Reviewer Guidance

Conducting a Clinical Safety Review of a New Product Application and Preparing a Report on the Review

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

February 2005
Good Review Practices

2005D-0057

GDL 1

Reviewer Guidance

Conducting a Clinical Safety Review of a New Product Application and Preparing a Report on the Review

Additional copies are available from:

Office of Training and Communication
Division of Drug Information, HFD-240
Center for Drug Evaluation and Research
Food and Drug Administration
5600 Fishers Lane
Rockville, MD 20857
(Tel) 301-827-4573
http://www.fda.gov/cder/guidance/index.htm

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

February 2005
Good Review Practices

TABLE OF CONTENTS

I. INTRODUCTION 1

II. GENERAL GUIDANCE ON THE CLINICAL SAFETY REVIEW 2

A. Introduction 2

B. Explanation of Terms 3

C. Overview of the Safety Review 4

D. Differences in Approach to Safety and Effectiveness Data 4

E. Identifying and Assembling Source Materials for the Safety Review 6

F. Identifying Major Concerns at the Outset 7

G. Auditing Source Materials 8

H. The Purpose of Individual Case Review/"Drug-relatedness" 8

III. SPECIFIC GUIDANCE ON THE CONTENT OF THE SAFETY REVIEW 9

7.1 Methods and Findings 11

7.1.1 Deaths 11

7.1.2 Other Serious Adverse Events 15

7.1.3 Dropouts and Other Significant Adverse Events 16

7.1.3.1 Overall Profile of Dropouts 16

7.1.3.2 Adverse Events Associated with Dropouts 18

7.1.3.3 Other Significant Adverse Events 19

7.1.4 Other Search Strategies 19

7.1.5 Common Adverse Events 19

7.1.5.1 Applicant's Approach to Eliciting Adverse Events in the Development Program 20

7.1.5.2 Establishing Appropriate Adverse Event Categories and Preferred Terms 20

7.1.5.3 Incidence of Common Adverse Events — Assessment of Various Databases 22

7.1.5.4 Common Adverse Event Tables 23

7.1.5.5 Identifying Common and Drug-Related Adverse Events 23

7.1.5.6 Additional Analyses and Explorations 24

7.1.6 Less Common Adverse Events 25

7.1.7 Laboratory Findings 25

7.1.7.1 Overview of Laboratory Testing in the Development Program 26

7.1.7.2 Selection of Studies/Analyses for Drug-Control Comparisons of Laboratory Values 26

7.1.7.3 Standard Analyses and Explorations of Laboratory Data 27
7.1.7.3.1 Analyses Focused on Measures of Central Tendency 27
7.1.7.3.2 Analyses Focused on Outliers or Shifts from Normal to Abnormal 27
7.1.7.3.3 Marked Outliers and Dropouts for Laboratory Abnormalities 28
7.1.7.4 Additional Analyses and Explorations 29
7.1.7.5 Special Assessments: Hepatotoxicity, QTc, Others 29
7.1.8 Vital Signs 30
7.1.8.1 Extent of Vital Signs Testing in the Development Program 30
7.1.8.2 Selection of Studies and Analyses for Overall Drug-Control Comparisons 30
7.1.8.3 Standard Analyses and Explorations of Vital Signs Data 30
7.1.8.3.1 Analyses Focused on Measures of Central Tendency 30
7.1.8.3.2 Analyses Focused on Outliers or Shifts from Normal to Abnormal 30
7.1.8.3.3 Marked Outliers and Dropouts for Vital Signs Abnormalities 30
7.1.8.4 Additional Analyses and Explorations 30
7.1.9 Electrocardiograms (ECGs) 30
7.1.9.1 Extent of ECG Testing in the Development Program, Including Brief Review of Preclinical Results 30
7.1.9.2 Selection of Studies and Analyses for Overall Drug-Control Comparisons 31
7.1.9.3 Standard Analyses and Explorations of ECG Data 31
7.1.9.3.1 Analyses Focused on Measures of Central Tendency 31
7.1.9.3.2 Analyses Focused on Outliers or Shifts from Normal to Abnormal 31
7.1.9.3.3 Marked Outliers and Dropouts for ECG Abnormalities 31
7.1.9.4 Additional Analyses and Explorations 31
7.1.10 Immunogenicity 31
7.1.11 Human Carcinogenicity 31
7.1.12 Special Safety Studies 32
7.1.13 Withdrawal Phenomena/Abuse Potential 32
7.1.14 Human Reproduction and Pregnancy Data 32
7.1.15 Assessment of Effect on Growth 32
7.1.16 Overdose Experience 33
7.1.17 Postmarketing Experience 34
7.2 Adequacy of Patient Exposure and Safety Assessments 34

7.2.1 Description of Primary Clinical Data Sources (Populations Exposed and Extent of Exposure) Used to Evaluate Safety 34
7.2.1.1 Study Type and Design/Patient Enumeration 35
7.2.1.2 Demographics 36
7.2.1.3 Extent of Exposure (Dose/Duration) 36
7.2.2 Description of Secondary Clinical Data Sources Used to Evaluate Safety 36
7.2.2.1 Other Studies 37
7.2.2.2 Postmarketing Experience 37
7.2.2.3 Literature 37
7.2.3 Adequacy of Overall Clinical Experience 37
7.2.4 Adequacy of Special Animal and/or In Vitro Testing 38
7.2.5 Adequacy of Routine Clinical Testing 38
7.2.6 Adequacy of Metabolic, Clearance, and Interaction Workup 39
7.2.7 Adequacy of Evaluation for Potential Adverse Reactions for Any New Drug and Particularly for Drugs in the Class Represented by New Drug; Recommendations for Further Study 39
7.2.8 Assessment of Quality and Completeness of Data 41
7.2.9 Additional Submissions, Including Safety Update 42
7.3 Summary of Selected Adverse Reactions, Important Limitations of Data, and Conclusions 42
7.4 General Methodology 44
7.4.1 Pooling Data Across Studies to Estimate and Compare Incidence 44
7.4.1.1 Pooled Data vs. Individual Study Data 44
7.4.1.2 Combining Data 45
7.4.2 Explorations for Predictive Factors 45
7.4.2.1 Explorations for Dose Dependency for Adverse Findings 45
7.4.2.2 Explorations of Time-Dependency for Adverse Findings 46
7.4.2.3 Explorations for Drug-Demographic Interactions 47
7.4.2.4 Explorations for Drug-Disease Interactions 48
7.4.2.5 Explorations for Drug-Drug Interactions 48
7.4.3 Causality Determination 49
List of Tables 51

Contains Nonbinding Recommendations

Reviewer Guidance[1]
Conducting a Clinical Safety Review of a New Product Application and Preparing a Report on the Review

This guidance represents the Food and Drug Administration's (FDA's) current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. You can use an alternative approach if the approach satisfies the requirements of the applicable statutes and regulations. If you want to discuss an alternative approach, contact the FDA staff responsible for implementing this guidance. If you cannot identify the appropriate FDA staff, call the appropriate number listed on the title page of this guidance.

I. INTRODUCTION

This good review practice (GRP) guidance is intended to assist reviewers conducting the clinical safety reviews as part of the NDA and BLA review process, provide standardization and consistency in the format and content of safety reviews, and ensure that critical presentations and analyses will not be inadvertently omitted. The standardized structure also enables subsequent reviewers and other readers to readily locate specific safety information.

This guidance is an expansion of section 7 of the clinical review template and is entirely compatible with that template. The structure of this guidance, as an annotated outline, is meant to correlate exactly with the section headings of the review template, providing the pertinent guidance under each heading. The guidance also provides, as attachments, illustrations of displays and graphs that have been used successfully in the past. These are not requirements, but examples, and reviewers can substitute for, or modify, them, or simply find them unnecessary in particular cases. It is expected that new attachments will be added as examples become available.

The commentary and suggestions under each section of the guidance, together with appended examples, provide suggested analyses, methods of presentations, and discussion of special cases and potential difficulties. Some flexibility in implementing the guidance will be needed, as different types of applications and datasets may require modifications to the structure outlined in

[1] This guidance has been prepared by the Integrated Summary of Safety group, a subcommittee of Good Review Practices Track 8. The Track 8 Committee has been charged with developing a guidance for the clinical review of a marketing application under the Good Review Practices (GRP) initiative.

this guidance. If sections are omitted, the review should briefly explain the reason for the omission.

FDA's guidance documents, including this guidance, do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

II. GENERAL GUIDANCE ON THE CLINICAL SAFETY REVIEW

A. Introduction

This Good Review Practice (GRP) Guidance provides an annotated outline of the safety component of a clinical review of an application (NDA, BLA) and guidance on how to conduct and organize the safety review. [2] It is usually most efficient and informative to include all the safety findings, whatever the source, in the safety section of the clinical review (i.e., apart from the description of individual studies in the efficacy review). In some cases, however, it may be more appropriate to discuss some or all aspects of safety as part of the discussion of individual efficacy studies and reference them in this section (e.g., studies with mortality outcomes, development programs in which most of the safety data come from one or two large multi-center studies, and when evaluation and review of safety data may be more convenient or informative study by study).

The safety review has two distinct components: (1) identification and assessment of the significance of the adverse events reported in clinical trials (controlled or uncontrolled) and (2) evaluation of the adequacy of the applicant's safety evaluation. This guidance describes an approach that integrates safety findings across all studies and other clinical experience. Consideration of the safety findings in individual studies, without a thoughtful integration of the overall safety experience, is not adequate for a safety review.[3]

Although much of the guidance in this document is directed primarily toward the clinical reviewer and toward the analysis of particular events, the evaluation of safety data also involves analyses of event rates, estimation of risk over time, exploration of possible subgroup differences, and identification of risk factors associated with serious events, all analyses that

---

[2] It is recognized that no drug is safe in the sense of being entirely free of adverse effects. Reference in the Food, Drug and Cosmetic Act to the "safety" of a drug for the uses recommended in labeling has been interpreted as meaning that the benefits of a drug outweigh its risks for those uses. The safety review, however, is not a risk benefit analysis, but rather, is the part of the NDA review that assesses and describes the risks of the drug.

[3] It is important to distinguish between the concept of performing an integrated safety review and the separate question of whether or not to pool data across studies in the conduct of that review. For the purpose of this document, an integrated safety review refers to the principle of bringing together in one place in the review all data and analyses pertinent to a particular safety issue (e.g., liver toxicity). Whether one looks primarily at data from individual studies or at datasets resulting from pooling of certain studies to address a particular safety concern is not critical to the concept of an integrated review. Either approach, or both approaches, will usually be used by a reviewer in carrying out an integrated review.

involve substantial knowledge of methods of validly quantifying risk and providing measures of uncertainty. Clinical reviewers should therefore collaborate with their biostatistical colleagues when necessary in the preparation of reviews and consider when it may be appropriate to conduct a joint statistical and clinical review for particularly important safety issues.

The conceptual framework of this guidance is similar to the framework used for advising manufacturers on submitting safety data in FDA's *Guideline for the Format and Content of the Clinical and Statistical Section for New Drug Applications* (Clinical/Statistical guidance)[4] as well as in the guidance *M4: The Common Technical Document for the Conduct of Human Clinical Trials for Pharmaceuticals — Efficacy.*

B. Explanation of Terms

Because several related terms are used in this guidance that could cause some confusion, the following explanations are intended as clarification.

For purposes of this guidance, the term ***adverse reaction*** is used to refer to an undesirable effect, reasonably associated with the use of a drug, that may occur as part of the pharmacological action of the drug or may be unpredictable in its occurrence. This term does not include all adverse events observed during use of a drug, only those for which there is some basis to believe there is a causal relationship between the drug and the occurrence of the adverse event. The term ***adverse event*** is used here to refer to any untoward medical event associated with the use of a drug in humans, whether or not it is considered drug-related. The phrases ***serious adverse drug experience*** and ***serious adverse event*** are used in this guidance to refer to any event occurring at any dose, whether or not considered drug-related, that results in any of the following outcomes:

- Death
- A life-threatening adverse experience
- Inpatient hospitalization or prolongation of existing hospitalization
- A persistent or significant disability/incapacity
- A congenital anomaly or birth defect.

Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered serious adverse drug events when, based upon appropriate medical judgment, they may jeopardize the patient or subject and may require medical or surgical intervention to prevent one of the outcomes listed in this definition. Examples of such medical events include allergic bronchospasm requiring intensive treatment in an emergency room or at home, blood dyscrasias or convulsions that do not result in inpatient hospitalization, or the development of drug dependency or drug abuse. Documents developed by the International Conference on Harmonisation (e.g., E2) add to serious events those that prolong hospitalization but do not include cancer and overdose.

[4] See http://www.fda.gov/cder/guidance/statnda.pdf.

Finally, the term ***adverse dropout*** is used in this guidance to refer to subjects who did not complete the study because of an adverse event, whether or not considered drug-related; adverse dropouts include subjects who received the test drug, reference drugs, or placebo.

C. Overview of the Safety Review

The safety review has four principal tasks:

1) To identify and closely examine serious adverse events that suggest, or could suggest, important problems with a drug-- specifically, adverse reactions severe enough to prevent its use altogether, to limit its use, or require special risk management efforts

2) To identify and estimate the frequency of the common (usually nonserious) adverse events that are, or may be, causally related to the use of the drug;

3) To evaluate the adequacy of the data available to support the safety analysis and to identify the limitations of those data. At a minimum, this includes assessments of whether the extent of exposure at relevant doses is adequate

4) To identify unresolved safety concerns that will need attention prior to approval or that should be assessed in the postmarketing period, including such concerns as the absence of data from high-risk populations or potential interactions

In addition, the safety review should:

- Identify factors that predict the occurrence of adverse reactions, including patient-related factors (e.g., age, gender, ethnicity, race, target illness, abnormalities of renal or hepatic function, co-morbid illnesses, genetic characteristics, such as metabolic status, environment) and drug-related factors (e.g., dose, plasma level, duration of exposure, concomitant medication)

- Identify, where possible, ways to avoid adverse reactions (dosing, monitoring) and ways to manage them when they occur

- For a drug that is to be approved, provide a comprehensive evaluation of risk information adequate to support a factual and sufficient summary of the risk information in labeling.

D. Differences in Approach to Safety and Effectiveness Data

Approaches to evaluation of the safety of a drug generally differ substantially from methods used to evaluate effectiveness. Most of the studies in phases 2-3 of a drug development program are directed toward establishing effectiveness. In designing these trials, critical efficacy endpoints are identified in advance, sample sizes are estimated to permit an adequate assessment of effectiveness, and serious efforts are made, in planning interim looks at data or in controlling multiplicity, to preserve the type 1 error (alpha error) for the main end point. It is also common to devote particular attention to examining critical endpoints by defining them with great care

and, in many cases, by using blinded committees to adjudicate them. In contrast, with few exceptions, phase 2-3 trials are not designed to test specified hypotheses about safety nor to measure or identify adverse reactions with any pre-specified level of sensitivity. The exceptions occur when a particular concern related to the drug or drug class has arisen and when there is a specific safety advantage being studied. In these cases, there will often be safety studies with primary safety endpoints that have all the features of hypothesis testing, including blinding, control groups, and pre-specified statistical plans.

In the usual case, however, any apparent finding emerges from an assessment of dozens of potential endpoints (adverse events) of interest, making description of the statistical uncertainty of the finding using conventional significance levels very difficult. The approach taken is therefore best described as one of exploration and estimation of event rates, with particular attention to comparing results of individual studies and pooled data. It should be appreciated that *exploratory analyses* (e.g., subset analyses, to which a great caution is applied in a hypothesis testing setting) are a critical and essential part of a safety evaluation. These analyses can, of course, lead to false conclusions, but need to be carried out nonetheless, with attention to consistency across studies and prior knowledge. The approach typically followed is to screen broadly for adverse events and to expect that this will reveal the common adverse reaction profile of a new drug and will detect some of the less common and more serious adverse reactions associated with drug use.

With respect to assessment of serious events, there are two distinct situations. First, there are the events readily recognized as consequences, or at least potential consequences, of the treatment (i.e., adverse reactions) because they would be unusual in the population under study. Second, and particularly critical, are serious events that are not so readily attributed to the drug because they can occur even without the drug, for example, because they are known to result from the underlying disease or are relatively common in the population being studied (e.g., heart attacks, strokes in an elderly population) and could therefore represent intercurrent illness. Adverse events that do not seem typical of what drugs do (that is, that are not hematologic, hepatic, renal, dermatologic or pro-arrhythmic) can be especially difficult to attribute to a drug. The history of the relatively late recognition of the practolol syndrome (sclerosing peritonitis, oculomucocutaneous syndrome), retroperitoneal fibrosis with methylsergide (Sansert), pulmonary hypertension with aminorex and other appetite suppressants, thromboembolic disease with oral contraceptives, endometrial cancer with post-menopausal estrogens, suicidal ideation with interferons, and more recently, cardiac valvular disorders with fenfluramine, illustrates this problem. Perhaps most difficult of all is the situation where the adverse event is, or could be, a consequence of the disease being treated. Thus, it was extremely difficult to discover that many drugs for heart failure (beta agonists, phosphodiesterase inhibitor inotropes, and a vasodilator, flosequinan) caused increased rates of the same kinds of death seen with the underlying disease (i.e., due to progressive heart failure or arrhythmias), that anti-arrhythmics could provoke new arrhythmias, and that interferon could cause depression in patients with cancer or multiple sclerosis, conditions that are themselves associated with mood alteration. Distinguishing the effects of a drug on the immune or other impaired systems in patients with cancer or HIV infection can also be difficult. Many years ago, a last resort drug for rheumatoid arthritis, azaribine (Triazure), was approved despite a number of arterial thrombi seen during development because those were thought to be more common in patients with RA (the drug was

removed from the market shortly after approval, however, when unusual thrombotic events (e.g., thrombosis of a digital artery) became apparent). Drugs for seizure disorders and schizophrenia can be difficult to assess with respect to causing sudden death because patients with the disorders they treat have a relatively high rate of this event. Usually, the only way to establish that these are adverse reactions is through controlled trials of significant size. Sometimes, the controlled trials to evaluate effectiveness will be large enough to address these issues, but sometimes, where there is a significant concern, special, large safety studies may be needed.

There is no simple answer to these difficult assessments, but this guidance, similar to section H of the Clinical/Statistical guideline *(Guideline for the Format and Content of the Clinical and Statistical Sections of New Drug Applications)*[5] and the guidance *M4 The CTD – Efficacy,* suggest an approach, namely, close examination of all patients who die or who leave a study prematurely because of any adverse event (whether or not thought drug-related),[6] with explicit consideration of the possibility that the event was drug-related (the "prepared mind" approach). With respect to discovering that a drug causes a modestly increased rate of serious events that are relatively common in the population, only large controlled trials can provide a satisfactory answer and the reviewer needs to consider whether such trials are needed. In some cases, there are reasonably well-established surrogate markers that can predict severe injury. For example, an increased rate of transaminase elevations accompanied by a small number of cases in which bilirubin elevation accompanies the transaminase elevation can predict the occurrence of more severe liver injuries in some patients, and visual field defects may portend irreversible peripheral vision loss. Similarly, substantial QT interval prolongation on the electrocardiogram predicts the occurrence of Torsade de Pointes-type ventricular tachycardia.

E. Identifying and Assembling Source Materials for the Safety Review

Before beginning the safety review, the reviewer should identify and assemble (or locate electronically) all available materials for the review. These materials include:

- The applicant's Integrated Summary/Analysis of Safety (ISS)
- Adverse event tables in the NDA/BLA submission[7]
- Case report forms (CRFs) for patients who experienced serious adverse events or who dropped out of a study because of an adverse event. The reviewer should request these CRFs

[5] See http://www.fda.gov/cder/guidance/statnda.pdf.

[6] 21 CFR 314.50(f)(2) requires that CRFs be submitted for each patient who died during a clinical study or who did not complete the study because of an adverse event, whether believed to be drug related or not. It should be clear from the application that the ISS and other safety reports include all adverse events that were seen during development, not just those judged by investigators or the applicant to have been potentially drug-related. This is also a useful point to make at a pre-NDA/BLA meeting.

[7] If the reviewer determines that adverse event tables provided by the applicant are accurate and fairly represent the data they purport to display, the tables may be included in the safety review as appendices. If applicant-generated tables are used, the review should identify the applicant as the source.

if the applicant does not include them in the submission (although they are required under 21 CFR 314.50). If the number of cases is very large (e.g., for dropouts) and many of the events are similar, it may be reasonable to request only a sample of CRFs.[8] Note that, in some cases, dropouts attributed to other reasons will upon review be associated with an adverse event.

- Individual patient adverse reaction data listings, laboratory listings, and baseline listings, usually accessible electronically[8]

- The applicant's narrative summaries of deaths, serious adverse events, and other events that resulted in dropouts

- If available, displays of individual patient safety data over time for patients who experienced serious adverse events

- The safety sections of the sponsor's proposed labeling

- Common Technical Document (CTD) safety-related sections (module 2, Sections 2.5.5, 2.7.4), which give an overview of the applicant's approach to the safety evaluation and a detailed summary of the safety data

- Any other safety-related documents, such as discussions of related drugs, descriptions of use of adverse drug reactions (ADR) coding dictionaries to combine data across studies, specific studies of safety hypotheses

F. Identifying Major Concerns at the Outset

Although the review will assess the data submitted, it may be useful to identify at the outset particular concerns that will be explored because they are suggested by the pharmacology of the drug or by safety concerns with pharmacologically related drugs. Thus, the clearance pathway of a drug will suggest certain potential drug-drug interactions or certain effects of decreased renal or hepatic function. Similarly, the pharmacologic class, and prior experience, could lead to focus on particular laboratory or clinical abnormalities (e.g., muscle or liver abnormalities with HMGCoA reductase inhibitors, QT prolongation with fluoroquinolone anti-infectives, gastrointestinal, renal, and cardiovascular effects of nonsteroidol anti-inflammatory drugs, liver abnormalities with endothelin receptor antagonists, cognitive impairment with sedating drugs, sexual dysfunction with selective serotonin reuptake inhibitors). These concerns are considered further in section 7.2.5.

---

[8]The reviewer should be able to easily access individual patient information. The reviewer may want to clarify formatting and accessibility concerns at the pre-NDA meeting. For hardcopy submissions, an index that directs the reviewer to the exact location (volume and page number) of the CRF, the narrative summary, and the individual patient safety data display is essential (for sample index see Table 8.0.1). For electronic submissions, the PDF files should have sufficiently detailed *bookmarks* to offer easy navigation by the reviewer. For example, narratives should be bookmarked by patient ID number, not just by study treatment or treatment assignment.

G. Auditing Source Materials

Although there are no established standards for auditing safety data in a submission, the review should describe efforts to assess consistency of the data provided (e.g., comparing information included in case report forms, case report tabulations, and narrative summaries for individual patients). For important adverse events, for example, it is generally important to consider not only the applicant's narrative description, but the associated CRF or hospital records and submitted laboratory, radiology, or pathology results.

H. The Purpose of Individual Case Review/"Drug-relatedness"

An important part of the safety review is reviewing individual cases of death, serious adverse events, adverse events leading to discontinuation (adverse drop-outs), and discontinued patients who are lost to follow-up. One reason to review the details of individual cases is to determine whether the event was coded to the correct preferred term. The assessment of causality for specific adverse events in NDAs/BLAs is heavily dependent on comparisons of event rates between treatment groups, and the numerator of these rate calculations includes events coded to the same preferred term. A case might be incorrectly included in the numerator of a rate calculation if the event is incorrectly coded to a specific preferred term. Events may be incorrectly coded to preferred term by the applicant when they summarize the data or because an investigator used a verbatim term incorrectly when recording the event in the case report form. An example of incorrect coding would be if an investigator used the verbatim term *acute liver failure* for a case of increased ALT and the applicant coded the event to acute liver failure. One would not want to include such a case in the numerator of a risk calculation for acute liver failure. Similarly, a case could be incorrectly excluded from a numerator. Inconsistent coding (e.g., peripheral edema coded as "heart failure" for one patient, but "metabolic abnormality" for another) could result in an inappropriately low numerator.

A second reason to conduct individual case review of deaths, serious adverse events, and adverse events leading to discontinuation is to determine whether there is a likely explanation for the event other than the drug that is the subject of the application, such as another drug or concomitant illness (e.g., documented acetaminophen overdose in a case of acute liver failure would argue against attribution to the test drug; documented cholecystitis would argue against attribution of cholestasis to the test drug). If there is no likely alternative explanation for the event, the event must be considered at least possibly drug-related, and should be included in a rate calculation.

A third reason for individual case review of deaths, serious adverse events, and adverse events leading to discontinuation is to look for other reasons that might exclude the drug as a cause of the event. One example would be when an adverse event occurred during a placebo washout period before exposure to study drug occurred. Events that occur prior to exposure would not be included in the numerator of risk calculations. Events that begin long after discontinuation of the drug might also be considered unlikely to be drug-related, but care must be taken in excluding them, as there are examples of such late drug-caused reactions (e.g., FIAU (fialuridine, a nucleoside analog) where liver failure was seen well after the drug was stopped, probably

because it induced mitochondrial DNA damage that became a problem only when mitochondria tried to replicate) and because some chronic reactions might not be detected immediately.

A fourth reason for individual case review of deaths, serious adverse events, and adverse events leading to discontinuation is to look for results of rechallenge. A potentially important source of information about causality is when an individual is rechallenged with drug, accidentally or deliberately. Recurrence with rechallenge is a potentially strong indicator of causality, but interpretation of the results of rechallenge is highly dependent on the natural course of the event being considered. For noncyclical events that are exceedingly rare in the background (e.g., acute liver failure, aplastic anemia) recurrence of the event upon rechallenge (i.e., positive rechallenge) provides strong evidence of causality. Positive rechallenges are less definitive for diagnoses/events that can occur in cyclical or recurrent fashion (e.g., worsening glucose control in a subject with diabetes mellitus), but close observation of the patient's whole course (i.e., both challenge periods and dechallenge periods) may be helpful. Rechallenges that do not result in recurrence of the event (i.e., negative rechallenge) suggest (but do not prove) that the drug did not cause the event. One must consider such factors as whether it was possible for the event to recur, the dose of drug and duration of exposure at which the subject was rechallenged, and whether the length of observation following rechallenge was sufficient to allow recurrence of the event of interest.

It is important to distinguish the processes described above from the causality analyses of drug-related events often provided by investigators and applicants in NDA/BLA submissions. The analyses of drug-related adverse events presented by applicants are usually based on assessments made by investigators at the time of an event, are highly dependent on information about the side effect profile of the drug available at the time of the study (e.g., what is in the investigator's brochure), and are not informed by awareness of the entire safety database. These analyses are generally not expected to provide much useful information in assessing causality.

Assessment of the drug-relatedness of an adverse event is fundamentally different for relatively frequent and relatively rare events. For the former, a reviewer would compare the incidence of adverse events occurring in the study drug group to that in the placebo (or other control) group (in RCT). For rare events, the expected rate in a clinical trial database would be zero. Thus, if even a few cases (sometimes even a single case) of a rare life-threatening event occurred when none was expected, that would represent a serious safety problem for a drug product that does not provide unique efficacy or some other advantage over available treatments.

III. SPECIFIC GUIDANCE ON THE CONTENT OF THE SAFETY REVIEW

The following sections bear the same names and numbers of the section of the clinical review template that contains the safety review (Section 7.0).

This guidance organizes the safety review into three main sections:

- Methods and Findings (Section 7.1)

This section contains 17 subsections. Overall, this section should describe the relevant data sources, the safety assessments that were carried out, and the major findings of the detailed safety review. Section 7.1 should use a systematic approach to describing available data. It should focus first on the serious and potentially serious reactions, the kind that can affect the approval decision or severely limit the use of the drug (see Subsections 7.1.1 to 7.1.4). Focus should then move to the more common reactions that rarely influence approval but are often critical to patient and physician acceptance of the drug. Section 7.1 should then consider less common events, laboratory findings, vital signs, electrocardiograms (ECGs), immunogenicity, human carcinogenicity, human reproductive toxicity, withdrawal phenomena, abuse potential, and overdose (Subsections 7.1.5 to 7.1.17).

- Adequacy of Patient Exposure and Safety Assessments (Section 7.2)

  This section should address the adequacy of exposure (e.g., overall patient numbers and numbers for specific demographic subsets, duration of exposure, the dose levels at which exposure took place,[9] the quality and completeness of safety evaluations, whether all necessary evaluations were conducted (e.g., animal tests, in vitro tests, long-term safety testing, specific assessments of ECG effects), and whether any additional safety assessments are needed (either pre- or postapproval). This section should also include a subsection on additional submissions of safety data, including safety update(s).

- Summary of Selected Drug-Related Adverse Events, Important Limitations of the Data, and Conclusions (Section 7.3)

  This section should identify and briefly summarize the critical findings of the safety review, including the adverse events the reviewer considers to be important and drug-related and any important limitations of the safety database.

This guidance also discusses general analytical methods that may be useful for multiple aspects of the safety assessment in section 7.4 (see page 45). This section discusses pooling of data, explorations for adverse reaction predictive factors, dose-dependency evaluations, time dependency evaluations, duration of adverse reactions, drug-demographic interactions, drug-disease interactions, and drug-drug interactions.

The annotated outline of the review begins here:

7.0 Integrated Review of Safety

[9] The proportion of patients exposed to the dose range that is effective should be considered. A total exposure consistent with ICH recommendations (1500 total with 300-600 for 6 months and 100 for one year), may, on examination, reveal far fewer who received an effective dose (i.e., the dose that would be used). ICH E-1 ([The *Extent of Population Exposure to Assess Clinical Safety: For Drugs Intended for Long-Term Treatment of Non-Life-Threatening Conditions, March 1995 (http://www.fda.gov/cder/guidance/iche1a.pdf)*]) is clear in its expectation that the 6-month and one year exposures should be at dosage levels intended for clinical use. Although it is silent with respect to the 1500 figure, exposure at lower doses would not be expected to be informative about the safety of the clinically useful dose.

7.1 Methods and Findings

This section consists of 17 subsections (e.g., death, other serious adverse events, laboratory findings). Each of these subsections is organized somewhat differently, depending on the content. In presenting analyses in the safety review, it is important to clearly distinguish between the applicant's analyses and conclusions and those of the reviewer.

In discussing serious adverse events and dropouts (Sections 7.1.1, 7.1.2, and 7.1.3), it is critical that the reviewer identify individual patients in a way that enables subsequent readers to readily access data and supporting information if needed (e.g., study #, investigator #, patient ID#).

7.1.1 Deaths

Identifying Deaths Relevant to the Safety Review

Deaths occurring during the following time periods or under the following conditions should be assessed:

- Deaths occurring during participation in any study, or during any other period of drug exposure
- Deaths occurring after a patient leaves a study, or otherwise discontinues study drug, whether or not the patient completes the study to the nominal endpoint, if the death:
  - is the result of a process initiated during the study or other drug exposure, regardless of when it actually occurs; or
  - occurs within a time period that might reflect drug toxicity for a patient leaving a study or otherwise discontinuing drug. For drugs with prompt action and relatively short elimination half-lives, 4 weeks is a reasonable time period. For drugs with particularly long elimination half-lives or drug classes with recognized potential to cause late occurring effects (e.g., nucleoside analogs, gene therapies, or cell transplants), deaths occurring at longer times after drug discontinuation should be evaluated.

The reviewer should consider all deaths that occurred in a drug's development program and any other reports of deaths from secondary sources (e.g., postmarketing or literature reports), without regard to investigator or applicant judgment about causality. It is also important to consider deaths on control treatments for comparison, even though they are obviously not related to the drug in the application. Individual deaths should be listed in a table (see Table 7.1.1.1), unless they are an effectiveness study outcome.

Applicants will provide line listings of all patients who died in studies, together with a brief narrative. (See *E3 Structure and Contents of Clinical Reports*, Section 12.3.[10] The narratives may also be placed in the Integrated Summary of Safety.)

[10] The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) *E3 Structure and Contents of Clinical Study Reports.*

Distinguishing Expected from Unexpected Deaths

Certain causes of death are sufficiently unusual in the absence of drug therapy, even in large databases, that they would almost always be considered unexpected (e.g., aplastic anemia or acute hepatic necrosis) and deserve detailed individual discussion. Other fatal events occur at such frequency in the general population that they would be expected to occur in any large database absent drug therapy (e.g., fatal strokes and heart attacks), especially in the elderly.[11] In most cases, these events need to be examined for frequency but discussion of individual cases is not helpful. Expected deaths would include:

- Deaths in studies in which mortality is an endpoint and the cause of death is expected for the disease or condition
- Deaths in studies in diseases where high mortality rates are expected and the cause of death is expected (e.g., cancers). Note, however, that early deaths in cancer studies are a concern as patients are usually selected for clinical trials because they were not expected to die soon.
- Coincidental deaths resulting from progression of underlying disease present at enrollment in a study (e.g., a patient who dies from progression of cancer or Alzheimer's Disease or an acute myocardial infarction attributed to underlying coronary artery disease present prior to study entry
- Deaths from intercurrent long-term illness. These include the wide variety of fatal events that can be seen in any population, especially a relatively elderly population, such as sudden death (presumably representing an arrhythmia), fatal infections, surgical emergencies, or intracranial hemorrhage).

Even though fatal events may be expected in a population, the reviewer should not without further consideration readily accept the conclusion that a fatal event is due to the underlying disease or an intercurrent illness and not the drug. For each fatal event, the reviewer should specifically consider the possibility that the event represents an as yet unsuspected adverse reaction. Even if there is nothing about these deaths to suggest a drug cause, it is critical to assess whether the rate of these events is increased. The best way to do this, of course, is by comparison with a control group (a single trial or pooled), but if no control group is available, it may be of value to look at databases of other drugs used in the same population.

When distinguishing between unexpected and expected deaths, the reviewer should make clear the bases for the distinctions (e.g., early deaths in cancer patients are unexpected if the entered patients were chosen because they were not expected to die soon; hematologic deaths are unexpected in a postinfarction study). For unexpected deaths, the individual medical events associated with the death should be carefully evaluated and discussed in detail in the review. Expected deaths should be classified as to type of death, but it is usually not necessary to discuss in detail the individual medical events associated with those deaths. What is critical is to consider whether there is a suggestion that their rate is increased, the adequacy of the data to evaluate this, and the need to know more (e.g., because of experience with related drugs).

---

[11]Note that *unexpected* is used differently from its use in 21 CFR 312.32, where it refers to adverse events not identified in the investigator's brochure and therefore reportable in an IND Safety Report.

Pooling of Relevant Data

Before conducting any mortality analyses, the reviewer must consider the poolability of the data pertinent to deaths. If data are not poolable, analyses should be conducted for separate databases, then examined together. See section 7.4.1. for discussion of pooling.

Overall Mortality Analysis

The review should include an analysis of overall mortality for all phase 2 and 3 exposures across treatment groups as well as cause-specific mortality to the extent possible. The *fineness* of classification depends on the quality of data and the number of events (cardiovascular (broad) vs. AMI, sudden death, CHF (more specific)), recognizing that assessing cause-specific mortality is very difficult even in the best circumstances, such as a study in which there is an attempt to describe such endpoints prospectively.[12] Death from an acute myocardial infarction can be indistinguishable from death resulting from an arrhythmia, for example. Analyses should be corrected for differences in drug exposure using person-time in the denominator to calculate mortality rates.[13] If person-time exposure is not included in the submission (ideally, it should be requested at the pre-NDA/pre-BLA meeting), it should be requested as soon as the need is recognized. This correction can be done only for those deaths for which person-time data are available. It may be useful to present both crude mortality and mortality expressed in person-time in an appendix table (see Table 7.1.1.2 for sample display). Life table approaches may be helpful in cases when there are more than a few deaths, and when the direction of different studies varies significantly. Ideally, one would have mortality data from other databases for comparison (e.g., from other drugs in the same class).

Discussion of Applicant's Assessment of Deaths

The reviewer should describe and evaluate the applicant's assessment of deaths, including the following:

- The applicant's criteria for including deaths in the NDA/BLA (e.g., whether the criteria were reasonable, whether the criteria were met)
- The methods used by the applicant to detect and classify deaths
- The applicant's method of analyzing overall mortality and cause-specific mortality
- The applicant's judgments on the drug-relatedness of events associated with deaths

Reviewer's Assessment of Deaths

---

[12] Temple R, G Pledger, The FDA's Critique of the Anturane Reinfarction Trial, *N Engl J Med* 303:1488-1492, 1980.

[13] Since placebo and active control patients can generally have had shorter durations of exposures than patients given the new drug, they may have had less opportunity for serious events to have occurred.

The reviewer's assessment of deaths, reflecting both the applicant's and the reviewer's analyses, should include the following:

- Listing of information upon which reviewer assessment is based (e.g., CRFs, narrative summaries, consultant reports, autopsy reports)

- Tabular Summary of Deaths. Deaths should be summarized in an appendix table, as illustrated in Table 7.1.1.1. It may be useful to distinguish between those deaths for which exposure data are available and those for which such data are unavailable (e.g., for postmarketing deaths, exposure data may never be available). In the table and subsequent discussion, there should be a clear identifier so that subsequent reviewers can identify the particular patient.

- Analysis of overall mortality for Phase 2 and 3 drug exposures across treatment groups (see Table 7.1.1.2)

- Analysis of cause-specific mortality across treatment groups (this could use a table similar to Table 7.1.1.2)

- The reviewer's overall judgment about the drug-relatedness of medical events associated with death (i.e., which deaths were probably explained by factors other than the study drug (e.g., another drug, underlying illness, another illness common in the population) and which could not reasonably be explained by such factors). Differences from the applicant's evaluation should be noted and discussed.

- Further discussion of the individual events associated with death and believed to be potentially drug-related, either because they are increased in rate compared to control or because of the nature of the event (e.g., events typically drug-related, such as aplastic anemia or acute hepatic necrosis, or events that would not be expected in the population studied such as sclerosing abdominal or pulmonary conditions or rapidly progressive unexplained renal failure). Any uncertainty about drug-relatedness should lead to inclusion of the event. For each of these individual events, brief narratives should be included in the review or (if numerous) attached in an appendix.

- Other relevant analyses, such as analysis of dose response (administered dose, body weight and surface area adjusted dose, cumulative dose, schedule (including duration of infusion for IV drugs) analysis of mortality within critical subgroups (e.g., demographic, disease severity, excretory function, concomitant therapy), drug-demographic, drug-disease, and drug-drug interactions (see section 7.4)

- When deaths occur in uncontrolled studies, best available estimates of mortality in the population studied, in the absence of the treatment (see Section 7.4.3)

- When deaths are relatively frequent, the reviewer should consider some of the approaches described for Common Adverse Events (see section 7.1.5)

7.1.2 Other Serious Adverse Events

Identification of Nonfatal, Serious Adverse Events

The reviewer should identify, without regard to the applicant's causality judgment, all serious adverse events that occurred in the drug's development program or were reported from secondary sources (e.g., postmarketing or literature reports). Serious adverse events may, in addition to signs, symptoms, and diagnosable events, include changes in laboratory parameters, vital signs, ECG, or other parameters of sufficient magnitude to meet the regulatory definition of a serious adverse drug experience.[14]

Applicants generally provide a line listing of all patients in phase 2 and 3 of the development program who had an event meeting FDA's criteria for a serious adverse event. For each such event, the applicant should also provide a brief narrative (see ICH *E3 Structure and Content of Clinical Study Reports*, Section 12.3.2). Because the definition of *serious* is subject to some interpretation, the reviewer should make clear how the applicant created the list. For example, applicants may include events considered serious by investigators, even if they do not technically meet the FDA or ICH definition for a serious event. If such events are included, the inclusion parameters should be noted.

Discussion of Nonfatal Serious Adverse Events

This section of the review should contain the following:

- A brief description of data sources used in the review of individual cases (e.g., case report forms, applicant's narrative summaries, hospital records).

- The Tabular Summary of serious adverse events (see sample listing, Table 7.1.2.1)

- An analysis of overall rate of serious events and rate of specific serious events, for each treatment group in critical subgroups (e.g., demographic, disease severity, excretory function, concomitant therapy), and by dose. The median duration of exposure should be examined across treatment groups. If there is a substantial difference in exposure across treatment groups, incidence rates should be calculated using person-time exposure in the denominator, rather than number of patients in the denominator, similar to the presentation for deaths in Table 7.1.1.2.

- Reviewer's overall assessment of which serious adverse events were probably explained by factors other than the study drug (e.g., another drug, underlying illness, another illness common in the population) and which could not reasonably be explained by such factors)

[14] See 21 CFR 312.32(a); 314.80(a); 600.80(a).

including any pertinent information from ***non-serious*** events (e.g., seizures not leading to hospitalization, all syncope) that may be related to the serious event.

- Further discussion of each individual serious adverse event judged to be drug related (i.e., each adverse reaction), as needed, including any relationship of the reaction to death. For each of these reactions, brief narratives should be included in the review or (if numerous) attached in an Appendix.

- A discussion or listing of serious events considered unlikely to be drug-related (may be identified in the Tabular Summary, illustrated in Table 7.1.1.1).

If serious nonfatal adverse events are relatively frequent, the reviewer should consider some of the approaches described for Common Adverse Events (see section 7.1.5).

7.1.3 Dropouts and Other Significant Adverse Events

FDA regulations require that the CRFs from patients who discontinue treatment in association with an adverse event (adverse drop-out) be submitted with the application (21 CFR 314.50(f)(2)) and their analysis constitutes a critical part of the safety evaluation (see Section 7.1.3.2).

ICH *E3* (*Guideline on Structure and Content of Clinical Study Reports*) defines a new category of *other significant adverse events*. It includes:

- Marked hematological or other lab abnormalities not meeting the definition of serious. This will need to be an individual judgment, probably depending on the drug (e.g., CPK elevation could have a different implication for a statin and a different drug)

- Any events that led to an adverse dropout or any other intervention such as dose reduction or significant additional concomitant therapy (an expansion of the *adverse dropout* concept that appears in 21 CFR 314.50(f)(2), and in the Clinical/Statistical guideline, *and M4: The CTD — Efficacy)*

- Potentially important abnormalities not meeting the above definition of serious and not leading to death or modification of therapy (e.g., a single seizure, syncopal episode, orthostatic symptoms)

If the applicant has included listings for *other significant adverse events* these may be described here, under a subsection separate from the discussion of dropouts (see Section 7.1.3.3). Alternatively, marked laboratory changes may be described under Laboratory Findings (Section 7.1.6).

7.1.3.1 Overall Profile of Dropouts

The review should contain an overall profile of dropouts from clinical trials. The profile should classify dropouts from the overall phase 2 and 3 study pool by reason for dropping out (e.g.,

adverse event, treatment failure, lost to follow-up). Where there are clinically relevant differences in dropout rates for certain subsets (e.g., dropouts in placebo-controlled trials vs. dropouts in other studies; dropouts in certain demographic or disease-related subgroups), the profile should also classify dropouts for those subsets. The reviewer should explain the basis for selecting identified subsets and provide mutually exclusive tabulations in which individual patients are counted only once.[15] Ordinarily, the dropouts should be categorized in a table or tables appended to the safety review (see Table 7.1.3.1.1). It can be useful to display, graphically or in tables, the cumulative dropout rates for each treatment group within each study, especially for cause-specific reasons, when this information is available and to assess patient baseline risk factors that contribute to differential cumulative dropout patterns. When pooling data, consideration of dropout patterns over all studies may reveal information that is useful to the overall safety evaluation.

When classifying dropouts, the reviewer should carefully examine the reasons identified by the applicant for subjects dropping out. Heightened scrutiny is warranted for:

- Dropouts classified as administrative, lost to follow-up, or a similar term
- Dropouts for which the applicant changed the investigator's determination of the reason for the drop out.

Discontinuations attributed to adverse events require submission of the corresponding CRF but CRFs may not be submitted for dropouts classified as administrative or lost to follow-up. If such CRFs are not available, the reviewer may need to request at least a sample of them to determine whether these dropouts may have occurred in association with an adverse event. Where dropouts are reclassified by an applicant (i.e., assigned a reason for dropping out other than the one given in the CRF), the review should indicate how and by whom such reclassifications were made and comment on the appropriateness of decisions.

Ordinarily, the reviewer should combine patients categorized as dropping out for intercurrent illness and patients categorized as dropping out for adverse drug reactions (if the applicant makes that distinction) under the general category of dropouts for adverse clinical events (CRFs should be provided for both of these categories). This categorization is neutral from the standpoint of causality judgment, recognizes the great difficulty in making distinctions between adverse drug reactions and intercurrent illness, and encourages the reviewer to consider the possibility that what seemed to be another illness or a consequence of the underlying illness was, in fact, an adverse drug reaction.

The reviewer should examine the number and distribution of dropouts to identify potential problems with study conduct or analyses (e.g., a substantial number of dropouts due to *lost to follow-up* and sites with disproportionately high dropout rates should be a sign of concern). For example, early dropouts generally, and differential (drug group vs placebo group) early dropouts

[15] *Mutually exclusive* refers to the reason for dropping out. Patients should be identified with only one of the reasons. However, patients may be represented in more than one column (treatment group) of a table (e.g., patients in a crossover study may have survived several treatment arms and then dropped out).

in particular, often present difficulties in conducting and interpreting the effectiveness analysis and may suggest breakdown of blinding. The review should discuss any concerns about dropouts and the methods employed by the reviewer to address them.

7.1.3.2 Adverse Events Associated with Dropouts

The analysis of adverse events associated with dropouts is important for two distinct reasons. First, it identifies the type and frequency of adverse events that patients were unable to tolerate even in a clinical trial setting, where there is arguably more support for enduring adverse events than in a clinical practice setting. This provides important prescribing information that can contribute to dose selection, and in some cases, to choosing a method of titration. In most cases, there will be little doubt about which of these events are attributable to the drug because the events will be of relatively high frequency, even if withdrawals because of them are not, and the main issue will be their frequency and importance. It is usually not necessary to review these events case by case.

Second, and the reason CRFs for dropouts due to adverse events are provided automatically to reviewers, these adverse dropouts may provide a clue to unexpected, but important, adverse reactions, (e.g., fibrosing intra-abdominal or pulmonary illnesses, progressive liver or kidney diseases, cardiac valve damage, neurological diseases, arteritis, thromboembolic diseases, all of which have been caused by drugs) that can easily be dismissed as intercurrent illness. The frequency of these events is likely to be very low, and the review should contain an analysis of each such adverse event that resulted in withdrawal from the study, whether or not the event was attributed to the drug. The reviewer should avoid dismissing such events as intercurrent illness and specifically consider the possibility that each dropout not due to a known effect of the drug might reflect an unexpected effect of the drug. The applicant will usually provide a line listing of adverse dropouts (Table 7.1.3.2.1) and this listing (which need not be attached to the review) can serve to identify events needing further scrutiny. Review of the CRFs can often provide critical insights. The reviewer should describe how she or he analyzed these events.

With respect to the more common adverse events leading to discontinuation of treatment, the review should present:

- The incidence of adverse events associated with dropouts. Ideally, incidence would be presented in a table or tables appended to the safety review with separate tables for subsets of the overall clinical data pool in which there were clinically meaningful differences in dropout incidence (Table 7.1.3.2.2). Tables displaying incidence of events should include each event that led to a dropout even if a single patient had more than one such event.

- Whether the event can reasonably be considered drug-related; this conclusion will be based on comparisons between treatment groups in controlled trials and can be informed by the overall rate of the adverse event (Section 7.1.5), and the known pharmacology of the drug.

- The dose response and time dependency of the drop-outs and drug-demographic, drug-disease, and drug-drug interactions (see section 7.4 General Methodology).

For the rarer events that could suggest an important adverse reaction, the critical review determination is whether any of these events suggest drug-induced injury. These events need to be considered individually, with narratives and reference to other databases as appropriate.

Where the review contains applicant-generated tables, it is important for the reviewer to determine and describe how the tables were created. A table may identify one or more adverse events as having caused a particular patient to withdraw, in which case it would represent the actual incidence of specific adverse events that led to dropout. This approach is preferred. Alternatively, a table may list the adverse events that a subject experienced at the time of dropout and not identify any event (or events) as causing the dropout. This approach does not provide the actual incidence of adverse events associated with dropouts and is of less value. The reviewer should make clear in the review which of these approaches was used, or whether an alternative approach was used.

7.1.3.3 Other Significant Adverse Events

If a submission separates out information on adverse events that led to dose reduction or significant additional concomitant therapy,[16] but not to discontinuation of treatment, those findings should be described using an approach similar to that proposed above for adverse dropouts.

7.1.4 Other Search Strategies

In addition to reviewing deaths, serious adverse events, and adverse events associated with dropouts, it may be useful to construct algorithms involving combinations of clinical findings that may be a marker for a particular toxicity (e.g. serotonin syndrome, cough, chest congestion and shortness of breath that may constitute drug-related bronchospasm, or drug-induced Parkinsonism). When such algorithms are used, the algorithm and results of the search using the algorithm should be described in the review. Generally, and where possible, such searches should be done while the reviewer is blinded to treatment, as this will minimize bias when identifying cases.

It should be noted that the causal relation of a drug to uncommon serious adverse events may be supported by less serious events that are more common. For example, the likelihood that a drug caused a small number of cases of serious liver toxicity may be supported by a higher rate of transaminase elevation.

7.1.5 Common Adverse Events

This section of the review focuses on establishing the common adverse reaction profile for the drug and determining the content of the adverse reaction table(s) to be included in labeling. NDAs typically contain numerous tables and analyses of adverse event incidence (e.g., by study, by various pools of studies, and for the overall database). In general, what are included are *TESS*, treatment emergent signs and symptoms (i.e., signs and symptoms not present at baseline,

[16] This is recommended in the ICH *E3* guidance.

or not present at the severity seen on treatment). To approach these data, the reviewer should generally go through the steps outlined in sections 7.1.5.1-7.1.5.5.

7.1.5.1 Applicant's Approach to Eliciting Adverse Events in the Development Program

Adverse events can be elicited by open-ended questions or checklists with varying degrees of specification. Each approach has advantages and disadvantages, but results can differ greatly and may lead to marked differences in reported adverse event rates across studies (it would not usually be appropriate to pool results obtained using both methods). The reviewer should describe the applicant's method or methods of eliciting adverse event data in clinical trials, including whether checklists were used, the frequency with which patients were assessed, and whether the approaches differed among studies. Identification of signs (abnormal findings observed by a clinician) would seem to be less of a problem, as these are elicited by physical examination, but use of a physician exam checklist could lead to a different result from a more general requirement for physical exam. If different approaches were used (e.g., checklists in studies conducted in the United States, open-ended inquiries in European studies), the reviewer should consider and discuss in the review the effect, if any, on the adequacy of adverse event information collected.

7.1.5.2 Establishing Appropriate Adverse Event Categories and Preferred Terms

Although investigator adverse reaction terms are provided as part of study reports and are listed in case report tabulations, the integrated analysis of the ISS requires the applicant to use some way of grouping closely related events to obtain an overall rate for a category of events. This is accomplished by using a so-called *dictionary* of preferred terms, such as COSTART,[17] MedDRA,[18] the latter a more granular listing developed under the auspices of ICH. These *dictionaries* are in fact lists of preferred terms and leave (especially COSTART and other older dictionaries) considerable discretion to the classifier to choose the term that best reflects the verbatim term reported by the investigator. The categorization of such systems, however, may not capture, or can dilute, the true meaning of certain events. In addition, terms used in COSTART may not be informative (e.g., pain, tooth disorder). It is expected that as MedDRA becomes more widespread, this will no longer be such a problem. It is critical that the reviewer assess the appropriateness of the applicant's categories and the coding of adverse event verbatim terms to preferred terms and understand how the verbatim terms (including terms in languages other than English) were classified.

In assessing the applicant's coding of events, the reviewer should compare the applicant's preferred terms to the verbatim terms used by investigators and patients, focusing on the events leading to dropouts or other changes in treatment as well as to serious adverse events. The applicant will usually provide (ideally this will have been agreed to at the pre-NDA/BLA meeting, but if not, it should be sought early in the review) the following tables and listings for

---

[17] Food and Drug Administration, *Coding Symbols for Thesaurus of Adverse Reaction Terms*, 5th ed., FDA, Rockville, MD, 1995.

[18] MedDRA (Medical Dictionary for Regulatory Activities), http://www.meddramsso.com/NewWeb2003/index.htm.

assistance with this assessment; they should be provided in a form the reviewer can manipulate, such as a SAS transport file, not just in PDF format:

- Adverse event tables in individual study reports based on investigator terms for events

- A comprehensive line listing of all adverse events in phase 2 and 3 studies with a column containing investigator terms coded under a preferred term (see Table 7.1.5.2.1. This table is for reference only; it would not be included in the review)

- Listing of preferred terms and the investigator and patient terms that were subsumed under the preferred term. This table is for reference only; it would not be included in the review, although parts of it might be).

The reviewer should consider the following:

- Whether terms are too narrow (*splitting*), resulting in an underestimation of the true incidence for a particular event or syndrome (e.g., somnolence, drowsiness, sedation, and sleepiness probably all refer to the same event)

- Whether the terms are too broad or over-inclusive (*lumping*), so that important events that should be examined separately are diluted by less important events (e.g., loss of consciousness and syncope subsumed under hypotensive events or hypotension)

- Whether terms used lack a commonly understood meaning (e.g., mouth disorder, tooth disorder, GI disorder) and, if so, whether the incidence of individual events subsumed under these terms should be expressed separately or mapped to a different preferred term

- Whether terms exaggerate a finding (acute liver failure for a transaminase elevation) or minimize the importance of an event (hypotension for a syncope episode)

- Whether the coding of adverse events is similar across treatment groups

In any of these cases, the reviewer (or the applicant at the request of the reviewer) may have to recalculate rates using alternative terms or different groups of terms.

Usually it will be impossible to evaluate all or even most adverse event terms in this much detail. However, certain preferred terms should always have the subsumed verbatim terms examined because they may conceal important events. For example, *accidental injury* often includes fractures and/or lacerations related to falls. The fall, itself, however, may not have been captured as an adverse event. Additionally, *edema* may sometimes include *facial edema*. Since facial edema often represents an allergic reaction, one would not want allergic events lumped together with peripheral edema events. In general, adverse event terms associated with discontinuation or serious consequences deserve the closest scrutiny, but other classifications should be at least spot-checked. The review should comment on how this issue was addressed.

7.1.5.3 Incidence of Common Adverse Events — Assessment of Various Databases

Applicants typically prepare a wide variety of tables of adverse event rates for individual studies and pools of various studies. Those tables generally include investigator causality assessments and severity ratings. The tables the reviewer considers useful should be appended to the review. Incidence rates for common adverse events may be estimated from the relatively small portion of the overall database that is contained in the controlled (especially placebo-controlled) trials. For these more common events, the ability to compare rates on drug with a control outweighs the disadvantage of basing the rate estimates on fewer subjects. In determining incidence rates for common adverse events, the reviewer should identify the subset of trials in the phase 2 and 3 database that will provide the best estimate of rates and develop tables of event rates based on that judgment.

- If possible, the reviewer should rely on pooled data from studies using the same comparator group (e.g., only placebo-controlled trials) and of roughly similar duration. If some of the trials also had an active control, rates for that group (pooled) can also be included (see Table 7.1.5.3.1). If different doses were used, both a pooled all doses group and individual dose groups can be shown. The best comparison is of the groups included in all studies (drug at a particular dose and placebo), but the others (active controls, individual dose groups) may also be useful (also see Section 7.4 for broader discussion of pooling).

- If there are not adequate numbers of patients in such trials to give meaningful rate estimates, the reviewer should consider pooling placebo-controlled trials, active control trials, and three arm trials (i.e., trials that do not all have the same control group). Even when this approach is needed overall, smaller subsets of studies, or even individual studies, can be used to examine high-frequency events.

- Most applicants will construct adverse event tables by compiling and presenting the numbers and/or percentages of patients experiencing an adverse event in a study (or the absolute number of adverse events experienced in a group), without regard to the duration of treatment received. This is often satisfactory for relatively short-term studies. If studies of significantly different durations are pooled, however, or if there is a different discontinuation rate in the treatment arms and the risk of the adverse reaction persists over time, one must consider these durations to understand the real occurrence rate that patients will experience. One way to deal with the problem of different durations is to use the total person-time exposure for each treatment group and calculate the rate of the adverse event per period of exposure (# of patients with adverse event total person-time exposure), rather than the risk (# of patients with adverse event total number of patients). This is particularly useful for the more important adverse reactions and reactions that occur at a fairly constant rate over time, but the person-time approach can also be used when the hazard rate changes over time. In this case, however, the observation period must be broken into component periods (e.g., evaluating person-time rates for each treatment for month 1, month 2, ...).

- If concurrently controlled data are unavailable, overall rates from well-monitored, single-arm databases can be used to provide some indication of rates that were observed in treated

patients, but there is little ability to establish causality except insofar as reactions are predicted by the known pharmacology of the drug.

For the most part, attributions of causality by the investigators should be discounted, and adverse events should be assessed without regard to attribution. Also, in general, tables should give rates for all severities of a given effect, although in some cases (notably cytotoxic drugs), it is important to distinguish more and less severe reactions, as the former may be therapy-limiting or may affect the overall benefit-risk conclusion for the drug. For events with high background rates (e.g., headache, fatigue, and other events that occur frequently independent of drug therapy), however, display of all reported events can result in a high event rate that obscures drug-relatedness. This can be a particular problem when time on drug is prolonged. For example, it is common for studies of 4 to 6 weeks duration to report headache at a high (20 to 25 percent) rate. In that case, considering the severity or causality assessment of such events may allow a better assessment (e.g., if severe headaches are found only in the drug-treated group). Events that are more severe and for which subjects have multiple occurrences while on drug therapy are more likely drug-related. In determining incidence, however, both single occurrence and multiple occurrence events should be counted as one event.

Some categories of adverse events (e.g., decreased cognitive or sexual function) are notoriously difficult to detect without special efforts, such as targeted questionnaires. If the database includes special studies intended to identify these events, they should generally be given more credence than nontargeted studies, which tend to substantially underestimate rates (See Section 7.1.9). Incidence rates should be based on findings from the targeted studies.

7.1.5.4 Common Adverse Event Tables

The review should contain a table (or tables) that presents the best overall display of commonly occurring adverse events, generally those occurring at a rate of 1 percent or more (but lower rates can be presented for very large databases). The table, or tables, will form the basis for the adverse reaction table in labeling. The table may use a higher cut off than 1 percent if doing this does not lose important information. Adverse events that are equally common on drug and placebo, or more common on placebo, are usually omitted. The frequency cut-off for inclusion of adverse events in the table (e.g., > 1%) is inherently arbitrary. If one is used, the review should explain how the threshold was determined. It may also be informative to include tables that distinguish between common adverse events on the basis of severity. It is most common to group adverse events within body systems, but a display by descending frequency may also be useful.

7.1.5.5 Identifying Common and Drug-Related Adverse Events

For common adverse events, the reviewer should attempt to identify those events that can reasonably be considered drug related. Although it is tempting to use hypothesis-testing methods, any reasonable correction for multiplicity would make a *finding* almost impossible, and studies are almost invariably underpowered for statistically valid detection of small differences. The most persuasive evidence for causality is a consistent difference from control across studies, and evidence of dose response. The reviewer may also consider specifying criteria for the

minimum rate and the difference between drug and placebo rate that would be considered sufficient to establish that an event is drug related (e.g., for a given dataset, events occurring at an incidence of at least 5 percent and for which the incidence is at least twice, or some other percentage greater than, the placebo incidence would be considered common and drug related). The reviewer should be mindful that such criteria are inevitably arbitrary and sensitive to sample size.

7.1.5.6 Additional Analyses and Explorations

For adverse events that seem drug related (the analyses suggested can have no value for unrelated events), the reviewer should perform the following additional analyses (see Section 7.4, General Methodology, for discussion of methods for the explorations and analyses identified below), as appropriate:

- Explorations for dose dependency. These are important. The reviewer should ordinarily rely on fixed dose studies, as titration studies tend to show that those who tolerate higher doses have lower adverse reaction rates, but in some cases titration studies may show a clearly increased rate of adverse reactions with dose. It may also be useful to evaluate safety as a function of weight-adjusted dose, body surface-adjusted dose, or cumulative dose. Dose increases may be associated with adverse reactions or the severity of adverse reactions.

- For events that occur commonly, explorations of time to onset

- For common, troublesome events (e.g., somnolence, nausea) explorations of adaptation to develop information on the time course of, and tolerance for, such events

- Explorations for demographic interactions (rates and comparisons with control for demographic and other subsets) for at least the more common and important adverse events. The applicant will have provided such analyses under 21 CFR 314.50(d)(5)(vi). Note that this analysis may require use of less optimal tables of pooled results (see section 7.1.5.3).

- Explorations for drug disease and drug-drug interactions if there is a strong signal for an interaction or good rationale for expecting an interaction

- Selective exploration of certain adverse events in an attempt to better characterize them. For example, if rash appears to be drug related, the reviewer may want to look more closely at individual cases of rash. The applicant's line listing of all adverse events across the entire phase 2 to 3 databases would be a good source for identifying individual cases of rash. If a subject dropped out because of rash, the applicant should have provided a narrative discussion of the event, which would also be a good source for attempting to better characterize the event. Although the data collected on nonserious adverse events is usually sparse, the reviewer could still request additional information from the applicant on commonly occurring adverse events that require further characterization.

- When adverse events of a given type vary markedly in severity, separate analyses of each severity may be useful.

A description of the methods used in such additional explorations should be provided, with all results, interpretations, and pertinent discussion. Where an applicant's analysis is considered inadequate, this should be noted and an alternate developed by the reviewer or requested from the applicant.

7.1.6 Less Common Adverse Events

In general, a fairly large database is needed to evaluate less common adverse events. To identify relatively rare events of significant concern, the reviewer has to examine the occurrence of adverse events over the entire phase 2 to 3 database, including data for which there is no useful concurrent control. The overall database is typically heterogeneous, including uncontrolled exposure for varying durations and at varying doses, and is unlikely to lend itself to meaningful estimates of rates or assessment of causality (except where there has been rechallenge). Thus, it may be sufficient for the reviewer to group these data in gross categories of incidence and by body system. For example, it may be useful to categorize less common events in order of decreasing frequency within the following incidence ranges:

- Adverse events occurring at rates less than or equal to 1/100
- Adverse events estimated to occur at rates between 1/100 and 1/1000
- Adverse events estimated to occur at rates less than 1/1000

The reviewer should then develop a condensed list of reactions to be included in the Adverse Reactions section of labeling.[19] This list should eliminate events that are common in the general population and not likely to be drug related and adverse events characterized by terms that are too vague to be helpful, unless the reviewer is able to identify a more meaningful term that was subsumed into the vague term when the adverse event was coded by the applicant (see Section 7.1.5.2 above).

Some of the reactions in the condensed list may be of particular concern, but insufficiently clear as to whether they are caused by the drug to lead to a Warning/Precaution in the labeling. In that case, it is useful to notify the safety evaluator in the Office of Drug Safety who will be monitoring the drug after marketing.

7.1.7 Laboratory Findings

The approach to review of laboratory findings (chemistry, hematology, and urinalysis) is generally similar to that suggested for the other categories of safety data. As considered in greater detail below, the review should identify laboratory tests performed in the clinical studies, describe the dataset from which laboratory findings information is obtained, describe the

[19] A draft guidance for industry and reviewers, *Content and Format of the Adverse Reactions Section of Labeling for Human Prescription Drugs and Biologics*, was issued in June 2000. Once finalized, it will represent the Agency's thinking on this topic.

methods used to assess findings, discuss pertinent findings, and review the more important findings in depth. Laboratory findings discussed in detail in other sections of the review (e.g., Section 7.1.2 Other Serious Adverse Events, Section 7.1.3 Dropouts and Other Significant Adverse Events) need not be discussed in detail in this section. This section should refer to the more detailed discussions of such findings elsewhere in the review.

7.1.7.1 Overview of Laboratory Testing in the Development Program

The review should provide an overview of what laboratory testing (chemistry, hematology, and urinalysis) was carried out. It is preferable to summarize the overall approach, rather than provide detailed comments about laboratory testing for each study. The review should contain the following to the extent relevant to the data:

- Discussion of any discrepancies between planned analyses and analyses that were done (e.g., tests omitted or added, changes in planned frequency of testing)

- Discussion of procedures used to evaluate abnormal values (e.g., whether patients were followed until their values normalized, whether any patients were rechallenged, the procedures used for sample analysis (i.e., central or local labs, *windows* of time in which lab values were considered[20])

- A summary table identifying the numbers of patients exposed to test drug who had baseline laboratory values and follow-up assessments

- Whether results of unscheduled lab tests were included in the principal analyses and tables

The reviewer should note that laboratory tests obtained at *unscheduled visits* (e.g., when a patient is hospitalized for an adverse event) are often not included in the NDA/BLA laboratory database. In those cases, the only place a reviewer would learn of an abnormal laboratory value might be a narrative summary (or occasionally a CRF). Too often, however, the narrative summary includes only a preferred or verbatim term (e.g., *acute renal failure*) and does not include the laboratory value of interest (e.g., BUN/creatinine). In such cases, the laboratory data of interest should be requested from the applicant.

7.1.7.2 Selection of Studies/Analyses for Drug-Control Comparisons of Laboratory Values

Controlled comparisons generally provide the best data for deciding whether there is a signal of an effect of a drug on a laboratory test. Placebo-controlled trials are generally short term, however, and therefore unsuitable for assessing late-developing abnormalities, so that longer term data also need to be examined. If there is no concomitant control, comparison may need to be made with similar populations outside the NDA (e.g., in other applications). In identifying the sample population for comparison of laboratory values, the reviewer should pool relevant

[20] Applicants may consider only lab values obtained within a certain window around the protocol-specified date for collection. In some cases, the laboratory data obtained outside the window may be available, but the applicant may choose not to include it.

studies. The review should explain how the studies to be pooled were selected. In comparing laboratory values, there are additional considerations when using pooled data (in addition to those discussed in Section 7.4.1 Methodology, Pooling), including:

- The methods of sample collection and handling in different studies
- The assay methods used in different studies
- The reference ranges used in different studies

Several analyses may be needed. Separate analyses should be performed for patients with normal values at baseline, for patients with abnormal values at baseline, and for patients without baseline values. In general, there will need to be at least one analysis that includes all data (data from planned or unplanned visits, values collected as follow-up to abnormal findings).

7.1.7.3 Standard Analyses and Explorations of Laboratory Data

This review should generally include three standard approaches to the analysis of laboratory data. The first two analyses are based on comparative trial data. The third analysis should focus on all patients in the phase 2 to 3 experience. Analyses are intended to be descriptive and should not be thought of as hypothesis testing. P-values or confidence intervals can provide some evidence of the strength of the finding, but unless the trials are designed for hypothesis testing (rarely the case), these should be thought of as descriptive. Generally, the magnitude of change is more important than the p-value for the difference.

7.1.7.3.1 Analyses Focused on Measures of Central Tendency

The central tendency analysis generally compares mean or median changes from baseline across treatment groups, and the review should contain the results of these analyses for all laboratory measurements. Although marked outliers are typically of greatest interest from a safety standpoint (see below), at times a potentially important effect may be revealed only in analyses looking at differences in mean change from baseline. For example, several drugs that cause modest decreases in uric acid because of a uricosuric effect have caused acute renal failure (ticrynafen, suprofen) in inadequately hydrated patients. Suprofen was withdrawn from the market for this reason. Mean changes in electrolyte levels can also signal risks.

It is generally useful to include as appendices tables providing data on central tendency (see Table 7.1.7.3.1.1). The reviewer should note and discuss signals that emerge from these tables and indicate those for which further study is needed, if any.

7.1.7.3.2 Analyses Focused on Outliers or Shifts from Normal to Abnormal

The review should focus on patients whose laboratory values deviate substantially from the reference range. Applicants usually include displays and analyses designed to detect such outliers. The relevant data would come from shift tables, scatter plots, box plots, cumulative distribution displays, and tables providing incidence of patients across treatment groups who had a potentially clinically important deviation from normal on one or more laboratory parameters

while on treatment (see Tables 7.1.7.3.2.1). In analyzing outliers, the reviewer should be aware of the following:

- Regression to the mean (and an apparent upward shift) can be expected if patients are screened for normality, giving a shift even if there is no drug effect; comparison with control groups is critical.
- If there are more measurements performed during treatment than baseline and abnormal values are randomly occurring, there is more opportunity for outliers during treatment. Again, comparison with a control group is critical.
- For important laboratory parameters, the reviewer should carefully consider the cut-points used by the applicant to define *normal* and *abnormal.*
- If values used to identify outliers are too extreme, important findings may not be identified.
- If values used to identify outliers are not large enough, important findings may be obscured by grouping important outliers and trivial findings (e.g., values greater than two times upper limit of normal for transaminase are common in many datasets and may not distinguish hepatotoxic from non-hepatotoxic agents; 3-fold and higher elevations appear to be more discriminating).

Decisions about what criteria to use to identify outliers should, if possible, be made at the pre-NDA meeting. Because it is not possible to know in advance what criteria will be optimal for detecting between-group differences, it may be useful to conduct analyses using cut points other than those chosen by the applicant. In addition, it may be useful to consider between-group comparisons of the following:

- Cumulative or other distributions of data, rather than solely proportions of patients meeting some arbitrary criterion
- Patients with large shifts within the normal reference range
- Patients who meet outlier criteria for more than 1 variable simultaneously (e.g., transaminase and bilirubin)
- Patients having persistent abnormalities (more likely to be real deviations)

Analyses of outliers should serve as a source of signals for events to explore in more depth. The reviewer should discuss signals that emerge and indicate those for which further exploration is needed. The details of the explorations carried out and the results should be provided in subsection 7.1.6.4 as described below.

7.1.7.3.3 Marked Outliers and Dropouts for Laboratory Abnormalities

The reviewer should carefully analyze individual patients with large changes in laboratory values. These changes are much more likely to identify significant problems than mean or median changes from baseline. Applicants typically provide a list that identifies patients with extreme changes, usually specified in advance. Individual patient data displays should be available to the reviewer for all such patients. Even for relatively uncommon events, it is helpful to compare rates in treatment and control groups.

Discontinuation of treatment for a laboratory abnormality may be considered a marker of perceived clinical importance of the finding. It is again useful to compare treatment groups, taking into account duration of treatments, for rates of discontinuation for particular laboratory abnormalities. Because of the importance of looking at dropouts for laboratory changes (even a small number of marked abnormalities, such as liver function or WBC count, may signal major problems), all such dropouts in the phase 2 to 3 population should be identified. The reviewer should generally analyze and comment on each individual patient identified as dropping out for any significant laboratory abnormalities. In some cases, it is critical to note whether appropriate testing has been carried out to rule out non-drug-related mechanisms (e.g., viral hepatitis serological testing in patients with transaminase elevation or more severe liver injury) and whether appropriate additional tests have been performed (e.g., bilirubin in patients with transaminase elevation).

7.1.7.4 Additional Analyses and Explorations

Additional analyses may be appropriate for certain laboratory findings, including analyses for dose dependency, time dependency, and also drug-demographic, drug-disease, and drug-drug interactions (see Section 7.4 Methodology). The review should discuss the rationale for additional explorations, the methods used, and the results and interpretations.

7.1.7.5 Special Assessments: Hepatotoxicity, QTc, Others

Certain laboratory assessments are so critical to the safety assessment that they deserve special attention in any review. For example, hepatotoxicity has been an important cause of drug marketing withdrawals from the 1950s (iproniazid) to the present (ticrynafen, benoxaprofen, troglitizone, bromfenac) and has led to important limitations on the use of many more drugs (isoniazid, labetalol, trovafloxacin, tolcapone, nefazodone, felbamate). At present, it appears that a potential for severe hepatotoxicity may be signaled by a set of findings sometimes called *Hy's Law*, based on the observation by Hy Zimmerman, a major scholar of drug-induced liver injury, that a pure hepatocellular injury leading to jaundice had serious implications, a 10 to 50 percent mortality. Over the years, this observation has led to the following proposition:
In a drug development database, a potential for severe hepatotoxicity is signaled by the following set of findings:

1. An increased rate of transaminase elevations (3x ULN, 5x ULN, 10x ULN, etc.) in treated patients compared to control

2. No significant evidence of obstruction (elevated AP), although some elevation may follow severe hepatocellular injury

3. A very small number of cases (two, perhaps even one) of transaminase elevation accompanied by a rise in bilirubin to 2x ULN

The explanation for the usefulness of this signal is the high capacity of the liver for bilirubin excretion; it takes a good deal of damage to the liver to impair bilirubin excretion (in the absence

of obstruction). This signal has been present for troglitizone, bromfenac, and dilevalol (never approved in the United States, but hepatotoxic in Portugal).

Table 7.1.7.5.1 is an outline of a comprehensive assessment of available data pertinent to potential hepatotoxicity. A similar outline will be developed for assessment of electrocardiographic QT abnormalities, a risk factor for potentially fatal arrhythmias (see section 7.1.9).

7.1.8 Vital Signs

Vital signs can be analyzed and reported using an approach essentially identical to that taken for laboratory data. This section should be organized in a similar manner to the laboratory section.

7.1.8.1 Extent of Vital Signs Testing in the Development Program

7.1.8.2 Selection of Studies and Analyses for Overall Drug-Control Comparisons

7.1.8.3 Standard Analyses and Explorations of Vital Signs Data

7.1.8.3.1 Analyses Focused on Measures of Central Tendency

7.1.8.3.2 Analyses Focused on Outliers or Shifts from Normal to Abnormal

7.1.8.3.3 Marked Outliers and Dropouts for Vital Signs Abnormalities

7.1.8.4 Additional Analyses and Explorations

7.1.9 Electrocardiograms (ECGs)

ECG data can be analyzed and reported using an essentially identical approach to that taken for laboratory data. The adequacy of the assessment (see 7.2) may be especially important in this case, given recent experience with drugs that prolong the QT interval and cause the ventricular tachycardia known as Torsade de Pointes (TdP). A guidance document on the design, conduct and interpretation of clinical studies assessing the effects of drugs on the QT interval is under development as a part of the ICH effort. The current version, a step 2 guidance (E14: Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs) can be found at http://www.ich.org. The safety review should provide in this section an overview of effects on the QT interval, organized in a similar manner to the laboratory section. This section of the safety review should summarize the results of any studies designed specifically to assess the effects of the drug on the QT interval.

7.1.9.1 Extent of ECG Testing in the Development Program, Including Brief Review of Preclinical Results

This section should describe the number of baseline and on-study ECGs obtained, who read the ECGs, and what methodology was used (e.g., automatic, blinded cardiologists).

7.1.9.2 Selection of Studies and Analyses for Overall Drug-Control Comparisons

7.1.9.3 Standard Analyses and Explorations of ECG Data

7.1.9.3.1 Analyses Focused on Measures of Central Tendency

7.1.9.3.2 Analyses Focused on Outliers or Shifts from Normal to Abnormal

7.1.9.3.3 Marked Outliers and Dropouts for ECG Abnormalities

7.1.9.4 Additional Analyses and Explorations

7.1.10 Immunogenicity

Data on the impact of immunogenicity (if applicable) on safety, efficacy, and/or clinical pharmacology and pharmacokinetics may be summarized in this section and referenced throughout the review.

All therapeutic proteins have the potential to elicit antibody responses. An antibody response to a protein may have no consequences or, in some cases, can lead to potentially serious sequelae. Adverse immune responses to a protein drug could result in one or more of the following outcomes:

- For a product that is intended as replacement for a missing endogenous substance, antibodies could neutralize the replacement product and generate a clinical deficiency syndrome.
- Neutralization of a protein product by *blocking* antibodies could reduce the efficacy of a life-saving product.
- Antibody development could result in a life-threatening hypersensitivity response.

Factors that tend to increase the likelihood of an immune response include whether the protein is highly conserved in nature (less likely if it is), whether the protein product is administered via the subcutaneous route (more likely if it is), and whether the protein intended for chronic use. This section of the review should assess the adequacy of the immunogenicity data provided to address these issues.

7.1.11 Human Carcinogenicity

Although formal studies in humans of the carcinogenic effects of drugs and biologics are uncommon, reflecting the expectation that induction of cancer would occur over a very long period of exposure, a systematic assessment of human tumors reported during drug development can provide useful safety information in some cases. Such an assessment would be appropriate where controlled studies are of long duration (e.g., more than a year), especially for drugs or biologics that have positive genotoxicity or animal carcinogenicity findings or are known immune modulators.

7.1.12 Special Safety Studies

The review should describe and discuss results of any studies designed to evaluate a specific safety concern or concerns. These studies may include:

- Studies to assess whether a drug has safety concerns common to its pharmacologic class (e.g., a study to assess effects of a benzodiazepine hypnotic on driving, respiration, memory, or next day psychomotor functioning)
- Studies in topical products (including systemic products delivered by a patch) to assess cumulative irritancy, contact sensitizing potential, photosensitivity, and photoallergenicity
- Studies to characterize a drug's effect on QT interval, part of most modern development efforts
- Studies intended to demonstrate a safety advantage over therapeutic alternatives (less extrapyramidal effect for an antipsychotic, less sedation for an anti-histamine, less cough from an angiotensin II blocker than an ACE inhibitor). Such studies must include the comparator agent (a failure to see the side effect in a placebo-controlled study is usually not informative without the active control to demonstrate assay sensitivity).
- Studies in special populations thought to be at increased risk and likely to use the drug.

In labeling, the results of these studies should, as appropriate, supersede data from less targeted studies (e.g., observational safety data collected from efficacy trials).

7.1.13 Withdrawal Phenomena/Abuse Potential

The review should contain a discussion of abuse potential and any apparent withdrawal symptoms. For therapeutic classes with a history of abuse potential and withdrawal phenomena (e.g., sedative/hypnotics and anxiolytics), studies are usually performed to assess these issues. The review should comment on the adequacy and findings of these studies. For other drugs, adverse events that emerge after discontinuation of the drug should be assessed to determine whether they may indicate a withdrawal phenomenon. If the applicant evaluated the potential for withdrawal phenomena, the review should indicate whether there was a prospective or post-hoc assessment of withdrawal emergent signs and symptoms (during drug taper or following discontinuation) and discuss the implications of the approach used on the reliability of the findings.

7.1.14 Human Reproduction and Pregnancy Data

Although formal studies in humans of the effects of drugs on reproduction, pregnancy, or lactation are uncommon, the review should summarize any drug exposure in pregnant or nursing women, including any inadvertent exposure during the drug's development and exposure identified from secondary sources (e.g., postmarketing surveillance). If there is no information on drug exposure in pregnant or lactating women, the review should acknowledge that fact. The review should discuss positive and negative findings.

7.1.15 Assessment of Effect on Growth

Increasingly, clinical reviewers are presented with analyses of height and weight data collected during studies of pediatric subjects. These data are generally inadequate to allow for definitive conclusions about an effect of drug on growth for several reasons. Assessment of the effect of drug on growth requires accurate measurements, particularly for height, and in most studies, height is not measured accurately. Growth is a process that occurs over long periods of time, and controlled trials of several weeks duration may not provide a sufficient period of observation to assess the effect of drug on growth. Open label studies can offer longer periods of time to observe effects on growth, but the lack of a control group limits the ability to separate the effect of drug and underlying disease on growth. Review of height and weight data for possible effects on growth makes use, in part, on approaches described above in the laboratory data section. Analysis of changes in central tendency and outlier analysis, for example, apply to the evaluation of the effect of a drug on growth. There are, however, some distinctive issues that must be considered.

First, the sponsor should describe how weight and height were measured. The manner in which these measurements were made will bear on how much confidence the reviewer can have in the data provided. For example, a development program in which the measurement schedule and methodology were standardized and in which the study staff were trained in measurement, will result in more reliable data than a development program that did not standardize procedures. The review should therefore include a description of the measurement methodology.

Second, growth is not constant throughout childhood and varies by age and sex. Without consideration of these factors at baseline, absolute mean changes in weight and height can give misleading results. Adjustment of growth for age and sex can be done by conversion of a child's height and weight to a z-score, which is the number of standard deviations that an individual's measurement is from the mean for age and sex matched children in the general population. A decrease in mean z-score for a group is interpreted as evidence of a lag in growth compared to what would be predicted using general population data. In a controlled trial, differences in mean z-score changes from baseline between treatment groups may provide evidence of an effect of drug on growth. Declines in mean z-scores in open label studies are less easily interpreted because these could result from the effect of drug or could be caused by the disease for which the treatment is being studied.

Sponsors should provide analyses of height and weight data that assess measures of central tendency and outlier analyses using height and weight z-scores. Although results from these analyses will not provide definitive proof of drug related effects on growth in most cases, they may help identify candidates for prospective studies of the effect of drug on growth in children. The review team should request such analyses at the pre-NDA meeting.

7.1.16 Overdose Experience

The review should summarize all overdose experience with a drug in humans (including both information provided by the applicant and information obtained from secondary sources) and describe the constellation of signs, symptoms, and other abnormalities one might expect to see in association with overdose. Phase 1 data should be reviewed to identify subjects who may have received higher doses than those used in later phases of study. In addition, patients with certain

physiological differences that would compromise their ability to clear a drug (e.g., renal impairment, hepatic impairment, limited CYP4502D6 activity for a drug cleared by this isozyme) may provide data relevant to the clinical implications of overdose.

7.1.17 Postmarketing Experience

Relevant findings from postmarketing experience, if any, should be described briefly here and referenced in the summary section (Section 7.3).

7.2 Adequacy of Patient Exposure and Safety Assessments

Section 7.1 is an assessment of the adverse events seen during the development program. Section 7.2 should provide the reviewer's comments on the adequacy of drug exposure and the safety evaluations performed as part of the development program. This section addresses the regulatory question of whether or not *all tests reasonably applicable* were conducted to assess the safety of the new drug. Was there adequate experience with the drug in terms of overall numbers of patients and in appropriate demographic subsets of patients? Were doses and durations of exposure appropriate? Were all (or not all) appropriate tests performed in the exposed patients? Were all necessary and appropriate animal tests performed? Were all the appropriate clinical tests carried out (e.g., electrocardiographic assessment of effects on QT interval)? Was the drug adequately worked up metabolically? Were appropriate in vitro studies of drug-drug interaction carried out according to current guidelines? Were all potentially important findings adequately explored: for example, to what extent was psychomotor impairment specifically assessed in a drug that is sedating?

If important data are missing, this could influence the regulatory action on the drug. A critical task of the reviewer in this section is identification of specific concerns that need to be addressed by the applicant, either before approval or postapproval. Even more than for most other parts of the review, the reviewer needs to be conscious of recent developments and discuss issues broadly. Finally, this section is the place for detailed comments on the quality and completeness of the data provided.

The review should clearly describe the studies and overall extent of the data supporting the evaluation of safety. The reviewer should then make a judgment about the adequacy of the clinical experience with the new drug for assessing safety.

7.2.1 Description of Primary Clinical Data Sources (Populations Exposed and Extent of Exposure) Used to Evaluate Safety

In this section the reviewer should identify and characterize the primary safety data sources used in conducting the review. If these are described elsewhere in the review, this section can reference those sections. The primary source is generally the database derived from the applicant's development program. Studies in this program will generally have full study reports related to safety, or studies that are grouped for analysis of safety in an Integrated Summary of Safety; case report forms will be available. These studies usually will have been closely monitored. Secondary sources may also be available and may be of critical importance (e.g., for

a drug already available in other countries), and there may be some parts of the database that have had limited analyses (i.e., only for deaths and adverse dropouts); these are described in section 7.2.2.

Tables and graphs are useful in describing the data sources for the safety review. Generally, the reviewer should use the tables and graphs in this section to characterize the overall database. The detailed tables and other displays for this subsection may be included in an appendix to the review, but summary tables and narrative statements should be included here. The reviewer should also characterize the per patient data (narratives, CRFs, CRTs and electronically accessible databases for baseline information. See section 7.4 for discussion of ability to link databases.

7.2.1.1 Study Type and Design/Patient Enumeration

The reviewer should include in an appendix a table, such as that illustrated in Table 7.2.1.1.1, enumerating all subjects and patients across the entire development program, phases 1 to 3. This is a critical table that identifies the important patient pools and denominators for subsequent analyses and incidence estimates.

The reviewer should also include an appendix table that provides brief descriptive information for all individual studies, including study design (fixed dose vs. flexible dose, parallel vs. crossover), dosing schedule, study location (foreign vs. domestic), treatment groups and doses, sample sizes, patient population (elderly). Studies that were designed to assess a particular aspect of safety (e.g., ECG, ophthalmic) should be noted. Most NDAs/BLAs will include a table of all studies, as such a table is called for in the Clinical/Statistical guidline *and in the Common Technical Document.*

Applicants sometimes segregate certain clinical trials from their primary source data (see 7.2.2, secondary source data), especially foreign data. This may be appropriate, especially if there is a basis for believing that these data differ substantially in quality and/or completeness or in critical aspects of investigator practice from the data included in the primary source database. This is a matter of judgment, however, and cannot be assumed to be valid. An explanation should be provided in the review describing the basis for decisions about what data were included and what excluded from the primary source data.

An NDA/BLA generally includes data from patient samples that are at different levels of completeness in terms of data entry, information collected, and validation. Table 7.2.1.1.1 should include patient counts (or estimates) from all studies contributing data, regardless of these factors. Data cutoff dates or database *lock dates* for the various databases comprising the NDA/BLA should be identified at this point in the review. For example, the cutoff date for the overall safety database derived from completed studies might be more distant, while the cutoff date for submitting serious adverse events from all studies may generally be more recent. These dates may likely need updating during the course of NDA/BLA review as more data become available.

7.2.1.2 Demographics

The reviewer should include appendix tables in a format similar to that illustrated in Table 7.2.1.2.1 (showing percent distribution within treatments of patients by age, gender, and race as well as weight in various groups), providing overall demographic information for phase 1 and phase 2 to 3 study pools separately. It may be appropriate to provide demographic displays for subsets within these larger pools at other points in the review.

7.2.1.3 Extent of Exposure (Dose/Duration)

There are many ways to summarize the dose and duration experience with a new drug. Either can be expressed as mean, median, maximum, with histograms or other displays that give the numbers exposed at various doses or for various durations. A particularly useful approach is to provide combined dose and duration information. It is suggested that the review contain tables in the format illustrated in Table 7.2.1.3, enumerating patients on the basis of mean daily dose of the NDA/BLA drug and duration of administration for phase 1 and phase 2 to 3 study pools separately. If the study used a titration design, the modal dose (if 2 different doses were used for the same duration, the larger, or maximal modal dose) may be the more useful summary statistic. It is particularly important to examine the subgroup of patients who received a dose at least as large as the dose intended for marketing.

It may also be useful to provide similar tables based on maximum dose, modal dose, dose expressed as mg/kg or mg/m^2 or even plasma concentrations, if such data are available.

It may also be useful to provide similar tables for various subgroups (e.g., males and females separately, various age groups separately, and patients with various comorbid illnesses of interest separately). There should be similar displays for active control drugs if any were included in trials for the new drug.

Finally, it may be useful for the review to include an appendix table providing total person time exposure data for the NDA/BLA drug, active control, and placebo, for the phase 2 to 3 database.

7.2.2 Description of Secondary Clinical Data Sources Used to Evaluate Safety

Secondary source data are (1) data derived from studies not conducted under the applicant's IND and for which CRFs and full study reports are not available,[21] or studies so poorly conducted (e.g., poor ascertainment for adverse events), that they cannot be reasonably included in the primary source database, (2) postmarketing data, and (3) literature reports on studies not conducted under the IND. Often the applicant may have made the distinction between the data considered primary source data and other data, and the reviewer needs to examine the rationale for this distinction.

[21]If CRFs are available from any such studies and the data quality is comparable to that of data from studies conducted under the applicant's IND, these data would ordinarily be included in the primary source database.

The secondary sources should be briefly described. It is worth emphasizing that secondary source data may be a critical source of information for review, despite the generally lower quality of these data, because they often provide the larger database needed to look for less common serious adverse events and may be reliable with respect to deaths and serious adverse events.

7.2.2.1 Other Studies

The NDA/BLA should be clear in describing exactly what other studies provided data and what the basis was for not integrating such data with the primary source data (e.g., no CRFs, no study reports, not adequately monitored). Lack of clarity in this should be noted by the reviewer.

7.2.2.2 Postmarketing Experience

If postmarketing data are available, this section should describe briefly the type of information available for review. An example of such a description would be a comment that a line listing for (a specified number of) spontaneous reports from marketing in (country) was provided, along with narrative summaries for the serious adverse events among the reports and an estimate of product use in (country) during that time period. As is the case for most spontaneous reports, these reports are likely to be difficult to interpret. Important events will be described in appropriate sections (e.g., 7.1.1 and 7.1.2, Deaths and Other Serious Events).

7.2.2.3 Literature

Relevant literature may be incorporated in various sections of the NDA, but is ordinarily included in section 5.4 of the Common Technical Document (*M4: The CTD – Efficacy*). The NDA/BLA may include a separate literature section or the literature may be provided or referenced as called for in various places in the Clinical/Statistical guidline in section II F, Other Studies and Information. The applicant should have provided a description of the search strategy to assess the world literature (e.g., databases used, key search words), the personnel who carried it out (their credentials) and whether the search relied on abstracts or full texts (including translations) of articles. A cutoff date for the literature search should also have been provided. A copy (translated as required) should have been submitted for any report or finding judged by the applicant to be potentially important.

This section of the review should describe what information from the literature search was provided for review, the extent to which the above description of an ideal presentation was met, and whether any missing information is important (and/or was obtained by the reviewer). Independent literature reviews conducted by the reviewer should be described here as well.

Actual safety findings should be described in appropriate sections of the safety review to present from the literature reports in this section of the review.

7.2.3 Adequacy of Overall Clinical Experience

In evaluating the adequacy of clinical experience with the drug, the reviewer should refer to current ICH guidance on extent and duration of exposure needed to assess safety[22] as well as the draft guidance on *Pre-Marketing Risk Assessment.*[23] The review should specifically address the following:

- Whether an adequate number of subjects were exposed to the drug, including adequate numbers of various demographic subsets and people with pertinent risk factors
- Whether doses and durations of exposure were adequate to assess safety for the intended use
- Whether the design of studies (open, active-control, placebo-control) was adequate to answer critical questions
- Whether potential class effects were evaluated (e.g., for anti-arrhythmic effects, evaluation of the potential for pro-arrhythmic effects) and whether problems suggested by pre-clinical data were assessed
- Whether patients excluded from the study limit the relevance of safety assessments (e.g., diabetics, people over 75, people with recent myocardial infarction, people with renal or hepatic functional impairment, or people on other therapy). This may depend on the signals of toxicity that were observed in the patients who were studied.

7.2.4 Adequacy of Special Animal and/or In Vitro Testing

The clinical reviewer should not attempt a general assessment of the preclinical program, but rather, comment on whether preclinical testing was adequate to explore certain potential adverse reactions, using preclinical models based either on a drug's pharmacology or on clinical findings that emerged early in clinical development. For example, for a drug anticipated to cause QT prolongation because of its drug class or because QT prolongation was seen in phase 1 studies, there are in vitro models to evaluate this potential. The reviewer should note whether such studies were done. If such studies were performed, the results would be summarized in the Pharmacology Review.

7.2.5 Adequacy of Routine Clinical Testing

The reviewer should comment on the adequacy of routine clinical testing of study subjects, including efforts to elicit adverse event data and monitor laboratory parameters, vital signs, and ECGs. In assessing the adequacy of clinical testing, the reviewer should consider the adequacy of the methods and tests used and the frequency of testing. The adequacy of specific testing intended to assess certain expected or observed reactions should be discussed under subsection 7.2.7.

---

[22] ICH *E1A The Extent of Population Exposure to Assess Clinical Safety: For Drugs Intended for Long-Term Treatment of Non-Life-Threatening Conditions* (http://www.fda.gov/cder/guidance/iche1a.pdf) recognizes possible differences in expected exposure (e.g., more patient exposure would be expected for drugs with small effects, or drugs that are used prophylactically in well populations, where only a small fraction of patients will benefit).

[23] A draft guidance *Premarketing Risk Assessment* was issued in May 2004. Once finalized, it will represent the Agency's thinking on this topic.

The reviewer should be alert to the absence of data in an NDA laboratory database for analytes that are typically included in routine laboratory monitoring. For example, it was discovered after approval that the NDA laboratory database for the anti-epileptic drug zonisamide did not have data on serum bicarbonate. It was later determined that this drug is associated with a non-anion gap metabolic acidosis. The serum bicarbonate data would have been helpful in identifying this adverse reaction earlier.

7.2.6 Adequacy of Metabolic, Clearance, and Interaction Workup

Knowledge of how a drug is metabolized and excreted is critical to anticipating safety problems in patients with impaired excretory or metabolic function and problems resulting from drug-drug interactions.

Drug-drug interaction assessment is a critical part of a modern drug development program and should evaluate the drug both as a substrate for interactions (interference with its clearance) and as an inducer or inhibitor of the clearance of other drugs. The reviewer should comment on the adequacy of in vitro and in vivo testing carried out by the applicant to identify the following:

- The enzymatic pathways responsible for clearance of the drug and the effects of inhibition of those pathways, notably CYP450 enzymes and p-glycoproteins
- The effect of the drug on CYP450 enzymes (inhibition, induction) and the effects of the drug on the PK of model compounds
- The major potential safety consequences of drug-drug interactions

Details of these assessments will be found in the Clinical Pharmacology Review and in the summary of that evaluation in the Medical Officer's Review.

7.2.7 Adequacy of Evaluation for Potential Adverse Reactions for Any New Drug and Particularly for Drugs in the Class Represented by New Drug; Recommendations for Further Study

The reviewer should discuss the adequacy of the applicant's efforts to detect specific adverse reactions that are potentially problematic and might be expected with a drug of any class (e.g., QT prolongation or hepatotoxicity) or that are predicted on the basis of the drug class (e.g., sexual dysfunction with SSRI antidepressants). The reviewer should also discuss whether the applicant should have made efforts to assess certain events that it did not assess. The reviewer should also discuss pertinent negative findings (absence of findings) for a drug in this section of the review (see examples below).

The adverse events that warrant specific attention will vary depending on the characteristics of the drug and the drug class. The known pharmacology of the drug would suggest some evaluations (e.g., first dose effects for peripheral alpha blockers, tolerance and withdrawal effects for central alpha agonists, urinary retention with anti-cholinergics, QT prolongation with type III anti-arrhythmics, extrapyramidal effects with antipsychotics, muscle pain with statins), while experience with other members of the class would suggest others (e.g., hepatotoxicity with thiazolidinedione PPAR gamma agonists (glitizones), tendon problems with fluoroquinolones).

There should be a subheading for each adverse reaction that warrants special consideration (even if not observed) and, under each subheading, a discussion of what was done to detect the reaction and the adequacy of the approach. The following list of potential adverse reactions, and some of the drug and therapeutic classes that might trigger higher interest in them, may be a useful starting point in assembling a list (it is also important to examine labeling for other members of the drug's pharmacologic class):

Hepatotoxicity (NSAIDs, thiazolidinedione PPAR gamma agonists)
Pancreatic toxicity
QT prolongation (any antiarrhythmic, antipsychotic, antihistamine, fluoroquinolone)
Vasodilator effects, such as hypotension (alpha blockers) or edema (dihydropyridine calcium channel blockers)
Withdrawal effects (beta blockers, central alpha agonists, SSRIs, narcotics)
Orthostatic hypotension (any antihypertensive, antipsychotics)
Hypertension (any sympathomimetic or phosphodiesterase inhibitor)
Tachycardia
Neutropenia (drugs related to ticlopidine, procainamide, clozapine)
Bleeding (drugs inhibiting any aspect of clotting or platelet function, NSAIDs)
Aplastic anemia
Increased coagulation times
Muscle injury (any HMG CoA Reductase Inhibitor (statin) or other lipid-lowering drug)
Sedation (any psychotropic drug)
CNS stimulation
Anticholinergic activity
Allergic reactions
Sexual dysfunction (any antidepressant, sedating drug)
Elevated intraocular pressure
Cataracts
Retinopathy
Worsening glucose tolerance/diabetes (diuretics, atypical antipsychotics)
Pro-arrhythmic effects and increased mortality (most nonbeta blocker anti-arrhythmics)
Increased CHF and SD mortality (any inotrope, some negative inotropes such as calcium channel blockers)
Nephropathy (NSAIDs)

Example 1: If orthostatic hypotension was an expected adverse reaction, but was not observed, the reviewer should determine whether the applicant made efforts to detect it and, if so, whether the applicant's approach (e.g., timing and frequency of vital signs testing) was adequate to detect it.

Example 2: If QT prolongation was observed in phase 1 studies, the reviewer should ascertain whether the applicant made efforts, beyond routine ECG testing, in phases 2 and 3 to explore the consequences in patients of the observed QT prolongation and, if so, whether those efforts were adequate, including adequate exposure to higher doses. For example, how did the applicant follow-up patients who experienced clinical events that may be manifestations of torsade de

pointes (e.g., syncope, dizziness, or palpitations)? Holter monitoring, for example, might have been appropriate in such patients.

7.2.8 Assessment of Quality and Completeness of Data

The reviewer should provide general overall assessments of the quality and completeness of the data available for conducting the safety review and describe the bases for these assessments. More than that, attention to completeness and quality of assessment is important throughout the review. The reviewer should recognize that quality may *differ from* the primary source data and for data over which the applicant had less control. The following examples illustrate some of the ways in which applicants can differ in the quality and completeness of data they provide:

- Applicants may differ in what they include in a CRF. For example, if additional laboratory data are collected at unscheduled visits or after the normal end of a trial, some do not include these data. Such data may be stored in some other place (a *correspondence file*). Sometimes additional information is attached to the front of the CRF as *queries*. If CRFs do not indicate any additional testing beyond the routine assessments, the reviewer should ascertain whether additional testing was done to reassess abnormal values before the next routine visit (e.g., at an unscheduled visit). If the CRFs do not indicate that additional testing was performed, the reviewer should ask the applicant if additional laboratory data are available.

- If it is apparent that the CRF contains insufficient information about an adverse event (e.g., if a patient was hospitalized for an adverse event), the reviewer needs to determine whether there is additional information available. Such an observation also raises the general question of whether all pertinent data have been included in CRFs.

- The reviewer should be concerned about patients with abnormal clinical or laboratory findings who are lost to follow up, particularly if there are significant numbers of such patients. In these situations, the reviewer may consider asking the applicant to attempt to obtain the needed follow up information. If the information cannot be obtained, it may be appropriate to perform sensitivity analyses to assess the possible impact of missing data, assuming a worst-case outcome.

- The reviewer should be particularly alert to situations in which applicants make changes in CRFs to reclassify adverse events or reasons for subjects dropping out without the investigator's agreement. There is greater concern where serious adverse events are reclassified and reclassifications are done without blinding. The reviewer should ask the applicants about procedures used (if unclear) and attempt to assess the impact of multiple changes on the safety evaluation.

- For electronic data, the reviewer should clarify what information is, and is not, included in the electronic files. For example, if a reviewer is relying on electronic files from the case report forms, it is important to know what, if any, information from the CRFs was not included. A separate file may be needed for any missing data.

7.2.9 Additional Submissions, Including Safety Update

The initial NDA/BLA submission may not contain all information pertinent to the safety evaluation. Further data submissions may be planned at the time of initial submission and filing (e.g., results of additional long-term follow-up), may represent responses to specific questions or discipline review letters, or may be part of the safety update required under regulations (21 CFR 314.50(d)(5)(vi)(b)). It is critical to review these data to determine whether safety conclusions are affected, particularly with respect to serious or fatal events.

This section should:

- Describe safety submissions, noting whether the results have been incorporated into the rest of the review or are considered in this section
- For those safety matters not incorporated into the rest of the review, discuss any data with important implications for safety. In general, this will involve deaths, adverse dropouts and other serious events, and these should be considered as in sections 7.1.1, 7.1.2, 7.1.3, as appropriate to the (usually) small numbers. Only if these events alter the overall safety picture will a more detailed discussion of the entire area (e.g., deaths, liver injury) be needed.

Any reports of important changes in foreign labeling or new studies that give insight into more common events should also be noted.

7.3 Summary of Selected Adverse Reactions, Important Limitations of Data, and Conclusions

This section of the review should briefly summarize each of the adverse reactions that the reviewer considers important and drug-related (i.e., this should constitute a *problem list* for the drug). For each adverse reaction, there should be a separate subheading followed by a brief summary of the reaction and references to sections of the review (e.g., other parts of the safety section, Clinical Pharmacology, studies described in the Efficacy section) containing more detailed information about the adverse reaction generally, or specific aspects of the reaction. The review should integrate by reference all relevant details about the reaction, including patient identifying numbers for certain patients (e.g., for deaths). Below is a sample summary section entry for QT prolongation:

QT Prolongation

Dose-related QT prolongation compared to control was seen in all controlled trials, with a mean change of 20 msec at 100 mg/day (peak), the recommended maximum dose, and smaller changes at lower doses; 5 percent of patients had QTc values over 500 msec at some point, compared with ___ percent on placebo. The drug's metabolism is predominately via CYP4503A4, so that moderate inhibitors of this enzyme could lead to greater QTc prolongation. The QTc effects of doses greater than 100 mg have not been studied.

- See Section 7.1.1 (Deaths) at page __ for discussion of deaths that may be related to QT prolongation and detailed discussion of the finding
- See Section 7.1.9 (ECGs) at page __ for discussion of ECG changes
- See Section 7.1.10 (Special Studies) at page __ for dose response study of QT prolongation (doses of 10, 40 and 100 mg)
- See Section 7.2.4 (Metabolic and Interaction Workup) at page __ for discussion of the adequacy of the applicant's in vivo and in vitro assessments of the metabolism of (Drug) and potential relation of drug-drug interactions to QT prolongation
- See Section 3.2 (Animal Pharmacology/Toxicology) at page __ for discussion of the animal models used to evaluate effects on K channels, and QT prolongation
- See Section 7.2.1.1.2.3 (Literature) at page __ for published articles about similar products and methodological suggestions.
- See Section 7.1.13 (Overdose) at page __.
- Patient ID numbers for possibly relevant deaths: _____, _____, _____.

As the QT prolongation example shows, it is useful to identify the various sections of the clinical review that can be referenced for additional details about an identified adverse event. If the review is converted into a PDF file, bookmarking can be used to electronically link the text in the problem list to earlier sections of the review.

In addition, in this section the reviewer should provide summary recommendations for further studies, with a reference to section 7.2 for more details.

The review should also provide overall conclusions about the safety of the drug, including:

- Overall assessment of the available safety information, referring both to what it has shown and its adequacy
- The limitations of the available data
- Additional information needed, including both further analyses and additional studies.

- Comparison, to the extent possible, of the safety of the drug under review to the safety of other available products, and the basis for that comparison (direct comparative data vs. clinical opinion)
- Whether a risk management program (beyond labeling) is needed and why
- Analysis of likely uses beyond labeling, (e.g., in more severe patients, in other diseases, in children)
- Whether there is a need for postmarketing safety studies

7.4 General Methodology

This section of the guidance describes analytical methods that have general application to the safety review and provides a location in the review for any general discussion of methodological issues not discussed elsewhere, organized by the subsections listed here, with additional sections as needed. It is important to consider early in the review whether the available patient level data will allow the analyses the reviewer intends. For example, in examining whether particular baseline risk factors are related to an adverse event, the reviewer will either need to extract the baseline characteristics from case report tabulations or be sure the information is available in a retrievable form. Similarly, it may be important to link individual safety observations with other on therapy data, such as dose, duration of treatment, concomitant therapy, other adverse events, lab data or effectiveness results (it is obviously best if such issues are considered at pre-NDA/BLA meetings).

7.4.1 Pooling Data Across Studies to Estimate and Compare Incidence

7.4.1.1 Pooled Data vs. Individual Study Data

Before estimating the incidence of adverse events, the reviewer must select the patient sample of interest. Pooling data from different studies can improve the precision of an incidence estimate (i.e., narrow the confidence intervals by enlarging the sample size). Better precision is particularly important for lower frequency events, which can be difficult to detect and may not occur in some studies. Pooling can also provide the larger database that will permit explorations of possible drug-demographic or drug-disease interactions in subgroups of the population. Pooling can also, however, obscure real potentially meaningful differences between studies. The review should explain why any pooling used in the review was chosen. When making decisions about pooling, the reviewer should consider the following:

- It is most appropriate to combine data from studies that are of similar design, that is, similar in dose, duration, choice of control, methods of ascertainment, and population (checklist vs. general inquiries vs. no prompt at all; in psychiatric drug trials it is typical for obsessive compulsive patients to spontaneously report adverse events more frequently than schizophrenic patients. It is also possible that different populations may have different vulnerabilities to a drug, and therefore, different risk profiles.) When the studies are similar in design but differ in duration, it may be critical to account for exposure duration and to look for time-dependent events.

- Even when the pooled analysis is the primary one, it is important to explore the range of incidences across the studies being pooled. For a specific adverse event, if the incidence differs substantially across the individual studies in a pool, the pooled value should not be used, as it is probably not meaningful and, in some cases, could obscure important information about predictors for that event. (In one case, for example, several studies were combined and a reassuringly low estimate of phototoxicity was obtained. Subsequent examination of individual study results found one study with a substantial rate of phototoxicity. The study was the only outpatient study done (i.e., the only one in which patients had an opportunity to be exposed to sunlight).) In some situations, the incidence may be best described by the range in the various studies. For the photoxicity example above, however, the most relevant data are those from the outpatient study, the only study that was conducted under conditions pertinent to intended use.

- In some cases, observed differences in rates in various studies can be explained (e.g., better ascertainment, different populations), so that a consistent rate can be determined from a subset of studies.

- Formal tests for extreme values may be useful to assess appropriateness of assay pooled data (e.g., test of heterogeneity such as the Breslow-Day Chi-Square test could be used). Alternatively, the reviewer might use a more subjective approach, such as determining if the direction of the difference is always the same across studies, or use a graphic display of incidence by study to informally consider the extent of variability and to identify outliers; outliers may be important in identifying subgroups of patients who are at particular risk for certain adverse reactions.

7.4.1.2 Combining Data

In pooling data, usually the numerator events and denominators for the selected studies are simply combined. Other more formal weighting methods can be used (e.g., weighting studies on the basis of study size or inversely to their variance). The review should describe how the pooling was performed, as well as the rationale for selection of the method used.

7.4.2 Explorations for Predictive Factors

Adverse reaction rates may differ considerably from one patient population to another and may change over time. Factors that may affect the safety profile of a drug should be explored during the review. Explorations for common predictive factors, such as dose, plasma level, duration of treatment and concomitant medications, and patient-predictive factors such as age, sex, race, concomitant illnesses, are considered below. In general, these explorations are meaningful only for adverse events that appear to be drug-related (see Section 7.4.3).

7.4.2.1 Explorations for Dose Dependency for Adverse Findings

If data from randomized, parallel, fixed-dose studies (or data from studies in which patients were randomized to fixed dose ranges), are available, they should be analyzed for evidence of dose dependency for any adverse reactions. If plasma concentration data are available, it may be

useful to explore plasma concentration effect relationships as well. It may also be useful to reconfigure dose as mg/kg or mg/m^2, to decrease the effect of size or weight differences on drug exposure. Dose-response relationships should also be examined in demographic subgroups (e.g., females, blacks, elderly patients). Dose-dependency analyses are usually performed by simple inspection of incidence rates across different doses or different weight or body surface area-adjusted doses. Formal statistical testing can also be used. If formal statistical tests are performed for a study that includes placebo control as well as different doses, and a drug-placebo difference is apparent, it may be desirable to focus on between-dose group differences.

Flexible Dose Titration Studies

Although it is tempting to try to extract dose-response or plasma level-response data from flexible dose (titration) studies, and the ICH dose-response guideline[24] encourages this, there are many potential problems with such analyses. In particular, many adverse reactions show considerable time dependency, some occurring early, some late. It is easy to confound dose (or plasma concentration) with duration when dose is increased over time. In some cases, such as anticancer drugs or drugs that are known to produce anti-cholinergic or sedating reactions, the drug is dose-adjusted to toxicity, which will often obscure any dose-response relationship. In addition, if dose is increased only in patients without adverse effects (i.e., subjects who are resistant to them), the higher doses will be associated with lower adverse effect rates. On the other hand, if dose is titrated to clinical effect, and adverse reactions occur late (so that they do not affect the dose given), analysis of the rate with respect to dose may be useful. For example, erythropoietin, used to treat anemia in patients with chronic renal failure or cancer, is titrated to maintain hemoglobin within a specific range. Given the delayed therapeutic response (erythropoesis), analysis of adverse events by dose or cumulative dose prior to a reaction can give insight into dose-related toxicity.

Cumulative Dose Dependency

For certain adverse reactions, it may be possible to demonstrate a relationship between cumulative dose and the occurrence of the reaction (e.g., liver fibrosis and cirrhosis with methotrexate, cardiotoxicity with doxorubicin, renal toxicity with Amphotericin B). For drugs that are used chronically, the reviewer should consider the possibility that cumulative dose may predict toxicity and discuss this in the review.

7.4.2.2 Explorations of Time-Dependency for Adverse Findings

The reviewer should explore time dependency of adverse reactions in two ways — time to onset of the finding and duration of the finding:

Time of Onset

[24] See ICH *E4Dose-Response Information to Support Drug Registration* (http://www.fda.gov/cder/guidance/iche4.pdf).

Although most adverse reactions occur early in treatment and may be best characterized by a crude incidence rate (number with the reaction divided by number exposed), others may occur only after some delay of weeks, months, or longer. A crude incidence rate, based on a patient population exposed predominantly for short periods, will understate the importance of such adverse reactions for chronically used drugs. For important adverse reactions that occur later in treatment, there should be explorations of the time dependency of the reaction. Possible methods include:

- A life table (Kaplan-Meier graph) describing risk as a function of duration of exposure (i.e., cumulative incidence)

- Plotting risk for discrete time intervals over the observation period (i.e., a hazard rate curve) reveals how risk changes over time.

- Adjusting for duration of exposure by expressing the adverse reaction rate in terms of person-time (person-time is duration of exposure summed across all patients, e.g., 2 patients each exposed for 6 months = 1 patient-exposure-year). This approach is useful only when one can safely assume that the hazard rate is constant over time.

Duration of Adverse Event

Certain adverse events that occur at initiation of treatment may ***appear*** to diminish in frequency with continued use. Possible explanations for this phenomenon include adaptation or tolerance, decreased reporting of the event even by patients though it is still occurring at the same rate, and reduced dose or dropping out in patients with the event. For drugs used chronically and for which there was an adverse event that seemed to diminish in frequency over time, it may be useful to characterize and quantify the change. It would be important, for adverse events of interest, to determine whether the decreased rate simply reflected discontinuation by affected patients or real adaptation. One way to make this distinction is to identify a cohort that experienced an event of interest during a specified period of a trial, but nonetheless completed the trial, and observe the rate of the event in that cohort over time. This cohort of survivors could be compared to a similar cohort of placebo recipients who experienced the same event at baseline. The same approach could be used for adverse events occurring later in treatment. It is usually sufficient to do such analyses for those adverse reactions that are relatively common and likely to be drug related (see Section 7.1.5.4 for methods to identify drug-related events).

7.4.2.3 Explorations for Drug-Demographic Interactions

Numerous methods can be used to analyze age, gender, and race implications for safety, and applicants must present analyses of safety information for these population subsets. In most cases, there will be pharmacokinetic information available for some or all of these subsets, which may help in interpreting adverse event rates. In some cases, it may be useful to construct subgroups based on more than one factor. For example, bleeding is the principal risk associated with use of thrombolytic agents in patients with acute myocardial infarction. Women tend to have more bleeding than men, and risk is inversely related to weight. Thus, an analysis by

gender weight subgroups can identify the group at greatest risk of bleeding (thin women). It may also be useful to consider age-gender or race-gender subgroups. Formal analysis should be limited to events considered common (e.g., occurring at an incidence of at least 2 percent) and that occur at a clearly greater rate on drug than placebo. In small studies or for low frequency events, there will usually not be sufficient power to detect differences between groups, so that these analyses will usually be based on pooled data. In general, these analyses are descriptive, comparing risk of an event in one subset with the risk in another (men vs. women, old vs. young, black vs. white); as these comparisons obviously do not reflect randomization to the subset (baseline characteristic) of interest, formal statistical comparisons are usually not warranted. For these descriptive comparisons, two approaches deserve consideration; when the control rates of adverse events differ for population subsets these approaches can provide quite different results: (1) evaluation of relative risk (RR) (cumulative risk on drug/cumulative risk on comparison drug or placebo) and (2) evaluation of attributable risk (AR) (cumulative risk on drug - cumulative risk on comparison drug or placebo).

When background event rates differ by demographic subgroup, relative risk analysis will provide a quantitative estimate of the difference in effect of the drug, but the attributable risk may be a better estimate of the importance of the risk in the subsets. To illustrate, consider a comparison of drug-induced nausea for males versus females. Suppose the rate of nausea on placebo is 1 percent for men and 3 percent for women and that on drug it is 3 percent for men and 9 percent for women. The risk ratios (RR) for both sexes are 3 and the relative risk for men and women (RRf/RRm) is one (no difference), yet the attributable risk is much greater for women than men (6 percent vs. 2 percent), a finding of possible importance in treatment. Such a difference has been observed for several adverse reactions of amlodipine, a calcium channel blocker, and is described in labeling as a gender difference, even though the RR's are the same.

7.4.2.4 Explorations for Drug-Disease Interactions

The reviewer should be alert to the possibility that co-morbidity will affect the adverse reaction profile of the drug (i.e., a drug-disease interaction). Such interactions can arise from abnormalities of excretory function (renal or hepatic disease), and typically, the applicant will have carried out formal pharmacokinetic studies in patients with hepatic and renal disease to indicate the potential for such reactions. The reviewer needs to consider, in that case, whether PK differences are manifested as differences in adverse reaction rates. Apart from differences in adverse reaction rates related to PK differences, differences in rates can also reflect true differences in susceptibility to adverse reactions (i.e., real pharmacodynamic differences). In general, the same methods described for exploring drug-demographic interactions can be applied here.

7.4.2.5 Explorations for Drug-Drug Interactions

The clinical reviewer should be alert to the potential of drug-drug interactions to affect the safety profile of the drug. Again, these interactions could be either pharmacokinetic (affecting elimination of the drug) or pharmacodynamic, in either case leading to observed differences in adverse reaction rates for the subgroups receiving or not receiving co-administered drugs. Typically, there will be formal interaction studies to evaluate potential pharmacokinetic effects

of concomitant therapy on drugs metabolized by CYP450 enzymes, but PK interactions can also occur through effects on renal excretion and transport (P-glycoprotein) proteins. True pharmacodynamic interactions are less frequently recognized but can be important (e.g., marked hypotension when sildenafil is given with organic nitrates). In general, the same methods described for exploring drug demographic interactions can be applied here.

7.4.3 Causality Determination

In assessing the critical question of whether an adverse event is caused by a drug, whether the drug is capable of causing that adverse event in the population is usually of greater interest than whether the drug caused the event in each patient who reported the event, but the approach to causality is distinctly different for relatively common events and relatively rare, serious events.

Common Events

Where events are common and occur in multiple patients in controlled trials, it is usually not necessary or helpful to consider each case individually. Rather, all reported cases can be considered potentially drug-related, and causality is assessed by comparing the rates of reports in patients treated with test drug and in control groups. If an event is clearly more frequent with test drug than the control, it can be attributed to treatment with the test drug.

Uncommon, Serious Events

Causality judgments are much more difficult for uncommon (e.g., < 1/1000) serious events where there are, in most cases, no useful comparisons to control groups. The reviewer therefore must form a judgment as to the plausibility of drug-relatedness for the individual cases.

- The following questions should be considered:

  1. Was the patient in fact exposed to drug and did the adverse event occur after drug exposure?
  2. Did the patient have a clinical experience that meets the criteria for the adverse event of interest? (Establishing a standard case definition may be helpful here.)
  3. Is there a reasonably compelling alternative explanation for the event? (For example, recent benzene exposure for a case of aplastic anemia; the event is a well-recognized consequence of the patient's underlying illness.)
  4. Is the adverse event of a type commonly associated with drug exposure, such as hematologic, hepatic, renal, dermatologic or pro-arrhythmic events? (But also see below caution about discarding events that do not seem plausibly drug related.)

- After assessing individual cases to identify events that could be drug-related and for which there are no compelling alternative explanations, the reviewer should compare the observed rate of occurrence of the event in the database with a best estimate about the background rate for the event for the population being studied. For an event like aplastic anemia, with a background rate of perhaps 1 per million person years, finding even one case suggests a causal relationship. For events that occur more frequently in the absence of drug therapy

(e.g., MI, stroke, sudden death, seizure, which could occur at rates of 0.1 to 1 percent, depending on the population), the finding of one or two cases may be very difficult to interpret in the absence of a substantial controlled trial database.

- The reviewer should also evaluate any other information about the drug that bears on causality including:

  1. Whether the drug is a member of a class of drugs known to be causally associated with the event of interest
  2. Presence of other adverse events in the database that may be associated with the event of interest (e.g., a general finding of drug associated transaminitis or animal findings suggestive of hepatotoxicity would substantially strengthen the signal generated by the finding of a single case of hepatic failure)
  3. Positive re-challenge with the drug (although it would be unusual to deliberately re-challenge for a serious event, there may occasionally be inadvertent re-exposures that are informative)

- Caution Concerning Relative Plausibility of Uncommon, Serious Events

The reviewer should be cautious about dismissing uncommon, serious events *that don't seem plausibly drug-related* and should consider differences in common less serious adverse reactions that might predict the uncommon serious reactions with longer use. There are numerous examples of uncommon, serious adverse reactions that are uniquely associated with a drug or drug class:

- tendon rupture associated with the quinolone antibiotics
- heart valve lesions associated with fenfluramine
- practolol syndrome
- retroperitoneal fibrosis with Sansert
- pulmonary hypertension with Aminorex (a European weight loss drug), and various other drugs
- suicidal ideation with interferons, Accutane
- intussusception with rotovirus vaccine
- pulmonary fibrosis with amiodarone

List of Tables

TABLE 7.0.1 **INDEX FOR LINKING IDENTIFIED PATIENTS WITH SUPPLEMENTARY PATIENT INFORMATION IN THE NDA (CRFS, NARRATIVE SUMMARIES, AND PATIENT DATA LISINGS)**..... ERROR! BOOKMARK NOT DEFINED.

TABLE 7.1.1.1 **DEATHS LISTING**..... ERROR! BOOKMARK NOT DEFINED.

TABLE 7.1.1.2 **MORTALITY BY TREATMENT GROUP FOR POOL OF PHASE 2-3 STUDIES WITH NEW DRUG**.. ERROR! BOOKMARK NOT DEFINED.

TABLE 7.1.2.1 **SERIOUS ADVERSE EVENT LISTING**..... ERROR! BOOKMARK NOT DEFINED.

TABLE 7.1.3.1.1 **DROPOUT PROFILE: INCIDENCE OF DROPOUT BY TREATMENT GROUP AND REASON FOR PHASE 2 TO 3 STUDIES WITH NEW DRUG**.. ERROR! BOOKMARK NOT DEFINED.

TABLE 7.1.3.2.1 **ADVERSE EVENT DROPOUT LISTING**. ERROR! BOOKMARK NOT DEFINED.

TABLE 7.1.5.2.1 **TREATMENT EMERGENT ADVERSE EVENT LISTING**.....ERROR! BOOKMARK NOT DEFINED.

TABLE 7.1.5.3.1 **TREATMENT EMERGENT ADVERSE EVENT INCIDENCE FOR POOL OF 6-WEEK PLACEBO-CONTROLLED TRIALS**.....ERROR! BOOKMARK NOT DEFINED.

TABLE 7.1.7.3.1.1 **MEAN CHANGE FROM BASELINE FOR SERUM CHEMISTRY PARAMETERS IN POOL OF PLACEBO-CONTROLLED STUDIES**. ..... ERROR! BOOKMARK NOT DEFINED.

TABLE 7.1.7.3.2.1 **INCIDENCE OF POTENTIALLY CLINICALLY SIGNIFICANT CHANGES IN SERUM CHEMISTRY PARAMETERS FOR POOL OF PLACEBO CONTROLLED STUDIES FOR NEW DRUG** .ERROR! BOOKMARK NOT DEFINED.

TABLE 7.1.7.5.1 **HEPATOTOXICITY EVALUATION** ..... ERROR! BOOKMARK NOT DEFINED.

TABLE 7.2.1.1.1 **ENUMERATION OF SUBJECTS/PATIENTS FOR NEW DRUG DEVELOPMENT PROGRAM**. ERROR! BOOKMARK NOT DEFINED.

TABLE 7.2.1.2.1 **DEMOGRAPHIC PROFILE FOR PHASE 2-3 STUDIES WITH NEW DRUG**..... ERROR! BOOKMARK NOT DEFINED.

TABLE 7.2.1.3.1 **NUMBER (PERCENT) OF PATIENTS RECEIVING NEW DRUG ACCORDING TO MEAN DAILY DOSE AND DURATION OF THERAPY IN PHASE 2-3 STUDIES (N=2500)**.. ERROR! BOOKMARK NOT DEFINED.

Table 7.0.1

Index for Linking Identified Patients with Supplementary Patient Information in the NDA (CRFs, Narrative Summaries, and Patient Data Listings[1]

Study Number[2]	Patient Number[3]	Case Report Forms		Narrative Summaries		Patient Data Listings	
		Volume[4]	Pages[5]	Volume	Pages	Volume	Pages

[1]Separate indices should be provided for patients exposed to new drug, active control drugs, and placebo.
[2]Study numbers should be numerically ordered and tabbed as separate sections within the index.
[3]Patient numbers should be numerically ordered within each study section.
[4]The volume number provided in this index should be the unique volume number assigned to the volume as part of the complete NDA, and not a separate volume number assigned to the volume as part of a section of the NDA.
[5]The page numbers provided in this index should be the unique page numbers assigned for the entire volume, and not separate page numbers assigned to the separate sections that might be included in any particular volume.

Table 7.1.1.1
Deaths Listing[1,2,3]
Treatment = New Drug[4]
Cutoff Date[5]

Trial	Center	Patient	Age (yrs)	Sex	Dose[6] (mg)	Time[7] (Days)	Source[8]	Person Time[9]	Description[10]

[1]A footnote should describe the rule for including deaths in the table (e.g., all deaths that occurred during a period of drug exposure or within a period of up to 30 days following discontinuation from drug and also those occurring later but resulting from adverse events that had an onset during drug exposure or during the 30-day follow up period). Other rules may be equally appropriate.
[2]Deaths occurring outside the time window for this table should be listed elsewhere.
[3]This table should be provided by the sponsor in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the sponsor and the reviewing division.
[4]Similar lists should be provided for patients exposed to placebo and active control drugs.
[5]This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed patients were not available for entry). Generally this date should be no more than several months prior to the submission date for an NDA. This date as well as this table may likely need to be updated during the course of NDA review as more data become available.
[6]Dose at time of death, or if death occurred after discontinuation, note that, as well as last dose before discontinuation.
[7]Days on drug at time of death; or if death occurred after discontinuation, note how many days on drug before discontinuation and also how many days off drug at time of death.
[8]This listing should include all deaths meeting the inclusion rule, whether arising from a clinical trial or from any secondary source (e.g., postmarketing experience). The source should be identified in this column (i.e., 1^0 for deaths arising from primary source clinical trials and 2^0 for those arising from secondary sources).
[9]This column should identify patients (yes/no) for whom person-time data are available, so the reviewer can know which patients were included in the mortality rate calculations.
[10]Since narrative summaries should be available for all deaths, the description can be very brief (e.g., myocardial infarction, stroke, pancreatic cancer, suicide by drowning).

Table 7.1.1.2
Mortality by Treatment Group
for Pool of Phase 2-3 Studies with New Drug[1,2,3]
Cutoff Date[4]

Treatment Group[5]	Total Number of Patients[6]	Total Number of Deaths[7]	Crude Mortality[8]	Patient Exposure Years (PEY)[9]	Total Deaths with Person-Time[10]	Mortality per 100 PEY[11]
New Drug						
Active Control						
Placebo						

[1]This table provides data comparing overall mortality across treatment groups for the pool of all phase 2 to 3 studies in the development program. Similar tables may be appropriate for other subgroups within the phase 2 to 3 program (e.g., a table may be provided for a pool of all similarly designed short-term placebo controlled trials). Similar tables may be appropriate for certain individual trials of interest. All deaths should be counted, regardless of the investigator's or the sponsor's judgment about causality, including (1) any deaths occurring during participation in any of the studies in the target pool, (2) any deaths occurring after a patient leaves any of the targeted studies, whether prematurely or after completion to the nominal endpoint, if the death is (a) the result of a process initiated during the study, regardless of when it actually occurs, or (b) occurs within 4 weeks of a patient leaving a study, or longer for drugs with particularly long elimination half-lives or from drug classes with known late occurring effects. The actual rule used for including deaths should be provided in a footnote to the table. In case there are substantial deaths of specific causes, it may be appropriate to provide data for cause specific mortality as well.

[2]Patients participating in crossover trials should be enumerated for each of the pertinent columns of the table (e.g., a patient receiving treatment in each of the three arms of a 3-way crossover study comparing new drug, active control, and placebo would be included in all three columns.

[3]This table should be provided by the sponsor in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the sponsor and the reviewing division.

[4]This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed patients were not available for entry. Generally, this date should, be no more than several months prior to the submission date for an NDA. This date as well as this table may likely need to be updated during the course of NDA review as more data become available.

[5]In the sample table, only 1 row is provided for an *active control* group. One such category may suffice for certain NDAs, but may not for others, and the decision regarding how to categorize active control patients should be made in consultation with the reviewing division. Similarly, for this table, only 1 row is provided for new drug, with the implication that all new drug patients, regardless of dose, should be included in the calculations for that column. Other approaches (e.g., distinguishing patients on the basis of dose) may be equally appropriate.

[6]The Ns in these rows should match the N's in Table 5.1.1.1., and if not, an explanation should be provided in a footnote.

[7]This is the total number of deaths for each group.

[8]This is simply the total number of deaths divided by the total number of patients exposed in each group.

[9]This column should provide person-time in patient exposure years (PEY). This table assumes a constant hazard rate; however, in certain situations, it may be appropriate to stratify by increments of exposure.

[10]This is the subset of total deaths for which person-time is available.

[11]This is the number of deaths for whom person-time is available divided by PEY for each group, and multiplied by 100.

Table 7.1.2.1
Serious Adverse Event Listing[1,2,3]
New Drug Clinical Trials
Source: Phase 2-3 Trials[4]
Sorting A: Randomized Treatment, Trial #, Investigator/Center #, Patient #[5]
Treatment = New Drug[6]
Cutoff Date[7]

Trial	Center	Patient	Age (yrs)	Sex	Dose[8] (mg)	Time[9] (days)	Body System	Prefrrd Term	Adverse Event[10]	W/D[11]

[1]This is a line listing of all reported adverse events that met the sponsor's definition of being a *serious* adverse event, regardless of whether or not considered drug related, for all patients participating in the phase 2 to 3 trials in the development program. This listing is a critical component of the integrated safety summary.
[2]The variables included in this listing include:

- -Trial #
- -Center #
- -Patient # (a unique number that identifies this patient in the NDA database)
- -Age
- -Sex
- -Dose (in mg) at time of event onset
- -Time, i.e., duration, of exposure (in days) at time of event onset
- -Body system category for event (using COSTART or other thesaurus)
- -Preferred term for event
- -Adverse event as reported by investigator and/or patient
- -An indication of whether or not the event led to withdrawal
- -Serious Adverse Event Type (e.g., fatal, life-threatening).

The following additional variables may be considered for inclusion as well:

- -Race
- -Weight
- -Height

-Dose expressed as mg/kg, mg/mm^2, or even plasma concentration, if available
-Other drug treatment
-Severity of adverse event (mild, moderate, severe)
-Action taken (e.g., none; decrease dose; discontinue treatment)
-Outcome
-Causality assessment by investigator (definitely; probably; possibly; unlikely related)
-Location in NDA of CRF (e.g., patient narrative summary)

[3]The exact design of the table and whether or not it needs to be provided in electronic format should be established in discussions between the sponsor and the reviewing division.
[4]Similar listings may be provided for individual studies as part of full reports for such studies, and possibly for other pools that are subsets of this larger pool.
[5]It is essential to provide this listing in two different forms (i.e., sorting A (by patient) and sorting B (by adverse event)). This listing is for sorting A, by patient, and permits the reviewer to explore all the serious adverse events reported for each individual patient. Sorting B (by adverse event) should be as follows: Randomized Treatment, Body System, Preferred Term, Adverse Event, Trial, Center, Patient #, Age, Sex, Dose, Time, W/D. Sorting B permits the reviewer to explore all the reported serious adverse events of a similar type.
[6]This sample listing is for all new drug patients across all studies in the phase 2 to 3 development program. Similar listings should be provided for active control and placebo patients.
[7]This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed patients were not available for entry). Generally, this date should be no more than several months prior to the submission date for an NDA. This date as well as this table may likely need to be updated during the course of NDA review as more data become available.
[8]This column should include the dose being administered (in mg/day) at the time the event occurred.
[9]This column should include the time (i.e., duration of exposure (in days)), at the time the event occurred. If the event occurred after discontinuation of drug, a footnote should note how long after discontinuation.
[10]This column should include the adverse event in the language reported by the investigator and/or patient (i.e., before coding).
[11]This column should include an indication of whether or not the adverse event led to discontinuation of the assigned treatment.

Table 7.1.3.1.1
Dropout Profile: Incidence of Dropout by Treatment Group and Reason for Phase 2 to 3 Studies with New Drug[1,2,3]
Cutoff Date[4]:

Reasons for Dropout[5]	Treatment Groups[6]		
	New Drug N =	Placebo N =	Active Control N =
Lack of Efficacy	%[7]	%	%
Adverse Event	%	%	%
Lost to Follow up	%	%	%
Other	%	%	%
Total Dropouts	%	%	%

[1]This sample table should be based on a pool of all trials in the phase 2 to 3 development program. Similar tables may be appropriate for other subgroups within the phase 2 to 3 program (e.g., a table should be provided for a pool of all similarly designed short-term placebo controlled trials). Similar tables may be appropriate for certain individual trials of interest.

[2]Patients participating in crossover trials should be enumerated for each of the pertinent columns of the table (e.g., a patient receiving treatment in each of the three arms of a 3-way crossover study comparing new drug, active control, and placebo would be included in all three columns).

[3]This table should be provided by the sponsor in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the sponsor and the reviewing division.

[4]This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed patients were not available for entry). Generally this date should be no more than several months prior to the submission date for an NDA. This date as well as this table may likely need to be updated during the course of NDA review as more data become available.

[5]This sample table includes 4 categories for dropout, but a more detailed breakdown may be of interest as well.

- The adverse event category here would include all patients identified as dropping out for adverse events, regardless of whether or not the events were judged by the investigator or sponsor to be drug related and regardless of what other reasons may have been identified in association with dropout. Patients identified as dropping out for intercurrent illness would ordinarily be included under this adverse event category. Similarly, a patient identified as dropping out for an adverse event and lack of efficacy would also ordinarily be included under this adverse event category.
- Lost-to-follow up is an important outcome to track, since it reflects on the overall conduct of the studies.
- The *other* category is intended to include all other reasons that may generally be considered nontreatment related. This category is often identified as *administrative*, and includes such reasons as patient refused further participation, patient moved away, patient improved, patient not eligible, protocol violation, unknown.
- Decisions about what categories to include should be made in consultation with the reviewing division.

[6]In the sample table, only 1 column is provided for an *active control* group. One such category may suffice for certain NDAs, but may not for others, and the decision regarding how to categorize active control patients should be made in consultation with the reviewing division. Similarly, for this table, only 1 column is provided for new drug, with the implication that all new drug patients, regardless of dose, should be included in the calculations for that column. Other approaches (e.g., distinguishing patients on the basis of dose) may be

equally appropriate. The N's in these column headings should match the N's in Table 5.1.1.1., and if not, an explanation should be provided in a footnote.

[7]Numbers for this table should be rounded to the nearest integer.

Table 7.1.3.2.1
Adverse Event Dropout Listing[1,2,3]
New Drug Clinical Trials
Source: Phase 2-3 Database[4]
Sorting A: Randomized Treatment, Trial #, Investigator/Center #, Patient #[5]
Treatment = New Drug[6]
Cutoff Date[7]:

Trial	Center	Patient	Age (yrs)	Sex	Dose[8] (mg)	Time[9] (days)	Body System	Prefrrd Term	Adverse Event[10]	Serious[11]	Outcome[12]

[1]This is a line listing of all reported adverse events identified as leading to discontinuation, regardless of whether or not they were considered drug related, for all patients participating in trials identified as sources for this listing. Thus, all events categorized as *intercurrent illness* leading to discontinuation would, nevertheless, be included in this listing, and any judgments about attribution can be included in the narrative summary. This listing is a critical component of the integrated safety summary.

[2]The variables included in this listing include:

- -Trial #
- -Center #
- -Patient # (a unique number that identifies this patient in the NDA database)
- -Age
- -Sex
- -Dose (in mg) at time of event onset
- -Time (i.e., duration, of exposure (in days) at time of event onset)
- -Body system category for event (using COSTART or other thesaurus)
- -Preferred term for event
- -Adverse event as reported by investigator and/or patient
- -An indication of whether or not the event met definition for serious
- -Outcome

The following additional variables may be considered for inclusion as well:

- -Race
- -Weight
- -Height
- -Dose expressed as mg/kg, mg/mm^2, or even plasma concentration, if available
- -Other drug treatment
- -Severity of adverse event (mild, moderate, severe)
- -Action taken (e.g., none; decrease dose, discontinue treatment)
- -Causality assessment by investigator (related, not related)
- -Location in NDA of CRF, patient narrative summary)

[3]The exact design of the table and whether or not it needs to be provided in electronic format should be established in discussions between the sponsor and the reviewing division.

[4]Similar listings may be provided for individual studies as part of full reports for such studies and, possibly, for other pools that are subsets of this larger pool.

[5]It is essential to provide this listing in two different forms (i.e., sorting A (by patient) and sorting B (by adverse event)). This listing is for sorting A, by patient, and permits the reviewer to explore all the adverse events reported as leading to discontinuation for each individual patient. Sorting B (by adverse event) should be as follows: Randomized Treatment, Body System, Preferred Term, Adverse Event, Trial, Center, Patient #, Age, Sex, Dose, Time, Serious. Sorting B permits the reviewer to explore all the adverse events of a similar type reported as leading to discontinuation.

[6]This sample listing is for all new drug patients across all studies in the phase 2 to 3 development program. Similar listings should be provided for active control and placebo patients.

[7]This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed patients were not available for entry). Generally this date should be no more than several months prior to the submission date for an NDA. This date as well as this table may likely need to be updated during the course of NDA review as more data become available.

[8]This column should include the dose being administered (in mg/day) at the time the event occurred.
[9]This column should include the time (i.e., duration of exposure (in days)), at the time the event occurred.
[10]This column should include the adverse event in the language reported by the investigator and/or patient, i.e., before coding.
[11]This column should include an indication of whether or not the adverse event met the criteria for *serious* as defined for the development program overall.
[12]This column should categorize the outcome upon follow up evaluation for the adverse event leading to discontinuation, as follows:

(R) Resolved
(P) Persisting
(U) Unknown

Table 7.1.5.2.1
Treatment Emergent Adverse Event Listing[1,2,3]
New Drug Clinical Trials
Source: Phase 2-3 Database[4]
Sorting A: Randomized Treatment, Trial #, Investigator/Center #, Patient #[5]
Treatment = New Drug[6]
Cutoff Date[7]:

Trial	Center	Patient	Age (yrs)	Sex	Dose[8] (mg)	Time[9] (days)	Body System	Preferred Term	Adverse Event[10]	Serious[11]	W/D[12]

[1]This is a line listing of all reported treatment emergent adverse events, regardless of whether or not they were considered drug related, for all patients participating in trials identified as sources for this listing. A footnote should identify all studies contributing to this pool.
[2]The variables included in this listing include:
- -Trial #
- -Center #
- -Patient # (a unique number that identifies this patient in the NDA database)
- -Age
- -Sex
- -Dose (in mg) at time of event onset
- -Time (i.e., duration, of exposure (in days) at time of event onset)
- -Body system category for event (using COSTART or other thesaurus)
- -Preferred term for event
- -Adverse event as reported by investigator and/or patient
- -An indication of whether or not the event met definition for serious
- -An indication of whether or not the event led to withdrawal

The following additional variables may be considered for inclusion as well:

-Race
-Weight
-Height
-Dose expressed as mg/kg, mg/mm^2, or even plasma concentration, if available
-Other drug treatment
-Duration of adverse event
-Timing of adverse event relative to last dose
-Severity of adverse event (mild, moderate, severe)
-Action taken (none, decrease dose, discontinue treatment)
-Outcome
-Causality assessment by investigator (definitely, probably, possibly, or unlikely related)
-Location in NDA of CRF, patient narrative summary

[3]The exact design of the table and whether or not it needs to be provided in electronic format should be established in discussions between the sponsor and the reviewing division.

[4]Similar listings may be provided for individual studies as part of full reports for such studies, and possibly for other pools that are subsets of this larger pool.

[5]It is essential to provide this listing in two different forms (i.e., sorting A (by patient) and sorting B (by adverse event)). This listing is for sorting A, by patient, and permits the reviewer to explore all the adverse events reported for each individual patient. Sorting B (by adverse event (i.e., 1 row for each occurrence of each adverse event)) should be as follows: Randomized Treatment, Body System, Preferred Term, Adverse Event, Trial, Center, Patient #, Age, Sex, Dose, Time, Serious, W/D. Sorting B permits the reviewer to explore all the reported adverse events of a similar type.

[6]This sample listing is for new drug patients (i.e., for all patients exposed to New Drug in the phase 2 to 3 studies that are part of the Integrated Primary Database). Similar listings should be provided for active control and placebo patients.

[7]This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed patients were not available for entry). Generally this date should be no more than several months prior to the submission date for an NDA. This date as well as this table may likely need to be updated during the course of NDA review as more data become available.

[8]This column should include the dose being administered (in mg/day) at the time the event occurred.

[9]This column should include the time (i.e., duration of exposure (in days)), at the time the event occurred.

[10]This column should include the adverse event in the language reported by the investigator and/or patient (i.e., before coding).

[11]This column should include an indication of whether or not the adverse event met the criteria for *serious* as defined for the development program overall.

[12]This column should include an indication of whether or not the adverse event led to discontinuation of the assigned treatment.

Table 7.1.5.3.1 Treatment-Emergent Adverse Event Incidence for Pool of 6-Week Placebo-Controlled Trials[1-10] Cutoff Date[11]:			
Body System/ Adverse Event[12-14]	Percentage of Patients Reporting Event[15]		
	New Drug N[16]=	Active Control N=	Placebo N=
Body as a Whole			
Headache			
Etc.			
Cardiovascular System			
Postural Hypotension			
Etc.			
Gastrointestinal System			
Constipation			
Etc.			
.			
.			
Urogenital System **Impotence**[17]			
Etc.			
.			
.			

[1]This table compares the incidence of treatment emergent adverse events across treatment groups for a pool of similarly designed placebo-controlled trials of new drug. Generally, an arbitrary threshold incidence for new drug patients is used as a criterion for selecting adverse events to include; $\geq$ 1% for new drug is a commonly used rule, but others may be equally appropriate. The criterion used should be noted in the table title or in a footnote.

[2]Study pools other than those described for this sample table may be equally appropriate, and similar tables useful for individual trials may also be of interest.

[3]In the sample table, only 1 column is provided for an *active control* group. One such category may suffice for certain NDAs, but may not for others, and the decision regarding how to categorize active control patients should be made in consultation with the reviewing division.

[4]Similarly, for this table, only 1 column is provided for new drug, with the implication that all new drug patients, regardless of dose, should be included in the calculations for that column. Other approaches (e.g., dividing patients on the basis of dose), may be equally appropriate. If the studies used were fixed-dose studies,

it is generally most informative to preserve the dose categories in constructing this table. However, dose categories that are not relevant to the doses that are being recommended for use may reasonably be omitted from this table. It is generally not useful to try to artificially construct dose categories from dose titration studies, since there is often confounding of dose and time.

[5]Data are often available on the investigator's opinion regarding whether or not any particular adverse event was in fact related to the drug being taken. Some reviewers consider this useful information and may construct tables that include only those events considered possibly, probably, or definitely drug-related by the investigator. Others ignore such judgments and include all reported adverse events, with the view that the control groups, especially placebo if present, should permit one to make causality decisions, regardless of the investigators' judgments about drug-relatedness. Either approach can be acceptable, but it is critical that a footnote indicate clearly when adverse events are not included due to investigators' judgments that they were not drug-related, since this approach may reduce the adverse event rates that appear in the table.

[6]Data are also often available on the intensity of the reported adverse events, generally including categories of *mild, moderate, or severe*. Adverse event tables may ignore such classifications and pool all events together, or some attempt may be made to focus only on a subset of reported events (e.g., only those classified as *severe*). Again, either approach is acceptable, but it is important to describe in a footnote what approach was taken.

[7]Not uncommonly, a new drug is developed for more than one indication. If adverse event rates appear to be to occur at similar rates across the indications, it may be reasonable to pool the data in creating an adverse events table, possibly one providing greater precision. However, it is not inconceivable that adverse event rates may vary depending on the population studied, and if this appears to be the case, pooling may not be appropriate.

[8]Adverse events that occur at a rate for placebo that is $\geq$ the rate for new drug should be removed from the table and noted only as a footnote.

[9]Patients participating in crossover trials should be included in the calculations for each of the pertinent columns of the table (e.g., a patient receiving treatment in each of the three arms of a 3-way crossover study comparing new drug, active control, and placebo would be included in the calculations for all three columns).

[10]This table should be provided by the sponsor in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the sponsor and the reviewing division.

[11]This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed patients were not available for entry). Generally this date should be no more than several months prior to the submission date for an NDA. This date as well as this table may likely need to be updated during the course of NDA review as more data become available.

[12]Adverse events should be organized under body system categories.

[13]Within each body system category, adverse events should be ordered according to decreasing frequency.

[14]Adverse events during exposure are generally obtained by spontaneous report and recorded by clinical investigators using terminology of their own choosing. Consequently, it is not possible to provide a meaningful estimate of the proportion of individuals experiencing adverse events without first grouping similar types of events into a smaller number of standardized event categories. Generally a table of this type should use these preferred adverse event terms, and a footnote should identify the system used for coding investigator terms. Adverse event terms that convey no useful information (e.g., joint disorder), should be replaced by more clinically useful terms or deleted

[15]Percentages should be rounded to the nearest integer. Although not strictly hypothesis testing, p-values give some feeling for the strength of the finding and should be produced for all new drug/placebo pairwise comparisons and any p-values meeting a $p < 0.05$ level of significance should be noted by an asterisk (*) as a superscript to the %.

[16]The N for each column should be provided at the column heading, so that only the percentage of patients having that adverse event need be included in the table, and not the actual number.

[17]The rates for gender specific adverse events (e.g., impotence) should be determined using the appropriate gender specific denominator, and this fact should be indicated with a footnote.

Table 7.1.7.3.1.1
Mean Change from Baseline for Serum Chemistry Parameters[1]
in Pool of Placebo-Controlled Studies[2,3,4]
Cutoff Date[5]:

Serum Chemistry Parameters and Units of Measure[6]	Treatment Groups[7,8]								
	New Drug			Placebo			Active Control		
	N[9]	BL[10]	Change from BL[11]	N	BL	Change from BL	N	BL	Change from BL
Albumin (g/dl)									
Alkaline Phosphatase (U/L)									
Bilirubin, total (mg/dl)									
BUN (mg/dl)									
CK (U/L)									
Calcium (mg/dl)									
Cholesterol (mg/dl)									
Creatinine (mg/dl)									
GGT (U/L)									
Glucose (mg/dl)									
LDH (U/L)									
Phosphorus (mg/dl)									
Potassium (mmol/L)									

Sodium (mmol/L)									
Triglycerides (mg/dl)									
Uric Acid (mg/dl)									

[1]This table provides data comparing the mean change from baseline across treatment groups for serum chemistry parameters. An acceptable alternative would be to provide median change from baseline. The postmeasurement is generally the worst value during treatment.
[2]This sample table is based on a pool of similarly designed placebo controlled trials. Other pools, as well as individual trials may also be of interest.
[3]Patients participating in crossover trials should be enumerated for each of the pertinent columns of the table (e.g., a patient receiving treatment in each of the three arms of a 3-way crossover study comparing New Drug, active control, and placebo would be included in all three columns).
[4]This table should be provided by the sponsor in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the sponsor and the reviewing division.
[5]This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed patients were not available for entry). Generally this date should be no more than several months prior to the submission date for an NDA. This date as well as this table may likely need to be updated during the course of NDA review as more data become available.
[6]The parameters included in this list are for illustration. In general, the list should include all those serum chemistry parameters measured in whatever pool of studies is the focus of the table. Similarly, the units of measure are for illustration, and these details should be worked out in consultation with the reviewing division.
[7]In the sample table, only 1 column is provided for an *active control* group. One such category may suffice for certain NDAs, but may not for others, and the decision regarding how to categorize active control patients should be made in consultation with the reviewing division.
[8]Similarly, for this table, only 1 column is provided for new drug, with the implication that all new drug patients, regardless of dose, should be included in the calculations for that column. Other approaches (e.g., dividing patients on the basis of dose), may be equally appropriate. If the studies used were fixed-dose studies, it is generally most informative to preserve the dose categories in constructing this table. However, dose categories that are not relevant to the doses that are being recommended for use may reasonably be omitted from this table. It is generally not useful to try to artificially construct dose categories from dose titration studies, since there is often confounding of dose and time.
[9]N represents the number of patients who had the serum chemistry parameter of interest assessed at baseline and at least one follow up time.
[10]This column should provide the baseline means for all the serum chemistry parameters of interest.
[11]This column should provide the mean change from baseline to patient's worst on drug value for each of the serum chemistry parameters of interest. While not hypothesis testing, p-values provide some measures of the strength of the finding and should be produced for all new drug/placebo pairwise comparisons and any p-values meeting a $p < 0.05$ level of significance criterion should be noted by an asterisk (*) as a superscript to the mean change from baseline.

Table 7.1.7.3.2.1
Incidence of Potentially Clinically Significant Changes in Serum Chemistry Parameters[1] for Pool of Placebo Controlled Studies for New Drug[2,3,4]
Cutoff Date[5]:

Serum Chemistry Parameters and PCS Criteria[7] L=Low; H=High; ULN=Upper Limits of Normal	Treatment Groups[6]								
	New Drug			Placebo			Active Control		
	Total Pts[8]	Abnormal		Total Pts	Abnormal		Total Pts	Abnormal	
		Nbr[9]	%[10]		Nbr	%		Nbr	%
Albumin-L (< 2.5 g/dl)									
Alkaline P'tase-H (> 400 U/L)									
Bilirubin, total-H (> 2 mg/dl)									
BUN-H (> 30 mg/dl)									
CK-H (> 3XULN)									
Calcium-L (< 7 mg/dl)									
Calcium-H (> 12 mg/dl)									
Cholesterol-H (> 300 mg/dl)									
Creatinine-H (> 2 mg/dl)									
GGT-H (> 3XULN)									
Glucose-L (< 50 mg/dl)									
Glucose-H (> 250 mg/dl)									

LDH-H (> 3XULN)									
Phosphorus-L (< 2.0 mg/dl)									
Phosphorus-H (> 5.0 mg/dl)									
Potassium-L (< 3.0 mmol/L)									
Potassium-H (> 5.5 mmol/L)									
SGOT/AST-H (> 3XULN)									
SGPT/ALT-H (> 3XULN)									
Sodium-L (< 130 mmol/L)									
Sodium-H (> 150 mmol/L)									
Triglycerides-H (> 300 mg/dl)									
Uric Acid (F)-H (> 8.0 mg/dl)									
Uric Acid (M)-H (> 10.0 mg/dl)									

[1]This table provides data comparing the incidence across treatment groups of patients who were normal at baseline meeting criteria of having had a change on any of the listed serum chemistry parameters of potential clinical significance (PCS). Separate listings should be provided for patients who were abnormal at baseline and met these PCS criteria.

[2]This sample table is based on a pool of similarly designed placebo controlled trials. Other pools, as well as individual trials may also be of interest.

[3]Patients participating in crossover trials should be enumerated for each of the pertinent columns of the table (e.g., a patient receiving treatment in each of the three arms of a 3-way crossover study comparing New Drug, active control, and placebo would be included in all three columns).

[4]This table should be provided by the sponsor in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the sponsor and the reviewing division.

[5]This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed patients were not available for entry. Generally this date should be no more than several months prior to the submission date for an NDA. This date as well as this table may likely need to be updated during the course of NDA review as more data become available.

[6]In the sample table, only 1 column is provided for an *active control* group. One such category may suffice for certain NDAs, but may not for others, and the decision regarding how to categorize active control patients should be made in consultation with the reviewing division. Similarly, for this table, only 1 column is provided for new drug, with the implication that all new drug patients, regardless of dose, should be included in the calculations for that column. Other approaches (e.g., distinguishing patients on the basis of dose), may be equally appropriate.

[7]The parameters included in this list are for illustration. In general, the list should include all those serum chemistry parameters measured in whatever pool of studies is the focus of the table. Similarly, the proposed criteria for *potentially clinically significant* are for illustration, and these details should be worked out in consultation with the reviewing division.

[8]The total number of patients for each parameter should represent the number of patients for the treatment group who (1) had that parameter assessed at baseline and at least one follow up time and (2) for whom the baseline assessment was normal.

[9]The number abnormal represents the subset of the total number who met the criterion in question at least once during treatment. A separate listing should provide patient identification for those patients meeting the criterion.

[10]Percentage of the total number meeting the criterion should be rounded to the nearest integer. While not strictly hypothesis testing, p-values should be produced for all new drug/placebo pairwise comparisons and any p-values meeting a $p < 0.05$ level of significance should be noted by an asterisk (*) as a superscript to the %.

Table 7.1.7.5.1
Hepatotoxicity Evaluation

I. Data Collection
   A. Overview of liver chemistry data (tests performed, frequency, specific follow-up plans for abnormal values)
   B. Specific follow-up plan if chemistry is elevated at end of treatment
   C. Re-challenge plan, if any
   D. Exclusions from studies because of liver chemistry abnormalities, if any

II. Observations

   A. Abnormal liver chemistries seen in controlled trials (separate for pooled placebo controlled, active controlled) with greater than two week exposure. Rates can be given as events/exposed; positive findings can be also analyzed as events per patient year and examined for rates over time.
      1. Rates of 3x, 5x, 10x, 20x ULN elevations of AST (SGOT), ALT (SGPT), and either ALT or AST
      2. Rates of any elevations of bilirubin; rate of elevated bilirubin to >1.5x ULN
      3. Rates of alkaline phosphatase (AP) $\geq$ 1.5x ULN
      4. Rates of elevated transaminase accompanied by elevated bilirubin.

      All rates should be given for both drug and control group.

   B. For total database with exposure $\geq$ two weeks (i.e., including uncontrolled)
      Same as for controlled database (1-4)
   C. Individual events
      1. Listing of patients with any elevated transaminase (>3x ULN), without more than slight AP elevation, associated with increase in bilirubin to > ULN.
      2. Show time course of enzyme and bilirubin elevations
      3. For such patients, review clinical situation
         a. Ethanol history
         b. Evidence viral hepatitis
         c. Symptoms and course – follow-up is particularly important to detect underlying liver disease
         d. Special studies, notably Bx
         e. Possible confounding, including concomitant illness, concomitant medications (known hepatotoxins, including acetaminophen)

III. Possible problems/signals
   A. Any patient with elevated transaminase (to at least 3x ULN, generally higher), no evidence of obstruction (elevated AP) and even modestly (2X ULN) elevated bilirubin. Greater elevation of bilirubin is stronger signal.
   B. Greater rate than control of 3x, 5x, 10x, etc. elevations of transamisne.

Table 7.2.1.1.1
Enumeration of Subjects/Patients for
New Drug Development Program[1,2,3,4]
Cutoff Date[5]:

Study Groups	Treatment Groups		
	New Drug	Active Control[6]	Placebo
Completed Phase 1 (Clinical Pharmacology)			
Single Dose	120	30	30
Multiple Dose	60	30	30
Ph 1 Subtotal	180	60	60
Completed Phase 2-3 (Studies of Proposed Indication)			
Placebo Control[7]			
Fixed Dose	500	150	150
Flexible Dose	100	100	100
Active Control			
Fixed Dose	200	100	0
Flexible Dose	100	100	0
Uncontrolled			
Short Term	100	0	0
Long Term	700	0	0
Ph 2-3 Subtotal	1200[8]	450	250
Ongoing Phase 2-3 Studies (Studies of Proposed Indication)			
Placebo Control			
Flexible Dose	150[9]	0	150[9]
SD Subtotal	120	30	30
MD Subtotal	1410	480	430
Grand Total	1530	510	460

[1]This table provides a count by study type of the subjects/patients exposed to new drug, active control, and placebo across the entire set of studies in the development program that contributed safety and efficacy data for new drug. It should include all subjects/patients known or assumed to have received even a single dose of assigned treatment. It should exclude subjects/patients who are known not to have received any of the assigned treatments or for whom no follow up information is available subsequent to the assumed receipt of

assigned treatment. A separate listing of all such patients should be provided. [Note: If this list includes more than a few patients, this may indicate a potentially important problem in the conduct of studies.]

In creating this table, it is necessary to classify and group studies on the basis of several characteristics. For the purposes of this table, the following characteristics and distinctions were deemed important:

-Phase 1 vs Phases 2 to 3
-Completed vs Ongoing and Blinded
-Single Dose vs Multiple Dose
-Controlled vs Uncontrolled
-Short-Term vs Long-Term
-Placebo-Controlled vs Active-Controlled
-Fixed Dose vs Flexible Dose

Obviously, there are other features that may be important as well, and that could lead to additional breakdowns within the table or to separate tables (e.g., different indications, inpatient vs outpatient status, differences in the quality and completeness of data collected across different studies, foreign vs domestic). The characteristics to be used in classifying studies for the purpose of this table should be decided in consultation with the designated reviewing division at FDA.

In addition to this table that enumerates patients by category of study, it would be useful to have a table that enumerates patients by each individual study in the development program. This would be an expanded version of the above table that enumerates patients for each study (i.e., each of the categories in the above table would identify and provide data for the individual studies comprising that category). Sponsors ordinarily provide such a table.

[2]Patients participating in crossover trials should be counted in each of the pertinent columns of the table (e.g., a patient receiving treatment in each of the three arms of a 3-way crossover study comparing new drug, active control, and placebo would be counted in all three columns).

[3]Footnotes to this table should identify by study number all those studies comprising the various study groupings for this table. For example, in the sample table, the fixed dose placebo controlled trials contributing to the counts for that category should be listed in a footnote, and similarly for all other categories.

[4]This table should be provided by the sponsor in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the sponsor and the reviewing division.

[5]This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed patients were not available for entry). Generally this date should be no more than several months prior to the submission date for an NDA. This date as well as this table likely need to be updated during the course of NDA review as more data become available.

[6]In the sample table, only 1 column is provided for an *active control* group. One such category may suffice for certain NDAs, but may not for others, and the decision regarding how to categorize active control patients should be made in consultation with the reviewing division.

[7]In this table, a decision was made to pool all studies having a placebo arm, whether or not an active control arm was also included. Thus, the active control category includes only those active control studies that did not have a placebo control arm. Other approaches to grouping studies may be equally appropriate.

[8]The intent of this table is to provide a count of unique subjects/patients exposed to new drug, etc. in the development program. Since patients often participate in more than 1 study in a development program, it is necessary to have an approach to avoid counting patients more than once for the subtotals and grand totals. The approach used in this table is to include in parentheses in the pertinent cells of the table a count of the patients in that cell total who have already been counted by virtue of having participated in a previous study (e.g., a patient in an open extension trial should have been previously counted in an acute, controlled phase). The subtotals of unique individuals exposed to the assigned treatment can then be calculated by subtracting the sum of all numbers in parentheses from the sum of all the cell totals for each column (e.g., in this table, the completed phase 2 to 3 subtotal for new drug is 1700 less the 500 patients already counted in short-term controlled trials, or 1200).

[9]Frequently, some studies may be ongoing and blinded at the time of NDA submission, even though some individual patients having experienced serious adverse events may have been unblinded. In these instances, the table should include estimates of the numbers of patients exposed to new drug, etc. from these studies, since exact counts may not be available. Footnotes should indicate when the table entries are based on estimates rather than exact counts.

Table 7.2.1.2.1
Demographic Profile for Phase 2-3 Studies with New Drug[1,2,3,4,5]
Cutoff Date[6]:

Demographic Parameters	Treatment Groups[7,8]		
	New Drug N =	Placebo N =	Active Control N =
Age (years)			
Mean			
Range			
Groups[9]			
< 40	%	%	%
40-64	%	%	%
≥ 65	%	%	%
Sex			
Female	%	%	%
Male	%	%	%
Race[10]			
Caucasian	%	%	%
Non-Caucasian	%	%	%
Weight (kg)			
Mean			
Range			

[1]This table should be based on a pool of all trials in the phase 2 to 3 development program. Similar tables may be appropriate for other subgroups within the phase 2 to 3 program and also for certain individual trials of interest. The specific trials included should be listed.
[2]Patients participating in crossover trials should be included in the calculations for each of the pertinent columns of the table (e.g., a patient receiving treatment in each of the three arms of a 3-way crossover study comparing New Drug, active control, and placebo would be included in the calculations for all three columns).
[3]Numbers for this table should be rounded to the nearest integer.
[4]This sample table includes 4 demographic categories of obvious interest, however, others may be of interest as well (e.g., height, severity on baseline measures of disease severity). It may also be of interest to look at combinations of characteristics, such as gender and age (e.g., women under 50).
[5]This table should be provided by the sponsor in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the sponsor and the reviewing division.

[6]This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed patients were not available for entry). Generally this date should be no more than several months prior to the submission date for an NDA. This date as well as this table may likely need to be updated during the course of NDA review as more data become available.

[7]In the sample table, only 1 column is provided for an *active control* group. One such category may suffice for certain NDAs, but may not for others, and the decision regarding how to categorize active control patients should be made in consultation with the reviewing division. Similarly, for this table, only 1 column is provided for new drug, with the implication that all new drug patients, regardless of dose, should be included in the calculations for that column. Other approaches (e.g., distinguishing patients on the basis of dose), may be equally appropriate.

[8]If, as is often the case, the Ns available for calculating any particular demographic parameter are less than the Ns in the column headings, these Ns should be provided, along with an explanation, in footnotes.

[9]If there are pediatric exposures, these should be broken out as well.

[10]Other approaches to racial categorization may be substituted for that proposed in this sample table.

Table 7.2.1.3.1
Number (Percent) of Patients Receiving New Drug According to Mean[1,2,3,4,5,6,7] Daily Dose and Duration of Therapy in Phase 2-3 Studies (N=2500) Cutoff Date[8]:

Duration (Weeks)	Dose[9] (mg)							
	0<Dos≤5	5<Dos≤10	10<Dos≤20	20<Dos≤30	30<Dos≤50	50<Dos	Total (AnyDos)	(%)
0<Dur≤1	6	19	31	31	25	13	125	(5%)
1≤Dur<2	6	19	31	31	25	13	125	(5%)
2≤Dur<4	13	37	62	63	50	25	250	(10%)
4≤Dur<12	31	94	156	156	125	63	625	(25%)
12≤Dur<24	25	75	125	125	100	50	500	(20%)
24≤Dur<48	25	75	125	125	100	50	500	(20%)
48≤Dur<96	13	37	62	63	50	25	250	(10%)
96≤Dur	6	19	31	31	25	13	125	(5%)
Total (AnyDur)	125	375	623	625	500	252	2500	(100%)
(%)	(5%)	(15%)	(25%)	(25%)	(20%)	(10%)	(100%)	

[1]This table is calculated by first categorizing patients on the basis of the interval of exposure for each (e.g., a patient exposed for 6 weeks would be counted in the 4<Dur≤12 row). The mean daily dose is then calculated for each patient for dose categorization (e.g., a 6-week patient with a mean daily dose of 15 mg would be counted in the 10<Dos≤20 column). Patients are enumerated in only 1 cell of the matrix (i.e., this is a mutually exclusive display). The dose and duration intervals need to be designed specifically for the drug of interest. The specific trials included should be listed. As with any table summarizing

data from disparate sources, it does not address all information needs, and it should be interpreted with caution (e.g., mean doses in the 4-12 row refer to mean doses over 0-12 weeks, not 4-12 as one might think). Nevertheless, the information provided provides useful information.

[2]Similar tables can be prepared for median, for modal, and for maximum dose.

[3]The same table can be generated for any individual study or for any pool of studies.

[4]The same table can be generated for any subgroup of interest (e.g., on the basis of age, sex, race, comorbid condition, concomitant medications, or any combination of these factors).

[5]Similar tables should be provided for active control drugs and placebo.

[6]If the total N for this table does not match the total N from Table 5.1.1.1, as may be the case (e.g., if dose or duration data are not available for all exposed patients counted in Table 5.1.1.1, a footnote should provide an explanation for the discrepancy).

[7]This table should be provided by the sponsor in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the sponsor and the reviewing division.

[8]This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed patients were not available for entry). Generally this date should be no more than several months prior to the submission date for an NDA. This date as well as this table may likely need to be updated during the course of NDA review as more data become available.

[9]Dose may also be expressed as mg/kg, mg/m^2, or in terms of plasma concentration if such data are available.