quantify spillage of food by test animals, and to determine if spillage of test diets is greater than spillage of control diets. Test animals should be weighed at least once a week. Petitioners should use this information to calculate intake of the test substance during each week of the study.

ii. Clinical Testing

Ophthalmological Examination: This examination should be performed on designated animals at the beginning of the study, every three months thereafter, and at the end of the study.

Hematology: Hematology tests should be conducted on all animals before dosing begins, at 3-month intervals during the study, and at the end of the study.

Clinical Chemistry: Clinical chemistry tests should be conducted on all animals in the study before dosing begins, at 3-month intervals thereafter, and at the end of the study.

Urinalyses: Microscopic evaluation of urine sediment and determination of specific gravity of urine samples are recommended before dosing, at three month intervals during the study, and at the end of the study. These tests should be performed on all animals in the study.

d. Necropsy and Histopathology Examination

See Chapter IV B 1 e for appropriate tissues and organs.

IV C 6. Carcinogenicity Studies with Rodents

Carcinogenicity studies (bioassays) in two rodent species (usually rats and mice) are recommended for substances in Concern Level III. One of the carcinogenicity studies (preferably in rats) should be combined with a chronic toxicity study (see Chapter IV C 7); one of the carcinogenicity studies (also preferably in rats) should include an *in utero* exposure phase (see Chapter IV C 8). These studies are designed to determine whether a substance possesses carcinogenic activity when administered to rodents in regularly repeated oral doses for the major portion of the lifetime of the test animal. For additional information on carcinogenicity studies the Agency refers the petitioner to several recent reviews.*

Unless specific exceptions are noted below, general recommendations for toxicity studies (see Chapter IV B 1) and for reporting the results of toxicity studies (see Chapter IV B 2) apply to carcinogenicity studies with rodents.

a. Experimental Animals

i. Age

In carcinogenicity studies without *in utero* exposure, dosing of rodents should begin as soon as possible after weaning and acclimation, and before they are 6 weeks old. In carcinogenicity studies with *in utero* exposure, dosing of rodents should begin at weaning.

ii. Species and Strain

In selecting rodent species and strains for carcinogenicity studies, it is important to consider the test animals' general sensitivity to carcinogenic chemicals and the responsiveness of particular organs and tissues of test animals to carcinogens. Preference should generally be given to species and strains with low incidences of spontaneous tumors.

At this time, there is no scientific basis for selecting among inbred, out-bred, or hybrid rodent strains for carcinogenicity studies. Instead, the important consideration is that test animals come from well-characterized and healthy colonies. A thorough understanding of the normal patterns of tumor development (background tumor incidence) throughout the lifespan of untreated test animals (historical and concurrent controls) is critical to the evaluation of the results of carcinogenicity bioassays. It should be noted that strains that are not inbred often have unpredictable background tumor incidences. Because recent information suggests survivability problems exist for some strains of rats (see Chapter IV C 6 a), test animals should be selected that are likely to achieve the recommended duration of this study.

iii. Number and Sex

Ideally, experimental and control groups should have a sufficient number of animals at the beginning of the study to ensure that at least 25 rodents per sex per group survive to the end of the study.

* Apostolou (1990);¹ Clayson and Clegg (1991);² Goodman and Wilson (1991);³ Parodi *et al.* (1991);⁴ Parry (1992);⁵ Pecerera (1991);⁶ Tomatis *et al.* (1992);⁷ Travis (1988);⁸ Travis *et al.* (1991);⁹ Vainio and Cardis (1992)¹⁰

IV C 6 Carcinogenicity Studies with Rodents^a Continued

However, an issue that has attracted considerable recent attention concerns the proportion of test animals surviving until the end of chronic studies, generally referred to as survivorship. Many toxicological guidelines (including the 1982 edition of these guidelines) have standards for valid negative carcinogenic bioassays that require 50% survival of rats until 24 months of age.^a This standard helps assure regulatory agencies that, when a substance is tested for carcinogenicity, it is tested for a sufficient period of time so that the tumorigenic potential of the substance can be adequately assessed from the results of the study.

Until recently, there was little or no indication that commonly used rat strains presented any problem in meeting the guidelines for survivorship. Within the past year or so, however, industry^b and the National Toxicology Program^c report difficulty in reliably achieving 50% survival at 24 months. It is not known whether the sensitivity of rats to chemical carcinogens has changed as survivorship has decreased.

FDA will be closely watching developments in this area of toxicity testing. If this is a continuing trend across time rather than a short-term problem, serious consideration will have to be given to developing means of addressing this problem. In that case, possible future recommendations for increasing survivorship include recommending diets for chronic studies that promote longevity (see Chapter IV C 6 a), recommending dietary restriction, and advising animal breeders to include adequate longevity as one characteristic for selecting future generations of rats (many breeders mainly select for fecundity and rapid growth in their breeding stocks).

Because survivorship of rats continues to change, FDA guidelines no longer require 50% survival (25/50 animals per sex per group) for carcinogenicity bioassays. However, the Agency recommends that petitioners carefully consider their choice of rat strains for carcinogenicity bioassays, since some strains have more serious problems with survivorship than other strains. FDA recommends that carcinogenicity studies begin with at least 50 animals per sex per group; the petitioner is encouraged to begin bioassays with more than 50 animals per sex per group if survivorship is expected to be a problem with the rat strain under study. If fewer than 25 animals per sex per group are expected to survive to the end of the study (24 months), petitioners should take particular care to ensure and document early detection of dead animals through attentive and frequent cage-side observations, thus minimizing the loss of animals to the study through autolysis. In addition, petitioners should consult with toxicologists and statisticians in the Agency as soon as a problem with survivorship in a carcinogenicity study is noticed.

If interim necropsies are planned, the total number of rodents of each sex per group should be increased by the number scheduled to be killed before completion of the study; at least 10 rodents per sex per group should be available for interim necropsy.

^a FDA (1982)¹¹

^b Burek (1990)¹²

^c Rao (1990)¹³

IV C 6 Carcinogenicity Studies with Rodents Continued

b. Administration of the Test Substance

i. Duration of Testing

Rats, mice and hamsters should be exposed to the test substance 7 days per week for 104 consecutive weeks (two years). If an *in utero* phase is added to this study, duration of dosing should be 104 consecutive weeks (two years) post-weaning.

In general, FDA does not recommend early termination of carcinogenicity studies due to decreased survivorship. Carcinogenicity bioassays should be conducted for a major portion of the test animal's lifetime. While it is desirable to have an optimum number of animals survive to the end of the study, the Agency believes there is more benefit, as well as added sensitivity, to be gained by conducting carcinogenicity bioassays for as long as possible, or for the full 24 months that is recommended in these guidelines.

ii. Dosed Groups

Information from subchronic toxicity studies should be used to identify dose levels of the test substance for carcinogenicity studies. At least three dose levels should be used (one dose level per group). No dose used in a carcinogenicity study should cause an incidence of fatalities high enough to prevent meaningful evaluation of the data from the study.

High Dose: The high dose should be the maximum tolerated dose (MTD).

It is not acceptable to select doses for carcinogenicity bioassays based on information unrelated to the toxicity of the test compound. For example, the highest dose in a carcinogenicity study should not be selected so as to provide a pre-determined margin of safety over the maximum expected human exposure to the test substance, assuming that the results of testing at that dose will be negative.

These guidelines recommend that the highest dose in carcinogenicity bioassays should be the MTD. In evaluating the results of carcinogenicity bioassays of direct food additives and color additives used in food, Agency scientists will consider the question of whether the substance was tested at the MTD as one of several factors that may affect interpretation of the results of the bioassay. The final report of the bioassay should include a description of the process used to select the MTD for the study.

The MTD is defined by the National Toxicology Program (NTP) as "that dose which, when given for the duration of the chronic study as the highest dose, will not shorten the treated animals' longevity from any toxic effects other than the induction of neoplasms". The Office of Science and Technology Policy provides similar advice, "The highest dose should be ... consistent with predicted minimal target organ toxicity and normal life span, except as a consequence for the possible induction of cancer." In addition, the NTP cautions that the MTD should not cause morphologic evidence of toxicity of a severity that would interfere with the interpretation of the study.⁴

⁴ National Toxicology Program Board of Scientific Counselors (1984)¹⁴

⁵ Anonymous (1985)¹⁵

⁶ National Toxicology Program Board of Scientific Counselors (1984)¹⁶

IV C 6 Carcinogenicity Studies with Rodents Continued

In general, the MTD is estimated following a careful analysis of data from appropriate subchronic toxicity tests. As the scientific community's experience with toxicity testing has accumulated, the need to consider a broad range of biological information when selecting the MTD has become increasingly clear. For example, data concerning changes in body and organ weight and clinically significant alterations in hematologic, urinary and clinical chemistry measurements, in combination with more definitive toxic, gross or histopathologic endpoints, can be used to estimate the MTD.^a

Although the high dose in a carcinogenicity study should be selected to achieve the MTD, the Agency recognizes that this goal may not always be met.^b There are uncertainties in predicting the MTD for long-term bioassays from the results of pre-chronic studies. Because working definitions of the MTD require the use of scientific judgment, it is sometimes possible for competent investigators looking at the same set of data to arrive at significantly different estimates of the MTD. Such disagreement may be based on different interpretations of the results of metabolic studies or different conclusions about whether an organ alteration is adaptive or toxicological. In situations such as these, when it is unclear what dose of the test substance is the MTD, petitioners should consult with the Agency to determine an appropriate high dose (MTD) for the carcinogenicity bioassay.

The Agency recognizes that use of the MTD in carcinogenicity bioassays has several advantages; these include:

- Compensating for the inherent lack of sensitivity of the bioassay, including the relatively small number of rodents used in the study;
- Providing consistency with other models used in toxicology (high enough doses must be used in order to elicit evidence of the presumed toxicity); and
- Permitting comparison of carcinogenic potencies of substances tested at the MTD, even when the data are collected from different studies.^c

However, the Agency acknowledges that its recommendation to conduct carcinogenicity studies at the MTD may result in the use of doses that are so high as to be unrepresentative of the toxicity of the test substance at lower doses in animals or humans (for example, excessively high doses of a test substance can saturate enzyme systems involved in detoxification of the test substance). Although other approaches to selecting the maximum doses for carcinogenicity studies are under consideration at the Agency, at the present time there is no acceptable alternative to the use of the MTD for the highest dose in these studies.

Low Dose: The low dose level should not interfere with the normal growth, development, and lifespan of test animals, nor should it produce any other signs of toxicity. In general, the low dose should not be less than 10% of the high dose.

Intermediate Dose: The exact dose selected as the intermediate dose may depend on the pharmacokinetic properties of the test substance.

^a Schwetz (1983)¹⁷

^b Schwetz (1983)¹⁷

^c McConnell (1989)¹⁸

IV C 6 Carcinogenicity Studies with Rodents Continued

Optional Fourth Dose Level: If significant differences exist in the pharmacokinetic or metabolic profiles of the test substance administered at high and low doses, an optional (fourth) dose level may be included in the study. This dose level should be the highest dose that produces a pharmacokinetic or metabolic profile similar to profiles obtained at lower doses. The number of test animals in the optional group should be selected to provide approximately the same sensitivity for the detection of the carcinogenic effects of the test substance as the high-dose group provides.

c. Observations and Clinical Tests

i. Observations of Test Animals

Body weight should be recorded weekly for all test animals throughout the study. Food consumption (or water consumption if the test substance is administered in the drinking water) should be measured every week during the carcinogenicity study; petitioners also should attempt to quantify spillage of food by test animals. Petitioners should use this information to calculate intake of the test substance for each week of the carcinogenicity study.

ii. Clinical Testing

Hematology: Erythrocyte counts and total and differential leukocyte counts for all test animals should be made before dosing, at 3, 6, 12 and 18 months during the study, and immediately prior to terminal necropsy.

Clinical Chemistry: Clinical chemistry tests should be performed on at least 10 animals per sex in each group before dosing, at 3, 6, 12, and 18 months during the study, and at the end of the study.

Urinalyses: Microscopic analysis of urine sediment and determination of specific gravity of urine samples are recommended before dosing, at 3, 6, 12, and 18 months during the study, and at the end of the study. These tests should be performed on at least 10 animals of each sex in each group of the study.

d. Necropsy and Histopathology Examination

See Chapter IV B 1 e for appropriate tissues and organs.

IV C 7. Combined Chronic Toxicity/Carcinogenicity Studies with Rodents

Carcinogenicity studies (bioassays) in two rodent species (usually rats and mice) are recommended for substances in Concern Level III (see Chapter IV C 6). One of the carcinogenicity studies (preferably in rats) should be combined with a chronic rodent toxicity study into a single, effective long-term study. Guidelines for the combined study are described in this chapter.

The Agency acknowledges that it is sometimes difficult to set appropriate dose levels for a combined chronic toxicity/carcinogenicity study with rodents. However, when pre-chronic studies permit reasonable estimates of toxicity in longer-term studies, the combined approach is recommended.

Unless specific exceptions are noted below, general recommendations for toxicity studies (see Chapter IV B 1) and for reporting the results of toxicity studies (see Chapter IV B 2) apply to combined chronic toxicity/carcinogenicity studies with rodents.

a. Experimental Animals

i. Age

In carcinogenicity studies without *in utero* exposure, dosing of rodents should begin as soon as possible after weaning and acclimation, and before they are 6 weeks old. In carcinogenicity studies with *in utero* exposure, dosing of rodents should begin at weaning.

ii. Species and Strains

In selecting rodent species and strains for combined chronic toxicity and carcinogenicity studies, it is important to consider the test animals' general sensitivity to carcinogenic chemicals and the responsiveness of particular organs and tissues of test animals to carcinogenic stimuli. Preference should generally be given to species and strains with low incidences of spontaneous tumors.

At this time, there is no scientific basis for selecting among inbred, out-bred or hybrid rodent strains for carcinogenicity studies. Instead, the important consideration is that test animals come from well-characterized and healthy colonies. A thorough understanding of the normal patterns of tumor development (background tumor incidence) throughout the lifespan of untreated test animals (historical and concurrent controls) is critical to the evaluation of the results of combined chronic toxicity and carcinogenicity studies in rodents. It should be noted that strains that are not inbred often have unpredictable background tumor incidences. Because recent information suggests there is decreased survivability for some strains of rats (see Chapter IV C 6 a), test animals should be selected that are likely to achieve the recommended duration of this study.

Rats generally are used for combined chronic toxicity and carcinogenicity studies; however, other rodent species may be used. If possible, the strain selected for this study should be susceptible to the carcinogenic or toxic effects of the class of substances to which the test compound belongs, unless the background tumor incidence in that strain is so high that a meaningful assessment of the effects of the test substance could not be made.

IV C 7 Combined Chronic Toxicity/Carcinogenicity Studies with Rodents Continued

iii. Number and Sex

Ideally experimental and control groups should have a sufficient number of animals at the beginning of the experiment to ensure that at least 25 rodents per sex per group survive to the end of the study. [Additional information on the subject of survivorship is contained in Chapter IV C 6 a.]

Satellite groups of test animals should be used to evaluate the chronic toxicity of the test substance; satellite experimental and control groups should consist of at least 10 rodents per sex.

If interim necropsies (other than those involving satellite animals) are planned, the total number of rodents of each sex per group should be increased by the number scheduled to be killed before completion of the study; at least 10 rodents per sex per group should be available for interim necropsy.

b. Administration of the Test Substance

i. Duration of Testing

Animals should be exposed to the test substance 7 days per week for 104 consecutive weeks (two years). If an *in utero* phase is added to this study, duration of dosing should be 104 consecutive weeks (two years) post-weaning.

Satellite groups of test and control animals used to assess the chronic toxicity of the test substance should be retained in the study for at least 12 months (one year).

ii. Dosed Groups

Information from subchronic toxicity studies should be used to identify dose levels of the test substance for combined chronic toxicity and carcinogenicity studies. At least three dose levels should be used (one dose level per group). No dose used in these studies should cause an incidence of fatalities high enough to prevent meaningful evaluation of the data from the studies.

a) Assessment of the Carcinogenicity of the Test Substance:

High Dose: The high dose should be the maximum tolerated dose (MTD). [Additional information on the selection of the MTD is discussed in Chapter IV C 6 b.]

Low Dose: The low dose level should not interfere with normal growth, development, and lifespan of test animals, nor should it produce any other signs of toxicity. In general, the low dose should not be less than 10% of the high dose.

Intermediate Dose: The intermediate dose level should be between the high and low doses of the test substance. The exact dose chosen as the intermediate dose may depend on the pharmacokinetic properties of the test substance.

IV C 7 Combined Chronic Toxicity/Carcinogenicity Studies with Rodents Continued

Optional Fourth Dose: If significant differences exist in the pharmacokinetic or metabolic profiles of the test substance administered at high and low doses, an optional (fourth) dose level may be included in the study. This dose level should be the highest dose that produces a pharmacokinetic or metabolic profile similar to profiles obtained at low doses. The number of test animals in the optional group should be selected to provide approximately the same sensitivity for the detection of carcinogenic effects as the high-dose group provides.

b) Assessment of the Chronic Toxicity of the Test Substance: Satellite control and dosed groups are included in the study to assess the chronic toxicity of the test substance. The highest dose for satellite animals should produce toxicity so that a toxicological profile of the test substance can be obtained. The lowest dose level for satellite animals should not cause any toxicity.

c. Observations and Clinical Tests

i. Observations of Test Animals

Body weight should be recorded weekly for all test animals throughout the study. Food consumption (or water consumption if the test substance is administered in the drinking water) should be measured every week during the combined chronic toxicity/carcinogenicity study; petitioners also should attempt to quantify spillage of food by test animals. Petitioners should use this information to calculate intake of the test substance for each week of the combined study.

ii. Clinical Testing

Ophthalmological Examination: This examination should be performed on all animals at the beginning of the study, every three months thereafter, and at the end of the study.

Hematology: Erythrocyte counts and total and differential leukocyte counts for all test animals in the principal experimental and control groups should be made before dosing, at 3, 6, 12, and 18 months during the study, and immediately prior to terminal necropsy.

Hematology tests also should be conducted on all rodents in the satellite groups of experimental and control animals. Hematology samples should be taken before dosing, at 3-month intervals during the study, and immediately before interim necropsy.

Clinical Chemistry: Clinical chemistry tests should be performed on at least 10 animals per sex in each principal experimental and control group before dosing, at 3, 6, 12, and 18 months during the study, and at the end of the study.

Clinical chemistry tests also should be conducted on all rodents in the satellite groups of experimental and control animals. Blood samples should be taken before dosing, at 3-month intervals during the study, and immediately before interim necropsy.

Urinalyses: Microscopic analysis of urine sediment and determination of specific gravity of urine samples are recommended before dosing, at 3, 6, 9, 12, and 18 months during the study, and at the end of the study. These tests should be performed on at least 10 animals of each sex in each principal experimental and control group in the study.

**IV C 7 Combined Chronic Toxicity/Carcinogenicity Studies
with Rodents Continued**

Urinalyses should also be conducted on all rodents in the satellite groups of experimental and control animals. Urine samples should be collected before dosing, at 3-month intervals during the study, and immediately before interim necropsy.

d. Necropsy and Histopathology Examination

See Chapter IV B 1 e for appropriate tissues and organs.

IV C 8. *In Utero* Exposure Phase for Addition to Carcinogenicity Study with Rodents

An *in utero* exposure phase should be added to one of two recommended carcinogenicity studies with rodents (see Chapters IV C 6 and 7). In general, the *in utero* phase should be added to the carcinogenicity bioassay with rats, because the rat is the recommended species for reproduction studies (see Chapter IV C 9) and the Agency has a larger database on carcinogenicity bioassays with *in utero* exposure in rats than in mice. The Agency recommends including an *in utero* exposure phase in carcinogenicity bioassays for direct food additives and color additives used in food because human fetuses will generally be exposed to these additives during *in utero* development.

a. Experimental Animals

i. Species and Strain Selection

This guideline is for use with the rat or mouse; if other species are used, modifications of this guideline will be necessary. Strains selected should not have low fecundity and should be sensitive to teratogens and embryotoxins.

ii. Age

All test and control parental animals should be weaned and acclimated before treatment begins.

iii. Number

The number of animals per sex recommended in the guideline to which the *in utero* phase is to be added should serve as a guide for determining the number of animals/group for mating. One male and one female per litter is preferred; no more than two males and two females per litter should be included in any group. For example, if the petitioner decides that each group in the combined chronic toxicity/carcinogenicity bioassay should contain 70 animals per sex, at least 70 litters/group should be produced in the *in utero* phase. Thus, for this example the number of parental animals per sex for the *in utero* phase should be sufficient to ensure at least 70 litters per group.

iv. Caging and Animal Maintenance

Animals should be single-caged for this phase, except during mating and lactation. Food and water should be provided *ad libitum*. The animals' diet should meet all nutritional requirements to support pregnancy in the test species. Special attention should be paid to diet composition when the test material itself is a nutrient, because such material may have to be incorporated into the diet at levels which may interfere with normal nutrition. Under these circumstances, an additional control group fed basal diet may be necessary.

IV C 8 *In Utero* Exposure Phase for Addition to Carcinogenicity Study with Rodents Continued

b. Dose Selection, Treatment Period, and Method of Dosing

i. Duration of Treatment

The parental animals (P) should receive the test substance for a minimum of four weeks prior to mating. Exposure should be continued throughout pre-mating, mating, gestation, and lactation until weaning of the F₁ animals.

ii. Route of Administration

The test compound or vehicle should be administered using the route which most closely approximates the pattern of human exposure (diet or drinking water). Oral intubation (gavage) may be appropriate in instances where human exposure is via a bolus dose or when it is essential for the animal to receive a specified amount of the test compound. The use of gavage may also be required when analysis of the agent in the diet is not possible, when the agent is not stable in the diet, or when the agent is not palatable. The maximum volume of solution that can be given by gavage in one dose depends on the test animal's size; for rodents, this should not exceed 1 ml/100 g body weight. If the test substance must be given in divided doses, all doses should be administered within a 6-hour period.

iii. Selection of Dose Levels

In general, the doses selected should be those that are recommended in the guideline to which the *in utero* phase has been added. However, as a result of maternal or fetal toxicity, it may be necessary to use lower doses during the *in utero* phase of chronic feeding studies in order to produce sufficient offspring for the post-weaning phase. Data justifying this protocol modification should be provided; it is strongly recommended that selections of doses for *in utero* phases of chronic feeding studies be based on the results of pilot studies. Results from metabolism and pharmacokinetic studies should also provide guidance in selecting an appropriate dosage regimen.

iv. Mating Procedures

For each mating, one or two females should be placed with one male. The following morning, each female should be examined for the presence of sperm in the vaginal lavage or the presence of a sperm plug. The day when sperm are found is considered day 0 of gestation. Sibling matings should be avoided.

v. Standardizing the Number of Pups per Litter

Standardization of the number of pups per litter through culling is optional. Litters may be standardized to 10 or 8 based on historical litter size for the strain. It is recommended that standardization be performed on postnatal day 4 by reducing all litters of more than 10 (or 8) to 10 (or 8) in a random manner. If possible, the retained litter-mates should consist of equal numbers of males and females; excess males or females should be randomly selected out. Random selection is important to guard against the human tendency to keep the most fit animals in the study.

IV C 8 *In Utero* Exposure Phase for Addition to Carcinogenicity Study with Rodents Continued

vi. Selection of F_1 Animals

One animal per sex per litter should be randomly selected.

c. Clinical Observations

i. Parental Animals

Parental animals should be observed carefully at least twice daily. Relevant behavior changes and all signs of toxicity, including mortality, should be recorded. Dams should be weighed immediately before the first dose of the test compound is administered, and weekly during gestation and lactation.

Optimally, animals should be weighed daily if the test compound is administered by gavage. Weekly measurements of food consumption should be made.

ii. F_1 Animals

These animals should be observed carefully at least twice daily. Observations of general appearance and the presence of dead pups should be recorded. Pups should be counted on days 0 (birth), 4, 7, 14, and 21 of lactation. Pups should be weighed as a litter on days 0 (birth), 4 (before and after culling, if appropriate), 7, and 14, but should be weighed individually on day 21. Number of pups per sex should be recorded on days 4 (before and after culling, if appropriate), 7, and 14; the sex of individual pups should be recorded on day 21.

d. Other Recommendations

i. Termination of P and F_1 Animals not Selected for the Post-Weaning Phase

These animals should be killed after weaning of the F_1 animals. If toxic signs or reproductive toxicity are observed, these animals should be subject to a complete gross necropsy.

ii. Data Reporting

Litter mates should be identified. Other data should be recorded as described for the toxicity test guideline used for the post-weaning phase (see Chapter IV C 9).

IV C 9. Reproduction and Developmental Toxicity Studies

The following guidelines for reproduction and developmental toxicity studies are presented in two parts, the first dealing with reproduction studies and the second with developmental toxicity (teratogenicity) studies. Reproduction studies evaluate adverse effects of agents on the reproductive systems of both males and females as well as the postnatal maturation and reproductive capacity of offspring. In addition, the cumulative effects of the test substance through two or three generations may be evaluated.

Developmental toxicity studies evaluate the effects of test compounds on a developing organism that result from exposure of either parent prior to conception, during prenatal development, or postnatally. The adverse effects are expressed as one or more end points that may be used to evaluate the toxic potential of an agent. The four major manifestations of an effect on the developing organism are: death, structural anomaly, altered or retarded growth, and functional deficiency. For many compounds, these manifestations are related to dosage. While high doses produce death, low doses that permit survival may produce malformed, retarded, or functionally deficient offspring.

a. Guideline for Reproduction Studies

This guideline for reproduction studies is for use with substances given orally to rodents. It is designed to provide information concerning the effects of a test substance on gonadal function, estrous cycles, mating behavior, conception, parturition, neonatal morbidity, mortality, lactation, weaning, and the growth and development of the offspring. The end points evaluated and the indices calculated must provide sufficient information and statistical power to permit the Agency to determine whether the chemical is associated with changes in reproduction and fertility. Additional information is found in the referenced material.*

Two generations, with one litter per generation, are recommended as the minimum reproduction study (see Figure 5). If results of developmental and other toxicity tests indicate that a test compound may be associated with developmental toxicity, the minimum reproduction study should be expanded. This guideline contains optional procedures for inclusion of additional litters per generation, additional generations, a test for teratogenic effects, and reproductive assessment by continuous breeding.

Unless specific exclusions are noted below, general recommendations for toxicity studies (see Chapter IV B 1) apply to reproduction studies. Recommendations that are unique to, or are particularly important for, reproduction studies are listed below.

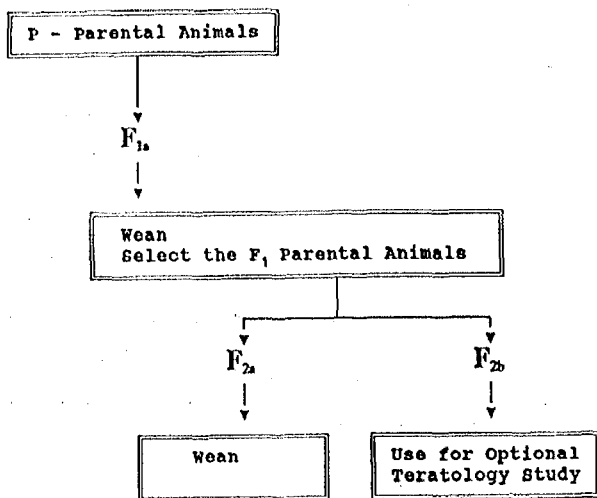
i. Experimental Animals:

Species and Strain Selection: Consideration should be given to the use of the most sensitive species based on the fact that, for the majority of known developmental toxicants, humans are as sensitive or more so than the most sensitive animal species. Because of the length of time and concomitant expense of multi-generation studies, the species selected for a reproduction study should be one that will yield the greatest amount of information per unit cost.

* Collins (1978);¹ Francis and Kimmell (1988);² EPA (1988 a,b);^{3,4} EPA (1991)⁵

Figure 5

2-Generation Reproduction and Teratology Study



IV C 9 Reproduction and Developmental Toxicity Studies Continued

Rodents, such as rats and mice, are usually selected for use in multi-generation studies because they are relatively small animals, gestation time is approximately three weeks, fertility rate is high, ovulation is spontaneous, litters are large enough to allow for inter- and intra-litter comparisons, and the animals are relatively easy to maintain under laboratory conditions. Strains with low fecundity should not be used.

Pedigrees for animals used in reproductive studies should be obtained from the supplier; parents (P or F_0) of the first generation animals should not be litter mates.

Number, Sex, and Age: All test and control animals should be weaned and acclimated to the study conditions before treatment begins. Each test and control group of animals should start with a number of animals sufficient to contain at least 20 males and 20 pregnant females near term. In order to achieve this number, it is usually necessary to start with 30 animals per sex per group in the first parental group (P or F_0) and 25 animals per sex per group in the parents (F_1) of the second generation. If a third generation is to be included in the study, there should be 25 animals per sex per group in the parents (F_2) of the third generation.

ii. Dose Selection, Treatment Period, and Method of Dosing:

Dose Selection: Several doses of the test compound should be used to facilitate the separation of dose-related responses from experimental variation. A minimum of three dose levels should be tested: a high dose, a low dose, and an intermediate dose. The high dose should produce some maternal toxicity (such as reduced body weight or weight gain) but no more than 10% maternal mortality. The lowest dose should not induce observable adverse maternal effects. The low dose may be a dose to which humans are expected to be exposed or a dose that gives measurable tissue levels but no measurable toxicity. Because the effects often vary linearly when plotted against the logarithm of the dose, the intermediate dose(s) should be evenly spaced, on a logarithmic scale, between the high and low doses.

Duration of Testing: Animals should be exposed to the test substance during the entire study. Exposure to the agent typically begins when the rats are 5 to 8 weeks of age. Generally, the first parental females (P or F_0) are exposed during two estrous cycles (two weeks before mating), through mating and pregnancy, to the weaning of the F_1 litter. Males of the first parental group should be dosed for at least one complete spermatogenic cycle (8-11 weeks) before mating and throughout the mating period, in order to detect adverse effects on spermatogenesis by the test substance. Litters (usually F_1 and F_2) should be exposed from the prenatal period throughout their entire postnatal lives. If a third generation is planned, these litters also should be exposed from the prenatal period throughout their entire lives.

Route of Administration: The test substance may be administered to rodents in the diet, by stomach tube (gavage), or in drinking water; the same method of administration should be used for all test animals throughout the study.

Mating Procedures: For each mating, one female should be placed with a male from the same dose group until pregnancy occurs or three weeks have elapsed. Each morning, all females should be examined for the presence of sperm in the vaginal lavage or the presence of a vaginal plug; the day when semen is confirmed is considered day 0 of gestation. Near parturition, pregnant females should be caged separately in delivery or maternity cages and may be provided with nesting materials.

IV C 9 Reproduction and Developmental Toxicity Studies Continued

For mating F_1 animals, a maximum of 2 males and 2 females are selected from each litter. The animals should be selected by random procedures or on the basis of the mean litter weight for each sex. Each F_1 female is mated with one F_1 male from the same dose group until pregnancy occurs or 3 weeks have elapsed. Sibling matings should be avoided. F_1 males and females not selected for mating should be sacrificed upon weaning.

If there is any indication that the test compound may be a reproductive toxicant in males, provisions for cross-mating treated males with untreated females should be made before the study begins.

Standardizing the Number of Pups per Litter: Standardization of the number of pups per litter through culling is optional. Litters may be standardized to 10 or 8 based on historical litter size for the strain. It is recommended that standardization be performed on postnatal day 4 by reducing all litters of more than 10 (or 8) to 10 (or 8) in a random manner. If possible, the retained litter-mates should consist of equal numbers of males and females; excess males or females should be randomly selected out. Random selection is important to guard against the human tendency to keep the most fit animals in the study.

Control Group(s): A concurrent control group is required. Control animals should be housed, fed, and handled the same as dosed animals and should be caged to preclude airborne or other contamination by the test substance.

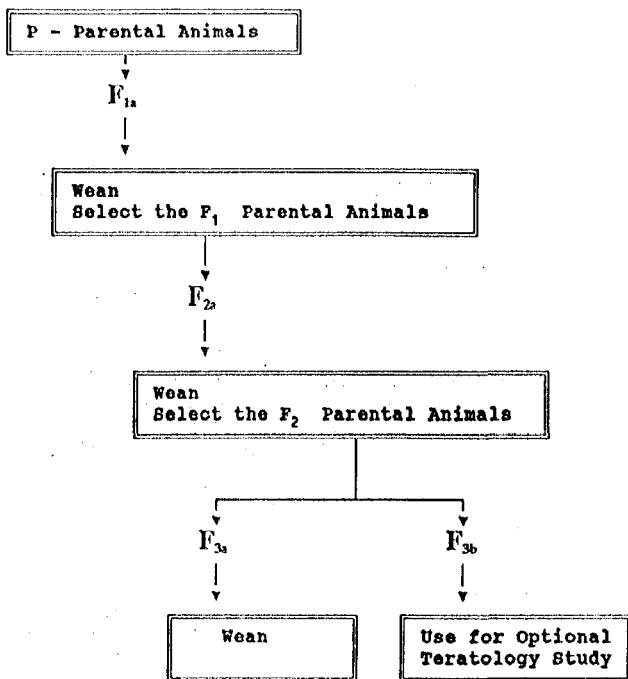
For dietary studies, the control group should be fed the basal diet. When a carrier vehicle for the test substance is used, the vehicle should be given to control rats at a volume equal to the maximum amount of vehicle given to any dosed group. If there is insufficient information on the toxic and carcinogenic properties of the vehicle used to administer the test substance, an additional control group that is not exposed to the vehicle should be included in the study.

Dosed Groups: Several doses of the test compound should be used to facilitate the separation of dose-related responses from experimental variation. A minimum of three dose levels should be tested: a high dose, a low dose, and an intermediate dose. The high dose should produce some maternal toxicity (such as reduced body weight or weight gain) but no more than 10% maternal mortality. The low dose should not induce observable adverse maternal effects. The low dose may be a dose to which humans are expected to be exposed or a dose that gives measurable tissue levels but no measurable toxicity. Because the effects often vary linearly when plotted against the logarithm of the dose, the intermediate dose(s) should be evenly spaced, on a logarithmic scale, between the high and low doses.

Optional Third Generation: If overt effects of a test substance on offspring are observed during the two-generation reproduction study, the study should be extended to a third generation to determine cumulative effects of the substance (see Figure 6 below). Selection of animals for mating and mating procedures for an additional generation should be carried out in the same manner as for the first generation: Randomly mated animals from F_2 should be mated to produce the third generation. F_3 animals are weaned and either necropsied or used for a longer-term toxicity study. F_3 animals are produced in the same manner as F_2 animals.

Figure 6

3-Generation Reproduction and Teratology Study



IV C 9 Reproduction and Developmental Toxicity Studies Continued

Optional Teratology Phase: Either the F_0 or the F_1 litter can be used to determine fetotoxic effects of the test substance. If a teratology phase is to be performed, pregnancy should be timed by the vaginal smear method. Approximately 24 hours prior to delivery, the dams are killed and caesarian sections are performed. The uterus is opened and examined for the presence of early and late deaths; corpora lutea are counted. The live fetuses are removed, weighed, sexed, and examined for gross malformations. To discover visceral abnormalities, half of the fetuses should be freshly dissected^a or should be preserved in Bouin's solution and sectioned by the Wilson technique.^b The remaining half of the fetuses should be stained for the detection of skeletal anomalies.

Optional Reproductive Assessment by Continuous Breeding: A test of fertility and total reproductive capacity that may be recommended in certain situations is that of fertility assessment by continuous breeding (see Figure 7 below). In this procedure, one male and one female per cage are housed as breeding pairs after one week of exposure to the test compound. The test compound is administered throughout the duration of the study, as it is in the multi-generation test. The offspring are removed from the cage when delivery has been completed, and the females can be re-impregnated immediately. Breeding continues for 14 weeks, after which time the pairs are separated for up to 3 weeks, and pregnant females are allowed to deliver their final litter. The offspring from the last litter are generally kept and bred once for evaluation of the second generation. The first litter of each mating pair also may be saved to ensure that a sufficient number of animals exist to perform follow-up tests in case fertility decreases in the parental animals. In that case, cross-over mating may be performed where treated animals are mated with untreated control animals in order to determine the affected sex. Treatment is discontinued during the cross-over mating period and resumed 7-days later. Parental animals are necropsied if an effect on that sex is observed. Continuous breeding permits the evaluation of approximately five litters per pair. If the compound affects only the early stages of spermatogenesis, toxic effects will be observed only in the last litters of the continuous breeding protocol, because the earliest matings of the males take place with sperm that have not been exposed to the test compound throughout all stages of spermatogenesis. Additional information is found in the referenced material.^c

iii. Clinical Observations and Histopathology Examination

Observations: Each animal should be observed at least twice a day; observation times should be selected to permit detection of the onset and progression of all toxic and pharmacologic effects of the test substance and to minimize the loss of animals and organs/tissues to the study because of management problems.

Individual records should be maintained for each animal. Toxicological and pharmacological symptoms and signs, including behavioral abnormalities, should be recorded daily; records should include the time of onset, duration, and intensity of symptoms and signs.

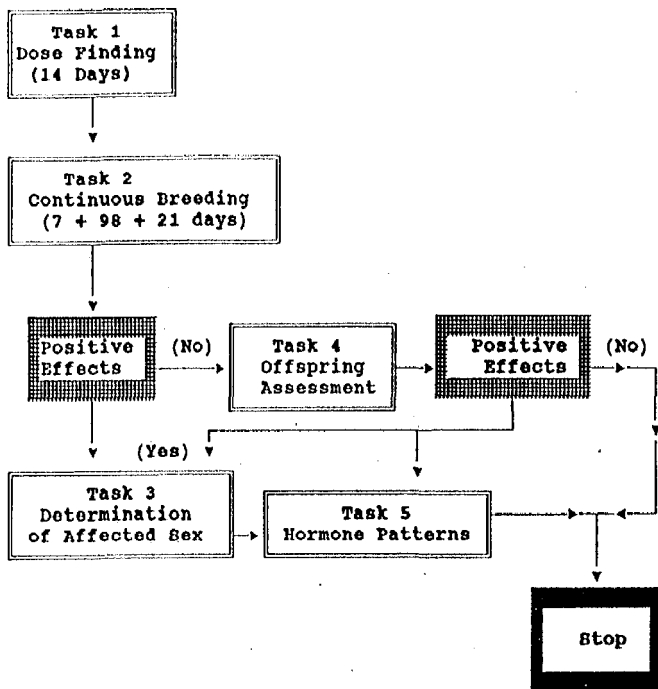
^a Staples (1977);^d Barrow and Taylor (1969)^e

^b Wilson (1965)^f

^c Lamb (1989);^g Reel *et al.* (1985)^h

Figure 7

Continuous Breeding Protocol



IV C 9 Reproduction and Developmental Toxicity Studies Continued

Dams should be weighed immediately before the test compound is administered for the first time, weekly until necropsy, and at necropsy; food and water consumption also should be recorded. For animals in the F₁ and F₂ generations selected for mating, the following determinations should be made at appropriate intervals: White blood cell counts, differential white blood cell counts, and levels of protein, albumin and globulin in serum; evaluations of these parameters are discussed in the section on immunotoxicity testing (see Chapter V D).

Any dam showing signs of imminent abortion or premature delivery should be necropsied on the day such signs are observed. Pregnant females in test and control groups should be allowed to litter normally.

The duration of gestation should be calculated from day 0 of pregnancy. Each litter should be examined as soon as possible after delivery for the number of pups, stillbirths, live births, and the presence of gross anomalies. Dead pups should be preserved and studied for possible defects and cause of death. The neonates should be carefully observed, sexed, and weighed on postnatal days 0 (the day of birth) and 4, and weekly thereafter (postnatal days 7, 14, 21, etc.).

Results from tests that are included in the list of primary indicators of immune toxicity (see Chapter V D) should also be evaluated as an immunotoxicity screen.

Neurotoxicity Screening: Multi-generation reproduction studies provide an excellent vehicle for testing compounds for potential neurotoxic effects. A neurotoxicity screening battery of tests, appropriate to the age of the test animals, should be conducted to detect neurological disorders, behavioral changes, autonomic dysfunctions and other signs of nervous system toxicity (see Chapter V C). Testing should be applied both to the dams and to the developing offspring (pre- and post-weanlings) and conducted at representative intervals throughout the duration of the study. For the neonates, the screen could include recording of physical landmarks of development (such as the appearance of fur, eye opening, genital development and incisor eruption) and functional measures of development (such as development of the righting reflex, the startle response and motility). All daily observations and all data derived from the neurotoxicity screening battery of tests, including positive and negative findings, should be recorded, analyzed using appropriate statistical procedures, and reported.

Gross Necropsy and Microscopic Examination: All test animals should be subjected to complete gross necropsy, including examination of external surfaces, orifices, cranial cavity, carcass, and all organs. The gross necropsy should be performed by, or under the direct supervision of, a qualified pathologist. Ideally, the pathologist who performs or supervises the gross necropsy should preferably perform the histopathological examination.

At necropsy, the uterus should be examined for the presence of implantation sites and resorptions. Reproductive organs, lymphoid tissue and organs (bone marrow, representative lymph nodes, Peyer's patches, spleen, stomach, and thymus) brain, peripheral nerve tissue, and target organs from all animals in a reproduction study should be preserved for future histopathological analysis. Microscopic examination should be made of all organs and tissues that show gross pathological changes.

Brain (at least three different levels), epididymides, ovaries, peripheral nerve, pituitary, prostate, seminal vesicles, spinal cord (at least two different locations), testes, uterus and vagina should be preserved for all dose levels. Histopathology should be performed on animals in all dose groups and for P and F₁ animals selected for mating. Histopathology on lymphoid organs and tissues should be performed for F₁ and F₂ animals selected for mating, as described in the section on immunotoxicity testing (see Chapter V D). Dead or moribund pups should be examined for defects.

IV.C 9 Reproduction and Developmental Toxicity Studies Continued

Results from necropsy examination and histopathology evaluation that are included in the list of primary indicators of immune toxicity (see Chapter V D) should also be evaluated as an immunotoxicity screen.

iv. Endpoints of Reproductive Toxicity

Because the maternal animal and not the developing organism is the individual treated during gestation, data generally should be calculated as incidence per litter or as number and percent of litters with particular end points. End points of reproductive toxicity are usually expressed as indices that encompass the animals' responses to the test compound from conception to weaning. The following indices should be calculated for each reproduction study: the female fertility index, two gestation indices, the weaning index or the lactation index, the sex ratio, and viability and growth indices at days 4, 7, 14, and 21 following birth. Supplemental end points of male reproductive toxicity may also be assessed if there is evidence of male-mediated effects on developing offspring.

Female Fertility Index: The female fertility index represents the percent of matings that result in pregnancies. It is calculated as follows: $[\text{number of pregnancies/number of matings}] \times 100$. This index reflects the total number of dams that have achieved pregnancy, including those that deliver at term, abort, or have fully resorbed litters. An accurate determination of the index requires a careful evaluation of the uterus at necropsy for the presence of implantation sites and resorptions. In a reproduction study with two litters per generation, calculations of this index should be performed only for the second litter. This index depends on male libido and fertility. Each female may be mated with up to two males; if a question arises about the fertility of the first male, a second male should be mated with the female.

Gestation Index: The gestation index evaluates the efficiency of pregnancy resulting in at least one live offspring. In this index, litters with only one live offspring are counted the same as those with more than one live offspring. The index is calculated as follows: $[\text{number of litters with live pups/number of pregnancies}] \times 100$. A related index, the live-born index, $[\text{number of pups born alive/total number of pups born}] \times 100$, is a measure of the total number of offspring lost, regardless of litter.

Weaning Index: The weaning index represents the ability of pups to survive from day 4 to day 21. It is calculated as follows: $[\text{number of pups alive at day 21/number of pups alive and kept on day 4}] \times 100$. This index corrects for the reduction of pups on day 4. If the pups are not reduced, a related index, the lactation index, is calculated: $[\text{number of pups alive on day 21/number of pups alive on day 4}] \times 100$. Regardless of the etiology, a decrease in the weaning index is considered to be indicative of adverse reproductive effects.

Sex Ratio: Determining the sex of pups at birth and verifying their sex at each weighing permits the relative fitness of each sex to be calculated as the offspring mature. The sex ratio is a particularly important parameter when one sex is expected to be affected by the test substance more than the other sex. This parameter is usually calculated as follows: $[\text{number of males/number of females}]$. The calculation $[\text{number of females or males/total number of animals}] \times 100$ yields the percentage of total animals that are male or female.

IV C 9 Reproduction and Developmental Toxicity Studies Continued

Viability Indices: The viability indices are measures of the offsprings' ability to survive during specific intervals of their lives, from birth (day 0) to day 4, day 4 to day 7, day 7 to day 14, and day 14 to day 21. For example, the day-7 viability index is calculated as follows: $\{\text{number of pups alive on day 7} / \text{number of pups alive and kept on day 4}\} \times 100$. The pups' ability to survive is important in reproduction studies because it may reflect the adequacy of postnatal nourishment, maternal neglect, and postnatal absorption of a toxic substance that is excreted in the mothers' milk. Regardless of etiology, decreases in viability indices are considered to be indicative of adverse reproductive effects.

Growth Indices: Growth indices are measures of the general physiological status of the pups. They are calculated for postnatal days 4, 7, 14, and 21 as follows (illustrated for postnatal day 7): $[\text{average weight of males or females in a test or control group on day 7} / \text{average weight of same animals on day 4}] \times 100$. Growth indices can be compared with viability indices for the same postnatal time: If pups exposed to a test compound show high survival at weaning but have a decreased average weaning weight, that may indicate impairment of nutrition or metabolism by the test compound.

Other Endpoints for Neonates: In special circumstances, for example, if screening data provide evidence of a potential for neurotoxicity, it may be recommended that more quantitative evaluation of the functional integrity of the developing nervous system be carried out. Such an evaluation may include, for example, quantitative measures of motor activity, sensorimotor reactivity, and cognitive function and a neuro-histochemical examination of *in situ* fixed nervous system tissue.

Maternal Toxicity: See the discussion of maternal toxicity and its significance in Chapter IV C 9 b 2.

Reproductive Endpoints for Males: Sperm evaluations may be recommended when reduced fertility rate or other information suggest that the test substance may be a reproductive toxicant in males. A sperm evaluation should include measures of sperm count, sperm motility, and sperm morphology.

Sperm counts from test species may be derived from ejaculated, epididymal, or testicular samples. Ejaculated sperm counts are influenced by several variables, including length of abstinence and the ability to obtain the entire ejaculate. If a pre-exposure baseline can be obtained for each male, then changes during exposure can be better defined. Epididymal sperm evaluations usually use sperm from the cauda portion of the epididymis. If sperm count is expressed on the basis of the weight of the cauda epididymis, absolute sperm count should also be reported in order to provide clarification of declines in sperm number. This is necessary because sperm contribute to the weight of the cauda. Sperm production may also be determined by enumerating elongated spermatid nuclei following homogenization of testes in a detergent-containing medium.

Sperm motility can be useful in identifying the changes that have occurred in the biochemical environment in the testes and epididymis. Motility estimates may be obtained on ejaculated, vas deferens, or cauda epididymal samples. Motility is influenced by many variables, including abstinence, the elapsed time between obtaining the sample and evaluation of motility, and the medium used to dilute the sample. Historical measures of motility have been obtained by using subjective, microscopic techniques. Recently introduced automated image-analysis techniques are more objective, provide a permanent record, and allow additional data to be obtained, such as swimming speed and swimming patterns.

IV C 9 Reproduction and Developmental Toxicity Studies Continued

Sperm morphology refers to the structural aspects of sperm. In most studies on test animals, only the head shape has been evaluated, but additional information may be gained from an evaluation of midpiece and tail morphology. The traditional approach to characterizing morphology has relied on subjective categorization of sperm shape from examination of stained slides. Individual sperm morphology profiles are stable over time. An increase in abnormal sperm morphology should be considered supportive evidence that the test compound has gained access to the germ cells.

v. Reporting the Results of Reproduction Studies

Reports of all reproduction studies should contain, in addition to the information required by the Good Laboratory Practice Regulations (see Chapter IV B 1 a), absolute values for all parameters, complete data (individual pups) and tables of data summarized and analyzed by litter. All major indices, discussed in the previous section, should be calculated. The dosage rate of test substance (doses) should be reported as mg/kg/day (milligrams of test substance per kilogram of body weight per day).

Problems commonly encountered in the review of multi-generation reproduction studies and developmental toxicity studies, include: Insufficient numbers of pregnant animals per control or treatment group, non-random selection procedures, and statistical analyses of data on a per-fetus basis instead of a per-litter basis. Careful consideration of recommended guidelines and the submission of protocols for review by the Agency prior to conducting the studies would help eliminate such problems.

In addition to the various indices in reproduction studies, data should also be examined as average number surviving to a given time period, such as average number surviving to day 4 or average number weaned. This analysis considers the total effect of the compound at all stages to that point and is a more sensitive indicator than each index separately.

b. Guideline for Developmental Toxicity Studies

The purpose of developmental toxicity studies is to provide data that can help determine if a test substance is embryotoxic or teratogenic. Treatment must begin early enough in gestation and continue long enough to include the major portion of organogenesis for the species used. This guideline may be used with substances given orally to the rat, mouse, hamster, and rabbit.

A developmental toxicity study may be combined with a multi-generation reproduction study, or it may be performed as a separate study. When combined with a reproduction study, the teratology assessment is usually performed on the last litter of the last generation, so as to maximize exposure to the test agent. If the test substance is believed to have the capacity to alter the rate of its own metabolism through production of metabolizing enzymes or as a result of damage incurred by the liver, then consideration should be given to evaluating the teratogenic potential of the compound using a separate study. If the results of a developmental toxicity study are positive, a second species may be tested in order to determine if the toxic effects of the test substance are limited to one species. Additional information is found in the referenced material.*

Unless specific exclusions are noted below, general recommendations for toxicity studies (see Chapter IV B 1) apply to developmental toxicity studies. Recommendations that are unique to, or are particularly important for, teratogenicity studies are listed below.

* EPA (1985)¹¹

IV C 9 Reproduction and Developmental Toxicity Studies Continued

i. Experimental Animals

Species and Strain Selection: When pharmacokinetic and metabolic data or other information suggest the most appropriate species for developmental toxicity testing, that species should be used. In the absence of such data, the most sensitive species should be used, based on the fact that for the majority of known human developmental toxicants, humans are as sensitive or more so than the most sensitive animal species. Commonly used species are the rat, mouse, hamster, and rabbit; preferred species are the rat and rabbit. Strains selected should not have low fecundity and should be sensitive to teratogens and embryotoxins.

Animals should be single-caged for this test, except during mating. Food and water should be provided *ad libitum*. The animals' diet should meet all nutritional requirements to support pregnancy in the test species. Special attention should be paid to diet composition when the test material itself is a nutrient, because such material may have to be incorporated into the diet at levels which may interfere with normal nutrition. Under these circumstances, an additional control group fed basal diet may be necessary.

Animals should be assigned to test and control groups in a stratified random manner to minimize inter-group weight differences and assure statistical comparability of relevant variables.

Number, Sex, and Age: All test and control animals should be young, mature, pregnant females of uniform age and size.

A sufficient number of females should be used so that each test and control group consists of at least 20 pregnant rats, mice, or hamsters, or 12 pregnant rabbits, at or near term. These are the minimum numbers of pregnant animals for developmental toxicity testing. The objective is to insure that enough litters are produced to permit effective evaluation of the teratogenic potential of the test compound.

ii. Dose Selection, Treatment Period, and Method of Dosing

Duration of Testing: The test substance should be administered daily throughout the treatment period. The minimum treatment period recommended for developmental toxicity studies includes the period of organogenesis of the species used. In rats and mice, this period includes days 6 through 15 of gestation; in hamsters, days 4 through 14; and in rabbits, days 7 through 18. Day 0 of gestation is considered as the day of finding a positive indication of mating. Alternatively, treatment may be extended to include the entire period of gestation, from fertilization to approximately one day before term.

If the developmental toxicity test is being conducted as part of a multi-generation reproduction study, the animals are dosed from before conception until they are necropsied.

Route of Administration: The test compound or vehicle should be administered using the route which most closely approximates the pattern of human exposure (diet or drinking water). Oral intubation (gavage) may be appropriate in instances where human exposure is via a bolus dose or when it is essential for the animal to receive a specified amount of the test compound. The use of gavage may also be required when analysis of the agent in the diet is not possible, when the agent is not stable in the diet, or when the agent is not palatable. The maximum volume of solution that can be given by gavage in one dose depends on the test animal's size; for rodents, this should not exceed 1 ml/100 g body weight. If the test substance must be given in divided doses, all doses should be administered within a 6-hour period.

IV C 9 Reproduction and Developmental Toxicity Studies Continued

Mating Procedures: For each mating, one or two females should be placed with one male. The following morning, each female should be examined for the presence of sperm in the vaginal lavage or the presence of a sperm plug. The day when sperm are found is considered day 0 of gestation (day 0 of gestation in rabbits is the day insemination is performed). Sibling matings should be avoided.

Pilot Study: To select appropriate doses of the test substance, a pilot or trial study is recommended, unless suitable information is available from other studies. It is not always necessary to carry out a trial study in pregnant animals. Comparison of the results from a trial study in non-pregnant animals and a main study in pregnant animals will establish whether the test substance is more toxic in pregnant animals. If a trial study is carried out in pregnant animals, the dose producing embryonic or fetal lethality should be determined.

Control and Dosed Groups: At least three test groups and one control group should be used in the primary developmental toxicity study. When the test substance is administered in a vehicle, the vehicle without the test substance should be administered to the control group. If there are insufficient data on the toxic properties of the vehicle used in administering the test substance, a sham control group should also be included. If no vehicle is used, then the controls should be sham treated. In all other respects, the control must be handled and maintained in a manner identical to that used with the groups given the test substance.

Unless limited by the physical or chemical nature or biological properties of the substance, the highest dose should ideally induce some overt maternal toxicity such as a statistically significant reduction in body weight, and yet still support reproduction. The highest dose should not cause a significant reduction in average litter size as compared to untreated controls, should not cause more than 10 percent maternal deaths, and should not exceed 5 percent of the diet. The low dose level should not induce observable effects attributable to the test substance. The intermediate dose(s) should be located logarithmically between high and low dose levels. The dosage administered by gavage should be based on the individual animal's weekly (preferably, the animal's daily) body weight.

Maternal Toxicity and Its Significance: End points which may serve as indicators of maternal toxicity include mortality, body weight, body weight gain, gestation length, organ weights, food and water consumption, clinical signs of toxicity, gross necropsy data and histopathology. The calculation of a corrected mean maternal weight gain (difference in an initial and terminal maternal body weight less the gravid uterus weight) may also be used as an index of maternal toxicity.

A variety of agents are known to have a selective toxic effect on the male, the female or on the offspring, while other chemicals exhibit a non-specific effect. When mother and offspring are adversely affected by a given agent, it can be very difficult to determine if the developmental toxicity is mediated by maternal toxicity or occurs independently of it. The sensitivity of the maternal system can vary significantly from that of the fetus due to differences in metabolism, distribution and elimination of agents.

At the same time, the response of the fetus can differ markedly from that of the mother, due to the sensitive developmental processes taking place in the offspring that have no counterpart in the adult.

Developmental effects in the absence of maternal toxicity are commonly regarded as the most serious manifestation of toxicity, due to the apparent increased sensitivity of the developing organism, as compared with the adult. When developmental effects are found in the presence of maternal toxicity, the primary cause of the effect is often left open to question. However, since

IV C 9 Reproduction and Developmental Toxicity Studies Continued

there is insufficient evidence to suggest that developmental toxicity is always a secondary toxic effect in the presence of maternal toxicity, a default assumption must be employed. It is assumed that developmental effects which occur in the presence of minimal maternal toxicity are considered to be evidence of developmental toxicity, unless it can be established that the developmental effects are unquestionably secondary to the maternal effects. In situations where developmental effects are observed only at doses where there is a substantial amount of maternal toxicity, then the possible relationship between the maternal toxicity and the developmental effects should be evaluated in order to make a proper assessment regarding the toxicity of a test compound.

The point at which a dose level would become unacceptable for evaluation due to the degree of maternal toxicity would vary on a case-by-case basis and such a determination would require scientific judgement. One generally accepted condition has been greater than 10% maternal mortality at the given dose level.

Clinical Observation and Pathological Examination of Dams, Fetuses, and Neonates:

Throughout the study, each animal should be observed at least twice daily. Relevant behavioral changes and all signs of toxicity, including mortality, should be recorded. Dams should be weighed immediately before the first dose of the test compound is administered (usually, day 6 or 7 of gestation), weekly until necropsy, and at the time of necropsy. Optimally, animals should be weighed daily if the test compound dose is administered by gavage. Weekly measurements of food consumption should be made; fluid consumption should be measured as appropriate. Any dam showing signs of imminent abortion or premature delivery during the study should be necropsied on the date such signs are observed.

The test should be terminated approximately one day before term, when the dams should be necropsied and examined microscopically for structural abnormalities or pathological changes that may have influenced pregnancies. Immediately after the dams are killed, fetuses should be delivered by hysterotomy. Care should be taken to insure that all fetuses (except those sacrificed before the end of the study) are delivered at approximately the same stage of fetal development. The uterus should be removed and the contents examined for embryonic or fetal deaths and for the number of live fetuses. For dead fetuses, it is usually possible to estimate the time of death in utero. In rats and rabbits, the number of corpora lutea should be determined. For dams that do not appear to be pregnant, a sodium sulfide or ammonium sulfide solution may be used to enhance the visibility of resorption sites. Evaluation of the females during cesarean sections and subsequent fetal analyses should be conducted blind in order to minimize unconscious bias.

Following removal from the uterus, each fetus should be examined externally, and all deviations from normal should be noted. The sex of each fetus should be determined. Each fetus should be weighed individually, and the mean fetal weight per sex per group should be calculated. Additional end points may be measured, such as the crown-to-rump distance on each fetus.

For rats, mice, and hamsters, one-half of each litter should be prepared and examined for skeletal anomalies (Alizarin Red single staining or Alizarin Red/Alcian Blue double staining are recommended). The remaining part of the litter should be prepared and examined for soft-tissue anomalies (Wilson sections or fresh-dissection techniques are recommended⁴). The percentage of fetuses designated for skeletal and soft-tissue analyses may be modified if the fresh dissection technique is used or if there is prior knowledge about the effects of the test compound indicating that toxic effects may be best identified by a particular technique. Each rabbit fetus should be carefully dissected and examined for visceral anomalies, then examined for skeletal anomalies.

* Wilson (1965);⁴ Staples (1977);⁴ Barrow and Taylor (1969)³

IV C 9 Reproduction and Developmental Toxicity Studies Continued

Reproductive organs, lymphoid organs and tissues (bone marrow, representative lymph nodes, Peyer's patches, spleen, stomach, and thymus), the brain, peripheral nerve tissue, and target organs from all animals in a developmental toxicity (teratogenicity) study should be preserved for future histopathological analysis, as necessary.

End Points Measured: Because the maternal animal, and not the developing organism, is the individual treated during gestation, data generally should be calculated as incidence per litter or as number and percent of litters with particular end points. Maternal toxicity is useful in assessing the validity of the high-dose level and the possibility that maternal toxicity is involved in subsequent developmental events. Parameters used to measure maternal toxicity include body weight, food and fluid consumption, daily clinical observations, and necropsy data, such as organ weights.

If treatment is given throughout gestation, implantation may be affected. If, however, treatment begins after implantation, conception and implantation rates should be the same in control and treated groups. End points to be measured per litter should include the number of implantations, corpora lutea, live fetuses, dead fetuses, and resorbed fetuses. For litters with live fetuses, mean male and female body weights and the incidence per litter of all divergences from normal fetal development should also be reported.

Analysis of Data: Values from control and test groups of animals should be compared statistically. The following techniques are recommended, but others may be substituted if they are appropriate. Maternal body weights should be compared by analysis of co-variance, adjusting for initial body weight, and then analyzed by protected least significant difference (LSD) tests. Fetal body weights should be evaluated using nested analysis of variance. Anomalies in litters should be compared by Fisher's Exact Test. Fetal survival and incidence of abnormalities per litter are compared by analysis of variance after transforming the data using the Freeman-Tukey Arc-Sine Transformation. When possible, the litter should be considered the statistical unit of measure.

Chapter V

Additional Recommended Studies

A. Introduction

The Agency recognizes that information about metabolism and pharmacokinetics, neurotoxicity, and immunotoxicity are significant endpoints in assessing the safety of direct food additives and color additives used in food. Recommended strategies for improving the ability to determine metabolism and pharmacokinetics and the neurotoxic and immunotoxic potentials of test substances are described in Chapters V B, C, and D, respectively. Because this chapter addresses toxicity studies that are recommended for the first time by FDA for assessing the safety of direct food additives and color additives used in food (see Figure 4, Chapter III C 1), they are discussed in greater detail than other recommended toxicity studies (see Chapter IV C).

1. Metabolism and Pharmacokinetics

FDA believes that data from studies on the adsorption, distribution, metabolism, and excretion of a chemical can provide insight into mechanisms of toxicity and are essential in the design and evaluation of results from other toxicity studies. Such data should be provided for all direct food additives and color additives used in food that are assigned to Concern Levels II or III. Recommendations for obtaining data on the metabolism and pharmacokinetics of these substances are presented in this document. In general, the Agency recommends that this information be obtained before subchronic and chronic toxicity tests are begun.

2. Neurotoxicity

It is recommended that the assessment of neurotoxic potential be carried out according to a process of tiered testing progressing from the identification of chemicals associated with neurotoxic effects (screening), through a characterization of the scope of nervous system involvement (characterization of effects), to the determination of dose response kinetics which includes the definition of the no-observed adverse effect level (dose-response). Screening for neurotoxic effects, which is considered to be one of the most critical steps in this tiered process, should be routinely and systematically carried out in short-term (see Chapter IV C 3), subchronic (see Chapter IV C 4), and reproductive and developmental toxicity (see Chapter IV C 9) studies. The neurotoxicity screen should include a specific histopathological examination of representative tissue samples of all major areas of the brain, spinal cord, and peripheral nervous system in conjunction with a functional evaluation battery of quantifiable observations and manipulative tests selected to detect signs of neurological, behavioral, and physiological dysfunctions. References to published literature that can guide the petitioner in selecting an appropriate neurotoxicity screen are included.

V A Introduction Continued

Study reports should include an integrated assessment of the potential for the test chemical to adversely affect the structural or functional integrity of the nervous system. This assessment should include results of the neurotoxicity screen and other toxicology data, as appropriate. Based on the assessment, an explicit statement should be made as to whether or not the test chemical represents a potential neurotoxic hazard which requires special testing. Recommendations about further neurotoxicity testing, if the results of the initial screens indicate the need for such testing, are included. However we urge petitioners to consult with Center scientists before undertaking additional neurotoxicity tests.

3. Immunotoxicity

An immunotoxicity screen should be routinely carried out in short-term (see Chapter IV C 3), subchronic (see Chapter IV C 4), and reproductive and developmental toxicity studies (see Chapter IV C 9). This screen consists of primary indicators of immunotoxicity described in Chapter V D 3; these indicators are a set of hematological, serum protein, histopathological, and body and organ weight endpoints that are routinely evaluated in standard toxicity tests.

Study reports should include an integrated assessment of the potential for the test chemical to adversely affect the immune system. This assessment should be based on results of the immunotoxicity screen (primary indicators of immunotoxicity) and other toxicology data, as appropriate. Based on the results of this assessment, an explicit statement should be made as to whether or not the test chemical represents a potential immunotoxic hazard which requires additional immunotoxicity testing (see Chapter V D 4 and 5).

If results of the immunotoxicity screen indicate the need for further testing, information that will help the petitioner choose additional immunotoxicity tests is provided. However, we urge petitioners to consult with Center scientists before undertaking additional immunotoxicity tests.

V B. Metabolism and Pharmacokinetic Studies

Results from animal toxicity studies are used by FDA to determine dose-response characteristics for any effects observed in the evaluation of the safety of food and color additives. Since the delivered dose of a substance to any affected tissue or organ is determined by the pharmacokinetics and metabolism of the substance in the test animal, toxicity studies are more easily interpreted, likely to achieve target doses, and avoid excessive toxicity if data from metabolic and pharmacokinetic studies are available during the planning of short-term, subchronic and/or chronic toxicity studies. Early determination of metabolic pathways and the rates of metabolism in different test species may provide explanations for species differences in any effects which are observed, and suggest biochemical or pharmacologic experiments which might be used to test explanations of such phenomena.

The Agency recommends that petitioners submit data that will enable our scientists to evaluate: 1) the extent of absorption, 2) tissue distribution, 3) pathways and rates of metabolism, and 4) rate(s) of elimination of the parent substance and any metabolites formed for all Concern Level II and III substances (see Figure 3). The Agency may recommend submission of additional metabolic and pharmacokinetic data based on the extent to which a chemical is metabolized, the potential toxicity of the metabolites, and the extent to which observed toxic effects seem to correspond to the presence of the parent substance or its metabolites.

1. Considerations in the Design of, Analysis of, and Use of Data from Metabolic and Pharmacokinetic Studies

Pharmacokinetic data can be used to predict plasma concentrations, target tissue doses, and the fate of the administered dose. This information can then help the petitioner and/or the Agency: 1) decide which toxicity studies should be conducted, 2) select doses for chronic toxicity and carcinogenicity studies, 3) determine the mechanism of toxicity and assist in the interpretation of toxicity data, and 4) improve the risk assessment process.

a. Design and Analysis of Metabolic and Pharmacokinetic Studies

Pharmacokinetic studies are most useful when they are performed early in the process of evaluating the toxicity of a chemical. However, additional metabolism and pharmacokinetic studies may be recommended after target organs have been identified in toxicity studies.

Whole animal (oral dosing) studies should be performed to determine gastrointestinal absorption and overall elimination rates for a compound. However, it is often most efficient to perform *in vitro* studies of metabolism before whole animal (oral dosing) studies to determine whether enzyme kinetics may explain known dose response curves or predict non-linear dose response curves. The results of early *in vitro* studies also can be used to optimize the choice of doses in whole animal pharmacokinetic studies.

Additional recommendations concerning the design and analysis of metabolism and pharmacokinetic studies are described below.

V B. Metabolism and Pharmacokinetic Studies Continued

i. Test Compound

In selecting the dosage form of a test compound to be administered in metabolic and pharmacokinetic studies, the chemical characteristics of the compound and its route of administration should be considered. The formulation of the test substance used for metabolic and pharmacokinetic studies should exhibit similar patterns of disintegration and/or dissolution as formulations used for toxicity studies. Chemical purity of the test compound should be established; impurities that may affect absorption, distribution, metabolism and excretion of the test compound should be identified. Stability of the compound in its carrier (i.e., food, water, or solvent) also should be determined. Chemical characteristics of the compound (i.e., low solubility, volatility) may make certain routes of administration impossible. It is critical that the dose absorbed into tissues be determined especially in studies where the test substance is added to the feed or water and is ingested *ad libitum*.

Use of radioactive substances facilitates mass balance determinations because radio-labels are relatively easy to detect in samples of tissues and body fluids. Determining the disposition pattern of the radio-label may be adequate for predicting doses that should be used in toxicity studies where the results of a test animal's overall exposure to the substance (parent compound and metabolites) is of concern. The radio-label should not be biologically labile; when a radioactive element is present at more than one position of the test compound, the radio-label should be uniformly distributed in the molecule.

The radiochemical purity of the test substance (radioactivity actually associated with the compound being tested) is another important consideration. If the test compound is not radiochemically pure and radio-labeled impurities are not identified, and if only the distribution of the radio-label in tissues and body fluids is determined, interpretation of the results may be difficult. For example, for a compound that is 95-96% radioactively pure and minimally absorbed (i.e., approximately 2% absorbed), it is impossible to unequivocally differentiate between 2% absorption of the test compound and 100% absorption of a radioactive impurity present at 2%.

ii. Animals

Metabolic and pharmacokinetic data from two rodent species (usually the rat and mouse) and a non-rodent species (usually the dog) are recommended. If a dose dependency is observed in metabolic and pharmacokinetic or toxicity studies with one species, the same range of doses should be used in metabolic and pharmacokinetic studies with other species. If human metabolism and pharmacokinetic data also are available, this information should be used to help select test species for the full range of toxicity tests, and may help to justify using data from a particular species as a human surrogate in safety assessment and risk assessment. (Human metabolism studies should be conducted according to the guidelines in Chapter VI B.)

Metabolism and pharmacokinetic studies have greater relevance when conducted in both sexes of young adult animals of the same species and strain used for other toxicity tests with the test substance. The number of animals used in metabolism and pharmacokinetic studies should be sufficient to reliably estimate population variability (see Chapter V B 1 e). A single set of intravenous and oral dosing results from adult animals, when combined with some *in vitro* kinetic results, may provide an adequate data set for the design and interpretation of short-term, subchronic and chronic toxicity studies.

Studies in multiple species may clarify what appear to be contradictory findings in toxicity studies (i.e., equal mg/kg bw doses having less effect in one species than in another). If disposition and metabolite profiles are found to be similar, then differences in responses among species could more reliably be attributed to factors other than differences in metabolism. Studies of the

V B. Metabolism and Pharmacokinetic Studies Continued

pharmacokinetics and metabolism of a substance in neonatal and adolescent animals provide information about any changes in metabolism associated with tissue differentiation and development. Animals with fetuses of known gestational age should be used for determining the disposition of the test substance in the fetus.

iii. Route of Administration

The most critical parameters required in assessing human exposure and target tissue dose are the gastrointestinal absorption rate and internal elimination rates (renal and hepatic) for the test compound. Without an intravenous (IV) dosing study, it is very difficult to determine what percentage of a chemical is absorbed, because the material excreted in the feces is composed of unabsorbed dose plus biliary and non-biliary (mucosal) elimination.

An intravenous study can provide accurate rates of metabolism-- without interference from intestinal flora--plus rates of renal and biliary elimination, if urine and bile are collected. This route also avoids the variability in delivered dose associated with oral absorption and ensures that the maximum amount of radiolabel is excreted in the urine or bile for purposes of detection. Once IV data and parameters are available, they can be used with plasma concentrations from limited oral studies to compute intestinal absorption via the ratio of Areas Under the (plasma and or urine) Curves or via simulations of absorption with gastrointestinal absorption models.

In single-dose pharmacokinetic studies of oral absorption, the primary concerns are with the extent of absorption and peak plasma or target tissue concentrations of the test substance. If the test vehicle affects gastric emptying, it may be necessary to use both fasted and non-fasted animals for pharmacokinetic studies.

iv. Dosage Regimen

Selection of the dosing regimen for metabolism and pharmacokinetic studies depends on the type of information that is needed. Metabolic and pharmacokinetic parameters are usually determined following a single administration of the test compound. Comparing parameters obtained from studies in which a range of single doses have been administered can be used to determine the doses at which saturation of absorption, distribution, metabolism or excretion occurs. Multiple dosing studies can be used to determine the potential of a compound to induce or inhibit its absorption, distribution, metabolism or excretion. Identification and quantification of the major metabolites following administration of single and multiple doses may indicate whether saturation or induction of a particular biotransformation pathway can occur.

In vitro experiments may be useful in screening for dose dependencies, and provide more accurate descriptions of the enzyme kinetics or other processes underlying dose dependencies observed in the whole animal. *In vitro* studies usually indicate identical metabolic pathways and metabolism rates comparable to those obtained from whole animal studies but require fewer animals to perform and can be completed in less time with fewer resources.*

*Bäärnhielm *et al.* (1986);¹ Green *et al.* (1986);² Lin *et al.* (1982)³

V B. Metabolism and Pharmacokinetic Studies Continued

v. Sampling

Blood (RBCs, plasma, and serum), urine, and feces are the most commonly collected samples. In addition, a few representative organ and tissue samples should be taken, such as liver, kidney, fat, and suspected target organs. Sampling times should depend on the substance being tested and the route of administration. In general, an equal number of blood samples should be taken in each phase of the concentration-versus-time curve. Intravenous (IV) studies usually require much shorter, and more frequent, sampling than is required for oral dosing. Time spacing of samples will depend on the rates of uptake and elimination. In a typical IV study, blood and tissue samples are taken in a "powers of 2" series, i.e. samples at 2, 4, 8, 16, and 30 (32) minutes, 1, 2, 4, 8, and 16 hours. Similar coverage could be obtained with only 7 time points by using a "powers of 3" series: 3, 9, and 30 (27) minutes, 1, 3, 9, and 24 (27) hours. Oral dosing studies usually extend to at least 72 hours, or 5 plasma half-lives, ensuring the excretion of 95% of the absorbed dose. The sampling schedule for an oral dosing experiment might be: 15 minutes, 30 minutes, 1, 2, 4, 8, 24, 48, and 72 hours. Such a sampling scheme would provide data coverage for evaluation of absorption, elimination, enterohepatic recirculation and excretion processes.

Whole Body Autoradiography (WBA) has been used with increasing frequency as a means of identifying tissues which concentrate test substances. This technique allows a small number of animals (5 - 10) to be used for screening purposes with a minimal investment in manual labor. FDA encourages the use of WBA with IV dosing, as a means of screening and selecting tissues of greatest relevance for later oral dosing studies. Animals used for WBA should be sacrificed during the elimination phase, between 1 and 5 plasma half-lives, since bioaccumulation at steady-state is the primary consideration in selecting specific tissues.

The number of animals used in metabolism and pharmacokinetic studies should be large enough to reliably estimate population variability. In the case of rats and mice, tissue and/or blood sample size is usually the limiting factor: analysis of the substance may require 1 ml or more blood, but it is difficult to obtain multiple blood samples of this size from one animal. As a consequence, a larger number of animals is required (3 - 4 per time point, 7 - 9 time points) when small rodents are used. Such an approach has the advantage of allowing limited sampling of critical tissues (e.g. liver, fat) at each time point, an option which is usually unavailable with large animals. The use of humans and large animals generally permits collection of multiple (serial) blood samples. For outcrossing populations like humans and large animals, individual differences in the rates of biotransformation are likely to be greater than those of inbred rodent populations; under these circumstances, more samples/sex/group may be needed to reliably estimate variability.

Individual metabolism cages are recommended for collecting urine and feces in oral dosing studies. Excreta should be collected for at least 5 elimination half-lives of the test substance. When urine concentrations will be used to determine elimination rates, sampling times should be less than one elimination half-life (taken directly from the bladder in IV studies); otherwise, samples should be taken at equal time intervals.

vi. In Vitro studies

In Vitro measurements employing enzymes, subcellular organelles, isolated cells and perfused organs may be used to augment the dose response information available from less extensive metabolic and pharmacokinetic studies. Because *in vitro* systems generally are less complex than whole animals, elucidation of a test compound's metabolic pathways and the pathways' kinetic characteristics may be facilitated. Such systems can be used to measure binding, adduct and conjugate formation, transport across cell membranes, enzyme activity, enzyme substrate specificity, and other singular objectives. Biochemical measurements that can be made using *in vitro* systems

V B. Metabolism and Pharmacokinetic Studies Continued

include: Intrinsic clearances of enzymes in an organ or tissue, kinetic constants for an enzyme, binding constants, and the affinity of the test compound and its metabolites for the target macromolecules. The activity of a hepatic drug-metabolizing enzyme *in vivo* may be approximated by kinetic constants that are calculated from *in vitro* studies; when a first-order approximation is used, the ratio of V_{max} to K_m is equal to the intrinsic clearance of the drug.^a *In vitro* measurements made using readily accessible tissues and body fluids from animals and man may also be useful in elucidating mechanisms of toxicity.

vii. Analysis of Data

Data from all metabolism and pharmacokinetic studies should be analyzed with the same pharmacokinetic model and results should be expressed in the same units. Concentration units are acceptable if the organ or sample size is reported, but percent of dose/organ is usually a more meaningful unit. In general, all samples should be analyzed for metabolites that cumulatively represent more than 1% of the dose.

A variety of rate constants and other parameters can be obtained from IV and oral dosing data sets, provided that good coverage of the distribution, elimination, and absorption (oral dose) phases is available. Typical parameters calculated to characterize the disposition of a test substance are: half-lives of elimination and absorption; area under the concentration-versus-time curve (AUC) for blood; total body, renal and metabolic clearances (Cl); volume of distribution (V_d); bioavailability (F); and mean residence and absorption times (MAT, MRT). Some of these parameters, such as half-lives and elimination rates, are easily computed from one another; the half-life is more easily visualized than the rate constant.^b

Computation of oral absorption (k_a) and elimination (E) rates is often complicated by the "flip-flop" of the absorption and elimination phases when they differ by less than a factor of 3.^c Because of these analysis problems, computation of absorption and elimination rates should not be attempted on the basis of oral dosing results alone.

Blood-tissue uptake rates (k_{12}) can often be approximated from data at early ($t < 10$ minutes) time points in IV studies, provided that the blood has been washed from the organ (e.g. liver) or the contribution from blood to the tissue residue is subtracted (fat). High accuracy is not usually required since these parameters can be optimized to fit the data when they are used in more complex models. Tissue-blood recycling rates (k_{21}) and residence times can be computed from partition coefficients if estimates of uptake rates are available.

Tissue/blood partition coefficients (R_p) should be determined when steady-state has been achieved. Estimates based on samples obtained during the elimination phase following a single dose of the test substance may lead to underestimates of this ratio in both eliminating and non-eliminating tissues unless its half-life is very long. Correction of these values for elimination has been described by several authors.^d

^a Rane *et al.* (1977);⁴ Gillette (1986)⁵

^b Wagner and Nelson (1963);⁶ Gibaldi and Perrier (1982)⁷

^c Notari (1987)⁸

^d Chen and Gross (1979);⁹ Lam *et al.* (1982)¹⁰

V B. Metabolism and Pharmacokinetic Studies Continued

It may be important to determine the degree of plasma protein and red blood cell binding of the test substance; calculation of blood clearance rates using plasma or serum concentrations of the substance that have not been adjusted for the degree of binding may under- or over-estimate the true rate of clearance of the test substance from the blood. This is usually done through experiments *in vitro*.

Two classical methods used in the analysis of pharmacokinetic data are the fitting of sums of exponential functions (2- and 3-compartment mammillary models) to plasma and/or tissue data, and less frequently, the fitting of arbitrary polynomial functions to the data (non-compartmental analysis).^a

Non-compartmental analysis is limited in that it is not descriptive or predictive; concentrations must be interpolated from data. The appeal of non-compartmental analysis is that the shape of the blood concentration-versus-time curve is not assumed to be represented by an exponential function and, therefore, estimates of metabolic and pharmacokinetic parameters are not biased by this assumption. In order to minimize errors in parameter estimates that are introduced by interpolation, a large number of data points that adequately define the concentration-versus-time curve are needed.

Analysis of data using simple mammillary, compartmental models allows the estimation of all of the basic parameters mentioned above, if data for individual tissues are analyzed with 1 or 2 compartment models, and combined with results from 2 - 3 compartment analyses of blood data. "Curve Stripping" analysis can be applied to such simple models through the use of common spreadsheet programs (i.e. LOTUS 1-2-3), as long as a linear regression function is provided in the program. Optimization of the coefficients and exponents estimated may require the use of more sophisticated software: a number of scientific data analysis packages such as RS/1 and SigmaPlot have the necessary capabilities. Specialized programs such as NONLIN^b, CONSAM,^c or SIMUSOLV^d will be needed when more complex models must be analyzed. Coefficients and exponents from mammillary models can be used to calculate other parameters; however, they should not be taken too literally, since mammillary models assume that all inputs are to a central pool (blood), which communicates without limitation into other compartments.^e This approach does not include details such as blood flow limitations, anatomical volumes or other physiological limits in the animal.

Physiologically based pharmacokinetic models (PB-PK) were developed to overcome the limitations of simple mammillary models. Physiologically based models describe the disposition of test substances via compartmental models which incorporate anatomical, biochemical and physiological features of specific tissues in the whole animal. The types of information added include organ-specific blood flows, volumes, growth models and metabolism rates. Metabolic parameters often are obtained from *in vitro* studies (i.e., enzyme reaction rates in cultured hepatocytes, plasma protein binding, etc.), while other parameters are becoming available as standard parameters in the literature. Parameters from mammillary models can be used to compute the value of parameters used in physiological pharmacokinetic models, using tissue-specific blood flows,

^aNotari (1987);^b Benet (1972);^c Vaughn and Trainor (1975)^d

^ePedersen (1977)¹³

^fBoston *et al.* (1988)¹⁴

^gSteiner *et al.* (1990)¹⁵

^hRoscigno and Segre (1966);ⁱ Fagarasan and DiStefano (1989)¹⁷

V B. Metabolism and Pharmacokinetic Studies Continued

anatomical volumes, and other information (literature values). Estimation of parameters for a simple mammillary model is often the first data reduction step in creating a physiological model.^a

Because PB-PK models are based on physiological and anatomical measurements and all mammals are inherently similar, they provide a rational basis for relating data obtained from animals to humans. Estimates of predicted disposition patterns for test substances in humans may be obtained by adjusting biochemical parameters in models validated for animals; adjustments are based on experimental results of animal and human *in vitro* tests and by substituting appropriate human tissue sizes and blood flows. Development of these models requires special software capable of simultaneously solving multiple (often very complex) differential equations, some of which were mentioned above. Several detailed descriptions of data analysis have been reported.^b

b. Use of Data from Metabolism and Pharmacokinetic Studies

Information from metabolism and pharmacokinetic studies can be used in the design and analysis of data from other toxicity studies. Some examples are described below.

Design of Toxicity Studies: The concentration-versus-time curve, peak, and steady-state concentrations of the test substance in blood or plasma provide information on the distribution and persistence of the substance in the animal which may suggest essential elements in the design of toxicity studies. For example, when metabolic and pharmacokinetic studies indicate that the test compound accumulates in the bone marrow, long-term toxicity tests should include evaluation of the test compound's effect on hematopoietic function and morphology. If a test compound is found to accumulate in milk, an investigator should plan to perform reproductive toxicity studies with *in utero* exposure and a nursing phase (cross-fostering study; see Chapter IV C 8). In addition, information from metabolic and pharmacokinetic studies can be used to predict the amount of test compound that enters biological compartments (tissues, organs, etc.) that may not suffer a toxic insult but may serve as depots for indirect or secondary exposure.

Setting Dose Levels: There is considerable debate about the use of metabolic and pharmacokinetic data in setting doses to be used in toxicity studies, particularly chronic toxicity and carcinogenicity studies. Current NTP policy for selecting the highest dose in carcinogenicity bioassays is described in Chapter IV C 6 b. In 1984, the NTP Ad Hoc Panel on Chemical Carcinogenesis Testing and Evaluation also recommended that pharmacokinetic data be considered along with subchronic toxicity data in setting all dose levels except the maximum-tolerated dose (MTD) in the carcinogenesis bioassay of chemicals.^c FDA agrees with these statements and recommends that pharmacokinetic data be used in conjunction with the results of short-term and subchronic toxicity studies to set appropriate dose levels for chronic toxicity, reproduction and teratology studies, and for setting dose levels below the MTD (highest dose) in carcinogenicity studies.

^aGillette (1986);^d O'Flaherty (1989)¹⁸

^bGibaldi and Perrier (1982)² Gerlowski and Jain (1983);¹⁹

^cNTP Report (1984)²⁰

V B. Metabolism and Pharmacokinetic Studies Continued

Determining Mechanisms of Toxicity: Information from metabolic and pharmacokinetic studies can be used to supplement conventional toxicology data in elucidating mechanisms of toxicity. Metabolites identified by a pharmacokinetic study can suggest mechanisms underlying a toxic response. Biologically reactive intermediates are often implicated in a toxic response; however, such metabolites are usually short-lived, reacting in the vicinity of their formation. The presence of potentially reactive intermediates can be deduced indirectly by measuring the formation of characteristic macromolecular (DNA, RNA, protein) adducts and metabolic conjugates. Measurement of metabolic conjugate vs adduct formation and the affinity of a compound and/or its metabolites for the target molecule may help identify mechanisms of toxicity and effective routes of detoxification.

Information from *in vitro* test systems concerning the formation of critically reactive metabolites may be used to establish the relationship between the formation of the reactive metabolite *in vivo* and duration of exposure to the test compound. This relationship is important in circumstances where critically reactive metabolites are only formed when the capacities of normal metabolic and other defensive or adaptive mechanisms are exceeded. Determining the concentrations of the test substance at which saturation of binding occurs may indicate at what concentration a compound is likely to deplete detoxifying conjugation pools and become available to react with target macromolecules.

Improving the Risk Assessment Process: Information from metabolic and pharmacokinetic studies increasingly is being incorporated into risk assessments. Conventional risk assessments typically involve linear extrapolation of external dose and an inter-species scale factor based on body weight or body surface area. Risks calculated by this approach may be under- or over-estimated. Many of the biological processes involved in the absorption, distribution, metabolism and excretion of a compound are dose dependent and, therefore, the toxicity observed may not be a simple function of administered dose. Development of appropriate pharmacokinetic models may enhance our ability to use metabolic and pharmacokinetic information in risk assessment.

2. Recommended Metabolism and Pharmacokinetic Studies

FDA believes that data from studies on the absorption, distribution, metabolism, and excretion of a chemical can provide insight into mechanisms of toxicity of chemicals and are essential in the design and evaluation of results from other toxicity studies. FDA believes that a set of basic pharmacokinetic and metabolism studies should be performed for all Concern Level II and III substances, but that additional studies may be recommended for a particular additive. Recommended studies should be performed with two rodent species (usually the rat and mouse) and one non-rodent species (usually the dog). In general, what constitutes an appropriate set of metabolism and pharmacokinetic studies will depend on the anticipated degree and type of toxic response to a test compound and by the estimated magnitude of human exposure to the compound. The recommended set of basic studies are:

- Intravenous studies using a tracer level dose should be conducted in adult male and female animals of species in which toxicity studies have already been conducted or in which chronic toxicity studies are contemplated. Blood, liver, and fat samples should be taken at all time points. The size and timing of urine and bile samples will depend on the dose of tracer and rate of excretion by each of these routes. Samples taken over periods of 30 min to 2 hours, at 2 or 3 time points, should be sufficient for computation of the cumulative excretion by these routes. Plasma, urine and bile should be analyzed for metabolites of the test substance that cumulatively represent more than 1% of the dose. Estimates of uptake and elimination rates should be made for each tissue sampled, using 2-compartment models.

V B. Metabolism and Pharmacokinetic Studies Continued

- Studies of the rate of metabolism (of the parent compound) as a function of dose (or concentration) should be conducted *in vivo* or *in vitro*, guided by results of metabolite analyses from the intravenous studies and available toxicology information. Hepatocytes or perfused livers will normally be used for such studies, but an examination of the distribution of metabolites between the plasma, bile and urine after IV dosing may indicate that the kidney is important in the metabolism of some chemicals. Enzyme kinetic parameters resulting from *in vitro* studies may be scaled up to whole organ rates and used to predict rates of metabolism in the whole animal as a function of dose.
- Oral dosing studies should be conducted in *ad libitum* fed animals, to determine the rate and cumulative absorption of the substance. Dosage and sampling times should be selected on the basis of results from toxicity tests, metabolic dose response data (II, above), and elimination rates determined from IV dosing studies. Bioaccumulative tissues should be sampled in addition to blood, urine and feces. A tissue that does not accumulate the substance should also be included for reference purposes. Whole Body Autoradiographic studies are recommended as a method for identifying bioaccumulative tissues prior to the initiation of oral dosing studies.

3. Additional Studies

Studies of enzyme induction and potential pharmacological adaptation should be conducted whenever chronic studies are recommended. The resulting information can be incorporated into multiple or continuous dosing models to simulate the plasma and tissue levels of test substance expected for a variety of doses in chronic studies being planned.

In cases where reproductive studies are recommended, pharmacokinetic experiments evaluating the distribution of the substance in the fetus, mother's milk, and neonates should be performed as an aid in selecting doses and designing reproductive toxicity studies. If the metabolic potential of the fetal and/or neonatal liver can be assessed in a preliminary *in vitro* study, this step is highly recommended.

Assuming that IV and oral dosing studies have already been completed for both male and female adult animals prior to the reproductive pharmacokinetic studies, sampling can be more limited, i.e. excretion studies combined with limited sampling of maternal blood, fetuses, milk, and neonatal tissues may be sufficient for characterization of the metabolic and pharmacokinetic processes of interest in pregnancy.

Depending on the types of toxic effects observed and the importance of understanding the mechanisms of these effects to the safety assessment of a direct food or color additive used in food, additional biochemical or *in vitro* experiments may be submitted by the petitioner in support of any mechanism proposed. Such studies should be substance-specific, and should be based on consultation with CFSAN, as appropriate.

V C. Neurotoxicity Studies

The nervous system regulates and maintains diverse biological processes that are essential not only for survival but also for maintaining an acceptable quality of life. The proper functioning of the nervous system enables an organism to receive information from its internal and external environments and to orchestrate appropriate adaptive physiological and behavioral responses. An extensive body of data demonstrates that diverse chemical substances can alter the structure and function of the nervous system in a variety of ways. Alterations which compromise the organism's ability to function appropriately in its environment are considered adverse. Neurotoxicity refers to any adverse effects of exposure to chemical, biological or physical agents on the structure or functional integrity of the developing or adult nervous system. The onset of neurotoxicity can vary from immediate to delayed following exposure to a toxic substance, and duration can vary from transient to persistent. Neurotoxicity may result from effects of the toxic substance directly on the elements of the nervous system or from effects of the toxicant on other biological systems which then adversely affect the nervous system. Neurotoxic effects are generally associated with a spectrum of biochemical, morphological, behavioral, and physiological abnormalities. Depending upon their severity, some of these abnormalities may have life-threatening consequences; more commonly, they result in diminished quality of life.

In 1985, FDA commissioned the Federation of American Societies for Experimental Biology (FASEB) to assess the ability of current FDA guidelines for toxicity testing of food and color additives used in food to detect neurotoxic hazards.^a One conclusion of the FASEB report was that information derived from conventional toxicity studies is so non-specific that it basically limits the detection of neurotoxicity to those disorders in which nervous system deficits are clearly evident.^b This type of approach minimizes the detection of subtle neurotoxic hazards and precludes an adequate assessment of the spectrum of potential neurotoxic effects on the structural and functional integrity of the nervous system.^c

Until recently, neurotoxicity was equated with neuropathy involving frank neuropathological lesions or overt neurological dysfunctions, such as seizure, paralysis or tremor. Examples of chemically induced neuropathy in humans (for example, from exposure to lead, organic mercury, hexane, carbon disulfide, and tri-ortho-cresylphosphate) emphasize the need for assessing the neurotoxic potential of chemicals to which humans may be exposed.^d Although neuropathy is appropriately recognized as a manifestation of neurotoxicity, it is now clear that there are numerous other signs of nervous system toxicity.^e Motor incoordination, sensory deficits, learning and memory dysfunctions, changes in emotion and altered states of arousal are also recognized as indices of neurotoxicity. Continued reliance on neuropathy as the primary criterion of neurotoxicity is overly simplistic and may significantly underestimate the neurotoxic potential of a chemical in adult or developing organisms.

^a Federation of American Societies for Experimental Biology Report (1986);¹ Leukroth (1987)²

^b Federation of American Societies for Experimental Biology Report (1986);¹ Leukroth (1987)²

^c McMillan (1987);³ Vorhees (1987);⁴ U.S. Office of Technology Assessment (1990)⁵

^d U.S. Office of Technology Assessment Report (1990);⁵ World Health Organization Report (1986);⁶ Spencer and Schaumburg (1980);⁷

^e Tilson (1987)⁸

V C Neurotoxicity Studies Continued

Ongoing research on nervous system toxicity continues to reveal the broad spectrum of biochemical, structural, and functional abnormalities that toxicants can elicit, both directly and indirectly.¹ Neurotoxic chemicals invariably initiate their effects at the molecular level, altering neurochemical processes. These changes may be of sufficient quality or sufficient magnitude to induce neuropathology. Neurochemical changes may also result in altered nervous system function and may be expressed as physiological or behavioral abnormalities.² Significant physiological or behavioral dysfunctions can occur prior to, or in the absence of, evident neuropathology or other signs of toxicity.³ This is exemplified by the marked behavioral dysfunctions associated with exposure to such neuroactive chemicals as barbiturates, amphetamines, ethanol, lead, and carbon monoxide at exposure levels that elicit no signs of neuropathy.⁴ This disassociation of neuropathology and functional changes may involve a number of factors, including the intrinsic toxicity of a chemical and, particularly, the dose and regimen of exposure.

Among the various approaches that can be used for assessing neurotoxicity, behavioral testing represents a practical means of obtaining a relatively comprehensive assessment of the functional development and integrity of the nervous system within the context of a standard toxicity study.⁵ Behavior is an adaptive response of an organism, orchestrated by the nervous system, to some set of internal and external stimuli. A behavioral response represents the integrated end product of multiple neuronal subsystems including sensory, motor, cognitive, attentional, and integrative components, as well as an array of physiological functions.⁶ As such, behavior can serve as a measurable index of the status of multiple functional components of the nervous system.

Behavioral testing typically is non-invasive and can be used repeatedly for longitudinal assessment of the neurotoxicity of a test compound, including persistent or delayed treatment-related effects.⁷ Furthermore, since neuronal function can be influenced by the status of other organ systems in the body (e.g. cardiovascular, endocrine, and immunologic systems), certain types of behavioral changes may indirectly reflect significant toxicity in other organ systems.

Behavioral testing has been established as a reliable toxicological index in safety assessment. Considerable progress has been made in the standardization and validation of neurobehavioral and

¹ Federation of American Societies for Experimental Biology Report (1986);¹ Leukroth (1987);² World Health Organization Report (1986);³ Spencer and Schaumburg (1980);⁷ Buelke-Sam *et al.* (1985);⁹ Reiter (1987)¹⁰

⁵ Federation of American Societies for Experimental Biology Report (1986);¹ Leukroth (1987);² Spencer and Schaumburg (1980);⁷ Reiter (1987);¹⁰ Anger and Johnson (1985)¹¹

⁶ Federation of American Societies for Experimental Biology Report (1986);¹ World Health Organization Report (1986);³ Reiter (1987);¹⁰ Riley and Vorhees (1986)¹²

⁴ Hutchings *et al.* (1987)¹³

³ Federation of American Societies of Experimental Biology Report (1986);¹ Vorhees *et al.* (1984)¹⁴

⁷ Mitchell and Tilson (1982)¹⁵

² Leukroth (1987)²

V C Neurotoxicity Studies Continued

neurodevelopmental testing procedures.^a As a result, a variety of behavioral methodologies is available for use in determining the potential of chemical substances to affect adversely the functional integrity of the nervous system in adult and developing organisms.^b Behavioral testing can be readily incorporated into toxicity testing protocols and can improve and expand current approaches to assessing neurotoxic hazard.

Because of the impact that nervous system toxicity can have on human health, assessing the neurotoxic potential of a chemical proposed for use as a food or color additive should be an essential element in that chemical's toxicological profile.^c As information from research in neurotoxicology continues to evolve, our understanding of the processes underlying neurotoxicity will become increasingly clear. This will enhance our ability to utilize more effectively the information about the neurotoxicity of test substances in support of regulatory decisions.^d

1. Evaluating Neurotoxicity

The reliability of assessing the neurotoxic potential of a test substance is directly related to the extent to which the detection and evaluation of neurotoxicity is included as a specific, defined objective of routine toxicity testing.^e A number of scientific panels and health-related organizations have recommended that the assessment of neurotoxic potential be carried out according to a structured process of tiered testing.^f Each tier would focus on a different objective. Testing should progress from the initial identification of chemicals that may be associated with neurotoxic effects (screening), through the subsequent characterization of the scope of nervous system involvement (characterization of effects), to the determination of dose response kinetics, including the no-observed-effect level (dose-response).

A tiered approach to neurotoxicity testing and evaluation allows for multiple decision points at which scientifically based decisions can be made about the adequacy of available information and the need for additional testing. To facilitate such decisions, specific summary statements regarding the neurotoxic potential of the test compound should be included in the evaluation of the results of each level of testing. Since toxicity to the nervous system should be evaluated within the context of a comprehensive assessment of all significant forms of toxicity for a test compound, the neurotoxicity summary statements should integrate all relevant toxicology data which are available. This includes information derived not only from tests specifically focused on the detection of nervous system

^a Federation of American Societies of Experimental Biology report (1986);¹ Leukroth (1987);² World Health Organization report (1986);³ Buelke-Sam *et al.* (1985);⁴ U.S. EPA report (1985);¹⁶ U.S. EPA report (1991);¹⁷ Kimmel *et al.* (1990)¹⁸

^b Leukroth (1987)²

^c Leukroth (1987);² U.S. Office of Technology Assessment report (1990);⁵ Reiter (1987);¹⁰ Nation Research Council, National Academy of Sciences report (1984);¹⁹ Sobotka (1986)²⁰

^d Reiter (1987);¹⁰ U.S. House of Representatives Committee on Science and Technology report (1986)²¹

^e Federation of American Societies of Experimental Biology report (1986);¹ Sobotka (1986)²⁰

^f Federation of American Societies of Experimental Biology report (1986);¹ World Health Organization report (1986);³ National Research Council, National Academy of Sciences report (1984);¹⁹ National Research Council, National Academy of Sciences report (1975)²²

V C Neurotoxicity Studies Continued

toxicity (e.g. neuropathology, behavioral dysfunctions, neurochemical alterations or physiological changes), but also from the more conventional testing that focuses on other toxic effects of the test compound, for example, adverse changes in growth, development, food or water intake, or endocrine status.

The neurobiological implications of some conventional toxic effects are certainly more evident than others. For example, a compound that induces specific teratogenicity of the nervous system, even at high dose levels, would be suspect for adversely affecting the development of nervous system function at lower doses. The neurotoxicological significance of other types of toxicity, however, may be less obvious. For instance, chemicals found to alter hormonal balance might also be suspected of affecting the structural or functional integrity of the nervous system, since endocrine status and the nervous system are interrelated. Altered growth, which is considered an index of general toxicity, may also signal the presence of neurotoxicity. In the developing organism, abnormal growth may reflect a treatment related neurotoxicity of the mother involving poor care of the nursing offspring. In the adult, altered growth stemming from changes in food or water intake may reflect underlying nervous system dysfunction, since both eating and drinking are consummatory behaviors with neuromuscular and physiological components under neuronal control. It should be clear, however, that such types of toxicology endpoints, by themselves, are not evidence of neurotoxicity. Rather, when viewed in conjunction with other available data, such effects may serve to indicate the possibility of treatment related effects on the nervous system.

a. Screening

The first stage in assessing neurotoxicity involves a process of screening to identify those chemicals that exhibit any potential for adversely affecting the nervous system. It should be clear that the primary objective of screening is detection. The information derived from screening is not intended, nor is it sufficient, to be used as the basis for determining the no-effect-level (NOEL) for neurotoxicity. The NOEL is more appropriately based on information derived from the later stage of dose-response testing.

There are basically three sources of neurotoxicity screening information. One involves the use of structure activity relationships (SAR), the second relies on published literature and other sources of documentation, and the third involves empirical testing. The usefulness and reliability of SAR for identifying potential neurotoxicants is, at the present time, rather limited due to the fact that SAR databases for neurotoxicity are still being developed. The use of published literature or other types of documented information, to the extent that this type of information is available and appropriate for regulatory application, can be of significant value in identifying chemicals that may affect the nervous system. However, this type of information is usually scattered and typically not available for many food-related chemicals. At the present time, the primary means of obtaining neurotoxicity screening data is through empirical testing. The experimental data needed to screen chemicals for potential adult and developmental neurotoxicity should be routinely obtained as part of standard general and developmental toxicity studies used for entrance-level testing of proposed food chemicals across the various levels of concern. Most appropriately, this would include both short-term (28 day) and subchronic (90 day) studies to screen for potential adult neurotoxicity following short-term and more prolonged exposures, and reproduction/teratology studies to screen for potential developmental neurotoxicity in the developing and mature offspring.

Screening for neurotoxicity involves the use of valid, cost-effective procedures which can be carried out rapidly and routinely on large numbers of chemicals to detect the presence or absence of

V C Neurotoxicity Studies Continued

immediate or delayed adverse effects on the nervous system.⁵ Neurotoxicity can appear as a wide range of morphological and functional abnormalities involving the nervous system at very specific or multiple levels of its organization.⁶ Under the previous guidelines for toxicity testing of proposed food and color additives used in food the identification of neurotoxic effects was based on information derived from a general pathological evaluation of a few sections of neuronal tissue and an unstructured casual observation of test animals in their cages for overt signs of toxicity.⁶ This approach limited the ability to detect anything but the most severe forms of neurotoxicity. To maximize the probability of detection, screening should be sufficiently comprehensive to enable the detection of a representative variety of pathological changes and functional disorders of the peripheral, central and autonomic segments of the nervous system.⁴ In developmental studies, neurotoxicity screening should enable the detection of treatment related effects not only in the immature developing offspring, but in the mature adult offspring as well.

An effective and comprehensive basic neurotoxicity screen would include both (1) a specific histopathological examination of tissue samples representative of all major areas and cellular elements of the brain, spinal cord, and peripheral nervous system, in conjunction with (2) a systematic examination of experimental animals inside and outside of their cages using a clearly defined functional evaluation battery of clinical tests and observations to provide a general assessment of the primary neurological, behavioral and physiological functions of the nervous system. Typically, a functional evaluation battery would include a variety of indices to detect significant behavioral changes (for example, in the level of activity and arousal, reactivity, motor coordination, gait, neurosensory function, and reflexes); physiological functions (including feed and water intake, body weight and autonomic signs); neurological disorders (such as paralysis, seizure, or tremor); and any other signs of abnormal behavior or nervous system toxicity.

To help ensure the complete and consistent application of the neurotoxicity screen throughout a particular study, each study protocol should include a detailed description of the particular screen to be used in that study, including its composition, the test procedures to be followed, the time periods at which the screen is to be carried out, the neuronal structures to be examined, the endpoints to be used, and the methods for recording and analyzing the data. Since neurotoxicity screening is intended to be a routine part of both general and developmental toxicity studies, the specific composition of the screen and the endpoints to be recorded should accommodate the focus of the different protocols and, specifically, be appropriate for the age of the animals to be tested. For example, when testing immature developing animals, it would be appropriate to include indices of the ontogenetic development of representative neuroanatomical, physical, and functional milestones as part of the neurotoxicity screen. There are a number of available publications to guide the design and conduct of neurotoxicity screens appropriate for the adult organism⁷ and for

⁵ Mitchell and Tilson (1982)¹⁵

⁶ Leukroth (1987);² Spencer and Schaumburg (1980);⁷ Reiter (1987);¹⁰ Anger and Johnson (1985);¹¹ Tilson and Mitchell (1992)²¹

⁷ U.S. FDA (1982)²³

⁸ Federation of the American Societies of Experimental Biology report (1986);³ Nelson (1991)²⁴

⁹ Leukroth (1987);² Vorhees (1987);⁴ U.S. Office of Technology Assessment report (1990);³ World Health Organization report (1986);⁶ Buelke-Sam *et al.* (1985);⁹ U.S. EPA report (1985);¹⁶ U.S. EPA report (1991);¹⁷ National Academy of Sciences report (1984);¹⁹ Tilson and Mitchell (1992)²¹ Tupper and Wallace (1980);²⁴ Deuel (1977);²⁷ Gad (1982);²⁸ Gad (1989);²⁹ Marshall *et al.* (1971)³⁰ Spencer *et al.* (1980);³¹ O'Donoghue (1988);³² Moser (1988);³³ Wier *et al.* (1989);³⁴ U.S. EPA

V C Neurotoxicity Studies Continued

developing and adult offspring.* As appropriate, more sensitive and objective indices of neurotoxicity, such as tests of cognitive function and automated measures of locomotor activity, could be included to supplement the basic neurotoxicity screen.⁵

b. Characterization

When a chemical is presumptively identified by SAR, empirical screening, or other sources of information as producing neurotoxicity, that chemical becomes a candidate for additional neurotoxicity testing. Chemicals not identified as having neurotoxic effects during screening will generally not be recommended for subsequent neurotoxicity testing, although exceptions may occur.

The next level of testing after screening focuses attention on determining the nature and extent to which the nervous system is affected by that chemical (characterization). At this level the neurotoxic effects found during screening are further characterized and studies are conducted to determine whether the test chemical has any other, possibly more subtle, effects on the structural and functional integrity of the nervous system in mature and developing organisms. The characterization of neurotoxicity should include information about the severity of effects, the temporal pattern of onset of effects (particularly when delayed neurotoxicity occurs), and the duration of effects. Neuropathological investigations should include *in situ* perfusion and a detailed histopathological examination (more detailed than the histopathology examination performed during screening) involving the use of special stains to highlight relevant neural structures.⁶

The neurofunctional assessment at this level should routinely include a core battery of behavioral and physiological tests designed to detect adverse changes to the primary subfunctions (e.g. cognitive, sensory, motor, and autonomic) of the nervous system in mature and developing nervous systems.⁶ The need for additional special tests may logically follow from information obtained during screening; for example, if a chemical is observed to induce convulsions during screening, the seizure potential and pro-convulsant properties of that chemical should be more specifically characterized during the second level of testing.

In concert with conventional toxicity testing protocols, routine neurotoxicity evaluation should generally be carried out using rodents as the primary species of choice. However, as part of the characterization level of testing for neurotoxicity, cross-species comparisons of the neurotoxic potential of the test compound should also be carried out if appropriate test systems are available. For such studies, alternate species should be non-rodent.

* Vorhees (1987);⁴ World Health Organization report (1986);⁶ Buelke-Sam *et al.* (1985);⁹ Vorhees *et al.* (1984);¹⁴ U.S.EPA report (1991);¹⁷ Kimmel *et al.* (1990);¹⁸ U.S. EPA report (1988);³³ Vorhees *et al.* (1979)³⁵

⁵ World Health Organization report (1986);⁶ U.S. EPA report (1985);¹⁶ National Academy of Sciences report (1984);¹⁹ U.S. FDA report (1982)²⁴

⁶ World Health Organization report (1986);⁶ U.S. EPA report (1985);¹⁶ U.S. EPA report (1991);¹⁷ Spencer *et al.* (1980)³¹

⁹ Leukroth (1987);² U.S. Office of Technology Assessment report (1990);³ World Health Organization report (1986);⁶ U.S. EPA report (1985);¹⁶ U.S. EPA report (1991);¹⁷ Wier *et al.* (1989);³⁴ Vorhees *et al.* (1979);³⁷ Geller *et al.* (1979)³⁸

V C Neurotoxicity Studies Continued

c. Dose-Response

A critical element in defining a chemical's neurotoxic hazard is determining the no-observed-effect level (NOEL). The NOEL should be based upon an accurate and reliable determination of the dose-response and dose-time relationships derived from repeated exposure studies, e.g. intermittent and continuous exposure regimes, typically using the most relevant and sensitive endpoint(s) identified in previous testing. As part of the definition of neurotoxic hazard, attempts should also be made to identify any unique factors that may affect the sensitivity of the experimental animals to the test compound and may help to identify uniquely susceptible human sub-populations. *Any additional experimental information about neurochemical changes that may underlie treatment-related structural and functional alterations or about the metabolism and physiologically based toxicokinetics of the test compound will help minimize the uncertainties involved in predicting human risk from neurotoxicity information derived from animal studies.*

2. Recommended Criteria for Designing a Neurotoxicity Screen

In developing any protocol for subacute (28-day), subchronic (90-day) or developmental toxicity studies, a specific design element of these studies should describe the collection of neurotoxicity screening data to detect any neuropathological or functional disorders of the central, peripheral or autonomic components of the nervous system. The neurotoxicity screen, at a minimum, should include both neurological and neurofunctional analysis.

- A specific histopathological examination should be made of tissue samples representative of all major areas and elements of the brain, spinal cord and peripheral nervous system.
- A functional evaluation battery of quantifiable observations and manipulative tests should be selected to detect signs of neurological, behavioral and physiological dysfunctions. In addition to the animal's appearance and body posture, a battery to screen for nervous system dysfunction would include sufficient information to assess such endpoints as the incidence and severity of seizure, tremor, paralysis or other signs of neurological dysfunction; the level of motor activity and arousal; the animals' level of reactivity to handling or other general stimuli; motor coordination and strength; gait; sensorimotor response to primary sensory stimuli; excessive lacrimation or salivation; piloerection; diarrhea; polyuria; ptosis; and any other sign of neurotoxicity.

In carrying out the functional evaluation screen, animals should be initially observed in their home cages and then removed to an open arena for the completion of the observations and testing.

To help ensure the complete and consistent application of the neurotoxicity screen throughout a particular study, each protocol should include a detailed description of the particular screen to be used in that study, including its composition, the test procedures to be followed, the time periods at which the screen is to be carried out, the neuronal structures to be examined, the endpoints to be used, and the methods for recording and analyzing the data generated. Since neurotoxicity screening is intended to be a routine part of both general and developmental toxicity studies, the specific composition of the screen and the endpoints to be recorded should accommodate the focus of the different protocols and, specifically, be appropriate for the age of the animals to be tested. For example, when testing immature developing animals, it would be appropriate to include indices of the ontogenetic development of representative neuroanatomical, physical and functional milestones as part of the neurotoxicity screen.

V C Neurotoxicity Studies Continued

There are a number of available publications to guide the design and conduct of neurotoxicity screens appropriate for the adult organism^a and the developing and adult offspring.^b The petitioner is encouraged to consider developing such protocols in consultation with FDA. The following points are of particular concern to the FDA:

- The neurotoxicity screen should consist of valid test methodology administered by trained personnel. Each testing laboratory should have demonstrated competence in the conduct of neurotoxicity evaluations. The testing laboratory should demonstrate the reliability and sensitivity of the screen to detect neurotoxic effects by the submission of historic or concurrent positive control data.
- The screen should be carried out with the experimenters blind to treatment condition, as appropriate, and should be applied to a sufficient number of male and female animals from each experimental and control group of the toxicity study to ensure valid statistical analyses. The latter involves consideration of the variability of the endpoints being measured. At the discretion of the test laboratory, satellite groups of animals may be used to carry out the neurotoxicity screen.
- During the conduct of the studies, the functional evaluation battery of observations and tests should be carried out systematically, using a prepared checklist when appropriate. All data should be recorded. Testing should be carried out at representative intervals throughout the duration of the study to provide information about the consistency of the neurotoxic effect(s), as well as their onset, duration and reversibility. The experimental design should appropriately control for potentially confounding variables, such as housing conditions, diet and nutritional status, circadian cycles, test to test interactions, handling, and environmental conditions.
- All of the data developed in the neurotoxicity screen, including positive and negative results, should be recorded, reported, and analyzed using appropriate statistical procedures. This information, together with any other pertinent toxicology data, should be incorporated into an integrated assessment of the potential for the test chemical to adversely affect the structural or functional integrity of the nervous system. Based on this assessment, an explicit statement should be made as to whether or not the test chemical represents a potential neurotoxic hazard which may require special neurotoxicity testing. Study protocols for additional neurotoxicity testing should utilize valid state-of-the-art methodology and should be developed in consultation with appropriate FDA personnel.

^a Leukroth (1987);² Vorhees (1986);⁴ U.S. Office of Technology Assessment report (1990);⁵ World Health Organization report (1986);⁶ Buelke-Sam *et al.* (1985);⁹ U.S. EPA report (1985);¹⁶ U.S. EPA report (1991);¹⁷ National Research Council, National Academy of Sciences report (1984);¹⁹ Tilson and Mitchell (1992);²¹ Nelson (1991)²⁵ Tupper and Wallace (1980);²⁶ Deuel (1977);²⁷ Gad (1982);²⁸ Gad (1989);²⁹ Marshall *et al.* (1971);³⁰ Spencer *et al.* (1980);³¹ O'Donoghue (1988);³² Moser (1988);³³ Wier *et al.* (1989);³⁴ U.S. EPA (1988);³⁵ Schultz and Boysen (1991);³⁶

^b Vorhees (1987);⁴ World Health Organization report (1986);⁶ Buelke-Sam *et al.* (1985);⁹ Vorhees *et al.* (1984);¹⁴ U.S. EPA report (1991);¹⁷ Kimmel *et al.* (1990);¹⁸ U.S. EPA report (1988);³⁵ Vorhees *et al.* (1979)³⁷

V D. Immunotoxicity Studies

Exposure to various chemicals has been associated with toxicity of the immune system in animals; these include environmental contaminants, chemicals in the occupational environment, and direct and indirect food additives. Regulatory agencies, including FDA,^a have recognized the importance of these types of effects for assessing the safety of chemicals to which humans may be exposed. Because of the rapid emergence of the field of immunotoxicology during the past two decades and the abundance of information that has accumulated with regard to the immune system as a target organ, various federal agencies and international organizations are preparing guidelines for the conduct of immunotoxicity studies.^b In addition, various testing approaches have been proposed by researchers in the field.^c

1. Immunity: A Brief Review

The immune system has been described in detail in a number of excellent reviews.^d Thus, only those aspects of immunity which are particularly relevant to immunotoxicity testing will be reviewed in this section. Immunological function encompasses a complex array of participating cell types and organ systems. Immunity may be defined in relation to the function of the various cellular components.

a. Humoral Immunity

Humoral immunity is defined in terms of the B-lymphocytes (B-cells), the antibody producing cells of the immune system. The B-cells, named because of their functional similarity to antibody-producing cells derived from the Bursa of Fabricius in birds, are found primarily in the spleen, lymph nodes, Peyer's patches in the gut, peripheral blood and bone marrow. The bone marrow is also the site of origination of B-cell precursors, the stem progenitor cells.

Immunoglobulins (Igs), the class of proteins that is comprised of the antibodies, are further classified with regard to particular peptide regions found on the light and heavy chains. At least five major classes of immunoglobulins have been defined for man and animals: IgA, IgD, IgE, IgG, and IgM. Antibodies function in concert with complement proteins that are produced in the liver and by macrophages to provide protection against bacterial and viral infections. Antibodies also help protect man and animals from agents that cause tumors and from some spontaneously occurring tumor cells.

Humoral immunity can be further classified with regard to the dependence of antibody production on T-lymphocyte help: T-cell dependent and T-cell independent immunities.

^a FDA (1978)¹

^b Luster *et al.* (1988);² U.S. EPA (1982);³ Sjoblad (1988);⁴ Tryphonas (1990);⁵ WHO (1986)⁶

^c Bick (1982);⁷ Dean *et al.* (1982);⁸ White *et al.* (1985);⁹ Exon *et al.* (1984);¹⁰ Vos *et al.* (1982);¹¹ van Loveren *et al.* (1988);¹² Schoental (1988)¹³

^d Dean and Thurmond (1987);¹⁴ Descotes and Mazuic (1987);¹⁵ Koller (1987);¹⁶ Irons (1989);¹⁷ Dean *et al.* (1986)¹⁸

V D Immunotoxicity Studies Continued

b. Cell-Mediated Immunity (CMI)

CMI derives its name from classical studies that demonstrated adaptive cell transfer of immunological function, graft v. host reactivity, etc. CMI is associated with the T-lymphocytes or T-cells (thymus-derived). Various classes of T-cells have been described, such as suppressors, helpers, inducers, and cytotoxic cells. Some of these T-cell types are involved in B-cell immunoregulation. T-cells secrete various peptide factors, referred to as lymphokines or cytokines, that modulate the activity of B- and T-cells. Cytotoxic T-cells participate in direct killing of invading microorganisms and tumor cells. T-cells are now commonly defined in terms of various membrane "antigens", such as T-4 (or CD4) for helper/cytotoxic cells and T-8 (or CD8) for suppressor/cytotoxic cells.

c. Non-Specific Immunity

Non-specific immunity is derived from other cell types that participate in the immune process. Natural killer (NK) cells are a group of cells that share certain properties with T-cells, but probably arise from different stem progenitor cells.^a These cells are known to play an important role in immune surveillance against spontaneous tumor formation. They also serve as a first line of defense, in cooperation with other phagocytic leukocytes (phagocytes or granulocytes), in the destruction of invading viruses and bacteria. Macrophages (activated monocytes) play a key role in antigen processing and presentation to lymphocytes; they interact with the T- and B-cells to facilitate antibody production. These cells also secrete cytokines, such as interleukin-1, which modulate certain T-cell functions.

Modulation of host resistance to infectious organisms can be the result of either direct or indirect effects on various cell components. Reduction in host resistance is referred to as immunosuppression. Severe or prolonged immunosuppression, as manifested in acquired immunodeficiency syndrome (AIDS), can result in an overwhelming number of infections, tumor formation, and death. Immune enhancement or hyperactivity of the immune system can result in hypersensitivities, such as allergic disorders and autoimmune diseases. The mechanisms of these disorders and diseases are complex and are dependent on factors such as genetic predisposition, age, medical condition, and environment. The development of autoimmunity, which has been associated with the use of various drugs,^b can have a pronounced toxic effect on a number of organ systems.

True allergic reactions, which are mediated mainly by IgE in man and certain animals, can result in a life-threatening condition known as anaphylactic shock. Certain food additives, such as sulfites, have been restricted in use because of their high sensitizing potential.^c Other food chemicals have been associated with hypersensitivity-like conditions such as the toxic oil syndrome^d and tryptophan-induced eosinophilia myalgia.^e

^a Herberman and Holden (1978)¹⁹

^b Bigazzi (1988)²⁰

^c Jacobsen and Gunnison (1987)²¹

^d Kammüller *et al.* (1988)²²

^e Belongia *et al.* (1990)²³

2. Key Concepts in Immunotoxicity Testing

These guidelines relate to the safety assessment of direct food additives and color additives used in food; such assessments are done on a case-by-case basis. The recommendations for immunotoxicity testing of food and color additives used in food presented in this section may or may not be relevant to those of other agencies and organizations. However, certain concepts from which these recommendations derive are shared by various others,^a including the World Health Organization.^b Other concepts may be unique to FDA, since these guidelines have been developed within the toxicity testing framework set forth in this book. These concepts are:

- Two types of immunotoxicity tests/procedures are defined: Type 1 Tests are those that do not require any perturbation of the test animal, such as immunization and challenge with an infectious agent.

- i) Primary indicators of immune toxicity are derived from Basic Type 1 Tests, such as hematology and serum chemistry profiles, routine histopathology examinations, and organ and body weight measurements from standard toxicity studies described in other sections of this book. Additional procedures, such as measurements of thymus weights and performance of more definitive histopathological evaluations of immune-associated organs and tissues, have been added.

- ii) Indicators of immune toxicity can also come from Expanded Type 1 Tests. These tests are logical extensions of Basic Type 1 tests; for example, Expanded Type 1 tests may extend the hematology, serum chemistry, and histopathology evaluations of standard toxicity studies. Many of these expanded tests can be performed with the same blood and tissue samples collected for the Basic Type 1 tests; in addition, many of the expanded tests can be performed retrospectively.

- Type 2 Tests include injections or exposure to test antigens, vaccines, infectious agents or tumor cells. If Type 2 tests are to be performed concurrently with a standard toxicity study, a satellite group of animals should be added to the recommended number of test animals in the study. Protocol designs for standard toxicity studies that include a satellite group of animals for Type 2 immunotoxicity tests will be recommended when available information indicates that a test compound may present an immunotoxic risk.

- Sets of Basic and Expanded Type 1 Tests are defined as Level I Immunotoxicity Tests. Some Level I tests screen for immunotoxic effects in test animals; others focus on defining an immunotoxic effect more specifically, such as determining the mechanism or cell types involved. Analogously, sets of Type 2 tests are defined as Level II Immunotoxicity Tests; Level II tests also can be used to screen for, or more specifically define, immunotoxic effects of food and color additives used in food.

^a Exon *et al.* (1984);¹⁰ Vos *et al.* (1982);¹¹ van Loveren *et al.* (1988);¹² Koller and Exon (1985)²⁴

^b WHO (1986)⁶

V D Immunotoxicity Studies Continued

3. Indicators of Possible Immune Toxicity

Basic Type 1 Tests: Primary Indicators

The primary indicators of possible immune toxicity are derived from routine measurements and examinations performed in toxicity studies recommended in other sections of this publication (Basic Type 1 tests). Indicators derived from short-term and subchronic toxicity studies, and developmental toxicity studies with rodents are listed below. If a substance produces one or more of these primary indicators of immune toxicity, more definitive immunotoxicity tests (Expanded Type 1 tests or Type 2 tests) may be recommended; such decisions will be made on a case-by-case basis.

a. Indicators from Short-Term and Subchronic Toxicity Studies

■ Hematology Indicators: Elevation or depression in white blood cell (WBC) counts; altered differential WBC counts; lymphocytosis and lymphopenia; and eosinophilia.

■ Clinical Chemistry Indicators: Elevated or reduced total serum protein in combination with an abnormal albumin-to-globulin (A/G) ratio. Other indicators often associated with immunologic dysfunction include abnormal levels of liver proteins and enzymes, such as albumin and the transaminases.

■ Histopathology Indicators: Abnormalities found during gross and routine histological evaluation of the lymphoid tissues, e.g. spleen, lymph nodes, thymus, gut-associated lymphoid tissue (GALT, in particular Peyer's patches), and bone marrow. Morphologic abnormalities such as scattered, focal mono-nuclear cell infiltrates in non-lymphoid organs (e.g. kidney and liver) may be relevant to autoimmune disease. If differences are seen in any lymphoid tissue, attention should be given to "cellularity" and prevalence of activated macrophages. The description could include *in situ* descriptions of the types of cells, density of the cell populations, lymphocyte distribution relative to distinguishing structures or defined areas of the organ. The histopathological analysis of routinely stained (hematoxylin and eosin) samples of the spleen should include descriptions of lymphocyte distribution and proliferation in known T- and B-cell areas, such as the germinal centers (for B-cells) and the periarteriolar lymphocyte sheath (PALS) for T-cells if abnormalities are observed. The histopathologic analysis of the lymph nodes and Peyer's patches should include a description of the immune activation (i.e. the relative number of follicles and germinal centers) when abnormalities or lesions are observed in these organs. When abnormalities of the thymus are observed, histopathologic analysis should be descriptive and quantitative as possible with regard to atrophy and necrosis and other observations. If the test compound is shown to either stimulate cell proliferation, or to cause atrophy and cell depletion in any lymphoid organ, the effect is likely to be viewed as a potentially immunotoxic effect requiring more definitive testing.

* In these instances, the effect does not need to be defined rigorously for each animal. The number of animals observed, however, should be a statistically significant sample size.

V D Immunotoxicity Studies Continued

■ Organ and Body Weight Indicators: Elevated or depressed spleen and thymus weights; elevated or depressed organ-to-body-weight ratios for the spleen and thymus (statistical treatment of the organ-to-body-weight ratios should include an analysis of co-variance, with body weight as the co-variant). Elevated or depressed body weights, although primarily an indicator of endocrine function, may also indicate indirect immunotoxic effects, since endocrine function can significantly effect the immune system.

b. Indicators from Developmental Toxicity Studies

■ Morbidity and Mortality Indicators: Unusual incidence of maternal infections.

■ Histopathology Indicators: Abnormalities found during gross evaluation of the fetal liver, spleen, and thymus.

■ For animals in the F₁ and F₂ generations:

i) Hematology Indicators: Elevation or depression in white blood cell (WBC) counts; altered differential WBC counts; lymphopenia and lymphocytosis; and eosinophilia.

ii) Clinical Chemistry Indicators: Elevated or reduced total serum protein in combination with an abnormal albumin-to-globulin (A/G) ratio.

iii) Histopathology Indicators: Abnormalities found during gross and routine histological evaluation of the lymphoid tissues, e.g. spleen, lymph nodes, thymus, gut-associated lymphoid tissue (GALT, in particular Peyer's patches), and bone marrow. Morphologic abnormalities such as scattered, focal mono-nuclear cell infiltrates in non-lymphoid organs (e.g. kidney and liver) may be relevant to autoimmune disease. If differences are seen in any lymphoid tissue, attention should be given to "cellularity" and prevalence of activated macrophages. The description could include *in situ* descriptions of the types of cells, density of the cell populations, lymphocyte distribution relative to distinguishing structures or defined areas of the organ. The histopathological analysis of routinely stained (hematoxylin and eosin) samples of the spleen should include descriptions of lymphocyte distribution and proliferation in known T- and B-cell areas, such as the germinal centers (for B-cells) and the periarteriolar lymphocyte sheath (PALS) for T-cells if abnormalities are observed. The histopathologic analysis of the lymph nodes and Peyer's patches should include a description of the immune activation (i.e. the relative number of follicles and germinal centers) when abnormalities or lesions are observed in these organs. When abnormalities of the thymus are observed, histopathologic analysis should be descriptive and quantitative as possible with regard to atrophy and necrosis and other observations. If the test compound is shown to either stimulate cell proliferation, or to cause atrophy and cell depletion in any lymphoid organ, the effect is likely to be viewed as a potentially immunotoxic effect requiring more definitive testing.

* In these instances, the effect does not need to be defined rigorously for each animal. The number of animals observed, however, should be a statistically significant sample size.

V D Immunotoxicity Studies Continued

iv) Organ and Body Weight Indicators: Elevated or depressed spleen and thymus weights; elevated or depressed organ-to-body-weight ratios for the spleen and thymus (statistical treatment of the organ-to-body-weight ratios should include an analysis of co-variance, with body weight as the co-variant). Elevated or depressed body weights, although primarily an indicator of endocrine function, may also indicate indirect immunotoxic effects, since endocrine function can significantly effect the immune system.

4. Expanded Type 1 Immunotoxicity Tests

Assessing the safety of food and color additives used in food usually requires the completion of various toxicity studies. In addition to the screen of primary indicators of possible immune toxicity provided by these toxicity studies and summarized above, additional tests for further evaluation of the immunotoxic potential of a test substance may be recommended by the Agency. The additional tests can be Expanded Type 1 Tests, discussed in this section, or Type 2 Tests, discussed in the next section. The Agency's recommendation that specific immunotoxicity tests be performed on test substances will be made on a case-by-case basis. Expanded Type 1 immunotoxicity tests include:

■ Hematology Tests: Flow cytometric analysis of B-lymphocytes, T-lymphocytes, and T-lymphocyte subsets (TH + TS or CD4 and CD8); immunostaining (immunoperoxidase or immunofluorescence) of B-lymphocytes, T-lymphocytes and T-lymphocyte subsets from peripheral blood or single cell suspensions from the spleen.^a

i) Hematology Indicators: Decreased or elevated percentages of any of the various lymphocytes relative to controls and abnormalities in the B-cell/T-cell and the TH/TS (CD4/CD8) cell ratios; these should be determined from differential counts of the immunostained preparations or from cytometric analysis.

■ Serum Chemistry Tests: Electrophoretic analysis of serum proteins to permit separation and quantification of the relative percentages of albumin and the α -, β -, and γ -globulin fractions; quantification of γ -globulin fractions (IgG, IgM, IgA, and IgE); analysis of total serum complement and components of complement (such as C3) from CH-50 determinations; immunochemical assay of serum cytokines, such as IL-2, IL-1, and γ -interferon; quantification of serum auto-antibodies, such as anti-nuclear, anti-mitochondrial, and anti-parietal cell antibodies.

i) Serum Chemistry Indicators: Statistically significant variations between experimental and control groups of animals for any of the parameters listed above.

■ Histopathology Tests: Immunostaining of B-lymphocytes in the spleen and lymph nodes, using polyclonal antibodies to IgG of the test animals;^b immunostaining of T-lymphocytes and T-lymphocyte subsets in the spleen, using monoclonal or polyclonal antibodies to various cell markers; micro-metric measurements of germinal centers and PALS of the spleen and the follicles and germinal centers of lymph nodes; morphometric analysis of the relative areas of the cortex and medulla of the thymus, using routinely stained histopathology sections.

^a Lovett *et al.* (1984);²⁵ Hudson *et al.* (1985);²⁶ Burchiel *et al.* (1987);²⁷ Hinton *et al.* (1973);²⁸ Falini and Taylor (1983)²⁹

^b Hinton *et al.* (1987);³⁰ Roginski and Hinton (1987)³¹

V D Immunotoxicity Studies Continued

i) Histopathology Indicators: Statistically significant variations between experimental and control groups of animals for any of the parameters listed above, using both analysis of variance (ANOVA) and a multiple comparison T-test, such as Dunnett's.^a

■ Tests for *In Vitro* Analysis of the Functional Capacity of Specific Cell Types:

i) Activity of Natural Killer (NK) Cells: The functional capacity of NK cells can be measured using the classical ⁵¹Cr chromium release assay;^b this assay is well standardized and has been used successfully with both mice and rats in various immunotoxicity studies.^c Of particular concern is reduced NK cell activity, which may be correlated with increased tumorigenesis and infectivity.

ii) Mitogenic Stimulation Assays for B- and T-Lymphocytes: Certain plant lectins stimulate blastogenesis and DNA synthesis of T- and B-lymphocytes: concanavalin-A (Con-A) and phytohemagglutinin (PHA) are known to preferentially stimulate T-lymphocytes, and an extract from pokeweed (PWM) as well as certain bacterial lipopolysaccharides (LPS) and protein extracts are known to preferentially stimulate B-lymphocytes *in vitro*. Since these assays are carried out *ex vivo*, they can be performed on preparations of peripheral blood. The assays are well characterized for use in various animals species (including man^d), can be performed on either peripheral blood or spleen-cell suspensions, and have been used in a number of immunotoxicity studies.^e Both reduced and elevated levels of blastogenesis or ³H incorporation into DNA are of interest in the evaluation of the immunotoxic potential of food and color additives used in food.

iii) Phagocytotic Index of the Macrophage: Various assays to determine the phagocytotic ability of macrophages have been described.^f These assays measure the ability of a macrophage to ingest particulate substances, such as plastic beads or iron filings, and can be performed on peripheral blood or single cell suspensions of lymphoid organs, such as the spleen. Other assays measure the capacity of the macrophage to destroy live bacteria through lysosomal enzyme activity.^g

^a Dunnett (1955)²²

^b Herberman and Holden (1978)¹⁹

^c Gorelik and Herberman (1981);³³ Smialowicz *et al.* (1989);³⁴ Holsapple *et al.* (1988)³⁵

^d Oppenheim and Rosentrich (1976)³⁶

^e Luster *et al.* (1988);² Dean *et al.* (1982);⁸ White *et al.* (1985);⁹ Exon *et al.* (1984);¹⁰ van Loveren (1988);¹⁷ Schoental (1988);³¹ Holsapple *et al.* (1988);³⁵ Cornacoff *et al.* (1988)³⁷

^f Koller and Exon (1985);³⁴ Loose *et al.* (1981);³⁸ Lewis and Adams (1985)³⁹

^g Keller (1978)⁴⁰

V D Immunotoxicity Studies Continued

iv) Stem Cell Assays: Bone marrow preparations can be used to investigate the pluripotent population or specific progenitor populations.^a Although these assays have not been used extensively in immunotoxicity evaluations, they may be recommended when histopathological evaluation indicates that the test substance may have caused changes in bone marrow.

5. Type 2 Immunotoxicity Tests

Evaluating the functional capacity of the immune system requires injecting a substance that elicits immunological reactivity in a test animal. Various antigens provide information about the types of immunity or cells that may be involved in an immune response. For example, protein antigens usually elicit T-dependent immune responses with subsequent production of antibodies to the protein. Polysaccharides elicit T-independent immune responses. Some antigens elicit cell-mediated immune responses, while immunogens such as complex bacteria and viruses may elicit humoral and cell-mediated responses. All of the antigens listed below have been tested in rodents; when an antigen has been used preferentially with a particular rodent species, this is noted.

■ T-Dependent Test Antigens: One of the most widely used antigens for rodents^b and non-rodents is sheep red blood cells (SRBC).^c For example, SRBCs have been widely used in mice in the Plaque-Forming Cell Assay:^d antibody-producing spleen cell suspensions are mixed with SRBCs, placed on covered slides, and incubated; each antibody-producing cell causes a small, clear area (plaque) to form on the slide; the plaques are then counted. Other T-dependent test antigens that have been widely used include keyhole limpet hemocyanin^e and bovine serum albumin.

■ T-Independent Test Antigens: Ficoll, a branched chain polysaccharide, haptenated ficoll, polyvinylpyrrolidone, and bacterial lipopolysaccharides have been used as T-independent test antigens with mice and rats.^f

^a Harigaya *et al.* (1982)⁴¹

^b Luster *et al.* (1988);² Bick (1982);⁷ Dean *et al.* (1982);⁴ White *et al.* (1985);⁹ Exon *et al.* (1984);¹⁰ Koller and Exon (1985);²⁴ Hinton *et al.* (1987);³⁰ Pestka *et al.* (1987);⁴² Reddy *et al.* (1987)⁴³

^c FDA, (1978);¹ Sjoblad (1990);⁴ Tryphonas (1990);⁵ WHO (1986);⁶ Bick (1982);⁷ Dean *et al.* (1982);⁸ White *et al.* (1985);⁹ Hinton *et al.* (1987);³⁰ Smialowicz *et al.* (1989);³⁴ Holsapple *et al.* (1988);³⁵ Pestka *et al.* (1987);⁴² Baecher-Steppan *et al.* (1989)⁴⁴

^d Jerne and Nordin (1963);⁴⁵ Cunningham (1965)⁴⁶

^e Exon and Koller (1984)¹⁰

^f Anderson and Blomgren (1971)⁴⁷

V D Immunotoxicity Studies Continued

■ Human Vaccines: Human T-dependent vaccines, such as tetanus toxoid, and the T-independent vaccine containing pneumococcal polysaccharide antigens have been used in both rats and mice.^a It is possible to compare responses of the test species to the vaccines with human responses, because standard human sera are available from FDA's Center for Biologics.^b

■ Test Antigens for Cell-Mediated Immune (CMI) Reactivity: Contact sensitizers such as dinitrochlorobenzene (DNCB) have been used to elicit delayed hypersensitivity (DTH) responses as a measure of CMI in animals. These assays can be performed in rodent^c as well as non-rodent species. The DTH assays are economical and correlate well with decreased CMI and host resistance to infectious agents in humans,^d as well as animals.^e The mixed-lymphocyte response (MLR) assay, which uses lymphocytes from a different strain, has been successfully used to evaluate CMI in mice.^f

■ Host Resistance Assays with Infectious Agents: A number of bacterial strains have been used to measure host resistance, including *Listeria monocytogenes*, various strains of *Streptococcus*, and *Escherichia coli*.^g Useful viral models^h include influenza, herpes, and cytomegalovirus.ⁱ A yeast infectivity model using *Candida albicans* has been described, as well as parasitic infectivity models using *Trichinella spiralis* and *Plasmodium yoelii*.^j

^a Vos (1977);⁴⁰ Spiers *et al.* (1979);⁴⁶ Benson and Roberts (1982)⁵⁰

^b Schiffman (1982)⁵¹

^c Godfrey and Gell (1978)⁵²

^d Maclean (1979)⁵³

^e Bradley and Morahan (1982)⁵⁴

^f Lauster *et al.* (1988)⁵

^g Bradley and Morahan (1982)⁵⁴

^h Dempsey and Morahan (1985);⁵⁵ Kern (1982)⁵⁶

ⁱ Selgrade *et al.* (1988)⁵⁷

^j Dempsey and Morahan (1985)⁵⁴ Dean *et al.* (1982)⁵⁸

V D Immunotoxicity Studies Continued

■ **Host Resistance Assays Using Syngeneic Tumor Cells:** Various assays of host resistance have been described using a number of cultured tumor cell lines.^a These assays, unlike those involving the infectious agents discussed above, do not require special barrier facilities to prevent infections from spreading throughout an animal colony. Two mouse assays have been validated: the PYB6 sarcoma assay and the B16F10 melanoma assay.^b An assay using a lung tumor model and the MADB106 tumor cell line also has been validated for use in immunotoxicity studies.^c

6. Relevance of Primary Indicators of Immune Toxicity to Health

a. Hematological Indicators

Hematologic screens recommended for toxicity studies are basically the same as those performed clinically as human health screens. Depressed or elevated WBC counts may be indicative of direct or indirect effects of the test substance on cellular proliferation and distribution. Total WBC counts are used clinically as a presumptive test for infection; they are also used to evaluate the severity of an inflammatory or allergic process. Routine differential WBC counts are used to differentiate among some types of infections and inflammatory responses; they also are used as a screen for toxicologic or pharmacologic effects: for example, immunosuppressive drugs may cause lymphopenia.

Altered lymphocyte counts may be relevant to immunodeficiency. Increased numbers of polymorphonuclear leukocytes can result from pathogenic infections and from pyrogenic and inflammatory processes. Eosinophilia is often associated with allergenic processes. It may also indicate an infectious, reactive, or neoplastic process. Altered red blood cell counts and platelet counts can be associated with autoimmune processes.

b. Serum Protein Indicators

Estimates of total serum proteins and the albumin/globulin ratio may give useful information about liver and lymphocyte function. The α - and γ -globulins (i.e. α - and γ -G) are primarily produced in the liver; γ -G are a product of the B-lymphocytes. Depressed α -G levels could lead to decreases in complement proteins that are required for phagocytosis; this could produce decreased resistance to bacterial infections. Reduced levels of γ -G also could mean reduced levels of antibodies necessary for humoral immunity to infectious agents. Altered levels of γ -G may indicate an effect on B-lymphocytes, T-lymphocytes, or simultaneous effects on both types of cells.

However, total globulin levels do not give specific information about which immunoglobulin classes are affected. Thus, when globulin levels are reduced, specific quantitative assays for the γ -G subclasses may be recommended. Electrophoretic and immunoelectrophoretic analyses of the serum γ -G subclasses or quantitative assays such as Enzyme-Linked Immunosorbent Assay (ELISA), Radioimmunoassay (RIA), or radial immunodiffusion may be recommended. This information may be important because reductions in γ -G and γ -M may be relevant to infection by opportunistic and

^a Dean *et al.* (1982);⁵⁸ Herberman (1985)⁵⁹

^b Murray *et al.* (1985)⁶⁰

^c Smialowicz *et al.* (1987)⁶¹

V D Immunotoxicity Studies Continued

pathogenic organisms, and changes in γ -A may indicate effects of the test substance on secretory immunity, such as gut-mediated immunity.

c. Histopathology Indicators

Abnormal results from gross and histological evaluation of the lymphoid organs (usually the spleen, thymus, and lymph nodes) are important indicators of various immunotoxic effects; histological evaluation of Peyer's patches and bone marrow also is recommended in screening for effects of a test substance on the immune system. Atrophy of the thymus gland with associated depletion of cortical thymocytes could be an indication of immunosuppression. Concomitantly, a similar effect on the lymphocytes in the periaarterial lymphocyte sheath of the spleen (PALS) would indicate an effect of the test substance on T-cells; both cell-mediated and humoral immunity can be affected. In the spleen and lymph nodes, defined regions are more densely populated with B-lymphocytes, with activated, antibody-producing B-cells, or with plasma cells. Effects on B-cell regions of these organs could be an indication of immunosuppression or immunoenhancement, depending on the result obtained.

d. Body and Organ Weights

Body and organ weights are generally recorded during toxicity studies. Spleen weights are usually recorded in all toxicity studies, but thymus weights may not be recorded in long-term studies. The thymus gland grows rapidly in young animals but begins to involute as the animals reach sexual maturity. In old animals, the thymus may be difficult to detect and measure because of the degree of involution.

Organ weights by themselves or in relation to body weights can be sensitive measures of organ atrophy or hypertrophy, but yield little information about immunotoxic effects. Reduced organ weights can result from direct effects on lymphocyte proliferation and differentiation and may be relevant for assessing immunosuppression. Hypertrophy of the lymphoid organs is usually associated with increased proliferation of cells (hyperplasia). Increased proliferation of lymphocytes can result from infections, stimulation by xenobiotics, altered metabolic processes, and certain forms of trauma, reactive, or autoimmune processes. In practice, however, changes in organ weights or organ-to-body-weight ratios are more relevant to immune toxicity when they are associated with appropriate histopathology findings.

7. Adequacy and Reliability of Primary Indicators of Immune Toxicity

If all primary indicators of possible immunotoxicity from toxicity studies are negative for a test substance, would this effectively rule out the possibility that the test substance produces significant immunotoxic effects? The answer to this question is complex; some of this complexity derives from the fact that the primary indicators of possible immune toxicity listed above are not sufficiently specific or sensitive to provide unambiguous answers. For example, it is not possible to differentiate B-lymphocytes from T-lymphocytes in routinely stained sections of lymphoid tissues, and standard hematology tests cannot distinguish among subcategories of T-lymphocytes. Special immunochemical stains, however, permit B- and T-cells to be visualized in tissue sections and blood smears, making available more information about the immunotoxic effects of the test substance.

If only short-term toxicity studies are performed on a particular test substance, concern about the adequacy and reliability of the immunotoxic indicators from these studies may be high. Subtle immunotoxic effects or immune toxicities that develop only after prolonged administration of the test

V D Immunotoxicity Studies Continued

substance may not be detected in short-term toxicity studies. Conditions of the longer-term toxicity studies, however, may make it difficult to detect some immune toxicities: the use of barrier facilities is common in carcinogenicity studies; because barrier facilities limit exposure of test animals to exogenous infectious microorganisms, detecting possible immunotoxic effects of a test substance in carcinogenicity studies may be compromised because spontaneous infection rates and mortality are evaluated as primary indicators of possible immunotoxicity in such studies.

Even with this disadvantage, many investigators and regulatory authorities recommend specific tests to identify and characterize immune system toxicities only when screening tests or indicators are positive.^a Additional rationale for this approach comes from the fact that most short-term toxicity studies incorporate at least one dose in the potentially highly toxic dose range. Additional tests for immunotoxicity should be performed to verify positive immunotoxic effects noted during screening studies or to determine if the positive result obtained for a primary indicator was a false positive indication of immunotoxicity. For example, certain test substances may cause increased or decreased food intake; nutritional deprivation from significantly decreased food intake has been shown to cause thymic and splenic atrophy.^b Effects on the endocrine system, such as stimulation of the production of growth hormone^c or prolactin^d and decreased levels of adrenocorticosteroids,^e can stimulate growth of the thymus. In response to such stimuli, involution of the thymus may proceed at a different rate in animals exposed to the test substance than in control animals. Therefore, measuring thymic weights at one specified time in a short-term toxicity study could give false positive or false negative indications of the test compound's immunotoxic potential. For this reason, the Agency recommends that a study of the effects of a test substance on thymic growth and involution be conducted at two or more time points during the study (such as midterm and final sacrifice). Because sex differences have been demonstrated for various immunologic studies,^f both sexes should be included in immunotoxicity evaluations.

There are data which suggest that the primary indicators do not evaluate toxic effects on all types of immune-related cells. Recent studies have shown that NK cell function may be affected without concomitant effects on either B- or T-lymphocytes.^g Other studies have shown that functional deficits of specific lymphocytes can occur without apparent changes in the proliferation or morphology of the cells as observed in standard histopathology preparations;^h the morphology of the cells is normal and a false negative result would be obtained in these instances.

^a WHO (1986);⁶ Vos *et al.* (1982)¹¹

^b Katz (1978)⁸²

^c Pierpaoli and Sorkin (1969)⁶¹

^d Berezi (1986)⁶¹

^e Claman (1975)⁶¹

^f Hinton *et al.* (1987);³⁰ Weimer and Roberts (1972)⁶⁶

^g Sniawicz *et al.* (1987)⁶¹

^h Vos *et al.* (1982);¹¹ Vos *et al.* (1983);⁶⁷ House *et al.* (1985)⁶⁸

V D Immunotoxicity Studies Continued

8. Recommendations for Further Immunotoxicity Testing when Primary Indicators of Immunotoxicity are Positive

Assessing the safety of food and color additives used in food usually requires the completion of various toxicity studies. In addition to the screen of primary indicators of possible immunotoxicity provided by these toxicity studies and summarized above, additional tests for further evaluating the immunotoxic potential of a test substance may be recommended by the Agency. In the sections that follow, the adequacy of primary indicators of immunotoxicity for test substances that have been assigned to each Concern Level will be discussed. The Agency's recommendation that specific immunotoxicity tests be performed on test substances that have been assigned to Concern Levels I, II, and III will be made on a case-by-case basis.

a. Immunotoxicity Studies for Compounds that have been Assigned to Concern Level III

Test compounds that have been classified as Concern Level III substances present the highest level of concern about their safe use as direct food additives and color additives used in food. When these substances undergo toxicity testing, primary immunotoxicity indicators may be negative, marginal, or positive. Immunotoxicity tests suitable for each of these situations will be discussed below.

i. Immunological Tests when Primary Indicators of Immunotoxicity are Negative or Marginal

If the primary immunotoxicity indicators from recommended toxicity tests are not positive, then no additional tests for the immunotoxic potential of the Concern Level III test compound would be recommended unless there were special circumstances. Such circumstances may include: 1) the rodent strains employed in toxicity testing were highly inbred and are known to be resistant to immunotoxic effects; 2) barrier or other facilities were used for long-term and short-term toxicity studies, which may have precluded exposure of the test animals to normal infectious agents present in the environment; and 3) omissions from the recommended guideline for standard toxicity tests, such as not measuring thymus weights during the growth phases of the test animals or omitting histopathological analysis of certain lymphoid organs. In these situations, some Type 1 immunotoxicity tests and a Type 2 immunotoxicity study of host resistance may be recommended, particularly if specific tests for immune toxicity had not been incorporated into subchronic toxicity studies.

ii. Immunological Tests when Primary Indicators of Immunotoxicity are Positive

When any of the primary indicators suggests that a Concern Level III test substance has an immunotoxic effect, additional testing will be recommended in order to assess the extent of risk to the immune system. In addition, positive effects on other target organs may indicate the need to assess the autoimmune potential of the compound.

Certain indicators may derive from effects on either B-cells or T-cells, or may be derived from effects on both types of cells. However, most of the primary indicators of immune toxicity are nondiscriminating with respect to specific lymphocytes involved and specific immune functions affected. Standard histopathology evaluation may provide some clues if there is an effect on the thymus or if areas in the spleen or lymph nodes are associated with specific types of lymphocytes. The objectives of expanded Levels I and II immunotoxicity tests are to delineate the specific cells

V D Immunotoxicity Studies Continued

type(s) which are affected, to evaluate the extent to which specific immune functions are impaired, and to relate these effects to risks such as infection, hypersensitivity, and carcinogenicity.

The immunotoxicity tests described in the following sections are for use with rats, and all tests should be conducted on each test animal. However, tests have been, or can be, adapted for use with mice or non-rodent species. When mice are used as test animals, serum from animals in each experimental group may need to be pooled if there is an insufficient quantity of serum from each animal to perform recommended hematology tests.

- a) Retrospective Level 1 Tests: No additional animals are needed for Retrospective Level 1 immunotoxicity tests when at least 10 animals of each sex are used in a standard toxicity study and appropriate samples of blood and tissues are properly treated and preserved. After removing blood cells, serum samples should be prepared by high-speed centrifugation, sterilized by filtration, and stored at 4-5°C in sealed containers. At least half of each lymphoid tissue/organ should be fixed briefly in Bouin's fixative (or other fixative shown to be appropriate) and stored in alcohol; sections from the tissue/organ can be processed for histopathological analysis by routine staining or by immunostaining.

If the standard toxicity study was a subchronic or chronic study (with exposure to the test substance for 90 to 120 days), and primary indicators suggested that the test material may be immunotoxic, the following Retrospective Level 1 tests should be performed on serum samples from the study:

- i) Electrophoresis of serum proteins.
- ii) Quantification of serum immunoglobulins (IgG, IgM, IgA, IgE).
- iii) Immunostaining for B- and T-lymphocytes in spleen and lymph nodes and micrometric analysis of the number of stained cells in specific regions of these organs.
- iv) Screening for serum autoantibodies to DNA, mitochondria, and other cell components in one or more tissues, such as liver and smooth muscle.* These tests should be performed when there is an indication that the test substance may affect B- or T-lymphocytes.
- v) Immunostaining for bound IgG may be recommended to determine if non-lymphoid organ toxicities noted during the standard toxicity study (particularly a long-term toxicity study) are due to an autoimmune reaction.

If the results of these Retrospective Level 1 tests demonstrate that the primary indicators of immune toxicity were false positives, then no further immunotoxicity testing would be recommended. However, if the results of these tests are inconclusive or confirm an immunotoxic effect of the test substance, additional testing would be recommended. The additional testing may include Type 1 and Type 2 immunotoxicity tests.

* Hinton (1993)⁶⁹

V D Immunotoxicity Studies Continued

b) Additional Level 1 Tests: Additional Level 1 immunotoxicity tests cannot be performed retrospectively, but must be incorporated into the protocol of standard toxicity studies. However, all of the tests described in this section can be performed on the same animals that are used in the standard toxicity study, provided that samples are processed appropriately. For example, half of the spleen can be used to make a cell suspension for cellular analysis immediately following sacrifice of the test animal; the remaining half can be processed for histopathology evaluation. Additional (non-retrospective) Level 1 tests that may be recommended include:

- i) Quantitative analysis of the B-cell to T-cell ratio (B/T) using either whole blood cells and spleen preparations or spleen preparations only.
- ii) Determination of spleen cellularity (the total number of white blood cells and lymphocytes per gram of wet tissue) and the total number of viable cells per gram of wet tissue or per million white blood cells.
- iii) Assay of mitogenic stimulation for B- and T-lymphocytes:
- iv) Analysis of NK cells using a suspension of spleen cells:
- v) Determination of the phagocytotic index of macrophages:
- vi) Electrophoresis of serum proteins: Although this test can be performed retrospectively, it is listed here because it is particularly useful for evaluating toxic immune effects on liver, macrophages, and lymphocytes.

c) Level II Tests: If primary indicators of immunotoxicity from standard toxicity studies suggest that a test compound may be immunotoxic, Level II tests may be recommended to identify specific functional immune defects. These tests may be performed on satellite groups of test animals in conjunction with a standard toxicity study or they may be performed on test animals in a separate immunotoxicity study. In the latter case, Level II tests should be performed with the same species, strain and age of test animals and the same doses of test substance used in the standard toxicity study of comparable duration. In addition, separate Level II immunotoxicity studies should be 3 to 6 weeks in duration so that test animals will be exposed long enough to enable primary and secondary immune effects to be identified. An additional period of time at the end of the study during which the test substance is not administered would permit evaluation of the reversibility of observed immune effects.

The following Level II tests may be recommended:

- i) Kinetic evaluation of primary and secondary immune responses of test animals to a T-dependent antigen, such as SRBC, tetanus toxoid, or KLH; serum antibody titers should be measured following initial and secondary injections of the antigen.
- ii) Evaluation of the primary humoral response to a T-independent antigen, such as pneumococcal polysaccharides; choice of the optimum challenge dose should be justified.
- iii) Evaluation of the delayed hypersensitivity response to a contact sensitizer during the second half of the study. Alternatively, evaluation of the mixed lymphocyte response can substitute for measurement of the DTH response as long as the assay has been validated with the particular rat strain used.

V D Immunotoxicity Studies Continued

d) **Enhanced Level II Tests:** These tests are designed to determine if a test substance that produces immune toxicity in Level I or Level II tests also affects host resistance to challenge with infectious agents or tumor cells. Enhanced Level II tests may be performed with either rats or mice, because many host resistance tests have been validated in mice. These tests would be recommended in a variety of circumstances; for example:

- i) If primary indicators of immunotoxicity from standard long-term toxicity studies showed increased mortality associated with administration of the test substance and effects on humoral immunity were identified from Level I and Level II tests, then bacterial (e.g. *Listeria monocytogenes*)^a or viral (e.g. *Influenza*)^b challenge tests associated with humoral immune protection would be recommended for evaluation of host resistance.
- ii) If there are indications that consumption of the test substance is associated with increased tumorigenesis and effects on phagocytosis, tumor challenge tests with PYB6 sarcoma, which tests cytolytic activity of T-cells and NK cells in mice,^c would be appropriate; a similar test for rats uses the MADB106 tumor line.^d
- iii) Finally, for test materials that have demonstrated T-cell or cell-mediated immune effects, challenge tests that use certain strains of *Streptococcus*^e or *Plasmodium yoelii*^f would be appropriate.

b. Immunotoxicity Studies for Compounds that have been Assigned to Concern Level II

Specific immunotoxicity tests generally are not recommended for test compounds that have been assigned to Concern Level II. However, if primary indicators of possible immunotoxicity from toxicity studies conducted on Concern Level II test substances are positive, additional Level I and Level II immunotoxicity tests may be recommended; such recommendations will be made on a case-by-case basis.

c. Immunotoxicity Studies for Compounds that have been Assigned to Concern Level I

Usually, short-term acute exposure studies (up to 30 days) are performed to assess the safety of Concern Level I compounds. Although guidelines for these studies (see Chapter IV C 3) do not recommend specific immunotoxicity tests, if primary indicators of possible immunotoxicity from short-term toxicity studies are positive, additional Level I and Level II immunotoxicity tests may be recommended. Such recommendations will be made on a case-by-case basis. One immunotoxicity test which measures the primary humoral response to the T-dependent antigen SRBC has been

^a Dean *et al.* (1982)⁵⁸

^b Dempsey and Morahan (1985)⁵⁵

^c Murray *et al.* (1985)⁶⁰

^d Smialowicz *et al.* (1987)⁶¹

^e Selgrade *et al.* (1988)⁵⁷

^f Dean *et al.* (1982)⁵⁸

V D Immunotoxicity Studies Continued

described for use with both rats^a and mice^b and has been recommended for use in short-term screening studies.^c

9. Animal Models for Immunotoxicity Tests

a. Rodent Models

These guidelines have focused on tests designed to assess immune toxicity in the rat. Specific strains have been used and validated by the Agency, including Sprague-Dawley, Spartan,^d and Osborne Mendel;^e the Fisher 344 rat has been recommended by others^f for studies with environmental compounds. Other strains of rat, such as the Buffalo strain, have been used in special studies to evaluate autoimmune disease potentiation.^g In addition, several mouse strains (mainly inbred strains) have been used to assess immune toxicity.

b. Non-rodent Models

Use of the dog for various immunopharmacologic studies has been described in the scientific literature.^h Level I immunotoxicity tests described in these guidelines can be performed on most large animal species; Level II immunotoxicity tests in other non-rodent models also may be acceptable, if validated: use of primates has been described.ⁱ Also, miniature swine have been shown to be an excellent non-rodent species for evaluation of various immune functions.^j The Agency has validated a number of immune function assays for use with this model.

^a Smialowicz *et al.* (1989)³⁴

^b Luster *et al.* (1988)³ Dean *et al.* (1982)⁴

^c Luster *et al.* (1988)²

^d Hinton *et al.* (1987)³⁰

^e Hinton (1991)⁶⁹

^f Smialowicz *et al.* (1987)⁶¹

^g Weening *et al.* (1978);⁷⁰ Michaelson *et al.* (1981);⁷¹ Silverman and Rose (1975)⁷²

^h Thiem *et al.* (1988)⁷³

ⁱ Tryphonas *et al.* (1989)⁷⁴

^j Siegel (1984)⁷⁵ Glocklin (1987);⁷⁶ Hinton and Kahn (1987);⁷⁷ Hinton and Kahn (1989);⁷⁸
Hinton *et al.* (1993)⁷⁹

V D Immunotoxicity Studies Continued

Immunomodulation of porcine as well as other food animals have been reviewed.^a Other perspectives on animals selection have been reviewed.^b

10. Recommended Strategy for Assessing the Immunotoxic Potential of Direct Food Additives and Color Additives Used in Food

- Primary indicators of immunotoxicity should be evaluated for short-term (28-day) toxicity studies, subchronic (90-day) toxicity studies, and developmental toxicity studies. Results of these evaluations should be incorporated into an integrated assessment of the potential for a test chemical to adversely affect the immune system. Based on this assessment, an explicit statement should be made as to whether or not the test chemical represents a potential immunotoxic hazard which requires further testing.
- Additional studies to assess the immune toxicity of food and color additives used in food will depend on the results of the evaluation of primary indicators of immune toxicity, the Concern Level to which the additive has been assigned, and other information available concerning the immunotoxicity of the additive.

11. Conclusion

The hierarchical grouping of recommended immunotoxicity tests by specificity and mechanics (e.g. tests that use injectable substances) can facilitate including immunotoxicity testing in standard toxicity studies. Expanded testing on existing samples is possible, and allows for a more definitive identification of potential immunotoxic effects. Such expanded testing may be necessary when additional information about a possible immunotoxic effect is important for the safety assessment of a direct food additive or color additive used in food. Immunotoxicity tests recommended in this section are summarized in Table 2 below.

^a Siegel (1984)²⁵

^b Glocklin (1987)²⁶

Table 2

Summary of Immunotoxicity Testing Recommendations for Direct Food Additives

Basic Testing (Rat Model)

- CBC, WBC differential
- Total serum protein, albumin-to-globulin ratio (A/G)
- Histopathology, gross and microscopic (spleen, thymus, lymph nodes, Peyer's patches, and bone marrow)
- Lymphoid organ and body weights

Retrospective Level I Testing: Included as a Possible Requirement in Standard Toxicity Study

- Electrophoretic analysis of serum proteins* (when positive or marginal effect noted in basic testing)
- Immunostaining of spleen and lymph nodes for B and T cells* (quantification of total Ig)
- Serum autoantibody screen and deposition of Ig (micrometry for semi-quantitation of the proliferative response)

Enhanced Level I Testing: Included as a Possible Requirement for More Complete Screening in the Standard Toxicity Study Core Group, with a Satellite Animal Group, or in a Follow-Up Study

- Cellularity of spleen (lymph nodes, thymus when indicated)
 - Quantification of total B and T cells (blood and/or spleen)
 - Mitogen stimulation assays for B and T cells (spleen)
 - NK functional analysis (spleen)
 - Macrophage quantification and functional analysis (spleen)
 - IL-2 functional analysis (spleen)
- When indicated or for more complete analysis, other endpoints such as total hemolytic complement activity or CH-50 assay with serum

Level II Testing: Includes a Satellite Group or Follow-Up Study for Screening of Functional Immune Effects

- Kinetic evaluation of the humoral response to a T-dependent AG (primary and secondary responses with either SRBC, TT, or other)
- Kinetic evaluation of the primary humoral response to a T-independent AG such as P_vax, TNP-LPS, or other recognized AG
- DTH response to known sensitizer of known T-cell affecter
- Reversibility evaluation

Enhanced Level II Testing: Includes a Satellite Group or Follow-Up Study For Evaluation of Potential Immunotoxic Risk

- Tumor challenge (MADB106 or other with the rat; PYB6 sarcoma with a mouse model)
- Infectivity challenge (*Trichinella*, *Candida* or other with the rat; *Listeria* or other with the mouse)

Abbreviations: CBC = complete blood count; WBC = white blood count; Ig = immunoglobulin; NK = natural killer; IL-2 = interleukin-2; SRBC = sheep red blood cells; and TNP-LPS = trinitrophenol lipopolysaccharide.

* Recommended for inclusion in basic testing.

Chapter VI

Human Studies:

This chapter presents general guidelines for the conduct of human clinical studies on foods and food ingredients. It also describes the types of human epidemiology data that may be useful to the Agency in assessing the safety of direct food additives and color additives used in food. Because human clinical studies were not included in the 1982 guidelines for direct food additives and color additives used in food, important issues related to these studies are discussed at length in this chapter.

The Agency does not require petitioners to conduct human clinical studies to support the safety of direct food additives and color additives used in food. However, petitioners may elect to perform such studies in certain circumstances, such as when the proposed additive will be consumed by humans at relatively high levels (see Chapter VII B). When petitioners conduct human clinical studies on substances intended for use as direct food additives and color additives used in food, however, the Agency recommends that the studies conform to the guidelines presented in this section. As usual, the Agency strongly recommends that petitioners planning to conduct human studies in support of the safety of direct food additives and color additives used in food consult with the Agency before the studies begin.

VI A. Clinical Evaluation of Foods and Food Additives

A major objective in the clinical testing of food and food additives is to assess aspects of safety that cannot be addressed adequately by non-human studies or by existing data on population exposure. For example, the Agency is now reviewing petitions for direct food additives that are intended to substitute for major nutrients such as fat and sugar. Because segments of our population may be exposed to large quantities of these additives for long periods of time, traditional methods of evaluating the safety of these substances may not be adequate. Testing these substances in animals at doses that greatly exaggerate their anticipated human exposures may not be possible. For these substances, human clinical studies may provide additional confidence in the safety of the food or food additive.

A food or food additive generally will be considered suitable for clinical testing if the substance is unlikely to produce significant toxic effects at the levels to which the subjects of the clinical study will be exposed. This usually is determined from the results of toxicity studies in animals or by examining existing data on population exposure. However, in cases where the type of toxic response associated with the consumption of a food or food additive by experimental animals is judged to be severe, exposure of subjects in clinical studies to the additive may need to be significantly below the level found to produce no toxic effects in an appropriate species.

Unlike patients participating in clinical trials of new drugs, no health benefit is anticipated for most test subjects in clinical studies of foods and food additives. Thus, the nature and weight of evidence required to establish the safety of these products for humans before clinical studies can begin may differ from that required to support testing under guidelines for investigational new drugs. Clinical studies of foods and food additives will focus on demonstrating safety; for example, the safety of an additive that may interfere with absorption of nutrients, whose status in the population is uncertain, may need to be evaluated in a clinical study.

1. General Considerations for Clinical Studies of Foods and Food Additives

Principles for the conduct of clinical trials are contained in the February 22, 1985 Federal Register¹ and codified in 21 CFR 314.126. The following guidelines identify general considerations for clinical studies of foods and food additives. Each consideration should be explicitly addressed in the clinical study's protocol.

- Before undertaking costly and time-consuming clinical studies as part of the safety assessment of a food or food additive, the investigator needs to formulate a defensible rationale for conducting human clinical studies and a clear set of objectives.
- Adequate preclinical investigations (including toxicity tests in animals) must have been completed. Results of these tests must establish that there is no expected toxicity to man at doses to be used in clinical studies. A clear, concise description of the design of pre-clinical studies and their results should be presented to FDA. Information about the history of use of the food or food additive outside the United States and documentation of the results of foreign clinical studies involving the food or food additive should also be presented for review.

¹ Anonymous (1985)¹

VI A Clinical Evaluation of Foods and Food Additives

- In designing protocols for clinical studies, the following should be considered: 1) the results of pre-clinical studies (including toxicity tests in animals) and foreign clinical studies; 2) the chemical nature of the proposed additive; and 3) all organs and organ systems that may be affected in man by consumption of the food or food additive under investigation.
- The sequence of clinical tests should be designed to maximize the safety of the research subjects.
- Guidelines for clinical trials of investigational new drugs should be followed in evaluating the qualifications of the principal investigator and investigating institution. In particular, careful consideration must be given to the qualifications of the investigator and the suitability of the investigating institution's facilities for conducting short- and long-term clinical trials.

FDA recognizes the need for the investigator to exercise sound clinical judgement based on his/her experience in an appropriate field of study. Studies involving healthy volunteers should be performed by investigators skilled in the evaluation of the safety of a variety of compounds. When subjects of a clinical study have a specific disease, as may be the case for clinical evaluation of foods for special dietary uses or special medical purposes, the investigators should be clinicians expert in the disease and disease process.

- The investigator should have high regard for the rights and safety of the test subject(s). The investigator is responsible for the administration of the food additive; thus, he/she must bear the ultimate responsibility for the welfare of the test subjects. All aspects of a clinical study generally are described in the study's protocol; however, because actions that have been identified as being in the best interests of the subjects at the beginning of a clinical study may change during the study, all aspects of the study must remain flexible and subject to modification. Aspects of the clinical study protocol subject to such modification include: 1) The nature and frequency of laboratory tests, 2) the duration of consumption of the food or food additive, and 3) the interval between test subjects' visits to the investigator.

Institutional review of research involving human subjects and the requirement for informed consent will provide additional safeguards for test subjects. Principles of institutional review and informed consent were set forth in the January 27, 1981 Federal Register,* codified in 21 CFR 50; and are summarized in Appendix A (see section VI A 5 below).

- There is some finite risk associated with the administration of every unapproved food and food additive to subjects of a clinical study; despite strict adherence to guidelines, the safety of subjects in the study cannot be guaranteed. Before beginning a clinical study, the investigator should consider what procedures will be used to detect adverse reactions to the test substance during the study. The investigator should establish criteria that will be used to decide when to discontinue the clinical study; these criteria may be changed during the study if the change is required to support the safety of the subjects.

* Anonymous (1981)²

VI A Clinical Evaluation of Foods and Food Additives

To further protect the safety of subjects of a clinical study, the sponsor of the study should provide appropriate follow-up after the study has ended. Such follow-up should be conducted or supervised by the investigator of the clinical study.

- Before a clinical study begins, the investigator should consider ways in which quality control of the study will be documented. Effective documentation of quality control will facilitate Agency review of the completed clinical study.
- FDA recommends that investigators use statistical expertise in the planning, design, execution, and analysis of pre-clinical and clinical studies. Such expertise will help ensure that the planned studies will provide the necessary information while minimizing the number of subjects (sample size estimation) and will strengthen the validity of estimates of safety obtained from the studies.

2. Specific Considerations for Clinical Studies of Foods and Food Additives

This section describes specific considerations concerning the protocol design, definition of study population, and statistical analysis of the results of human clinical studies with foods and food additives. These considerations should be explicitly addressed in the clinical study protocol.

a. Protocol Design

Protocols for clinical studies of foods and food additives should be described clearly and in sufficient detail to permit effective review and evaluation by CFSAN. In general, the protocol should be strictly adhered to throughout the clinical study; if the protocol is not adhered to, documentation of necessary modifications should be made (see item 7 in section 1 above). While it is rational and desirable to design studies to obtain specific information about the test substance, the generation of data justifying conclusions other than those originally anticipated can be a valuable result of clinical investigation.

The following are additional recommendations for the design of clinical study protocols for foods and food additives:

- A clear statement of objectives should be provided for each protocol. Good planning usually produces research questions that can be answered by direct inference from the study data. Since studies are frequently designed to answer more than one question, it is useful to list the questions to be answered in order of their priority.
- The rationale for conducting a clinical study should be presented. In addition, pre-clinical and clinical data relevant to the compound being studied and to the proposed protocol should be reviewed.
- A statement explaining the reasons for deciding on a particular length for the clinical study should be included in the protocol. In general, a clinical study should be of sufficient length to permit the demonstration of the safety (or lack of safety) of a food or food additive.
- A statement explaining the reasons for selecting particular dietary levels (dosages) of the food or food additive being tested should be included.
- Experimental design should include appropriate controls. When feasible, studies should be performed blind to avoid selection bias and bias in patient and physician responses.

Draft

VI A Clinical Evaluation of Foods and Food Additives

- Investigators should describe proposed methods of randomization and should present analyses that demonstrate the effectiveness of these methods.
- Objective observation methods should be used when possible and appropriate, observational endpoints should be rigorously defined, and methodology that will be used to quantify endpoints should be described. A statement describing quality control and frequency of data collection (endpoint monitoring) also should be included.
- Limitations that may be imposed on the clinical study because of protocol design or the failure of subjects to comply with the written protocol (such as withdrawals from the study, failure to randomize subjects effectively, technological limits of observations, etc.) and the possible effects these limitations may have on the outcome of the study should be addressed.

b. The Study Population

Clinical studies identify physiological responses to test substances in well-defined, small populations. These results are used to make inferences about responses to the test substance in larger, target populations. Study protocols should specify how subjects will be selected, their assignment to alternative test regimens, the specific conditions under which the trial will be conducted, and the nature of the target population to which the subjects' responses will be extrapolated. The following are additional recommendations for defining and selecting subjects for the clinical study:

- Each study protocol must be reviewed and approved by the appropriate Institutional Review Board; written, informed consent must be obtained for each subject in the study (see Appendix A in section VI A 5 below).
- Protocols should clearly define the selection criteria for subjects, including diagnostic criteria and reasons for exclusion from the study, and should compare and contrast the study population with the larger population likely to consume the food or food additive.
- Criteria for discontinuing the study should be stated clearly.
- Doses of the test substance should be selected so that a range of subject responses to the substance can be observed and the highest safe dose of the proposed additive can be determined. When individual subjects' responses are expected to be quite variable, testing at multiple doses in a double-blind, placebo-controlled study is recommended.
- A serious problem in clinical studies is determining the degree of subject adherence to the assigned protocol. Careful attention to subject compliance with the protocol is particularly important in outpatient studies. Protocols should state clearly how subjects' compliance will be monitored and should indicate when noncompliance will result in discontinuing the subject in the study. In general, data on subject compliance and noncompliance enhance the credibility of a study.

If it becomes apparent during the study that subjects are not complying with the study protocol, reasons for their noncompliance should be determined. All subjects initially included in a study must be reported on in the study's results, regardless of the degree of their compliance. Some noncompliance may necessitate identifying subgroups for evaluation, such as subjects who fail to consume foods containing the additive and subjects who report excessive use of alcohol or medication.

52454 1990

VI A Clinical Evaluation of Foods and Food Additives

- The number of subjects to be included in the study should be sufficient to be able to determine the safety of the test substance. Statistical estimates of the required number of subjects will depend upon: 1) The desired limit of detection of subjects' responses to the test substance; 2) the desired assurance against a false positive result; and 3) the acceptable risk of a false negative result.
- While it is desirable that placebo groups be included in early clinical studies of proposed foods and food additives (see page 17), this is not a requirement. Goals of early clinical studies may be 1) to gradually increase the dose of the test substance until physiological effects are observed or 2) to determine absorption and metabolism in humans in an effort to assess the adequacy of animal models used in safety assessments of the test compound. Therefore, subjects must be under careful observation during these studies.

The goals of early clinical studies often can be achieved effectively with an open (non-blind) study protocol. When clinical studies using blind comparisons of the test substance and a placebo or positive control substance should begin varies with the nature of the test material. During all phases of clinical investigation, the objective in using a placebo is to provide an adequate control for the compound under study. However, other methods of adequately controlling clinical studies exist. For example, the use of an active control compound or demonstration of a positive dose response to the food or food additive may constitute adequate control in some studies. For situations in which the natural course of a disease or condition is predictable and for which objective measurements of therapeutic or prophylactic response to the test compound can be made, results of carefully executed, open (non-blind) studies may be compared to historical data.

- Food additives should be studied in all age groups that may be significantly exposed, including, as appropriate, children, women of childbearing potential, older populations, and populations with specific disease conditions. The latter category includes populations that may be particularly exposed to, positively affected by, or at risk from a particular food or food additive.

Pregnancy tests should be administered to women of childbearing potential before the introduction of the test substance and the subject should be advised about suitable contraceptive measures. In general, women of childbearing potential should be excluded from the earliest clinical studies of a test substance. Once an adequate baseline of clinical information about the safety of a food or food additive has been obtained, however, women of childbearing potential may be included in clinical studies. For example, women of childbearing potential may participate in clinical studies when the teratogenic potential of the test substance has been determined to be negative in animals.

Follow-up to detect possible effects of the test substance on the fetus should be provided to women who become pregnant while on the study. Under these circumstances, transplacental passage of the substance and its secretion in milk should be assumed until proven otherwise.

VI A Clinical Evaluation of Foods and Food Additives

■ If the proposed food or food additive has a significant potential for use in children, its safety should be evaluated in children. Usually, studies in children are not attempted until there has been considerable clinical experience with the additive in adults. For certain proposed food additives, however, early clinical study in children may be warranted; in such cases, it is preferable to begin with older children, followed by younger children, infants, and premature infants. Detailed comments on pediatric studies are contained in "General Considerations for the Clinical Evaluation of Drugs in Infants and Children."^a Additional examples of guidelines concerning the clinical testing of foods or food additives in children are provided by the American Academy of Pediatrics.^b

■ Generally, physical examinations and laboratory tests should be performed to screen individuals with medically significant abnormalities from the clinical study. Laboratory tests should include the following: 1) Electrocardiograph; 2) urinalysis; 3) various tests on blood samples (for example, complete blood counts including platelet estimates, blood urea nitrogen, serum creatinine, tests of liver function, fasting blood sugar or 2-hour postprandial blood sugar, electrolytes, protein, and albumin); and 4) other tests that may be indicated by the nature of the test compound or from the results of previous animal and human clinical studies (for example, tests of vitamin status, prothrombin time, and blood lipid profiles).

■ In early clinical studies, when feasible, all subjects should refrain from taking medication (including over-the-counter drugs) for at least two (and preferably four) weeks before the study begins, unless interactions of the test substance with medication are the focus of the study. In some cases, a longer "washout" period will be required for return to a normal physiologic state before the clinical study begins. In later clinical studies, it may be desirable to examine the safety of combinations of the test substance and medication(s).

■ Post-study physical examinations for subjects of clinical studies often are necessary to ensure the subjects' safety. The results of these examinations should be fully documented.

c. Statistical Analyses

The following are general recommendations for statistical analyses in clinical studies of foods and food additives. Additional recommendations are contained in Chapter IV B 4.

■ Investigators are encouraged to seek expert biostatistical assistance prior to formulating the study design.

■ *A priori* description of the statistical methods to be used in analyzing data from a clinical study should be provided in the study's protocol.

■ Estimates of statistical power should be used to help determine the optimal number of subjects for a clinical study.

^a Reference (1977)³

^b American Academy of Pediatrics Committee on Nutrition report (1988);⁴ American Academy of Nutrition report (1987)⁵

VI A Clinical Evaluation of Foods and Food Additives

3. Sequence of Clinical Studies for Foods and Food Additives

The rationale behind serially conducted studies is that results of each study may influence the plan of succeeding studies. Investigators are encouraged to discuss data from animal studies and early clinical studies with CFSAN before conducting additional clinical studies.

a. Early Clinical Studies

The purpose of these studies is to determine the metabolism and the level of the food or food additive that gives an adverse or toxic response in man. Physiologic processes that are of primary interest in early clinical studies include: 1) Disposition (absorption, biotransformation, and excretion) of the food or food additive and its metabolites; 2) the potential of the food or food additive to induce enzyme levels or increase activity; 3) interactions between the food or food additive and nutrients that may necessitate balance studies; and 4) interactions between the food or food additive and medications that may necessitate drug bioavailability or drug metabolism studies. Information about the potential use of the test substance and all preclinical information about the test substance should factor into decisions about the appropriate sequence of early clinical studies.

For both ethical and scientific reasons, the initial introduction of a food or food additive into humans should be done with carefully selected subjects. Subjects for early clinical studies should be "normal" volunteers. *"Normal" generally means volunteers who are free from health problems that would complicate the interpretation of the study or increase the sensitivity of the subject to the toxic potential of the food or food additive.* Children, pregnant women, and women of childbearing potential usually should be excluded from early clinical studies.

Within the limitations described in the preceding paragraph, subjects of early clinical studies should be selected to accurately reflect the general population. Thus, individuals with mild but stable illnesses such as uncomplicated hypertension or arthritis may be considered for inclusion in initial clinical studies on a food or food additive. It also may be permissible--and even desirable--to include subjects with abnormalities for which consumption of the food or food additive may be particularly beneficial. For example, subjects with hyperlipoproteinemia may be included in an early clinical study on a food or food additive that functions as a non-caloric fat substitute. Additional examples include: (a) A food or food additive that will be used in the dietary management of organ failure should be tested in a population with failure of the organ under study; (b) a food or food additive designed to be deficient in a particular nutrient should be tested in a population that is unable to metabolize the nutrient in question (in fact, such a food or food additive may be harmful to a population with normal metabolism).

Most early clinical studies are sub-chronic (relatively short-term) and are generally less than 4 weeks in duration. These studies vary from single exposure to multiple exposures and examine a range of levels (doses) of the food or food additive. When several doses are being tested in a study, no research subject should be given the next-higher dose until sufficient exposure has occurred with the immediately preceding dose to be certain that serious adverse effects have not occurred.

For each food and food additive subjected to clinical investigation, it is also important to consider the appropriate frequency of laboratory tests and, when indicated by the results of previous studies, tests for specific organ or organ system effects. Independent of the outcome of clinical studies, thorough physical examinations and blood screening should be part of the follow-up for all subjects.

When unanticipated side effects occur in clinical studies, the investigator should determine the time required for elimination of the compound from the subject's system and reversal of the effects.

VI A Clinical Evaluation of Foods and Food Additives

b. Further Clinical Studies

Additional clinical studies may be designed to determine the safety of the proposed food additive during chronic intake (relatively long-term) and to gather more information about the food additive's adverse effects in humans. These studies should be performed after the general safety of the food or food additive in humans has been established in early, short-term clinical studies. The duration of exposure to the food or food additive in these studies will vary with the nature of the additive. Chronic administration in humans usually means continuous consumption for at least 8 to 12 weeks, unless contraindicated by adverse side-effects.

Relatively long-term clinical studies of food and food additives may emphasize the physiologic processes of enzyme induction or interaction of the additive with other substances (such as nutrients, medications, and other food additives). In addition, when designing studies to determine the safety of chronically consumed food additives, investigators should consider conducting nutrient balance studies; these studies help determine end-organ (or end-organ system) responses to the additive, including neurobehavioral changes.

Finally, clinical studies may be performed to obtain information about adverse effects of the food or food additive on specific subpopulations. For these studies, appropriate subpopulations may include children, pregnant women, women of childbearing potential, and older subjects. These studies may also include subjects with concomitant diseases who are undergoing therapy for the disease, particularly if such subjects represent segments of the population who are likely to consume the food or food additive after it has been approved.

Relatively long-term clinical studies should include a limited number of closely monitored subjects (rarely exceeding several hundred). In the clinical studies described above, the frequency of physical examinations and laboratory tests for subjects will depend upon the nature and relative safety of the food additive. For some subjects, daily supervision may be necessary. Early periods during a study will typically involve more frequent supervision of subjects than later periods. An example of a graded supervision plan would be one in which a test subject is seen by the investigator at least once a week for 2 to 4 weeks, once every other week for 6 to 8 weeks, at monthly intervals for 2 to 3 months, and bimonthly until the end of the follow-up period. Routine laboratory tests should be performed at frequent intervals; frequency and type of special laboratory tests should be determined by the nature of the food or food additive and its intended use.

In both early and chronic clinical studies of food additives, it is particularly important that a single formulation of the test substances be used throughout the study; in addition, investigators should test the compounds that will be marketed. Consideration should be given to relative exposures for particular food uses when such uses may alter the structure or effects of the test substance. A significant change in the formulation or manufacture of the food or food additive during chronic clinical studies may indicate the need for bioavailability studies on the (presumably changed) food or food additive. Results of these studies will enable meaningful comparisons to be made among clinical studies performed with different formulations of the test substance. When the petitioner intends to market a family of formulations and only a limited number of the formulations will be tested in clinical studies, petitioners should be prepared to demonstrate that the test compounds are fully representative of the family of formulations intended for marketing, particularly with respect to questions of safety.

VI A Clinical Evaluation of Foods and Food Additives

4. Submitting Reports of Clinical Studies on Foods and Food Additives to CFSAN

In submitting reports of clinical studies to CFSAN, particular emphasis should be placed on clear and concise: 1) statement of study objectives, 2) description of protocols, and 3) presentation of significant findings. Presentation of the results of a series of clinical studies on an proposed food additive should be scientifically logical and should specify the order in which the studies were conducted.

Early, relatively short-term clinical studies include tolerance studies. In reporting the results of tolerance studies, information on dose schedules and range of doses should be included. For relatively short-term clinical studies, the following questions should be answered in determining the safety of the proposed additive:

- What are the absorption, metabolism, tissue deposition, and major routes of excretion of the food or food additive?
- What is the half-life of the food or food additive in the human body? (Analysis of turnover and of other pharmacokinetic parameters of the test substance or its metabolites in various physiological compartments may aid in the interpretation of the results of toxicity studies.) (see Chapter V B);
- How may interactions between the food or food additive and nutrients or medications compromise the availability of any of these substances?
- How does the food or food additive affect the function of human organs and organ systems?
- What are the possible adverse reactions to the food or food additive in the general population of individuals who are likely to use the substance and in special (more sensitive) populations?

Reports on relatively long-term clinical studies should emphasize specific organ or organ system responses to the food or food additive and nutrient imbalances that occur with chronic use of the food or food additive.

Finally, the safety of a food or food additive may continue to be monitored after the substance has been approved. This can be accomplished by further clinical testing or by establishing a surveillance system and documenting adverse reactions to the food additive. The need for such a system is expected to vary with the nature and use of the approved food additive. Clinical testing and surveillance also may be useful in establishing the safety of expanded uses of the food or food additive or the safety of an altered food or food additive; these changes may occur as the result of changes in patterns of food consumption or food processing.

VI A Clinical Evaluation of Foods and Food Additives

5. Appendix A

The following principles are general guidelines for institutional review of, and conformed consent of subjects for, clinical studies. Additional information can be found in the references for this chapter.

a. Principles of Institutional Review

- An Institutional Review Board must be composed of no fewer than 5 persons from various backgrounds to assure complete and adequate review of clinical research activities commonly conducted by the institution. In addition to possessing the scientific competence necessary to review such institutional activities, the Board must be able to evaluate research applications and proposals in terms of institutional commitments and regulations, applicable law, standards of professional conduct and practice, and community attitudes.
- No member of a Board shall be involved in the initial or continuing review of an activity in which he has a conflicting interest, except to provide information requested by the Board.
- No Board shall consist entirely of persons who are officers, employees, or agents of, or are otherwise associated with the institution, apart from their membership on the Board.

b. Principles of Informed Consent

All subjects in a clinical evaluation are entitled to:

- a fair explanation of the procedures to be followed and the purposes of the procedures, including identification of any procedures that are experimental;
- a description of attendant discomforts and risks that may be reasonably expected;
- a description of benefits they may reasonably be expected;
- disclosure of appropriate alternative procedures that may be advantageous to the subject;
- an offer to answer any inquiries concerning the procedure; and
- instruction that the subject is free to withdraw his consent and discontinue participation in the project at any time, without prejudice to the subject.

IX References Continued

Chapter VI A: Guidelines for the Clinical Evaluation of Foods and Food Additives

1. Anonymous (1979) Refusal to approve the application (21 CFR 314.111 Section A5ii) Statement concerning adequate and well-controlled clinical investigations. *Fed. Reg. (April 1)* 107-110.
2. Anonymous (1988) Protection of human subjects (21 CFR Part 50) *Fed. Reg. (April 1)* 215-229.
3. Health, Education and Welfare (HEW) (1977) In, General considerations for the clinical evaluation of drugs in infants and children. HEW, PHS, and FDA Report. U.S. Printing Office.
4. American Academy of Pediatrics Committee on Nutrition [AAPCN] (1988) In, Clinical Testing of Infant Formulas with Respect to Nutritional Suitability for Term Infants. Prepared under FDA contract 223-86-2117. pp.1-36.
5. American Academy of Pediatrics Committee on Nutrition [AAPCN] (1987) Guidelines for the clinical evaluation of new products used in the dietary management of infants, children and pregnant women with metabolic disorders. A report by the Task Force on Clinical Evaluation of Products for Metabolic Disorders. pp.1-53.

Chapter VI B: Epidemiology Studies

1. Last, J.M. eds. (1983) In, A Dictionary of Epidemiology. (J.H. Abramson, S. Greenland, M.C. Thuriaux, associate eds.; J. Amsel, assistant ed.) Oxford University Press, New York, N.Y.
2. Anonymous. (1981) Reviews and Commentary. Guidelines for documentation of epidemiologic studies. A report by the Epidemiology Work Group of the Interagency Regulatory Liaison Group. *Am.J.Epidemiol.* 114(5):609-613.
3. Tollefson, L. (1988) Monitoring adverse reactions to food additives in the U.S. Food and Drug Administration. *Reg.Toxicol.Pharmacol.* 8(4):438-446.

Chapter VII C: Safety of Foods and Food Additives Developed by Biotechnology

1. Kessler, D. A., Taylor, M. R., Maryanski, J. H., Flamm, E. L., and Kahl, L.S. (1992) The Safety of foods developed by biotechnology. *Science* 256:1747-1749 (&1832).
2. Anonymous (1992) Statement of Policy: Foods derived from new plant varieties. *Fed. Reg.* 57:22984.

VI B. Epidemiology Studies

Epidemiology is the study of the distribution and determinants of health-related states and events in specified populations, and the application of this study to the control of health problems.^a The goal of all epidemiology studies is to uncover relationships between exposure to a specific agent and changes in health status.

Epidemiologic data are important to CFSAN in assessing food additive safety and have been used by the Agency as indicators where avenues of research and further human studies would be most productive. Guidelines for the proper conduct and documentation of epidemiology studies, such as selection of the study population, selection of appropriate controls, exposure assessment, methods used to adjust or control for confounding variables, and statistical analyses will not be discussed here. Appropriate guidelines have been published elsewhere,^b and should be consulted by the petitioner before submitting epidemiology data for consideration by the Agency.

There are two main categories of epidemiology studies, descriptive and analytic. Descriptive studies are concerned with the existing distribution of variables; they do not test hypotheses or make inferences concerning causality. Analytic studies are designed to examine associations, particularly hypothesized causal relationships, and focus on identifying or measuring the effects of specific risk factors.

1. Descriptive Epidemiology Studies

Descriptive epidemiology studies are relatively inexpensive to conduct and are usually of short duration. However, such studies are limited in their usefulness since no inferences can be made concerning causality. Generally, descriptive epidemiology studies are sentinel devices used to generate hypotheses or to provide evidence that indicates whether there is sufficient cause for conducting a lengthier and costlier analytic study.

a. Correlational Studies

Correlational studies, also called ecological studies, use grouped population data to relate exposure patterns of whole populations to disease incidence or mortality rates for whole populations. Because these studies do not examine the relationship between exposure and disease among individuals, the studies have been traditionally regarded as useful for generating, rather than definitively testing, a scientific hypothesis. Thus, the results of correlational studies would be insufficient to demonstrate a relationship without other types of data to support them.

^a Last (1983)¹

^b Anonymous (1981)²

VI B Epidemiology Studies Continued

b. Case Reports

Case reports are a type of descriptive epidemiology study frequently evaluated by CFSAN. Strongly suggestive anecdotal or clinical observations may indicate a possible causal relationship. Analytic epidemiology studies can then be designed to verify and quantify the risks, and to determine the role of confounding factors. Case reports of allergic reactions have been relied upon in safety evaluations, for example, in requiring label declarations for FD&C Yellow Nos. 5 and 6.

CFSAN also maintains the Adverse Reaction Monitoring System (ARMS), which is concerned with spontaneous reports from consumers and health professionals regarding alleged adverse effects from food products.^a It is a form of passive surveillance which was designed as a sentinel system to identify specific areas for focused clinical investigations on potentially causal associations. The ARMS is playing an increasingly important role in the post-marketing safety assessment of food-related products regulated by FDA, including the more ubiquitous food and color additives used in food, contaminants, vitamin/mineral supplements, and dietary supplements. In addition, the system has been useful in monitoring cases of infectious diseases transmitted through the food supply.

2. Analytic Epidemiology Studies

Although analytic epidemiology studies are more informative than descriptive studies, they are expensive and time-consuming to conduct. The types of analytic epidemiology studies commonly considered by CFSAN in safety evaluations include cross-sectional, prospective, and retrospective studies. Analytic epidemiology studies actually play a lesser role than descriptive studies in assessing food additive safety at CFSAN, primarily because well-designed and -conducted analytic epidemiology studies are not available for most products which FDA regulates. However, results from such studies, when available, are used in the overall safety evaluation of a compound. In addition, analytic epidemiology studies constitute the scientific base for the Agency's regulation of health claims on food and food labeling authorized by the Nutrition Labeling and Education Act of 1990.

a. Cross-Sectional Studies

Cross-sectional studies are those in which individuals are observed at only one point in time; such studies are commonly known as surveys. The presence or absence of disease and the presence or absence of suspected etiologic factors are determined in each member of the study population or in a representative sample at one particular time. The advantages of cross-sectional studies are that they are relatively inexpensive to conduct, and can be completed relatively quickly. However, cross-sectional studies reveal nothing about the temporal sequence of exposure and disease, and necessarily use current exposure as a surrogate for past exposure. Also, cross-sectional studies can only measure disease prevalence rather than incidence.

b. Prospective Studies

In prospective studies, also called cohort or follow-up studies, the investigator selects a study population of exposed and non-exposed individuals and follows both groups to determine the incidence of disease. The group can be characterized by factors thought to influence the development or course of the disease and by the presence or absence of risk factors (e.g. exposure or nonexposure to some agent). Prospective studies generally imply study of a large population, study

^a Tollefson (1988)³

VI B Epidemiology Studies Continued

for a prolonged period of years, or both. This type of study design is effective when there is good evidence of an association of the disease with a certain exposure (from clinical observations or from descriptive epidemiology studies), when exposure is rare, but incidence of disease among the exposed is high, and when the time between the exposure and disease is short. The major advantage of prospective studies is that the incidence rates of the disease under study can be measured directly; therefore, absolute and relative risks also can be measured directly. In addition, it is possible to analyze the association of a particular exposure with several diseases, and a temporal relationship between exposure and disease can be established.

There are a number of disadvantages to prospective studies, including: 1) The difficulty and expense of conducting the studies, since both large study populations and long periods of observation are required for definite results; 2) bias may be introduced if every member of the cohort is not followed; 3) the length of the study may be less than the latency period of the disease; for example, if the study is stopped before old age, many important diseases such as cancer may be missed; and, most importantly, 4) prospective studies are very inefficient for studying rare diseases.

Results of prospective studies have been used at CFSAN in assessing the potential carcinogenic risk of some compounds; for example, occupational cohort studies and studies of human populations accidentally exposed to a carcinogen have been used in safety assessments of benzene, dioxin, and methylene chloride. FDA has also provided financial support for prospective studies on accidental exposure to PBB's in a Michigan cohort, and exposure to methylmercury in fish in a cohort of pregnant women (and their offspring) in the Seychelles Islands.

c. Retrospective Studies

In retrospective studies, also known as case-control studies, the investigator selects cases with a specific disease, and appropriate controls without the disease, and obtains data regarding past exposure to possible etiologic factors in both groups. The rates of exposure of the two groups are then compared. A case-control approach is indicated when studying rare diseases, such as most cancers, since a very large number of individuals would be needed to produce enough cases of disease so that conclusions can be drawn in a prospective study. Although it is possible to detect the association of multiple exposures or factors with a particular disease, retrospective studies are generally used to study diseases that have some unique and specific cause, such as infectious agents, in order to avoid the problem of confounding etiologic factors.

Case-control studies can not estimate absolute risk or relative risk because the incidence of disease is not known in either the exposed or unexposed population as a whole. However, the relative risk can be estimated in retrospective studies by the odds ratio, which is the ratio of the odds of getting the disease to the odds of not getting the disease. The odds ratio is a good approximation of the relative risk when the subject cases are representative of all cases with regard to exposure, the controls are representative of all controls with regard to exposure, and the disease being studied is rare.

Retrospective studies are much less expensive and less time consuming to conduct than are prospective studies; usually, a relatively small population is needed for the study. Also, since the study selects only cases of the disease of interest, there is no bias incurred in determining the endpoint. However, bias is frequently incurred during detection and selection of cases, and during assessment of exposure. Controls should be identical to the exposed cases except for the factor under investigation, a requirement which is often difficult to achieve in practice. As with prospective studies, problems are frequently encountered in attempting to control for competing risk factors and confounders. The investigators can adjust for known confounders either by matching when selecting controls, statistically by stratification, or by use of regression models.

Draft

VI B Epidemiology Studies Continued

Results of case-control studies have been frequently used in safety evaluations at FDA, primarily to add further information to the overall assessment of safety. In the past, FDA has supported case-control studies on compounds of interest, such as the National Bladder Cancer Study and the use of artificial sweeteners.

d. Meta-Analyses

Meta-analysis is the reanalysis of pooled data from several distinct epidemiology studies. Meta-analyses are conducted to compensate for deficiencies in individual studies, particularly those involving study size, thereby providing a stronger case to prove or disprove a hypothesis. Where FDA evaluates a meta-analysis, the Agency considers such an analysis primarily as supporting evidence, rather than as primary evidence, that can confirm the validity of data concerning a hypothesis. The Agency must carefully scrutinize each meta-analysis to assess the soundness of its design and the quality of the data from individual studies to determine the significance of the data. Such scrutiny requires review of the original studies used for the meta-analysis.

52454 2001

Chapter VII

Emerging Issues in Safety Assessment of Food Additives and Color Additives Used in Food

A. Introduction

This section discusses approaches to testing that may be useful in assessing the safety of macro-additives (see Chapter VI B), bioengineered additives (see Chapter VII C), additives that are enzymes (see Chapter VII D), and microbially-derived additives (see Chapter VI E). This section also discusses the use of alternatives to whole (vertebrate) animal testing in safety evaluation (see Chapter VII F) and FDA's recognition of the potential for direct food additives and color additives used in food to cause both heritable and somatic genetic toxicity (see Chapter VII G).

Because the Agency's approaches to determining the safety of these additives will continue to evolve for some time, it is not yet appropriate to provide separate guidelines for acquiring toxicology information on the types of additives in this document. In general, the Agency recommends that petitioners follow guidelines for toxicity tests presented in other sections of this publication. In addition, this section suggests some important issues to consider when planning a program of toxicity testing designed to demonstrate the safety of unique additives. As always, we strongly recommend that petitioners discuss planned testing programs and protocols for toxicity tests with Center scientists before tests begin.

B. Macro-Additives

Macro-additives are a class of food additives that are intended to be replacements for conventional macro-nutrients such as fats, proteins, and carbohydrates and are intended for use at relatively high levels in food. Macro-additives may be nutritive or non-nutritive; they may be reasonably pure, well characterized chemicals or they may be complex mixtures whose complete characterization is not feasible; they may be well absorbed from the gastrointestinal tract or poorly absorbed; they may be manufactured from unusual or novel food sources or obtained by chemical synthesis.

The common characteristic of macro-additives is that they will be consumed in large quantities compared to conventional food additives and, as a consequence, they will present testing problems that require "customized" approaches. For example, it may not be feasible to calculate safety factors in the conventional way, that is, as a fraction of the highest oral dose that has no adverse effects in animals. Other means of providing margins of safety for macro-additives will have to be used; these may include information derived from metabolic, pharmacokinetic, and human clinical studies.

VII B Food Substitutes (Macro-Additives) Continued

1. Nutritional Concerns in Animal Toxicological Tests

Because of the expected high level of human consumption of these additives, animal test doses that are orders of magnitude greater than the Expected Daily Intake (EDI) for humans will often not be feasible. Attempts to achieve very high doses in the animal studies might result in nutritional imbalances or caloric deprivation that could confound interpretation of the toxicity studies. In order to test the highest dose feasible and yet avoid nutritional problems, it may be necessary for toxicity testing to be preceded by nutritional studies to determine adequate test diets and appropriate control diets for animals in toxicity studies.

If appropriate dietary controls include nutrient enhancement, care should be taken to avoid over-enriching the diet or changing nutrient ratios that would mask toxicological endpoints under consideration. For example, mineral oil as a test material would be mostly unabsorbed in the intestine where it would solubilize fat-soluble vitamins, leading to deficiencies of these nutrients. This effect may be eliminated by appropriate fortification of the diet with vitamins A, D, E, and K. Quantities of nutrients to be used for fortifying the diet should be determined experimentally, in relation to the amount of mineral oil (test substance) used. Under-fortification could fail to protect against nutrient deficiencies and over-fortification could lead to altered toxicological responses to xenobiotics and "background" pathology rates. Sufficiently great over-fortifications could produce hypervitaminosis.

Control and test diets should be of the same caloric density and nutritionally (micronutrients) equal to test diets. Selection of appropriate control diets may present particular problems when testing non-caloric food substitutes or food substitutes that interfere with absorption of nutrients. Due to nutrient variations in chow diets from batch to batch, it is preferable to use a semi-purified diet base in these studies.

Additional information can be found in Chapter IV B 5, Diets for Toxicity Studies and in Chapter IV B 1, General Guidelines for Toxicity Studies.

2. Absorption, Metabolism, Distribution, and Elimination Studies

Studies designed to follow the metabolic path and fate of macro-additives take on particular importance in providing assurance of safety if the conventionally calculated safety factor cannot be used. Greater understanding of the disposition and pharmacokinetics of the additive should help to diminish uncertainties regarding safety. Questions of the following types should be answered through appropriate studies:

- Does the product or its metabolites alter or interfere with absorption, metabolism, or excretion of normal nutrients or metabolic intermediates?
- Does the product or its metabolites alter the action of commonly used drugs?
- Is the product absorbed, metabolized, distributed, stored or excreted differently in man than in test animals?
- Does the product or its metabolites accumulate in tissues, and what are the toxicological consequences if there is accumulation?

VII B Food Substitutes (Macro-Additives) Continued

- If the product is poorly absorbed, does the high concentration in the gut affect gut morphology, physiology, or biochemistry? Are any changes in the gut morphology or biochemistry associated with the development of neoplasms of the gut?
- Does the product alter the composition or nature of the gut flora? If it does, what are the toxicological consequences of the changes?

3. Impurities and By-products

Because of the anticipated high human consumption levels of macro-additives, there is a concomitant high potential intake of impurities and by-products. Therefore, every effort should be made to identify and quantify the chemical constituents of the product. If any of these raise particular concerns, toxicity testing of the impurity or by-product itself may be recommended. Limits for impurities such as heavy metals, natural toxins, and anti-nutrition factors may need to be specified for the marketed product.

4. Clinical Studies

When animal studies have been completed or when there is reasonable assurance of safety of the macro-additive from animal studies, clinical studies with human subjects may be useful for increasing confidence in the safety of the product for human consumption. For example, humans may suffer subtle adverse effects not detected in animal studies due to differences in physiology or metabolism between animals and humans; human subpopulations (the old, young, and chronically ill) may each react differently to the food substitute. In addition, human studies may help compensate for the fact that conventional methods of calculating the Acceptable Daily Intake (ADI) may not be applicable to the results of standard toxicity studies on macro-additives.

VII C. Safety of Foods and Food Additives Derived from New Plant Varieties by Biotechnology

The regulatory framework and the FDA approach to assessing the safety of foods developed by biotechnology is discussed by the Commissioner^a and presented in detail in the Agency's *"Statement of Policy: Foods Derived from New Plant Varieties."*^b (It should be noted that the agency's statement only pertains to foods derived from new plant varieties, including those developed through biotechnology; the policy does not address all food additives developed through biotechnology nor is it strictly limited to foods derived from new plants developed solely through biotechnology.)

The following information provides a summary of the safety assessment of foods derived from new varieties of plants, and the FDA's approach to non-clinical safety testing.

FDA's science-based approach for ensuring the safety of foods from new plant varieties focuses safety evaluations on the objective characteristics of the food: The safety of any newly introduced substances and any unintended increased concentrations of toxicants beyond the range of known to be safe in food or alterations of important nutrients that may occur as a result of genetic modification. Substances that have a safe history of use in food and substances that are substantially similar to such substances generally would not require extensive pre-market safety testing. Substances that raise safety concerns would be subjected to closer inquiry. This approach is both scientifically and legally sound and should be adequate to fully protect public health while not inhibiting innovation.^c

Figure 8 summarizes the safety assessment of new plant varieties.

The Agency's approach to non-clinical safety testing of foods and food additives derived from new plant varieties has also been described.^d

Animal feeding trials of foods derived from new plant varieties are not conducted routinely. However, in some cases testing may be needed to ensure safety. For example, substances with unusual functions or that will be new macronutrients of the diet may raise sufficient concern to warrant testing. Tests could include metabolic, toxicological, or digestibility studies, depending on the circumstances.

^a Kessler *et al.* (1992)¹

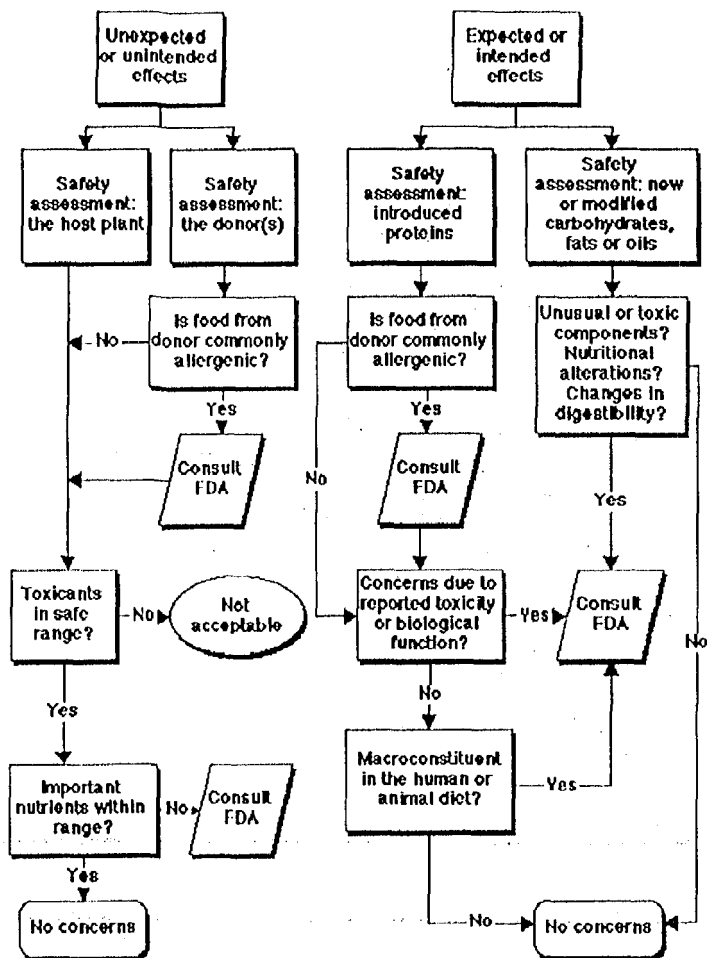
^b Anonymous (1992)²

^c Kessler *et al.* (1992)¹

^d Kessler *et al.* (1992)¹

Figure 8

Safety Assessment of New Varieties: Summary



Developers may also need to conduct tests on the "wholesomeness" of foods derived from new plant varieties as a means of ensuring that the food does not actually contain high levels of unexpected, acutely toxic substances. Such tests may provide additional assurance to consumers that food developed by new technology is as safe as food derived from varieties already in their grocery stores. However, animal tests on whole foods, which are complex mixtures, present problems that are not associated with traditional animal toxicology tests designed to assess the safety of single chemicals. Potential toxicants are likely to occur at very low concentrations in the whole food, and the tests may therefore be inadequately sensitive to detect toxicants. Efforts to increase the amount of whole food ingested by the test animals in order to increase the sensitivity and attempt to establish a traditional margin of safety (for example, a 100-fold safety factor) may not always be possible. When tests are contemplated, careful attention should be paid to the test protocol, taking into account such issues as nutritional balance and sensitivity.¹

¹ Kessler *et al.* (1992)¹

VII D. Enzymes

Commercial enzyme products may be obtained from edible plants and animals and from non-toxicogenic, non-pathogenic microorganisms. Questions about the microbial source of the enzyme (see Chapter VII E) and the nature and level of enzyme preparation in the food are of concern in evaluating the safety of commercial enzyme products because they influence the type and level of contaminating impurities in the food.

In general, enzyme preparations from organisms with a history of safe use do not require the same level of toxicological testing as enzymes from sources without a history of safe use in food. The safety of commercial enzyme products from sources without a history of safe use in food usually is evaluated on a case-by-case basis, but some generalizations about toxicology tests for these food additives can be made.

Because of the protein nature of enzymes and their susceptibility to digestion when consumed, residues of pure enzymes in processed food would be expected to have only limited toxic potential. If highly purified preparations of microbial enzymes are used in food processing, exposure to the enzymes is usually reduced to the parts-per-billion range. Such a level of exposure would ordinarily be too low to pose a safety concern, and toxicological testing may not be required. An exception to this generalization may occur if review by the Center's chemists results in concern for the presence in the enzyme preparation of a toxic material used in the purification process; however, this is unlikely because of the requirement that food grade chemicals be used in purification.

In most cases, however, commercial enzyme products from microbial sources are only partially purified. A variety of uncharacterized extraneous substances ("impurities") of biological origin may be present in the enzyme preparation at levels comparable to the active ingredient. These substances have no technical effect in food processing, but are allowed to remain in the enzyme products because the impurities do not interfere with enzyme function. In addition, the enzyme preparation may contain multiple enzyme activities that serve a variety of useful functions in processing food. When the types and levels of impurities in commercial enzyme products from microbial sources are considered to be significant, the Agency may recommend that safety be established by appropriate toxicity testing. Such a requirement usually can be met by 90-day toxicity studies in the rat and the dog. However, if review of the safety of the enzyme preparation raises questions about chemical contaminants, stability of the microbial strain, production of toxic products, etc., additional studies may be needed.

Enzyme products may be added directly to the food to be processed (e.g., rennet) or they may be immobilized on an insoluble matrix for use in processing liquid foods. Enzymes are immobilized by secure bonding (usually by means of a chemical reaction) to an insoluble matrix. Liquid food products (e.g., corn syrup) may be processed by passage over a column of the immobilized enzyme. Only negligible amounts of the immobilized enzyme are expected to enter the processed food. Depending on the nature of the immobilization matrix, however, some potential exists for contamination of the processed food by chemicals used in the immobilization process. If the Agency decides that information about the nature of the fixing agent and its potential migration to food raise questions of safety for foods processed by passage over an immobilized microbial enzyme, the Agency will recommend that the immobilized enzyme be subjected to 90-day toxicity studies in the rat and the dog or other appropriate study.

As described in the preceding paragraphs, a variety of factors will be taken into account by the Agency in deciding what information is needed to assess the safety of additives that are enzymes. Before conducting toxicity studies to assess the safety of such additives, petitioners should consult with Agency scientists. A comprehensive review of the safety concerns relating to additives that are enzymes will be issued in a separate publication.

VII E. Microbially Derived Food Ingredients

Microbially derived food ingredients may be food additives (including enzymes), color additives used in foods and substitute foods. A unique concern about the safety of microbially derived food ingredients is the microbial source; except for this concern, the safety of these ingredients will be evaluated as for analogues, non-microbially derived ingredients. A variety of factors will be taken into account by the Agency in deciding what information is needed to assess the safety of microbially derived food ingredients. Before conducting toxicity studies to assess the safety of such ingredients, petitioners should consult with Agency scientists. A comprehensive review of the safety concerns relating to microbial sources will be issued in another publication.

VII F. Advances in the Development of Alternatives to Whole Animal (Vertebrate) Testing

Because animal experimentation has become an emotional issue, it is important to recognize the growing impact of *in vitro* toxicology on the practice of toxicology. Although the field is often termed "alternative," experimental models have been applied to the three "R's" of Russel and Burch:^a to replace animal models, to reduce the number of animals used, or to refine test methods to minimize stress and suffering to animals.

This section is not intended as a guideline but serves to identify a future direction in methodology. In the context of this document, "alternatives to whole animal (vertebrate) experimentation" refers to *in vitro* tests for potential toxicity that substitute for or replace *in vivo* (whole animal) studies. "*In Vitro*" literally means "in glass", and is interpreted to mean "in a test tube" or "outside of the body".^b Alternative tests include short-term tests using isolated cells, tissues, and organs and studies involving mathematical modelling, epidemiology, or the use of human volunteers; short-term tests for genetic toxicity (see Chapter IV C 1) are excluded.

In practice, alternative tests are used to support the planning and interpretation of whole animal toxicity studies and are not yet used as substitutes for toxicity studies using whole animals. For example, an alternative test may be used 1) to determine the relative biological potency of a series of toxicants at the cellular level, 2) to select the animal model in which to conduct an *in vivo* test by comparing the metabolic properties of a toxicant at the cellular level in several species, and 3) to identify mechanism(s) of toxicity by defining the relationship between exposure to a toxicant and development of various toxicological endpoints at the cellular, subcellular and molecular levels of organization.

Recent advances that have been made in *in vitro* studies with isolated cells, tissues, and organs have directed the scientific community toward developing, validating, and evaluating alternative test systems. The predictive value of a standardized test must be assessed by means of a series of validation studies. Validation can demonstrate that the use of an *in vitro* test is equivalent to the use of an established *in vivo* test or that the *in vitro* test accurately predicts human toxicity. Anticipating a continued increase in the development and use of alternative *in vitro* test systems,^c the Agency encourages the development of approaches that can provide information relevant to the assessment of human risks.

1. Reasons for Developing Alternative Tests

Several reasons to encourage the development of alternative *in vitro* tests are listed below:

- **Economy and efficiency:** Once established, *in vitro* tests may provide toxicity information in a cost-effective and time-saving manner. Information generated from *in vitro* test systems can be used to increase the efficiency of whole-animal studies and decrease the number of animals used in toxicity testing. The relative simplicity and space-saving characteristics of *in vitro* methods also are viewed as advantages.

^a Russel and Burch (1959)¹

^b Schaeffer (1990)²

^c Goldberg and Frazier (1989);³ McKeegan *et al.* (1990)⁴

VII C Advances in the Development of Alternatives to Whole Animal (Vertebrate) Testing Continued

■ Information about human risk: Human cells, ethically obtained and successfully established *in vitro*, may provide information about a toxicant that is relevant to human risk. For example, a toxicant's mechanism of action or metabolism in human cells can provide the basis for selecting a suitable animal model for long-term toxicity studies.

2. Possible Applications of Alternative Tests

■ Isolated cells, tissues, and organs can be prepared and maintained in culture by methods that preserve properties characteristic of the same cells, tissues, and organs *in vivo*. Using such *in vitro* systems will permit data to be generated under controlled experimental conditions and in the absence of many complicating factors characteristic of experiments with whole animals. For example, the use of cell culture systems will enable the metabolism of a toxicant that occurs in one type of cell (*i.e.*, hepatocyte cells) to be studied separately from a toxic endpoint that occurs in a different cell type.

■ Several toxic endpoints may lend themselves to quantification in an *in vitro* test system. Relevant endpoints could be identified by comparing the action of a toxicant at cellular, subcellular or molecular sites with the toxic effects observed in the target organ or tissue *in vivo*. Analysis of a broad spectrum of *in vitro* cellular events may provide information about the *in vivo* progression of a toxic response as a function of toxicant concentration and time.

■ Because *in vitro* procedures have the potential to yield reproducible measurements, they theoretically lend themselves to standardization. However, interpreting data obtained from a standardized *in vitro* toxicity test with a reasonable degree of confidence can only occur after potential confounding factors, such as interactions between the test agent and non-cellular components of the test system, have been identified or eliminated.^a

■ The process of validation appears to be key to the full acceptance of alternative tests where the reliability and relevance of procedures are established for specific purposes.^b While there is much discussion about the framework for this process, several components appear essential to the overall coordination of the validation process, including: scientific consensus on the definition of a validated test, reference chemicals with defined toxicity and general availability, a central repository for test performance data and protocols, an established network of laboratories with the capabilities of method validation, and scientific understanding of the mechanistic basis of the toxicological process involved. An impartial and competent group of scientists from regulatory agencies and the research community could facilitate the implementation of the validation process.

^a Frazier and Bradlaw (1989)⁵

^b Balls, *et al.* (1990)⁶

VII C Advances in the Development of Alternatives to Whole Animal (Vertebrate) Testing Continued

3. Limitations of Alternative Tests

Limitations of *in vitro* tests are well known. For example:

- *In Vitro* test systems are not available for all tissues and organs. In addition, normal systemic mechanisms of absorption, penetration, distribution, and excretion are absent from *in vitro* test systems. *In Vitro* systems lack the complex, interactive effects of the immune, blood, endocrine systems, nervous system, and other integrated elements of the whole animal. Thus, *in vitro* tests cannot be used to study the complex nature of systemic toxicity.
- Validation of new methods is time-consuming and expensive; acceptance of *in vitro* tests as alternatives to traditional toxicity testing in whole animals is expected to be slow.^a While many schemes have been proposed to expedite these processes, no alternative *in vitro* test presently can replace an *in vivo* toxicity study.

4. Current Use of *In Vitro* Tests

Numerous & diverse *in vitro* tests have been developed. Their importance and use have been discussed in many publications.^b Many of these tests will be improved over time by the introduction of new scientific information and technological advances in *in vitro* toxicology and related fields, such as molecular biology and biotechnology. The Agency encourages the development and use of *in vitro* test systems for planning and interpreting the results from whole animal toxicity studies.

Significant advances have been made in the development of *in vitro* alternatives for ocular safety testing.^c Other *in vitro* systems have been proposed which measure a broad range of endpoints and are now in various stages of validation. The Agency is currently part of an interagency regulatory groups evaluating these proposed alternative test methods.

In Vitro approaches to toxicity testing can provide useful data when integrated with other information about the toxicity of food and color additives used in food. Results of *in vitro* tests can be used to optimize the design of conventional toxicity tests for a particular test substance by helping to determine appropriate dose levels and by helping to decide which species is the best model for man. Such improvements in the design of whole animal toxicity tests may reduce the number of test animals required to produce useful information about the safety of proposed food and color additives used in food.

^a Frazier (1990)⁷

^b McKeenhan *et al.* (1990);⁴ Frazier and Bradlaw (1989);⁵ Frazier (1990);⁷ Atterwill and Steele (1987);⁸ Balls *et al.* (1983);⁹ Berkly and Sherrod (1977);¹⁰ Bradlaw (1986);¹¹ Goldberg (1983);¹² Goldberg (1984);¹³ Goldberg (1985);¹⁴ Goldberg (1987);¹⁵ Grishman and Smith (1984);¹⁶ Rofe (1971);¹⁷ Rowan and Stratmann (1980);¹⁸ Rowan and Goldberg (1985);¹⁹ Stammati *et al.* (1981);²⁰ Tardiff (1978);²¹ Zucco and Hooisma (1980);²² Zucco and Hooisma (1982)²³

^c Frazier *et al.* (1987);²⁴ Nardone and Bradlaw (1983);²⁵ Frazier (1988);²⁶ Wilcox and Bruner (1990)²⁷

VII C Advances in the Development of Alternatives to Whole Animal (Vertebrate) Testing Continued

In Vitro tests can help elucidate the nature of the interaction between test substance and organism at the cellular, subcellular, and molecular levels. Thus, once the critical target organ or organ system has been identified in whole animal studies, *in vitro* tests can focus on the mechanism of action of the test substance at the target site. Information from these studies can assist the Agency in making decisions about the safety of proposed food and color additives used in food by comparing responses observed in human and animal cells and by facilitating extrapolation from high-dose to low-dose responses.

At present, in evaluating a petition for the use of a food or color additive, the Agency considers *in vitro* tests to be useful in helping to identify the mechanism(s) of action of the test substance and to provide information about subtle effects observed *in vitro* that may not be observed in *in vivo* studies.

VII G. Heritable and Somatic Genetic Toxicity

This chapter discusses FDA's interest in direct food additives and color additives used in foods that can cause both heritable and somatic genetic toxicity. While the FDA currently neither recommends specific tests to determine somatic and heritable genetic toxicity, nor regulates food and color additives used in food on the basis of such activities, the Agency has an heightened interest in this area.

1. Rationale for Testing for Heritable and Somatic Genetic Toxicity

Heritable genetic toxicity is chemically-induced damage to the DNA of male and female germ-line cells that is not correctly repaired, so that the damaged gene(s) can be inherited. The consequences of this genetic toxicity has been well documented, and a number of different genetic diseases have been characterized. Somatic genetic toxicity is chemically-induced damage to the DNA of dividing and non-dividing somatic cells (i.e. non-germ-line cells). The consequence of somatic genetic toxicity is that chemicals may alter gene functions in rapidly dividing somatic cells (e.g. intestinal lining and bone marrow) and in quiescent cells which may be forced to replicate in response to a regenerative or mitogenic stimulus (e.g. G_0G_1 peripheral lymphocytes). Genetic damage to these cells can lead to cancer and alteration of critical cellular functions (e.g. altered hormone and receptor site functions).

2. Rationale for Selecting a Specific Test Battery

Currently the Agency recommends the use of a battery of genetic toxicity tests (see Chapter IV C 1 c) for all chemicals that are direct food additives or color additives used in foods, including chemicals with structures assigned to all three structure categories (see Chapter III B 2), as well as chemicals associated with Concern Levels I, II, and III (see Figure 4 in Chapter III B 1). These tests are recommended to evaluate the genetic toxicity of chemicals in order to identify those chemicals that may be direct acting carcinogens (see Chapter IV C 1).

Short-term tests for genetic toxicity can also be conducted to evaluate the effects of chemicals on the genetic material of both somatic and germ-line cells, and the tests used for these purposes can overlap those used for predicting carcinogenicity. For example, the data obtained from the *Salmonella typhimurium* reverse mutation assay is not only useful in predicting the potential carcinogenicity of test substances,^a but it is also an important means of determining whether a chemical has the potential to damage the genetic material in both germ-line and somatic cells. Although FDA considers the information obtained from the test battery recommended in Chapter IV C 1 to be useful in assessing a chemical's potential to cause heritable and somatic genetic toxicity, the scientific community has not yet reached a consensus that these indicators are reasonably predictive of human responses.

While FDA does not recommend a unique battery of tests for determining heritable and somatic genetic toxicity, the Agency recognizes that certain types of tests may be useful for this purpose.

^a Tennant *et al.* (1987);¹ Ashby and Tennant (1988);² Ashby and Tennant (1991)³

VII G Heritable and Somatic Genetic Toxicity Continued

Historically, gene mutations in germ line cells have been detected using *in vivo* tests such as the sex-linked recessive lethal assay in *Drosophila melanogaster* and rodents.^a Unfortunately, the standard classical assay procedures are not completely satisfactory; each of these tests has one or more of the following limitations:

- standard procedures have a very low sensitivity for detecting known mutagenic chemicals, and the assays fail to detect dose-related increases in chemical activities;
- standard protocols have many deficiencies (e.g. they frequently lack concurrent positive controls, multiple test chemical doses are rarely used, etc.);
- standard protocols for heritable genetic toxicity cannot simultaneously measure somatic cell toxicity in the same animals; and
- standard methods require large numbers of animals and are very time consuming and expensive.

Thus, two groups of tests may provide a sensitive method for detecting heritable and somatic cell genetic toxicity. First, a battery of tests for germ-line and somatic cell genetic toxicity should include the same short-term genetic toxicity tests used to predict potential carcinogenicity (e.g. *Salmonella typhimurium* reverse mutation assay, *in vitro* ML mutation assay and an *in vivo* cytogenetics assay (see Chapter IV C 1)). Second, a battery of tests for germ-line and somatic cell genetic toxicity also should include the use of transgenic mice. The Agency recognizes that current genetic toxicity tests using transgenic animals do not directly demonstrate heritable genetic toxicity effects; however, chemical-induced genetic toxicity to germ cells demonstrates the potential for this to occur. Since research with several different experimental rodent models has been progressing rapidly, and a variety of transgenic rodents are now commercially available, it may be possible in the future to simultaneously assess chemically-induced genetic damage to germ line cells and to a variety of somatic tissues. The transgenic test system should have several advantages over classical tests for heritable genetic toxicity:

- the investigator can easily manipulate the treatment conditions so that tissue-specific toxicological effects can be compared for different assay protocols;
- the test requires relatively few animals (i.e. 2 or 3 animals per treatment group); and
- the test is relatively inexpensive and can be performed in a matter of days.

FDA continues to encourage the scientific community to develop sensitive assays for detecting germ-line and somatic cell genetic toxicity.

^a Abrahamson *et al.* (1980);⁴ Lee *et al.* (1983);⁵ Mason *et al.* (1987)⁶

Chapter VIII

Glossary

Acronym	Definition
Act	"the Act", i.e. Federal Food, Drug, and Cosmetic Act (1958)
ABS	chromosome aberration(s)
ADI	acceptable daily intake
A/G	albumin-to-globulin
ANOVA	analysis of variance
ARMS	adverse reaction monitoring system
B-cells	B lymphocytes
B/T	ratio of B to T lymphocytes
CAC	Cancer Assessment Committee
CAS	Chemical Abstract Service
CFR	Code of Federal Regulations
CFSAN	Center for Food Safety and Applied Nutrition
CIO	Chinese hamster ovary [cell(s)]
CMI	cell mediated immunity
CSO	Consumer Safety Officer
DNA	deoxyribonucleic acid
DTH	delayed type hypersensitivity
EAFFUS	everything added to food in the United States
EDI	estimated daily intake
ELISA	enzyme-linked immunosorbent assay
EPA	Environmental Protection Agency (United States)
FAP	food additive petition
FASEB	Federation of American Societies for Experimental Biology
FASP	Food Additive Safety Profile
FDA	Food and Drug Administration
GLP	good laboratory practices
GRAS	Generally Recognized as Safe
HGPRT	hypoxanthineguanine phosphoribosyl transferase activity
HTD	highest treatment dose
IARC	International Agency for Research on Cancer
Igs	immunoglobulins
LOEL	lowest observed effect level
LPS	lipopolysaccharide

52454 2016

IX Glossary Continued

Acronym	Definition
MFO	mixed function oxidase
ML	1.5178Y mouse lymphoma cell
MLR	mixed lymphocyte response
MTD	maximum tolerated dose
NCI	National Cancer Institute (United States)
NIEHS	National Institute of Environmental Health Sciences
NK	natural killer
NOAEL (NOEL)	no observed adverse effect level (no observed effect level)
NTP	National Toxicology Program
QAU	Quality Assurance Unit
QRAC	Quantitative Risk Assessment Committee
QRAs	quantitative risk assessments
PAFA	Priority-Based Assessment of Food Additives
PAIS	periarterial lymphocyte sheath
PB-PK	physiologically based pharmacokinetic model
PHA	phytohemagglutinin
PWM	pokeweed mitogen
RBC	red blood cells
Redbook	<u>Toxicological Principles for the Safety Assessment of Direct Food Additives and Color Additives Used in Food</u>
RIA	radio immunoassay
RNA	ribonucleic acid
R value	ratio of human consumption (mg/kg bw/day) to the lowest dose producing a compound-related adverse effect in the longest duration, highest quality study available
SAR	structure activity relationship
SCE	sister chromatid exchange
SHH	Syrian hamster embryo cell
SOP	standard operating procedure
SRBC	sheep red blood cells
T-cells	T lymphocytes, or thymus derived cells
TK	thymidine kinase
UDS	unscheduled DNA synthesis
WBC	white blood cells
WBA	whole body autoradiography

Chapter IX

References

Chapter I: Introduction

1. Committee of Labor and Public Welfare [CLPW] (1988) A legislative history of the Federal Food, Drug, and Cosmetic Act and its amendments. A report submitted by Mr. Hill from the Committee on Labor and Public Welfare entitled "Food Additives Amendment of 1958. U.S. Senate report No. 2422 to accompany H.R. 13254. Vol.14:913-937.
2. Code of Federal Regulations [CFR] (1981) Eligibility for classification as generally recognized as safe (GRAS). U.S. Government Printing Office *CFR Title 21, Section 170.3. (April 1):12-14.*
3. Code of Federal Regulations [CFR] (1981) General principles for evaluating the safety of food additives. U.S. Government Printing Office. *Title 21, Section 170.20 (April 1):12.*
4. U.S. Food and Drug Administration [FDA], Bureau of Foods. (1982) In, Toxicological principles for the safety assessment of direct food additives and color additives used in food. National Technical Information Service. Springfield, VA.

Chapter I A (1): Major Changes in the Revised Guidelines (Introduction)

1. U.S. Food and Drug Administration [FDA], Bureau of Foods. (1982) In, Toxicological principles for the safety assessment of direct food additives and color additives used in food. National Technical Information Service. Springfield, VA.
2. U.S. Food and Drug Administration [FDA], Center for Food Safety and Applied Nutrition. (1992) In, FDA/CFSA's Priority-Based Assessment of Food Additives Database. Toxicology Information Management Team (HFF-156), FDA, 200 C Street S.W., Washington, D.C. 20204

Chapter I B: Flexibility and Consistency in Guidelines for Toxicity Testing

1. Elkes, M. (1989) Europe 1992: Its impact on nontariff trade barriers and trade regulations with the United States. *Fd. Drug Cosm.L.J.* 44: 563-591.
2. Grignolo, A. (1989) How drug companies are preparing for 1992. *Fd. Drug Cosm.L.J.* 44:557-562.
3. Schneebaum, S.M. (1989) Products liability in the european community: What does it mean for U.S. companies? *Fd. Drug Cosmet.L.J.* 44:283-289.

IX References Continued

Chapter II A: Introduction

1. Food Additives Amendment, Section 409, Federal Food Drug, and Cosmetic Act (1958)
2. Federal Food Drug, and Cosmetic Act (1958)
3. Committee of Labor and Public Welfare [CLPW] (1988) A legislative history of the Federal Food, Drug, and Cosmetic Act and its amendments. A report submitted by Mr. Hill from the Committee on Labor and Public Welfare entitled "Food Additives Amendment of 1958. U.S. Senate report No. 2422 to accompany H.R. 13254. Vol.14:913-937.
4. Color Additive Amendment, Section 706, Federal Food Drug, and Cosmetic Act (1958)
5. Rulis, A.M. (1990) The Food and Drug Administration's Food Additive Petition Review Process. *Food Drug Cosmetic Law J.* 45:533-544.
6. Code of Federal Regulations 21 § 170.22 (1992)

Chapter II C (5): Evaluating the Safety of Food Additives and Color Additives Used in Food (Carcinogenic Risk Assessment)

1. Lorentzen, R.J (1984) FDA procedures for carcinogenic risk assessment. *Fd.Technol.* 38(10):108-111.
2. National Research Council [NRC]. (1983) In, Risk assessment in the federal government: Managing the process. Washington, D.C. National Academy of Sciences [NAS] Press. Washington D.C.
3. Office of Science and Technology (1985) In, Chemical carcinogens: A review of the science and its associated principles.
4. Anonymous. (1986) Chemical carcinogens: A review of the science and its associated principles. A report by the U.S. Interagency Staff Group on Carcinogenesis. *Environ. Health Perspect.* 67:201-282.
5. Gaylor, D.W. and Kodell, R.L. (1980) Linear interpolation algorithm for low dose assessment of toxic substances. *J.Environ.Path.Toxicol.* 4:305-312.
6. Dua, P.N. and Jackson, B.A. (1988) Review of pathology data for regulatory purposes. *Toxicol.Path.* 16:443-450.
7. Peto, R., Pike, M.C., Day, N.E., Gray, R.G., Lee, P.N., Parish, S., Peto, J., Richards, S. and Wahrendorf, J. (1980) Guidelines for simple sensitive significance tests for carcinogenic effects in long-term animal experiments. In, Long-term and short-term screening assays for carcinogens: A critical appraisal. International Agency for Cancer Research [IARC]. Lyon, France. *Monograph Suppl.* 2:311-426.

IX References Continued

Chapter II C (5): Evaluating the Safety of Food Additives and Color Additives Used in Food (Carcinogenic Risk Assessment) Continued

8. Bickis, M. and Krewski, D. (1985) Statistical design and analysis of the long-term carcinogenicity bioassay. In, Toxicological risk assessment, Vol. 1, Biological and statistical criteria. (D.B. Clayson, D. Krewski, and I. Munro eds.) CRC Press, Inc. Boca Raton, FL. pp.125-147.
9. McKnight, B. (1988) A guide to the statistical analysis of long-term carcinogenicity assays. *Fund.Appl.Technol.* 10:355-364.

Chapter III B (1): Determining Concern Levels

1. Rulis, A.M., Hattan, D.G., and Morgenroth, V.H. (1984) FDA's priority-based assessment of food additives. I. Preliminary Results. *Regulatory Toxicol.Pharmacol.* 4:37-56.

Chapter III B (2): Concern Levels (Structure Category Assignment)

1. Dehn, R.L. and Helmes, C.T. (1974) An automatic procedure for assessing possible carcinogenic activity of chemicals prior to testing. Stanford Research Institute. ____ Press. Menlo Park, CA: pp.128-____.
2. Gori, G.B. (1977) Ranking of environmental contaminants for bioassay priority. In, Air pollution and cancer in man. (U. Mohr, D. Schmahl, L. Tomatis, and W. Davis, eds.) International Agency for Research on Cancer [IARC]. Lyon, France. *Mono.16:99-111*.
3. Cramer, G.M., Ford, R.A. and Hall, R.L. (1978) Estimation of toxic hazard--A decision tree approach. *Fd.Cosmet.Toxicol.* 16:255-276.
4. Lutz, W.K. (1984) Structural characteristics of compounds that can be activated to chemically reactive metabolites: Use for a prediction of a carcinogenic potential. *Arch Toxicol. Suppl.* 7: 194-207.
5. Woo, Y.T., Arcos, J.C., Lai, D.Y. (1985) Structural and functional criteria for suspecting chemical compounds of carcinogenic activity: State-of-the-art of predictive formalism. In, Handbook of carcinogenic testing. (H.A. Milman and E.K. Weisburger, eds.) Noyes Publications, Park Ridge, N.J. pp.2-25.
6. Enslein, K. (1984) Estimation of toxicological endpoints by structure-activity relationship. *Pharmacol.Rev.* 36:131S-135S.
7. Ramiller, N. (1984) Computer-assisted studies in structure-activity relationship. *Am.Lab.* 78-88.
8. Enslein, K. and Craig, P.N. (1978) A toxicity estimation model. *J.Environ.Pathol.Toxicol.* 2:115-121.
9. Enslein, K., and Craig, P.N. (1982) Carcinogenesis: A predictive structure-activity model. *J.Toxicol.Environ. Health.* 10:521-530.

IX References Continued

Chapter III B (2): Concern Levels (Structure Category Assignment) Continued

10. Enslein, K., Lander, T.R. and Strange, J.R. (1983) Teratogenesis: A statistical structure-activity model. *Teratog.Carcinog.Mutagen.* 3:289-309.
11. Enslein, K., Lander, T.R., Tomb M.E. and Landis, W.G. (1983) Mutagenicity (Ames): A structure-activity model. *Teratog.Carcinog.Mutagen.* 3:503-513.
12. Enslein, K., Blake, B.W., Tomb, M.E., and Borgstedt, H.H. (1987) Prediction of Ames test results by structure-activity relationships. *In Vitro Toxicol.* 1:33-44.
13. Enslein, K., Borgstedt, H.H., Tomb, M.E., Blake, B.W., and Hart, J.B. (1987) A structure-activity prediction model of carcinogenicity based on NCI/NTP assays and food additives. *Toxicol. and Ind. Health.* 3:267-287.
14. Chou, J.T. and Jurs, P.C. (1979) Computer-assisted structure-activity studies of chemical carcinogens. An N-Nitroso compound data set. *J.Med.Chem.* 22:792-797.
15. Jurs, P.C., Chou, J.T. and Yuan, M. (1979) Computer-assisted structure-activity studies of chemical carcinogen. A heterogeneous data set. *J.Med.Chem.* 22:476-483.
16. Yuan, M. and Jurs, P.C. (1980) Computer-assisted structure-activity studies of chemical carcinogens: A polycyclic aromatic hydrocarbons data set. *Toxicol.Appl.Pharmacol.* 52:294-312.
17. Tinker, J. (1981) Relating mutagenicity to chemical structure. *J.Chem.Inf.Comput.Sci.* 21:3-7.
18. Yuta, K. and Jurs, P.C. (1981) Computer-assisted structure-activity studies of chemical carcinogens. Aromatic amines. *J.Med.Chem.* 24:241-251.
19. Stouch, T.R. and Jurs, P.C. (1985) Computer-assisted studies of molecular structure and genotoxic activity by pattern recognition techniques. *Environ. Health Perspect.* 61:329-343.
20. Rohrbaugh, R.H., Jurs, P.C., Ashman, W.P., Davis, E.G., and Lewis, J.H. (1988) A structure-activity relationship study of organophosphorus compounds. *Chem.Res.Toxicol.* 1:123-127.
21. Randic, M. and Jurs, P.C. (1989) On a fragment approach to structure-activity correlations. *Quant.Struct.Act.Relat.* 8:39-48.
22. Klopman, G. (1984) Artificial intelligence approach to structure-activity studies. Computer automated structure evaluation of biological activity of organic molecules. *J.Am.Chem.Soc.* 106:7315-7321.
23. Klopman, G. and Rosenkranz, H.S. (1984) Structural requirements for the mutagenicity of environmental nitroarenes. *Mutat.Res.* 126:227-238.
24. Rosenkranz, H.S., Klopman, G., Chankong, V., Pei-Edwards, J. and Haimes, Y.Y. (1984) Prediction of environmental carcinogens: A strategy of the mid-1980s. *Environ.Mutagen.* 6:231-258.

IX References Continued

Chapter III B (2): Concern Levels (Structure Category Assignment) Continued

25. Klopman, G., Contreras, R., Rosenkranz, H.S. and Waters, M.D. (1985a) Structure-genotoxic activity relationships of pesticides: Comparison of the results from several short-term assays. *Mutat.Res.* 147:343-356.
26. Klopman, G., Friction, M.R., and Rosenkranz, H.S. (1985b) Computer analysis of toxicological databases: Mutagenicity of aromatic amines in *Salmonella* tester strains. *Environ.Mutagen.* 7:625-644.
27. Rosenkranz, H.S., Mitchell, C.S., and Klopman, G. (1985) Artificial intelligence and Bayesian decision theory in the prediction of chemical carcinogens. *Mutat.Res.* 150:1-11.
28. Rosenkranz, H.S. and Klopman, G. (1990) Structural basis of carcinogenicity in rodents of genotoxicants and non-genotoxicants. *Mutation Res.* 228:105-124.

Chapter III B (3): Concern Levels (Estimation of Human Exposure to Direct Food Additives and Food Ingredients)

1. Federation of American Societies for Experimental Biology [FASEB] (1988) In, Estimation of Exposure to Substances in the Food Supply. Life Sciences Research Office, FASEB. (July).
2. Thompson, K.M., Burmaster, D.E., and Crouch, E.A. (1992) Monte Carlo techniques for quantitative uncertainty analysis in public health risk assessments. *Risk Anal.* 12:53-63.
3. Burmaster, D.E. and von Stackelberg, K. (1991) Using Monte Carlo simulations in public health risk assessments: Estimating and presenting full distributions of risk. *J.Exp.Anal.Env.Epi.* 1:491-512.
4. McKone, T.E. (1990) Dermal uptake of organic chemicals from soil matrix. *Risk Anal.* 10:407-419.
5. Finkel, A.M. (1990) Confronting uncertainty in risk management, a guide for decision makers. Resources for the Future Publishers. Washington, D.C.
6. Rubenstein, R.Y. (1981) In, Simulation and the Monte Carlo method. J. Wiley and Sons Publishers. New York, N.Y.
7. Dinovi, M.J. (1992) In, Principles for estimating exposure to direct food additives and chemical contaminants in the diet. Food and Color Additives Branch, Center for Food Safety and Applied Nutrition, FDA. 200 C Street S.W., Washington, D.C. 20204

IX References Continued

Chapter IV A: Introduction

1. U.S. Food and Drug Administration [FDA], Bureau of Foods (1982) In, Toxicological principles for the safety assessment of direct food additives and color additives used in food. National Technical Information Service. Springfield, VA.

Chapter IV B (1 d): Guidelines to Common Issues (Observations and Clinical Tests)

1. Gaylor, D.W., Suber, R.L., Wolff, G.L. and Crowell, J.A. (1987) Statistical variation of selected clinical pathological and biochemical measurements in rodents (42555). *Proc. Soc. Exp. Biol. Med.* 185:361-367.

Chapter IV B (3): Guidelines to Common Issues (Pathology Considerations in Toxicology Studies)

1. Dua, P.N. and Jackson, B.A. (1988) Review of pathology data for regulatory purposes. *Toxicol.Path.* 16(4): 443-450

Chapter IV B (4): Guidelines to Common Issues (Statistical Considerations in Toxicity Studies)

1. Dubey, S.D. [HFN-713] (1985) In, Draft Guideline for the Format and Content of the Statistical Section of an Application. [Docket #85D-0246] Center for Drug Evaluation, FDA, 5600 Fishers Lane, Rockville, Maryland. (June)

Chapter IV B (5): Guidelines to Common Issues (Diets for Toxicology Studies)

1. Belinsky, S.A., Kauffman, F.C., and Thursman, R.G. (1987) Effect of nutrition on monooxygenation and conjugation in the liver. In, Nutritional toxicology. (J.N. Hathcock, ed.) Academic Press. Orlando, FL. pp.41-61.
2. Clayson, D.B. (1975) Nutrition and experimental carcinogenesis: A review. *Cancer Res.* 35:3292-3300.
3. Conner, M.W. and Newberne, P.M. (1984) Drug-nutrient interactions and their implications for safety evaluations. *Fund.Appl.Toxicol.* 4:S341-S356.
4. Meydani, M. (1987) Dietary effects on detoxification processes. In, Nutritional toxicology. (J.N. Hathcock ed.) Academic Press. Orlando, FL. 2:1-39.
5. National Research Council (1982) In, Diet, nutrition, and cancer. National Academy Press. Washington, D.C.
6. Parke, D.V. and Ioannides, C. (1981) The role of nutrition in toxicology. *Ann.Rev.Nutr.* 1:207-234.

IX References Continued

Chapter IV B (5) Guidelines to Common Issues (Diets for Toxicology Studies)
Continued

7. Mitchell, G.V., Dua, P.N., Jenkins, M.Y., Grundel, E., and Arzadon, J.M. (1987) The effect of modifications of the AIN-76A diet on selected nutritional parameters and the incidence of pathology. *Fed.Am.Soc.Exp.Biol.* 46:915.
8. Rader, J.I. (1989) Considerations in designing and using standardized diets in toxicological experiments. In, Nutritional toxicology. (J.N. Hathcock ed.) Academic Press. Orlando, FL. 3:123-139.
9. Kritchevsky, D., Weber, M.M., Buck, C.L., and Klurfeld, D.M. (1986) Calorics, fat and cancer. *Lipids* 21:272-274.
10. Omaye, S.T. (1986) Effects of diet on toxicity testing. *Fed.Amer.Soc. Exp.Biol.* 45:133-135.
11. Ip, C. (1987) Fat and essential fatty acid in mammary carcinogenesis. *Am.J.Clin.Nutr.* 45:218-224.
12. Kritchevsky, D. (1985) Influence of dietary fiber on xenobiotics. In, Xenobiotic metabolism: Nutritional effects. American Chemical Society [ACS]. *ACS Sym. Series No. 277*.
13. Newberne, P.M. and Rogers, A.F. (1986) Labile methyl groups and the promotion of cancer. *Ann.Rev.Nutr.* 6:407-32.
14. Rao, G.N. (1988) Rodent diets for carcinogenesis studies. *J.Nutr.* 118:929-931.
15. Reddy, B.S., Cohn, L.A., McCoy, G.D., Hill, P., Welsburger, J.H., and Snyder, E.L. (1980) Nutrition and its relationship to cancer. *Adv. Cancer Res.* 32:237-345.
16. American Institute of Nutrition (1977) Report of the American Institute of Nutrition Ad Hoc Committee on Standards for Nutritional Studies. *J.Nutr.* 107:1340-1348.
17. Wise, A. and Gilbert, D.J. (1980) The variability of dietary fibre in laboratory animal diets and its relevance to the control of experimental conditions. *Fd.Cosmet.Toxicol.* 18:643-648.
18. Rader, J.I., Wolnik, K.A., Gaston, C.M., Celeski, E.M., Peeler, J.T., Fox, M.R.S., and Fricke, F.L. (1984) Trace element studies in weanling rats: Maternal diets and baseline tissue mineral values. *J.Nutr.* 114:1946-1954.
19. Fox, J.G., Aldrich, F.D., and Boylen, G.W. (1976) Lead in animal foods. *J.Toxicol.Environ. Health* 1:461-467.
20. National Research Council (1978) In, The nutrient requirements of laboratory animals. No. 10, 3rd revised edition. National Academy of Science Press. Washington, D.C.
21. Ross, R.H., Ryon, M.G., Daugherty, M.W., Drury, J.S., and Ensinger, J.T. (1980) Scientific rationale for the selection of toxicity testing methods. In, Human Health Assessment. National Technical Information Service. Springfield, VA.

IX References Continued

Chapter IV B (5) Guidelines to Common Issues (Diets for Toxicology Studies)
Continued

22. Cruse, J.P., Lewin, M.R., Ferulano, G.P., and Clarl, C.G. (1978) Co-carcinogenic effects of dietary cholesterol in experimental colon cancer. *Nature* 276:822-825
23. Cruse, J.P., Lewin, M.R., and Clark, C.G. (1978) Failure of bran to protect against experimental colon cancer in rats. *Lancet* 2:1278-1279.
24. Hamm, T.E., Raynor, T., and Caviston, T. (1982) Unsuitability of the AIN-76 diet for male F-344 and CD rats and improvement by substituting starch for sucrose. *Lab. Animal Sci.* 32:414-415.
25. Nguyen, H.T. and Woodard, T.C. (1980) Intraneuronic calculosis in rats. *Am.J.Path.* 100:39-56.
26. Harwood, E.J. (1982) The influence of dietary magnesium on reduction of nephrocalcinosis in rats fed purified diet. *Lab. Animal Sci.* 16:314-318
27. Medinsky, M.A., Popp, J.A., Hamm, T.E., and Dent, J.G. (1982) Development of hepatic lesions in male Fischer-344 rats fed AIN-76A purified diet. *ToxicolApplPharmacol.* 62:111-120.
28. Munro, H.N. (1985) Nutrient needs and nutritional status in relation to aging. *Drug-Nutrient Interact.* 4:55-74.
29. Garattini, S. (1985) Drug metabolism and actions in the aged. *Drug-Nutrient Interact.* 4:87-97.
30. Welling, P. (1985) Nutrient effects on drug metabolism and action in the elderly. *Drug-Nutrient Interact.* 4:173-207.
31. Levcille, G.A. and Cloutier, P.F. (1987) Isocaloric diets: Effects of dietary changes. *Am.J.Clin.Nutr.* 45:158-163.
32. Nutrition Foundation (1983) In, Report of the Ad Hoc Working Group on Oil/Gavage in Toxicology. The Nutrition Foundation. Washington, D.C.

IX References Continued

Chapter IV C (1): Guidelines for Oral Toxicity Studies (Short-Term Tests for Genetic Toxicity)

1. Tennant, R.W., Margolin, B.H., Shelby, M.D., Zeiger, E., Haseman, J.K., Spalding, J., Caspary, W., Resnick, M., Stasiewicz, S., Anderson, B. and Minor, R. (1987) Prediction of chemical carcinogenicity in rodents from *in vitro* genetic toxicity assays. *Science* 236:933-941.
2. Ashby, J. and Tennant, R.W. (1988) Chemical structure, *Salmonella* mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. *Mutation Research* 204:17-115.
3. Ashby, J. and Tennant, R.W. (1991) Definitive relationships among chemical structure, carcinogenicity and mutagenicity for 301 chemicals tested by the U.S. NTP. *Mutation Research* 257:229-305.
4. Ames, B.N., McCann, J. and Yamasaki, E. (1975) Methods for detecting carcinogens and mutagens with the *Salmonella*/mammalian-microsome mutagenicity test. *Mutation Res.* 31:347-364.
5. Maron, D.M. and B.N. Ames (1983) Revised methods for the *Salmonella* mutagenicity test. *Mutation Res.* 113:173-215.
6. Mitchell, I. de G. (1978) Microbial assays for mutagenicity: a modified liquid culture method compared with the agar plate system for precision and sensitivity. *Mutation Res.* 54:1-16.
7. McGregor, D.B., Martin, R., Cattanaach, P., Edwards, I., McBride, D., and Caspary, W.J. (1987): Responses of the L5178y tk⁺/tk mouse lymphoma cell forward mutation assay to coded chemicals. I: Results for nine compounds. *Environ.Mutagen.* 9(Issue 2):143-160.
8. McGregor, D.B., Brown, A., Cattanaach, P., Edwards, I., McBride, D., and Caspary, W.J. (1988): Responses of the L5178y tk⁺/tk mouse lymphoma cell forward mutation assay. II. 18 coded chemicals. *Environ.Mol.Mutagen.* 11:91-118.
9. McGregor, D.B., Brown, A., Cattanaach, P., Edwards, I., McBride, D., Riach, C., and Caspary, W.J. (1988): Responses of the L5178y tk⁺/tk mouse lymphoma cell forward mutation assay. III. 72 coded chemicals. *Environ.Mol.Mutagen.* 12(Issue 1):85-154.
10. Mitchell, A.D., Rudd, C.J., and Caspary, W.J. (1988): Evaluation of the L5178Y mouse lymphoma cells mutagenesis assay: intralaboratory results for sixty-three coded chemicals tested at SRI International. *Environ.Mol.Mutagen.* 12(Suppl. 13):37-101.
11. Myhr, B.C. and Caspary, W.J. (1988): Evaluation of the L5178Y mouse lymphoma cell mutagenesis assay: Intralaboratory results for sixty-three coded chemicals tested at Litton Bionetics Inc. *Environ.Mol.Mutagen.* 12(Suppl. 13):103-194.
12. Myhr, B., McGregor, D., Bowers, I., Riach, C., Brown, A.G., Edwards, I., McBride, D., Martin, R., and Caspary, W.J. (1990): L5178Y Mouse lymphoma cell mutation assay results with 41 compounds. *Environ.Mol.Mutagen.* 16(Suppl. 18):138-167.
13. Tindall, K.R., Stankowski, L.F., Machanoff, R., and Hsie, A.W. (1984) Detection of deletion mutations in pSV2gpt transformed cells. *Mol.Cell.Biology* 4:1411-1415.

IX References Continued

Chapter IV C (1): Guidelines for Oral Toxicity Studies (Short-Term Tests for Genetic Toxicity) Continued

14. Stankowski, L.F. and Tindall, K.R. (1987) Characterization of the ASS2 cell line for use in mammalian cell mutagenesis studies. In, *Mammalian Cell Mutagenesis*. Banbury Report. Cold Spring Harbor Laboratory, N.Y. (M.M. Moore, D.M. DeMarini, F.J. De Serres and K.R. Tindall, eds.) pp. 71-79.
15. Tindall, K.R. and Stankowski, L.F. (1987) characterization of the ASS2 cell line for use in mammalian cell mutagenesis studies. In, *Mammalian Cell Mutagenesis*. Banbury Report. Cold Spring Harbor Laboratory, N.Y. (M.M. Moore, D.M. DeMarini, F.J. De Serres and K.R. Tindall, eds.) pp. 283-292.
16. Crespi, C.L. and Thilly, W.G. (1984) Assay for gene mutation in a human lymphoblast line, A1111-1, competent for xenobiotic metabolism. *Mutation Res.* 128:221-230.
17. Crespi, C.L., Langenbach, R., and Penman, B.W. (1989) The development of a panel of human cell lines expressing specific human cytochrome p450 c DNAs. *Prog. Clin.Biolog.Res* 340B:97-106.
18. Liber, H.I. and Thilly, W.G. (1982) Mutation assay at the thymidine kinase locus in diploid human lymphoblasts. *Mutation Res.* 94:467-485.
19. Hsieh, A.W., Casciano, D.A., Couch, D.B., Krahn, D.F., O'Neill, J.P. and Whitfield, B.L. (1981) The use of Chinese hamster ovary cells to quantify specific locus mutation and to determine mutagenicity of chemicals. *Mutation Res.* 86:193-214.
20. Li, A.P., Aaron, C.S., Auletta, A.E., Dearfield, K.L., Riddle, J.C., Slesinski, R.S., Stankowski, L.F., (1991) An evaluation of the roles of mammalian cell mutation assays in the testing of chemical genotoxicity. *Res.Tox.Pharmacol.* 14:24-40.
21. Bradley, M.O., Bhuyan, B., Francis, M.C., Langenbach, R., Peterson, A. and Huberman, E. (1981) Mutagenesis by chemical agents in V79 Chinese hamster cells: A review and analysis of the literature. *Mutation Res.* 87:81-142.
22. Clive, D. and Spector, J.F.S. (1975) Laboratory procedures for assessing specific locus mutations at the TK locus in cultured L5178Y mouse lymphoma cells. *Mutation Res.* 31:17-29.
23. Clive, D., Johnson, K.O., Spector, J.F.S., Batson, A.G. and Brown, M.M.M. (1979) Validation and characterization of the L5178Y/tk⁺ mouse lymphoma mutagen assay system. *Mutation Res.* 59:61-108.
24. Clive, D., Moore, M.M., Mitchell, A.D., Myhr, B., Ray, V., Auletta, A.E., Cole, J., Kirby, P., Combes, R., Dearfield, K. and Harbell, J. (1992) Recommendations for the performance and evaluation of the L5178Y/tk⁺ mouse lymphoma assay. *Manuscript in Preparation*.
25. Preston, R.J., Au, W., Bender, M.A., Brewen, J.G., Carrano, A.V., Heddl, J.A., McFee, A.F., Wolff, S., and Wasson, J.S. (1981) Mammalian *in vivo* and *in vitro* cytogenetic assays. A report of the U.S. EPA's Gene-Tox Program. *Mutation Res.* 87:143-188.

IX References Continued

Chapter IV C (1): Guidelines for Oral Toxicity Studies (Short-Term Tests for Genetic Toxicity) Continued

26. Preston, R.J., Dean, B.J., Galloway, S., Holden, H., McFee A.F. and Shelby, M. (1987) Mammalian *in vivo* cytogenetic assays. Analysis of chromosome aberrations in bone marrow cells. *Mutation Res.* 189:157-165.
27. Heddle, J.A., Hite, M., Kirkhart, B., Mavournin, K., MacGregor, J.T., Newell, G.W. and Salamone, M.F. (1983) The induction of micronuclei as a measure of genotoxicity. A report of the U.S. Environmental Protection Agency Gene-Tox Program. *Mutation Res.* 123:61-118.
28. McGregor, J.T., Heddle, J.A., Hite, M., Margolin, B.H., Ramel, C., Salamone, M.F., Tice R.R. and Wild, D. (1987): Guidelines for the conduct of micronucleus assays in mammalian bone marrow erythrocytes. *Mutation Res.* 189:103-112.
29. Klopman, G., Frierson, M.R., Rosenkranz, H.S. (1990): The structural basis of the mutagenicity of chemicals in *Salmonella typhimurium*: The Gene-Tox data base. *Mutation Res.* 228:1-50.
30. Rosenkranz, H.S. and Ennever, F.K. (1988): Quantifying genotoxicity and non-genotoxicity. *Mutation Res.* 205:59-67.
31. Rosenkranz, H.S. and Klopman, G. (1990): Structural basis of carcinogenicity in rodents of genotoxins and non-genotoxins. *Mutation Res.* 228:105-124.
32. Rosenkranz, H.S. and Klopman, G. (1990): The structural basis of the mutagenicity of chemicals in *Salmonella typhimurium*: The National Toxicology Program data base. *Mutation Res.* 228:51-80.
33. Rosenkranz, H.S., Ennever, F.K., Dimayuga, M. and Klopman, G. (1990): Significant differences in the structural basis of the induction of sister chromatid exchanges and chromosomal aberrations in Chinese hamster ovary cells. *Environ.Mol.Mutagen.* 16:149-177.
34. Kakunaga, T. (1973): A quantitative system for assay of malignant transformation by chemical carcinogens using a clone derived from BALB/3T3. *Int.J.Cancer* 12: 463-473.
35. IARC/NCI/EPA Working Group (1985) Cellular and molecular mechanisms of cell transformation and standardization of transformation assays of established cell lines for the prediction of carcinogenic chemicals: Overview and recommended protocols. *Cancer Res.* 45:2395-2399.
36. Matthews, E.J. (1986) Assessment of chemical-induced transforming activity using BALB/c-3T3 cells. *J. Tissue Culture Methods* 10:157-164.
37. Matthews, E.J., Spalding, J.W. and Tennant, R.W. (1993) Transformation of BALB/c-3T3 cells IV. Rank-ordered potency of 24 chemical responses detected in a sensitive new assay procedure. *Environ. Health Perspectives* 103: In Press.
38. Matthews, E.J. (1993) Transformation of BALB/c-3T3 cells III. Development of an assay for quantification of chemical cytotoxicity of cells in high density cultures. *Environ. Health Perspectives* 103: In Press.

IX References Continued

Chapter IV C (1): Guidelines for Oral Toxicity Studies (Short-Term Tests for Genetic Toxicity) Continued

39. Matthews, E.J., Spalding, J.W. and Tennant, R.W. (1993) Transformation of BALB/c-3T3 cells V. Transformation responses of 168 chemicals compared with mutagenicity in *Salmonella* and carcinogenicity in rodent bioassays. *Environ. Health Perspectives* 103: In Press.
40. Matthews, E.J., Rosenkranz, H.S., Spalding, J.S., and Tennant, R.W. (1993) Transformation of BALB/c-3T3 cells VII. Structural basis of the chemical-induced transformation response. *Manuscript in Preparation*.
41. Williams, G.M., Laspia, M.F. and Dunkel, V.C. (1982) Reliability of the hepatocyte primary culture/DNA repair test in testing of coded carcinogens and noncarcinogens. *Mutation Res.* 97:359-370.
42. Butterworth, B.E., Ashby, J., Bermudez, E., Casciano, D., Mirsalis, J., Probst, G. and Williams, G. (1987) A protocol and guide for the *in vitro* rat hepatocyte DNA-repair assay. *Mutation Res.* 189:113-121.
43. Butterworth, B.E., Ashby, J., Bermudez, E., Casciano, D., Mirsalis, J., Probst, G. and Williams, G. (1987) A protocol and guide for the *in vivo* rat hepatocyte DNA-repair assay. *Mutation Res.* 189:123-133.
44. Shaddock, J.G., Heflich, R.H., McMillan, D.C., Hinson, J.A., and Casciano, D. (1989) Pretreatment with mixed-function oxidase inducers increases the sensitivity of the hepatocyte/DNA repair assay. *Environ.Mol.Mutagen.* 13:281-288.
45. Shaddock, J.G., Robinson, B.Y. and Casciano, D. (1990) Effect of pretreatment with hepatic mixed-function oxidase inducers on the genotoxicity of four rat carcinogens in the hepatocyte/DNA repair assay. *Mutagenesis* 5:387-391.

Chapter IV C (2): Guidelines for Oral Toxicity Studies (Acute Toxicity Studies)

1. Gad, S.C. and Chengelis, C.P. (1988) In, Acute toxicity testing perspectives and horizons. (S.C. Gad and C.P. Chengelis eds.) The Telford Press. Caldwell, N.J. pp. 2-4, 318^a, 156^b, 165-167^c, 159^d.
2. Food and Drug Administration [FDA] (1988) LD₅₀ Test Policy. *Fed.Reg.* 53 (October 11 Issue):39650-39651.
3. Litchfield, J.T., Jr., Wilcoxon, F. (1949) A simplified method of evaluating dose-effect experiments. *J.Pharmacol.Exp.Therap.* 96:113.
4. Anonymous. (1989) Comments on the LD₅₀ and acute eye and skin irritation tests. In, a position paper reported by the Animals in Research Committee of the Society of Toxicology (SOT) and approved by the SOT Council. *Fund.Appl.Toxicol.* 13:621-623.
5. Zbinden, G. and Flury-Roversi, M. (1981) Significance of the LD50 test for the toxicological evaluation of chemical substances. *Arch.Toxicol.* 47:77-99.

IX References Continued

Chapter IV C (2): Guidelines for Oral Toxicity Studies (Acute Toxicity Studies)
Continued

6. Lorke, D. (1983) A new approach to practical acute toxicity testing. *Arch.Toxicol.* 54:275-287.
7. Schutz, E. & Fuchs, H. (1982). A new approach to minimizing the number of animals used in acute toxicity testing and optimizing the information of test results. *Arch.Toxicol.* 51:197-220.
8. Deichmann, W.B. and LeBlanc, T.J. (1943) Determination of the approximate lethal dose with about six animals. *J.Indust. Hygiene Toxicol.* 25:415-417.
9. Bruce, R.D. (1985) An up-and-down procedure for acute toxicity testing. *Fund.Appl.Toxicol.* 5:151-157.
10. Gad, S.C., Smith, A.C., Cramp, A.L., Gavigan, F.A., and Derelanko, M.J. (1984) Innovative Designs and Practices for Acute Systemic Toxicity Studies. *Drug Chem.Toxicol.* 7:423-434.
11. Muller, H. and Kley, H.P. (1982) Retrospective study on the reliability of an "approximate LD₅₀" determined with a small number of animals. *Arch.Toxicol.* 51:189-196.

Chapter IV C (6): Guidelines for Oral Toxicity Studies (Carcinogenicity Studies with Rodents)

1. Apostolou, A. (1990) Relevance of maximum tolerated dose to human carcinogenic risk. *Reg.Toxicol.Pharmacol.* 11:68-80.
2. Clayton, D.B. and Clegg, D.J. (1991) Classification of carcinogens: polemics, pedantics or progress? *Reg.Toxicol.Pharmacol.* 14:147-166.
3. Goodman, G. and Wilson, R. (1991) Predicting the carcinogenicity of chemicals in humans from rodent bioassay data. *Environ.Hlth.Perspect.* 94:195-218.
4. Parodi, S., Malacarne, D., Romano, P. and Taningher, M. (1991) Are genotoxic carcinogens more potent than nongenotoxic carcinogens? *Environ.Hlth.Perspect.* 95:199-204.
5. Parry, J.M. (1992) Prediction of the carcinogenic potential of chemicals. *Mutagen.* 7:1-2.
6. Perera, F.P. (1991) Perspectives on the risk assessment for nongenotoxic carcinogens and tumor promoters. *Environ.Hlth.Perspect.* 94:231-235.
7. Tomatis, L., Narod, S. and Yamasaki, H. (1992) Transgeneration transmission of carcinogenic risk. *Carcinogenesis* 13:145-151.
8. Travis, C.C., ed. (1988) In, Carcinogen Risk Assessment. Plenum Press, New York, N.Y.
9. Travis, C.C., Wang, L.A. and Waehner, M.J. (1991) Quantitative correlation of carcinogenicity potency with four different classes of short-term test data. *Mutagen.* 6:353-360.
10. Vainio, H. and Cardis, E. (1992) Estimating human cancer risk from the results of animal experiments: relationships between mechanism and dose-rate and dose. *Am.J.Int.Med.* 21:5-14.

IX References Continued

Chapter IV C (6): Guidelines for Oral Toxicity Studies (Carcinogenicity Studies with Rodents) Continued

11. U.S. Food and Drug Administration [FDA], Bureau of Foods (1982) In, Toxicological principles for the safety assessment of direct food additives and color additives used in food. National Technical Information Service, Springfield, VA. p.126.
12. Personal communication between D. Hattan and J.D. Burek (1990) Toxicology Forum Summer Meeting, Aspen, Colorado (July 16-20).
13. Personal communication with G.N. Rao (1990) Toxicology Forum Summer Meeting, Aspen, Colorado (July 16-20).
14. National Toxicology Program, Board of Scientific Counselors (1984) In, report of the NTP Ad Hoc Panel on Chemical Carcinogenesis testing and evaluation. National Toxicology Program. p.126.
15. Anonymous. (1985) Chemical carcinogens; A review of the science and its associated principles. Part II. Office of Science and Technology Policy. *Fed.Reg.* 15(Thursday, March 14):whole issue.
16. National Toxicology Program, Board of Scientific Counselors (1984) In, report of the NTP Ad Hoc Panel on Chemical Carcinogenesis testing and evaluation. National Toxicology Program. pp.125-140.
17. Schwetz, B.A. (1983) In, Prechronic toxicological tests scope and use. A report by the National Toxicology Program, National Institute of Environmental Health Sciences. Research Triangle Park, NC.
18. McConnell, E.E. (1989) The maximum tolerated dose: the debate. *J.Am. College Toxicol.* 8:1115-1120.

Chapter IV C (9): Guidelines for Oral Toxicity Studies (Reproduction and Developmental Toxicity Studies)

1. Collins, T. F. X. (1978) Multigeneration reproduction studies. In, Handbook of Teratology. Vol 4. (T. G. Wilson and F. C. Fraser eds.) Plenum Publishing Co. New York, N.Y. PP. 191-214.
2. Francis, E. Z. and Kimmel, G. L. (1988) Proceedings of the workshop on one- vs two-generation reproductive effects studies. *J.Am.Coll.Toxicol.* 7:911-925.
3. Environmental Protection Agency. (1988) Part II. Proposed guidelines for assessing female reproductive risk; notice. *Fed.Reg.* 53(126):24833-24847.
4. Environmental Protection Agency. (1988) Part III. Proposed guidelines for assessing male reproductive risk and request for comments. *Fed.Reg.* 53(126):24849-24869.
5. Environmental Protection Agency. (1991) Guidelines for developmental toxicity risk assessment. *Fed.Reg.* 56(234):63798-63826.

IX References Continued

Chapter IV C (9): Guidelines for Oral Toxicity Studies (Reproduction and Developmental Toxicity Studies) Continued

6. Staples, R. E. (1977) Detection of visceral alterations in mammalian fetuses. *Teratology* 9:437-438.
7. Barrow, M. V. and Taylor, W. J. (1969). A rapid method for detecting malformations in rat fetuses. *J. Morphol.* 127:291-305.
8. Wilson, J. G. (1965) Methods for administering agents and detecting malformations in experimental animals. In, Teratology Principles and Techniques. (J. G. Wilson and J. Warkany eds.) University of Chicago Press. Chicago, Ill. pp. 262-277.
9. Lamb, J. C. (1989) Design and use of multigeneration breeding studies for identification of reproductive toxicants. In, Toxicology of the Male and Female Reproductive Systems. (P. K. Working, ed.) Hemisphere Publishing Co. Washington, D.C. pp. 131-155.
10. Reel, J. R., Lawton, A. D., Wolkowski-Tyl, R., Davis, G. W. and Lamb, J. C. (1985) Evaluation of a new reproductive toxicology protocol using diethylstilbestrol (DES) as a positive control compound. *J. Am. Coll. Toxicol.* 4:147-162.
11. Environmental Protection Agency. (1985) Hazard evaluation division standard evaluation procedure. Teratology studies. U. S. EPA, Office of Pesticide Programs. Washington, D. C. EPA-540/9-85-018. p. 31.

Chapter V B: Metabolism and Pharmacokinetic Studies

1. Bäärnhielm, C., Dahlback, H. and Skanberg, I. (1986) *In vivo* pharmacokinetics of felopidine predicted from *in vitro* studies in rat, dog, and man. *Acta Pharmacol. Toxicol.* 59:113-122.
2. Green, C.E., LeValley, S.E. and Tyson, C.A. (1986) Comparison of amphetamine metabolism using isolated hepatocytes from five species including human. *J. Pharmacol. Exp. Ther.* 237(3):931-936.
3. Lin, J.H., Sugiyama, Y., Awazu, S. and Hanano, M. (1982) *In vitro* and *in vivo* evaluation of the tissue-to-blood partition coefficient for physiological pharmacokinetic models. *J. Pharmacokin. Biopharm.* 10:637-647.
4. Rane, A., Wilkinson, G.R. and Shand, D.G. (1977) Prediction of hepatic extraction ratio from *in vitro* measurement of intrinsic clearance. *J. Pharmacol. Exp. Ther.* 200:420-424.
5. Gillette, J.R. (1986) On the role of pharmacokinetics in integrating results from *in vivo* and *in vitro* studies. *Fd. Chem. Toxicol.* 24:711-720.
6. Wagner, J.G. and Nelson, E. (1963) Percent absorbed time plots derived from blood level and/or urinary excretion data. *J. Pharm. Sci.* 52:610-611.
7. Gibaldi, M., and Perrier, D., eds. (1982) In, Pharmacokinetics. 2nd edition. Marcel Dekker Publishing Co. New York, N.Y.

IX References Continued

Chapter V B: Metabolism and Pharmacokinetic Studies Continued

8. Notari, R.E. (1987) In, Biopharmaceutics and Clinical Pharmacokinetics. Marcel Dekker Publishers, Inc. New York, N.Y. Chapter 4.
9. Chen, H.S. and Gross, J.F. (1979) Estimation of tissue-to-plasma partition coefficients used in physiological pharmacokinetic models. *J.Pharmacokin.Biopharm.* 7(1):117-125.
10. Lam, G., Chen, M-L. and Chiou, W.L. (1982) Determination of tissue-to-blood partition coefficients in physiologically-based pharmacokinetic studies. *J.Pharm.Sci.* 71:454-456.
11. Benet, L.Z. (1972) General treatment of linear mammillary models with elimination from any compartment as used in pharmacokinetics. *J.Pharm.Sci.* 61:536-541.
12. Vaughan, D.P. and Trainor, A. (1975) Derivation of general equations for linear mammillary models when the drug is administered by different routes. *J.Pharmacokin.Biopharm.* 3:203-218.
13. Pedersen, P.V. (1977) Curve fitting and modeling in pharmacokinetics and some practical experiences with NONLIN and a new program FUNFIT. *J.Pharmacokin.Biopharm.* 5:513-531.
14. Boston, R.C., Palmer, M.S., Golden, P.D. and Ramberg, C.F. (1988) Users notes for enhancement to SAAM/CONSAM Version 30 resource facilities for kinetic analysis. Contact the authors at the National Institute of Health, Building 10, Bethesda, Maryland.
15. Steiner, E.C., Rey, T.D., McCroskey, P.S. (1990) Simusolv modeling and simulation software reference. Dow Chemical Co. Midland, Michigan.
16. Rescigno, A. and Segre, G. (1966) In, Drug and Tracer Kinetics. 1st American edition. Blaisdale Publishing Co. Waltham, MA. p. 91.
17. Fagarasan, J.T. and DiStefano, J.J. (1989) Hidden oscillations in generalized linear mammillary compartmental models. *Math.Biosci.* 93:79-95.
18. O'Flaherty, F.J. (1989) Interspecies conversion of kinetically equivalent doses. *Risk Anal.* 9:587-598.
19. Gerlowski, L.F. and Jain, R.K. (1983) Physiologically based pharmacokinetic modeling: principals and applications. *J.Pharmacol.Sci.* 72(10):1103-1127.
20. National Toxicology Program Ad Hoc Panel (1984) In, report of the National Toxicology Program Ad Hoc Panel on Chemical Carcinogenesis testing and evaluation. pp.125-127.

IX References Continued

Chapter V C: Neurotoxicity Studies

1. Federation of American Societies for Experimental Biology [FASEB] (1986) Predicting neurotoxicity and behavioral dysfunction from preclinical toxicologic data. In, final report, Task Order #3, Contract No. [REDACTED], Food and Drug Administration [FDA], Washington, D.C.
2. Leukroth, R.W. (1987) Predicting neurotoxicity and behavioral dysfunction from preclinical toxicologic data. *Neurotoxicology and Teratology* 9:395-471.
3. McMillan, D.E. (1987) Risk assessment for neurobehavioral toxicity. *Environ. Health Perspect.* 76:155-161.
4. Vorhees, C.V. (1987) Reliability, sensitivity and validity of behavioral indices of neurotoxicity. *Neurotoxicology and Teratology* 9:445-464.
5. Office of Technology Assessment [OTA] (1990) In, Neurotoxicity: Identifying and controlling poisons of the nervous system. Government Printing Office, Washington, D.C. Document OTA-B4-436.
6. World Health Organization [WHO] (1986) In, Principles and methods for the assessment of neurotoxicity associated with exposure to chemicals. World Health Organization Press, Geneva, Switzerland. *Environmental Health Criteria Document* #60.
7. Spencer, P.S. and Schaumburg, H.H., eds. (1980) In, Experimental and clinical neurotoxicology. Williams and Wilkins Press, Baltimore, MD.
8. Tilson, H.A. (1987) Behavioral indices of neurotoxicity: What can be measured? *Neurotoxicology and Teratology* 9:427-443.
9. Buelke-Sam, J., Kimmel, C.A. and Adams, J. eds. (1985) Design considerations in screening for behavioral teratogens: Results of the collaborative behavioral teratology study. In, Neurobehavioral Toxicology and Teratology. Conference Proceedings 7(6):537-789.
10. Reiter, L.W. (1987) Neurotoxicology in regulation and risk assessment. *Dev. Pharmacol. Therapeutics* 10:354-368.
11. Anger, W.K. and Johnson, B.L. (1985) Chemicals affecting behavior. In, Neurotoxicity of Industrial and Commercial Chemicals. (J.L. O'Donoghue ed.) CRC Press, Boca Raton, FL. Vol. 1:51-148.
12. Riley, E.P. and Vorhees, C.V. eds., (1986) Handbook of behavioral teratology. Plenum Press, New York, N.Y.
13. Hutchings, D.E., Callaway, C.W. and Sobotka, T.J. (1987) Predicting neurotoxicity and behavioral dysfunction from preclinical toxicologic data, Symposium Introduction. *Neurotoxicology and Teratology* 9:397-401.
14. Vorhees, C.V., Butcher, R.E. and Brunner, R.L. (1984) Developmental toxicity and psychotoxicity of potassium iodide in rats: A case for the inclusion of behavior in toxicological assessment. *Fd. Chem. Toxicol.* 22:963-970.

IX References Continued

Chapter V C: Neurotoxicity Studies Continued

15. Mitchell, C.I. and Tilson, H.A. (1982) Behavioral toxicology in risk assessment: Problems and research needs. *CRC Crit.Rev.Toxicol.* 10:265-274.
16. U.S. Environmental Protection Agency [EPA] (1985) Toxic Substances Control Act Test Guidelines Final Rule, 40 CFR Part 798, Health effects guidelines, Subpart G - Neurotoxicity. *Fed. Reg.* 50:39458-39470.
17. U.S. Environmental Protection Agency [EPA] (1991) Pesticide Assessment Guidelines, Subdivision F Hazard Evaluation: Human and Domestic Animals. Addendum 10: Neurotoxicity, Series 81, 82, and 83, EPA-540/09-91-123. Health Effects division, Office of Pesticide Programs. EPA. Washington, D.C.
18. Kimmel, C.A., Rees, D.C. and Francis, E.Z. (1990) Qualitative and quantitative comparability of human and animal developmental neurotoxicity. *Neurotoxicology and Teratology* 12(3):171-292.
19. Nation Research Council, National Academy of Sciences [NRC/NAS] (1982) In, Strategies to determine needs and priorities for toxicity testing. Vol. 2. National Academy of Sciences Press. Washington, D.C.
20. Sobotka, T.J. (1986) The regulatory perspective of diet-behavior relationships. *Nutrition Rev.* 44:241-245.
21. U.S. House of Representatives Committee on Science and Technology (1986) In, Neurotoxins: At home and in the workplace. 99th Congress, Second Session. U.S. Government Printing Office. Washington, D.C. Report 99-827.
22. National Research Council, Committee on Toxicology, National Academy of Sciences [NRC/NSA] (1975) In, Principles for evaluating chemicals in the environment. National Academy of Sciences Press. Washington, D.C. pp.198-216.
23. U.S. Food and Drug Administration [FDA] (1982) In, Toxicological principles for the safety assessment of direct food additives and color additives used in foods. Document PB83-170696. National Technical Information Service. Springfield, VA.
24. Nelson, B.K. (1991) Selecting exposure parameters in developmental neurotoxicity assessments. *Neurotoxicol.Teratol.* 13:569-573.
25. Tilson, H.A. and Mitchell, C.I. (eds.) (1992) In, Neurotoxicology. Raven Press. New York, N.Y.
26. U.S. Environmental Protection Agency [EPA] (1988) Diethylene glycol ethers: test standards and requirements final rule. 40 CFR Parts 79 and 799. *Fed. Register* 53:5931-5953.
27. Vorhees, C.V., Butcher, R.E., Brunner, R.J. and Sobotka, T.J. (1979) A developmental test battery for neurobehavioral toxicity in rats: A preliminary analysis using monosodium glutamate, calcium carrageenan, and hydroxyurea. *Toxicol.Appl.Pharmacol.* 50:267-282.

IX References Continued

Chapter V C: Neurotoxicity Studies Continued

28. Spencer, P.S., Bischoff, M.C. and Schaumburg, H.H. (1980) Neuropathological methods for the detection of neurotoxic disease. In, Experimental and clinical neurotoxicology. (P.S. Spencer and H.H. Schaumburg, eds.) Williams and Wilkins Press. Baltimore, MD. pp.743-757.
29. Wier, P.J., Guerriero, F.J. and Walker, R.F. (1989) Implantation of a primary screen for developmental neurotoxicity. *Fund.Appl.Toxicol.* 13:118-136.
30. Geller, I., Stebbins, W.C. and Wayner, M.J., eds. (1980) Test methods for definition of effects of toxic substances on behavior and neuromotor function: Conference Proceedings. *NTIS Report No. ISBN-O-916086-02-2*.
31. Tupper, D.E. and Wallace, R.B. (1980) Utility of the neurological examination in rats. *Acta Neurobiol.Exp.* 40:999-1003.
32. Deuel, R.K. (1977) Determining sensory deficits in animals. In, Methods in psychobiology. (R. Myers, ed.) Academic Press. New York, N.Y. pp. 99-125.
33. Gad, S.C. (1982) A neuromuscular screen for use in industrial toxicology. *J.Toxicol.Environ. Health* 9:691-704.
34. Gad, S.C. (1989) Principles of screening in toxicology with special emphasis on applications to neurotoxicology. *J.Am. College Toxicol.* 8:21-27.
35. Marshall, J.F., Turner, B.H. and Teitelbaum, P. (1971) Sensory neglect produced by lateral hypothalamic damage. *Science* 174:523-525.
36. O'Donoghue, J.L. (1989) Screening for neurotoxicity using a neurologically based examination and neuropathology. *J.Am. College Toxicol.* 8:97-115.
37. Moser, V.C., McCormick, J.P., Creason, J.P., and MacPhail, R.C. (1988) Comparison of chlordimeform and carbaryl using a functional observational battery. *Fund.Appl.Toxicol.* 11:189-206.
38. Schulze, G.E. and Boysen, B.G. (1991) A neurotoxicity screening battery for use in safety evaluation: Effects of acrylamide and 3,3-iminodipropionitrile. *Fund.Appl.Toxicol.* 16:602-615.

IX References Continued

Chapter V D: Immunotoxicity Studies

1. U.S. Food and Drug Administration [FDA]. (1978) In, Inadvertent modification of the immune response: The effect of food, drugs and environmental contaminants. Proceedings of the fourth FDA science symposium. (I.M. Asher ed.) U.S. Government Printing Office. Washington, D.C.
2. Luster, M.I., Munson, A.H., Thomas, P.T., Holsapple, M.P., Fenters, J.D., White Jr., K.L., Lauer, I.D., Germolec, D.R., Rosenthal, G.J. and Dean, J.H. (1988) Development of a testing battery to assess chemical-induced immunotoxicity: National Toxicology Program's guidelines for immunotoxicity evaluation in mice. *Fund.Appl.Toxicol.* 10(1):2-19.
3. U.S. Environmental Protection Agency [EPA] (1982) In, Pesticides assessment guidelines, subdivision F. Hazard evaluation: Human and domestic animals. Document No. EPA-540/9-82-025. National Technical Information Services Accession #PB83-153916. U.S. Department of Commerce, Washington, DC.
4. Sjoblad, R.D. (1988) Potential future requirements for immunotoxicology testing of pesticides. *Toxicol.Ind.Health* 4(3):391-395.
5. Tryphonas, H. (1990) *personal communication*.
6. United Nations Environment Programme International Labor Office, World Health Organization [WHO] (1986) In, Immunotoxicology: Development of predictive testing for determining the immunotoxic potential of chemicals. Report of a Technical Review Meeting (November 3-5). WHO Press. London, U.K.
7. Bick, P.H. (1982) Immune system as a target organ for toxicity. *Environ. Health Perspect.* 43:3-7.
8. Dean, J.H., Luster, M.I., Boorman, G.A. and Lauer, I.D. (1982) Procedures available to examine the immunotoxicity of chemicals and drugs. *Pharmacol.Rev.* 34:137-148.
9. White, K.L., Sanders, V.M., Barnes, D.W., Shopp Jr., G.M. and Munson, A.E. (1985) Immunotoxicological investigations in the mouse: General approach and methods. *Drug Chem.Toxicol.* 8(5):299-331.
10. Exon, J.H., Koller, I.D., Henningsen, G.M. and Osborne, C.A. (1984) Multiple immunoassay in a single animal: A practical approach to immunotoxicologic testing. *Fund.Appl.Toxicol.* 4:278-283.
11. Vos, J.G., Krajnc, E.I., Beekhof, P.K. and van Logten, M.J. (1982) Methods for testing immune effects of toxic chemicals: Evaluation of the immunotoxicity of various pesticides in the rat. In, Pesticide Chemistry, Human Welfare and the Environment. Vol. 3. Mode of Action, Metabolism and Toxicology. (J. Miyamoto and P.C. Kearney eds.) Pergamon Press, Oxford, U.K. pp. 497-504.
12. van Loveren, H., Krajnc, E.I. and Vos, J.G. (1988) Development of a testing battery to assess chemical-induced immunotoxicity [letter to the editor]. *Fund.Appl.Toxicol.* 11(4):731-733.
13. Schoental, R. (1988) Immunotoxicology, AIDS and toxic chemicals [letter]. *J.Appl.Toxicol.* 8(4):295.

IX References Continued

Chapter V D: Immunotoxicity Studies Continued

14. Dean, J.H. and Thurmond, L.M. (1987) Immunotoxicology: An overview. *Toxicol.Path.* 15(3):265-271.
15. Descotes, G. and Mazauc, G. (1987) Immunotoxicology. *Adv.Vet.Sci.Comp.Med.* 31:95-119.
16. Koller, L.D. (1987) Immunotoxicology today. *Toxicol.Pathol.* 15(3):346-351.
17. Irons, R.D. (1989) Immunotoxicology--1987. *Year Immunol.* 4:296-299.
18. Dean, J.H., Murray, M.J. and Ward, E.C. (1986) Toxic responses of the immune system. In, Casarett and Doull's, Toxicology: The Basic Science of Poisons. 3rd edition. (C.D. Klassen, M.O. Amdur and J. Doull eds.) MacMillan Press. New York, N.Y. pp. 245-285.
19. Herberman, R.B. and Holden, H.T. (1978) Natural cell mediated immunity. *Adv. Cancer Res.* 27:305-377.
20. Bigazzi, P.E. (1988) Autoimmunity induced by chemicals. *J.Toxicol.Clin.Toxicol.* 26(3-4):125-156.
21. Jacobsen, D.W. and Gunnison, A.F. (1987) Sulfite hypersensitivity. A critical review. *CRC Crit.Rev.Toxicol.* 17(3):185-214.
22. Kammuller, M.E., Bloksma, N. and Seinen, W. (1988) Chemical induced autoimmune reactions and Spanish toxic oil syndrome: Focus on hydantoins and related compounds. *J.Toxicol.Clin.Toxicol.* 26(3-4):157-174.
23. Belongia, E.A., Hedberg, C.W., Gleich, G.J., White, K.E., Mayeno, A.N., Loegering, D.A., Dunnette, S.L., Pirie, P.L., MacDonald, K.L. and Osterholm, M.T. (1990) An investigation of the cause of the eosinophilia-myalgia syndrome associated with tryptophan use. *New Eng.J.Med.* 323(6):357-365.
24. Koller, L.D. and Eron, J.L. (1985) The rat as a model for immunotoxicity assessment. In, Immunotoxicology and immunopharmacology. (J.H. Dean, M.I. Luster, A.E. Munson and H. Amos eds.) Raven Press. New York, N.Y. pp. 99-111.
25. Lovett, E.J., Schnitzer, B., Karen, D.F., Flint, A., Hudson, J.L. and McClatchey, K.D. (1984) Application of flow cytometry to diagnostic pathology. *Lab.Invest.* 50:115-140.
26. Hudson, J.L., Duque, R.E. and Lovett, E.J. (1985) Applications of flow cytometry in immunotoxicology. In, Immunotoxicology and immunopharmacology. (J.H. Dean, M.I. Luster, A.E. Munson and H. Amos eds.) Raven Press, New York, N.Y. pp. 159-177.
27. Burchiel, S.W., Hadley, W.M., Cameron, C.L., Fincher, R.H., Lim, T.W., Elias, L. and Stewart, C.C. (1987) Analysis of heavy metal immunotoxicity by multiparameter flow cytometry: correlation of flow cytometry and immune function data in B6CFl mice. *Int.J.Immunopharmacol.* 9(5):597-610.
28. Hinton, D.M., Petrali, J.P., Meyer, H.G. and Sternberger, L.A. (1973) The unlabeled antibody enzyme method of immunohistochemistry: Molecular immunocytochemistry of antibodies on the erythrocyte surface. *J.Histochem.Cytochem.* 21(11):978-998.

IX References Continued

Chapter V D: Immunotoxicity Studies Continued

29. Falini, B. and Taylor, C.R. (1983) New developments in immunoperoxidase techniques and their application. *Arch.Pathol.Lab.Med.* 107:105-117.
30. Hinton, D.M., Jessop, J.J., Arnold, A., Albert, R.H. and F. A. Hines (1987) Evaluation of immunotoxicity in a subchronic feeding study of triphenyl phosphate. *Toxicol.Ind.Health* 3(1):71-89.
31. Roginski, H.B. and Hinton, D.M. (1987) Applications of immunoperoxidase histochemistry in immunotoxicology. *Fed.Am.Soc.Exp.Biol.* 46:544.
32. Dunnett, C.W. (1955) A multiple comparison procedure for comparing several treatments with a control. *J.Am.Stat.Assoc.* 50:1096-1121.
33. Gorelik, E. and Herberman, R.B. (1981) Inhibition of the activity of mouse natural killer (NK) cells by urethan. *J.Natl.Cancer Inst.* 66:543-548.
34. Smialowicz, R.J., Andrews, J.E., Riddle, M.M., Rogers, R.R., Luebke, R.W. and Copeland, C.B. (1989) Evaluation of the immunotoxicity of low level PCB exposure in the rat. *Toxicol.* 56(2):197-211.
35. Holsapple, M.P., White Jr., K.L., McCay, J.A., Bradley, S.G. and Munson, A.E. (1988) An immunotoxicological evaluation of 4,4'-thiobis-(6-1-butyl-m-cresol) in female B6C3F1 mice. 2. Humoral and cell-mediated immunity, macrophage function, and host resistance. *Fund.Appl.Toxicol.* 10(4):701-716.
36. Oppenheim, J.J. and D. L. Rosenstreich, D.L., eds. (1976) In, Mitogens in immunobiology. Academic Press, New York, N.Y.
37. Cornacoff, J.B., Lauer, L.D., House, R.V., Tucker, A.N., Thurmond, L.M., Vos, J.G., Working, P.K. and Dean, J.H. (1988) Evaluation of the immunotoxicity of beta-hexachlorocyclohexane (beta-HCH). *Fund.Appl.Toxicol.* 11(2):293-299.
38. Loose, L.D., Silkworth, J.B., Charbonneau, T. and Blumenstock, F. (1981) Environmental chemical-induced macrophage dysfunction. *Environ. Health Perspect.* 39:79-91.
39. Lewis, J.G. and Adams, D.O. (1985) The mononuclear phagocyte system and its interaction with xenobiotics. In, Immunotoxicology and immunopharmacology. (J.H. Dean, M.I. Luster, A.E. Munson, and H. Amos eds.) Raven Press. New York, N.Y. pp. 23-44.
40. Keller, R. (1978) Macrophage-mediated natural cytotoxicity against various target cells *in vitro*. I. Macrophages from diverse anatomical sites and different strains of rats and mice. *Brit.J.Cancer* 37:732-741.
41. Harigaya, K., Cronkite, E.P., Miller, M.E., and Shadduck, R.K. (1981) Murine bone marrow cell line producing colony-stimulating factor. *Proc.Natl.Acad.Sci* 78:6963-6966.
42. Pestka, J.J., Tai, J.H., Witt, M.F., Dixon, D.E., and Forsell, J.H. (1987) Suppression of immune response in the B6C3F1 mouse after dietary exposure to the *Fusarium* mycotoxins deoxynivalenol (vomitoxin) and zearalenone. *Fd.Chem.Toxicol.* 25(4):297-304.

IX References Continued

Chapter V D: Immunotoxicity Studies Continued

43. Reddy, R.V., Taylor, M.J. and Sharma, R.P. (1987) Studies of immune function of CD-1 mice exposed to aflatoxin B1. *Toxicol.* 43(2):123-132.
44. Baccher-Steppan, L., Nakaue, H.S., Matsumoto, M., Gainer, J.H. and Kerkvliet, N.I. (1989) The broiler chicken as a model for immunotoxicity assessment. 1. Standardization of *in vitro* immunological assays. *Fund Appl Toxicol.* 12(4):773-786.
45. Jerne, N.K. and Nordin, A.A. (1963) Plaque formation in agar by single antibody-producing cells. *Science* 140:405.
46. Cunningham, A.J. (1965) A method of increased sensitivity for detecting single antibody-forming cells. *Nature* 207:1106-1107.
47. Andersson, B. and Blomgren, H. (1971) Evidence for thymus-independent humoral antibody production in mice against polyvinylpyrrolidone and *E. coli* lipopolysaccharide. *Cell Immunol.* 2:411-424.
48. Vos, J.G. (1977) Immune suppression as related to toxicology. *CRC Crit.Rev.Toxicol.* 5(1):67-101.
49. Speirs, R.S., Benson, R.W., and Roberts, D.W. (1979) Modification of antibody response to type III pneumopolysaccharide by route of injection of pertussis vaccine. *Infect.Immun.* 23(3):675-680.
50. Benson, R.W. and Roberts, D.W. (1982) Evaluation of the immune response to type III pneumococcal polysaccharide as a means to evaluate T-independent immune function in the rat. *J.Toxicol.Environ.Health* 10:859-870.
51. Schiffman, G. (1982) Immune response to pneumococcal vaccine: Two uses. *Clin.Immunol. News L.* 3(6):33-37.
52. Godfrey, H.P. and Gell, P.G. (1978) Cellular and molecular events in the delayed-onset hypersensitivities. *Rev.Physiol.Biochem.Pharmacol.* 84:1-92.
53. Maclean, L.D. (1979) Host resistance in surgical patients. *J.Trauma* 19:297-304.
54. Bradley, S.G. and Morahan, P.S. (1982) Approaches to assessing host resistance. *Environ.Health Perspect.* 43:61-69.
55. Dempsey, W.L. and Morahan, P.S. (1985) Immune mechanisms of host resistance to viruses and assessment of immunotoxicologic effects. In, Immunotoxicology and immunopharmacology. (J.H. Dean, M.I. Luster, A.E. Munson and H. Amos eds.) Raven Press. New York, N.Y. pp. 55-68.
56. Kern, E.R. (1982) Use of viral infections in animal models to assess changes in the immune system. *Environ. Health Perspect.* 43:71-79.

Chapter V D: Immunotoxicity Studies Continued

57. Selgrade, M.K., Daniels, M.J., Burleson, G.R., L. D. Lauer, L.D. and Dean, J.H. (1988) Effects of 7,12-dimethylbenz[a]anthracene, benzo[a]pyrene and cyclosporin A on murine cytomegalovirus infection: Studies of resistance mechanisms. *Int.J.Immunopharmacol.* 10(7):811-818.
58. Dean, J.H., Luster, M.I., Boorman, G.A., Luebke, R.W. and Lauer, L.D. (1982) Applications of tumor, bacterial, and parasite susceptibility assays to study immune alterations induced by environmental chemicals. *Environ. Health Perspect.* 43:81-88.
59. Herberman, R.B. (1985) Immunological mechanisms of host resistance to tumors. In, Immunotoxicology and immunopharmacology. (J.H. Dean, M.I. Luster, A.E. Munson, and H. Amos eds.) Raven Press. New York, N.Y. pp. 69-77.
60. Murray, M.J., Kerkvliet, N.I., Ward, E.C. and Dean, J.H. (1985) Models for the evaluation of tumor resistance following chemical or drug exposure. In, Immunotoxicology and immunopharmacology. (J.H. Dean, M.I. Luster, A.E. Munson and H. Amos eds.) Raven Press. New York, NY. pp. 113-122.
61. Smialowicz, R.J., Rogers, R.R., Rowe, D.G. and Luebke, R.M., and Riddle, M.M. (1987) The effects of nickel on immune function in the rat. *Toxicol.* 44:271-281.
62. Katz, M. (1978) Nutrition and the immune response. In, Proceedings of the fourth annual FDA science symposium, inadvertent modification of the immune response, Effects of Foods, Drugs and environmental contaminants. HHS publication no. (FDA) 80-1074. (I. M. Asher ed.) Health and Human Services [HHS]. Washington, D.C. pp.35-36.
63. Pierpaoli, W. and Sorkin, E. (1969) Effect of growth hormone and anti-growth hormone serum on the lymphatic tissue and the immune response. *Antibiot.Chemother.* 15:122-134.
64. Berczi, I. (1986) Immunoregulation by pituitary hormones. In, Pituitary Function and Immunity. (I. Berczi ed.) CRC Press. Boca Raton, FL. pp. 227-240.
65. Claman, H.N. (1975) How corticosteroids work. *J.Allergy Clin.Immunol.* 55:145-151.
66. Weimer, H.E. and Roberts, D.M. (1972) Sex differences in the immunologic and phlogistic responses to *Salmonella typhosa* H-antigen in the albino rat. *Comp.Biochem.Physiol.* 41B(4):713-722.
67. Vos, J.G., Grouwer, G.M.J., van Leeuwen, F.X.R. and Wagenaar, S. (1983) Toxicity of hexachlorobenzene in the rat following combined pre- and postnatal exposure: Comparisons of effects on the immune system, liver, and lung. In, Immunotoxicology. (R. Hubbard and D.V. Parke eds.) Academic Press. London, U.K. pp. 219-234.
68. House, R.V., Lauer, L.D., Murray, M.J., Ward, E.C. and Dean, J.H. (1985) Immunological studies in B6C3F1 mice following exposure to ethylene glycol monomethyl ether and its principal metabolite methoxyacetic acid. *ToxicolAppl.Pharmacol.* 77(2):358-362.
69. Hinton, D.M. (1993) Unpublished observations.

IX References Continued

Chapter V D: Immunotoxicity Studies Continued

70. Weening, J.J., Fleuren, G.J. and Hoedemacker, Ph.J. (1978) Demonstration of antinuclear antibodies in mercuric chloride-induced glomerulopathy in the rat. *Lab. Invest.* 39:405-411.
71. Michaelson, J.H., McCoy, J.P.Jr. and Bigazzi, P.E. (1981) Antibodies to glomerular basement membrane: Detection by an enzyme-linked immunosorbent assay in the sera of Brown-Norway rats. *Kidney Int.* 20:285-288.
72. Silverman, D.A. and Rose, N.R. (1975) Spontaneous and methylcholanthrene enhanced thyroiditis in BUF rats. I. The incidence and severity of the disease, and the genetics of susceptibility. *J. Immunol.* 114:145-147.
73. Thiem, P.A., Halper, L.K. and Bloom, J.C. (1988) Techniques for assessing canine mononuclear phagocyte function as part of an immunotoxicologic evaluation. *Int. J. Immunopharmacol.* 10(6):765-771.
74. Tryphonas, H., Hayward, S., O'Grady, L., Loo, J.C., Arnold, D.L., Bryce, F. and Zawadzka, Z.Z. (1989) Immunotoxicity studies of PCB (Aroclor 1254) in the adult rhesus (*Macaca mulatta*) monkey--Preliminary report. *Int. J. Immunopharmacol.* 11(2):199-206.
75. Sigel, M.M. (1984) Overview of selected immunological components of the immune systems in food animals with consideration of immune modulation. In, Chemical Regulation of Immunity in Veterinary Medicine. (M. Kende, J. Gainer and M. Chirigos eds.) Alan R. Liss Press. New York, N.Y. pp. 3-26.
76. Glocklin, V.C. (1987) Current FDA Perspectives on Animal Selection and Extrapolation. In, Human risk assessment - The role of animal selection and extrapolation. (M.V. Roloiff, A.G.E. Wilson, W.E. Ribelin, W.P. Ridley and F.A. Ruecker eds.) Taylor S. Francis Press. London, U.K. and Philadelphia, PA. pp.15-22.
77. Hinton, D.M. and Kahn, M.A. (1987) Basal immune responses in miniswine: A nonrodent species for immunotoxicity assessment. *Fed. Am. Soc. Exp. Biol.* 46(4):1129.
78. Hinton, D.M. and Kahn, M.A. (1989) Evaluation of the miniature swine for use in immunotoxicity assessment: I. Quantitation of humoral responses to sheep red blood cells and effects of cyclophosphamide. *J. Am. Coll. Toxicol.* 8(5):1039-1052.
79. Hinton, D.M., Kahn, M.A. and Roberts, D.W. (1993) *Manuscript in Preparation*.

IX References Continued

Chapter VI A: Guidelines for the Clinical Evaluation of Foods and Food Additives

1. Anonymous (1985) New Drug and Antibiotic Regulations. [Adequate and Well-Controlled Studies, Section 314.126, p. 7486, p.7506] *Fed.Reg.* 50:(February 22) 7452-7519.
2. Anonymous (1981) Protection of human subjects *Fed.Reg.* 46:(January 27) 8942-8979.
3. Health, Education and Welfare (HEW) (1977) In, General considerations for the clinical evaluation of drugs in infants and children. HEW, PHS, and FDA Report. U.S. Printing Office.
4. American Academy of Pediatrics Committee on Nutrition [AAPCN] (1988) In, Clinical Testing of Infant Formulas with Respect to Nutritional Suitability for Term Infants. Prepared under FDA contract 223-86-2117. pp.1-36.
5. American Academy of Pediatrics Committee on Nutrition [AAPCN] (1987) Guidelines for the clinical evaluation of new products used in the dietary management of infants, children and pregnant women with metabolic disorders. A report by the Task Force on Clinical Evaluation of Products for Metabolic Disorders. pp.1-53.

Chapter VI B: Epidemiology Studies

1. Last, J.M. eds. (1983) In, A Dictionary of Epidemiology. (J.H. Abramson, S. Greenland, M.C. Thuriaux, associate eds.; J. Amsel, assistant ed.) Oxford University Press, New York, N.Y.
2. Anonymous. (1981) Reviews and Commentary. Guidelines for documentation of epidemiologic studies. A report by the Epidemiology Work Group of the Interagency Regulatory Liaison Group. *Am.J.Epidemiol.* 114(5):609-613.
3. Tollefson, L. (1988) Monitoring adverse reactions to food additives in the U.S. Food and Drug Administration. *Reg.Toxicol.Pharmacol.* 8(4):438-446.

Chapter VII C: Safety of Foods and Food Additives Developed by Biotechnology

1. Kessler, D. A., Taylor, M. R., Maryanski, J. H., Flamm, E. L., and Kahl, L.S. (1992) The Safety of foods developed by biotechnology. *Science* 256:1747-1749 (&1832).
2. Anonymous (1992) Statement of Policy: Foods derived from new plant varieties. *Fed. Reg.* 57:22984.

IX References Continued

Chapter VII F: Advances in the Development of Alternatives to Whole Animal Testing

1. Russell, W.M.S. and Burch, R.L., eds. (1959) In, The principles of humane experimental technique. Methuen Press. London, U.K.
2. Schaeffer, W.I. (1990) Terminology associated with cell, tissue, and organ culture, molecular biology and molecular genetics. *In Vitro Cell Dev.Biol.* 26:97-101.
3. Goldberg, A.M. and Frazier, J.M. (1989) Alternatives to animals in toxicity testing. *Sc.Am.* 261:24-30.
4. McKeehan, W.L., Barnes, D., Reid, L., Stanbridge, E., Murakami, H. and Sato, G. (1990) Frontiers in mammalian cell culture. *In Vitro Cell Dev.Biol.* 26:9-23.
5. Frazier, J.M. and Bradlaw, J.A., eds. (1989) In, Technical problems associated with in vitro toxicity testing systems. John's Hopkins University Press. Baltimore, MD. *Tech. Report No.1*.
6. Balls, M., Blaauboer, B., Brusick, D., Frazier, J., Lamb, D., Pemberton, M., Reinhardt, C., Roberfroid, M., Rosenkranz, H., Schmid, B., Spielmann, H., Stammati, A. and Walum, E. (1990) Report and recommendations for the CAAT/ERGATT Workshop on the validation of toxicity test procedures. *ATLA* 18:313-337.
7. Frazier, J.M. (1990) Scientific criteria for validation of *in vitro* toxicity test. Organization for Economic Co-Operation and Development. *Environment Monograph No.36*.
8. Atterwill, C.K. and Steele, C.E., eds. (1987) In, In Vitro methods in toxicology. Cambridge University Press. New York, N.Y.
9. Balls, M., Riddell, R.J., and Worden, A.N. eds. (1983) In, Animals and alternatives in toxicity testing. Academic Press. New York, N.Y.
10. Berk, J. and Sherrod, C., eds. (1977) In, In Vitro testing 1975-1976. Franklin Institute Press. Philadelphia, PA.
11. Bradlaw, J.A. (1986) Evaluation of drug and chemical toxicity with cell culture systems. *Fund.Appl.Toxicol.* 6:598-606.
12. Goldberg, A.M., ed. (1983) In, Product safety evaluation: Alternative methods in toxicology. Vol 1. Liebert Press. New York, N.Y.
13. Goldberg, A.M., ed. (1984) In, Acute toxicity testing: Alternative methods in toxicology. Vol. 2. Liebert Press. New York, N.Y.
14. Goldberg, A.M., ed. (1985) In, In Vitro toxicology: Alternative methods in toxicology. A progress report from the Johns Hopkins Center for Alternatives to Animal Testing. Vol. 3. Liebert Press. New York, N.Y.
15. Goldberg, A.M., ed. (1987) In, In Vitro toxicology: Approaches to validation, Alternative methods in toxicology. Vol. 5. Liebert Press. New York, N.Y.

IX References Continued

Chapter VII F: Advances in the Development of Alternatives to Whole Animal Testing Continued

16. Grisham, J.W. and Smith, G.J. (1984) Predictive and mechanistic evaluation of toxic responses in mammalian cell culture systems. *Pharmacol.Rev.* 36:151s-171s.
17. Rofo, P.C. (1971) Tissue culture and toxicology. *Fd.Cosmet.Toxicol.* 9:683-696.
18. Rowan, A.N. and Stratmann, C.J., eds. (1980) In, The use of alternatives in drug research. MacMillan & Co. Press. London, U.K.
19. Rowan, A.N. and Goldberg, A.M. (1985) Perspectives on alternatives to current animal testing techniques in preclinical toxicology. *Ann.Rev.Pharmacol.Toxicol.* 25:225-247.
20. Stammati, A.P., Silano, V., and Zucco, F. (1981) Toxicology investigations with cell culture systems. *Toxicol.* 20:91-153.
21. Tardiff, R.G. (1978) *In Vitro* methods of toxicity evaluation. *Ann.Rev.Pharmacol.Toxicol.* 18:357-369.
22. Zucco, F. and Hooisman, J., eds. (1980) In, Proceedings of the international workshop on the application of tissue culture in toxicology. *Toxicol.* 17:101-272.
23. Zucco, F. and Hooisman, J., eds. (1982) In, Proceedings of the second international workshop on the application of tissue culture in toxicology. *Toxicol.* 25:1-78.
24. Frazier, J.M., Gad, S.C., Goldberg, A.M. and McCulley, J.P., eds. (1987) In, A critical evaluation of alternatives to acute ocular irritation testing: Alternative methods in toxicology. Vol. 4. Liebert Press. New York, N.Y.
25. Nardone, R.M. and Bradlaw, J.A. (1983) Toxicity testing with *in vitro* systems: I. Ocular tissue culture. *J. Toxicol.-Cutaneous & Ocular Toxicol.* 2:81-98.
26. Frazier, J.M. (1988) Update: A critical evaluation of alternatives to acute ocular irritancy testing. In, Progress in *In Vitro* toxicology. (A. M. Goldberg, ed.) Liebert Press. New York, N.Y. pp.67-75.
27. Wilcox, D.K. and Bruner, I.H. (1990) *In vitro* alternatives for ocular safety testing: an outline of assays and possible future developments. *Altern. to Lab. Animals (ATLA)* 18:117-128.

IX References Continued

Chapter VII G: Heritable and Somatic Genetic Toxicity

1. Tennant, R.W., Margolin, B.H., Shelby, M.D., Zeiger, E., Haseman, J.K., Spalding, J., Caspary, W., Resnick, M., Stasiewicz, S., Anderson, B. and Minor, R. (1987) Prediction of chemical carcinogenicity in rodents from *in vitro* genetic toxicity assays. *Science* 236:933-941.
2. Ashby, J. and Tennant, R.W. (1988) Chemical structure, *Salmonella* mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. *Mutation Research* 204: 17-115.
3. Ashby, J. and Tennant, R.W. (1991) Definitive relationships among chemical structure, carcinogenicity and mutagenicity for 301 chemicals tested by the U.S. NTP. *Mutation Research* 257: 229-306.
4. Abrahamson, S., Wurgler, F.E., DeJongh, C. and Meyer, H.U. (1980) How many loci on the x-chromosome of *Drosophila melanogaster* can mutate to recessive lethals? *Environ.Mutagen.* 2:447-453
5. Lee, W.R., Abrahamson, S., Valencia, R., von Halle, E.S., Wurgler, F.E. and Zimmering, S. (1983) The sex-linked recessive lethal test for mutagenesis in *Drosophila melanogaster*. A report of the U.S. Environmental Protection Agency Gene-Tox Program. *Mutation Res.* 123:183-279.
6. Mason, J.M., Aaron, C.S., Lee, W.R., Smith, P.D., Thakar, A., Valencia, R., Woodruff, R.C., Wurgler, F.E. and Zimmering, S. (1987) A guide for performing germ cell mutagenesis assays using *Drosophila melanogaster*. *Mutation Res.* 189:93-102.