

**Attachment B: Clinical Safety Review of an NDA or BLA**

**TABLE OF CONTENTS**

**1. GENERAL ADVICE ON THE CLINICAL SAFETY REVIEW ..... 1**

1.1 INTRODUCTION ..... 1

1.2 DEFINITIONS ..... 2

1.3 PRINCIPAL TASKS OF THE SAFETY REVIEW ..... 3

1.4 DIFFERENCES IN APPROACH TO SAFETY AND EFFECTIVENESS DATA ..... 4

1.5 IDENTIFYING AND ASSEMBLING SOURCE MATERIALS FOR THE SAFETY REVIEW ..... 6

1.6 IDENTIFYING MAJOR CONCERNS AT THE OUTSET ..... 7

1.7 AUDITING SOURCE MATERIALS ..... 8

1.8 THE PURPOSE OF INDIVIDUAL CASE REVIEW/DRUG-RELATEDNESS ..... 8

**7. REVIEW OF SAFETY ..... 10**

SAFETY SUMMARY ..... 10

7.1 METHODS ..... 13

7.1.1 Studies/Clinical Trials Used to Evaluate Safety ..... 13

7.1.2 Categorization of Adverse Events ..... 15

7.1.3 Pooling of Data Across Studies/Clinical Trials to Estimate and Compare Incidence ..... 16

7.2 ADEQUACY OF SAFETY ASSESSMENTS ..... 18

7.2.1 Overall Exposure at Appropriate Doses/Durations and Demographics of Target Populations.. 19

7.2.2 Explorations for Dose Response ..... 20

7.2.3 Special Animal and/or In Vitro Testing ..... 20

7.2.4 Routine Clinical Testing ..... 21

7.2.5 Metabolic, Clearance, and Interaction Workup..... 21

7.2.6 Evaluation for Potential Adverse Events for Similar Drugs in Drug Class..... 21

7.3 MAJOR SAFETY RESULTS..... 24

7.3.1 Deaths ..... 26

7.3.2 Nonfatal Serious Adverse Events..... 30

7.3.3 Dropouts and/or Discontinuations..... 31

7.3.4 Significant Adverse Events ..... 34

7.3.5 Submission-Specific Primary Safety Concerns..... 35

7.4 SUPPORTIVE SAFETY RESULTS ..... 37

7.4.1 Common Adverse Events..... 37

7.4.2 Laboratory Findings..... 42

7.4.3 Vital Signs..... 46

7.4.4 Electrocardiograms (ECGs)..... 47

7.4.5 Special Safety Studies/Clinical Trials ..... 48

7.4.6 Immunogenicity ..... 48

7.5 OTHER SAFETY EXPLORATIONS..... 49

7.5.1 Dose Dependency for Adverse Events..... 49

7.5.2 Time Dependency for Adverse Events..... 50

7.5.3 Drug-Demographic Interactions..... 51

7.5.4 Drug-Disease Interactions..... 52

7.5.5 Drug-Drug Interactions ..... 52

7.6 ADDITIONAL SAFETY EXPLORATIONS..... 52

7.6.1 Human Carcinogenicity ..... 53

7.6.2 Human Reproduction and Pregnancy Data ..... 53

7.6.3 Pediatrics and Assessment of Effects on Growth..... 53

7.6.4 Overdose, Drug Abuse Potential, Withdrawal and Rebound ..... 54

7.7 ADDITIONAL SUBMISSIONS/SAFETY ISSUES..... 55

**SAMPLE TABLES..... 56**

TABLE 1: ENUMERATION OF SUBJECTS FOR NEW DRUG DEVELOPMENT PROGRAM ..... 56

TABLE 2: TREATMENT EMERGENT ADVERSE EVENT LISTING ..... 59

TABLE 3: DEMOGRAPHIC PROFILE ..... 61

TABLE 4: NUMBER (PERCENT) OF SUBJECTS RECEIVING NEW DRUG ..... 63

TABLE 5: DEATHS LISTING ..... 65

TABLE 6: MORTALITY BY TREATMENT GROUP..... 66

TABLE 7: SERIOUS ADVERSE EVENT LISTING ..... 68

TABLE 8: DROPOUT PROFILE ..... 70

TABLE 9: ADVERSE EVENT DROPOUT LISTING ..... 71

TABLE 10: TREATMENT EMERGENT ADVERSE EVENT INCIDENCE..... 74

TABLE 11: MEAN CHANGE FROM BASELINE FOR SERUM CHEMISTRY PARAMETERS ..... 76

TABLE 12: INCIDENCE OF POTENTIALLY CLINICALLY SIGNIFICANT CHANGES IN SERUM CHEMISTRY  
PARAMETERS ..... 78

**GLOSSARY OF ACRONYMS ..... 82**

This Attachment B is an expansion of the description in Attachment A for section 7, Review of Safety, of the clinical review template. The structure of this attachment, as an annotated outline, is meant to correlate exactly with the section headings of the clinical review template, providing the pertinent instructions for the review under each heading. The attachment also provides sample tables, including suggested displays and graphs that have been used successfully in the past. These displays are examples not requirements, and reviewers can modify or omit them as needed.

The commentary and suggestions under each section, together with appended examples, provide suggested analyses, methods of presentation, and discussion of special cases and potential difficulties. This attachment offers the reviewer general advice on the purpose and philosophy of performing a safety review before discussing the template sections.

## **1. GENERAL ADVICE ON THE CLINICAL SAFETY REVIEW**

### **1.1 Introduction**

This attachment is a good review practice (GRP) that provides an annotated outline of the safety review of a primary clinical review of a new drug application (NDA) or biologics license application (BLA). It also provides recommendations on how to conduct and organize the safety review.<sup>1</sup>

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/clinical trials 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/clinical trials and reference them in this section (e.g., studies/clinical trials with mortality outcomes, development programs in which most of the safety data come from one or two large multicenter studies/clinical trials, and when evaluation and review of safety data may be more convenient or informative clinical trial by clinical trial).

The safety review has two distinct goals: (1) reviewer assessment of the adequacy of the applicant's safety evaluation; and (2) identification and assessment of the significance of the adverse events reported in studies/clinical trials (controlled or uncontrolled). This GRP describes an approach that integrates safety findings across all clinical trials and other clinical experience including observational and pharmacoepidemiologic studies. Consideration of the safety findings in individual studies/clinical trials, without a

---

<sup>1</sup> It is recognized that no drug is safe in the sense of being entirely free of adverse effects. Reference in the Federal 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/BLA review that assesses and describes the risks of the drug.

thoughtful integration of the overall safety experience, is not adequate for a safety review.<sup>2</sup>

Much of the instruction in this GRP is directed toward the clinical reviewer and toward the analysis of particular events. However, clinical reviewers should collaborate with their biostatistical colleagues and others when necessary in the preparation of reviews to evaluate safety data that involve analyses of event rates, estimation of risk over time, exploration of possible subgroup differences, and identification of risk factors associated with serious events. These analyses require substantial knowledge of methods of validly quantifying risk and providing measures of uncertainty. Reviewers should also consider when it may be appropriate to conduct a joint statistical and clinical review for particularly important safety issues.

## 1.2 Definitions

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

- ***Adverse dropout*** refers to a subject who did not complete the study/clinical trial because of an adverse event, whether or not considered drug-related; adverse dropouts include subjects who received the test drug, reference drugs, or placebo.
- ***Adverse event*** refers to any untoward medical event associated with the use of a drug in humans, whether or not it is considered drug-related.
- ***Adverse reaction*** refers 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 or after 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.
- ***Serious adverse drug experience*** and ***serious adverse event*** refer to U.S. regulations (21 CFR 312.32(a); 314.80(a)), and may include any event occurring at any dose, whether or not considered drug-related, that results in any of the following outcomes:

---

<sup>2</sup> 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/clinical trials in the conduct of that review. For the purpose of this GRP, 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/clinical trials or at datasets resulting from pooling of certain studies/clinical trials 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.

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

Appropriate medical judgment of an important medical event may find that the subject was jeopardized and required medical or surgical intervention to prevent one of the outcomes listed in this definition. Such an event may be considered a serious adverse event although it did not result in death, was not life-threatening, and did not require hospitalization. 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.

- **Severe adverse event** refers to an adverse event that has been categorized to be of great intensity (as opposed to *mild* or *moderate*).
- **Significant adverse event** refers to the definition in the ICH guidance for industry *E3 Structure and Content of Clinical Study Reports* that includes: Marked haematological and other laboratory abnormalities (other than those meeting the definition of serious) and any events that led to an intervention, including withdrawal of test drug/investigational product treatment, dose reduction, or significant additional concomitant therapy, other than those reported as serious adverse events.<sup>3</sup>

### 1.3 Principal Tasks of the Safety Review

The safety review has four principal tasks:

- (1) 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 evaluation includes assessments of whether the extent of exposure at relevant doses is adequate.
- (2) To identify and closely examine serious adverse events that could signal 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.
- (3) 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.

---

<sup>3</sup> See <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm065004.htm>.

- (4) To identify unresolved safety concerns that will need attention before 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 subject-related factors (e.g., age, sex, ethnicity, race, target illness, abnormalities of renal or hepatic function, comorbid illnesses, genetic characteristics (e.g., 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 (e.g., 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 in labeling

#### **1.4 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 clinical trials in phase 2 and phase 3 of a drug development program are directed toward establishing effectiveness. When designing these clinical trials, applicants identify critical efficacy endpoints in advance, estimate sample sizes to permit an adequate assessment of effectiveness, and make serious efforts to plan interim data review or to control multiplicity to control the probability of a type 1 error (alpha error) for the main endpoint. 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 and phase 3 clinical trials are not designed to test specified hypotheses about safety nor to measure or identify adverse reactions with any prespecified 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 often will be large simple safety studies/clinical trials with primary safety endpoints that have all the features of hypothesis testing, including blinding, control groups, and prespecified statistical plans.

In standard cases, however, apparent findings emerge from an assessment of dozens of potential endpoints (adverse events) of interest, making it difficult to describe the statistical uncertainty of a finding using conventional significance levels. Therefore, the approach taken is best described as one of exploration and estimation of event rates, with particular attention to comparing results of individual studies/clinical trials 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/clinical trials 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, the serious events can occur 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 (i.e., that are not hematologic, hepatic, renal, dermatologic, or proarrhythmic) 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 postmenopausal estrogens, suicidal ideation with interferons, and cardiac valvular disorders with fenfluramine, illustrates this problem. Perhaps most difficult is the situation where the adverse event is, or could be, a consequence of the disease being treated. For example, it was extremely difficult to discover that many drugs for heart failure (e.g., 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., because of progressive heart failure or arrhythmias), that antiarrhythmics 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 rheumatoid arthritis (the drug was removed from the market shortly after approval, however, when unusual thrombotic events became apparent (e.g., thrombosis of a digital artery)). 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 events are adverse reactions is through controlled clinical trials of significant size. Sometimes, the controlled clinical trials to evaluate effectiveness will be large enough to address these issues, but sometimes, where there is a significant concern, special, large safety studies/clinical trials may be needed.

There is no simple answer to these difficult assessments, but this GRP suggests using the *prepared mind* approach of close examination of all subjects who die or who leave a study/clinical trial prematurely because of any adverse event (whether or not thought to be drug-related), with explicit consideration of the possibility that the event was drug-related.<sup>4</sup> 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 studies/clinical trials can provide a satisfactory answer and the reviewer needs to consider whether such studies/clinical 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 (ECG) predicts the occurrence of Torsade de Pointes-type ventricular tachycardia.

### **1.5 Identifying and Assembling Source Materials for the Safety Review**

Before beginning the safety review, the reviewer should identify all printed and electronic materials necessary for the review. The applicant should be contacted early in the review cycle to obtain any missing or supplementary materials. Review materials should include:

- Common technical document (CTD) safety-related sections (i.e., module 2, sections 2.5.5, Overview of Safety, and 2.7.4, Summary of Clinical Safety), which give an overview of the applicant's approach to the safety evaluation and a detailed summary of the safety data, respectively.
- The applicant's integrated summary of safety (ISS) (CTD section 5.3.5.3, Reports of Analyses of Data from More than One Study (Including Any Formal Integrated Analyses, Meta-Analyses, and Bridging Analyses)).
- Adverse event tables in the NDA/BLA submission.<sup>5</sup>

---

<sup>4</sup> 21 CFR 314.50(f)(2) requires that case report forms (CRFs) be submitted for each subject who died during a study/clinical trial or who did not complete the study/clinical trial because of an adverse event, whether believed to be drug-related or not. It should be clear from the application that the integrated summary of safety (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/pre-BLA meeting.

<sup>5</sup> 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 in lieu of reviewer-designed tables. Extra tables appended to a review for no particular purpose is strongly discouraged. If applicant-generated tables are used, the applicant should be identified as the source in the review.



- Case report forms (CRFs) for subjects who experienced death or nonfatal serious adverse events, or who dropped out of a study/clinical trial because of an adverse event. The reviewer should request these CRFs 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 large (e.g., for dropouts) and many of the events are similar, it may be reasonable to request only a sample of CRFs.<sup>6</sup> Review of CRFs for dropouts caused by, in the applicant's assessment, *other reasons* will occasionally identify an associated adverse event. CRFs for these dropouts should be included in the submission or requested of the applicant.
- Individual subject adverse reaction data listings, laboratory listings, and baseline listings, usually accessible electronically.
- The applicant's narrative summaries of deaths, serious adverse events, and other events that resulted in dropouts.
- If available, displays of individual subject safety data over time for subjects who experienced serious adverse events.
- The safety sections of the applicant's proposed labeling.
- Any other safety-related documents, such as discussions of related drugs, descriptions of use of adverse drug reactions coding dictionaries (e.g., Medical Dictionary for Regulatory Activities (MedDRA)) used to combine data across studies/clinical trials, specific studies/clinical trials of safety hypotheses.

## 1.6 Identifying Major Concerns at the Outset

It may be useful to identify at the outset of the review major concerns that may be suggested by the pharmacology of the drug or by safety concerns with pharmacologically related drugs. Thus, the metabolic and elimination pathways of a drug will suggest certain potential drug-drug interactions or certain effects of decreased renal or hepatic function. Similarly, prior experience with the pharmacologic class can lead to a focus on particular laboratory or clinical abnormalities (e.g., muscle or liver abnormalities with HMG-CoA reductase inhibitors; QT prolongation with fluoroquinolone anti-infectives; gastrointestinal, renal, and cardiovascular effects of nonsteroidal anti-inflammatory drugs (NSAIDs); liver abnormalities with endothelin receptor antagonists; cognitive impairment with sedating drugs; and sexual dysfunction with selective serotonin reuptake inhibitors (SSRIs)).

---

<sup>6</sup> The applicant's submission should allow the reviewer to easily access individual subject information. The reviewer may want to clarify formatting and accessibility concerns at the pre-NDA/pre-BLA meeting. 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 subject ID number, not just by study/clinical trial treatment or treatment assignment.

## 1.7 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 CRFs, case report tabulations, and narrative summaries for individual subjects). For example, for important adverse events, 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. In addition, reviewers can request audits by the Division of Scientific Investigation and review the reports.

## 1.8 The Purpose of Individual Case Review/Drug-Relatedness

It is important to review individual cases of death, serious adverse events, adverse events leading to discontinuation (adverse dropouts), and discontinued subjects who are lost to follow-up, for the following four reasons:

**(1) 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 a preferred term by the applicant when it summarizes the data or because an investigator used a verbatim term incorrectly when recording the event in the CRF. An example of incorrect coding is if an investigator used the verbatim term *acute liver failure* for a case of increased alanine aminotransferase (ALT) and the applicant coded the event to acute liver failure. Such a case should not be included in the numerator of a risk calculation for acute liver failure. Similarly, a case can be incorrectly excluded from a numerator. Inconsistent coding (e.g., peripheral edema coded as *heart failure* for one subject, but *metabolic abnormality* for another) can result in an inappropriately low numerator.

**(2) To determine whether there is a likely explanation for the event other than the drug that is the subject of the application.** Examples are 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.

**(3) To look for other reasons that might exclude the drug as a cause of the event.**

One example is when an adverse event occurred during a placebo washout period before exposure to drug occurred. Events that occur before 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.

- (4) To look for results of rechallenge.** A potentially important source of information about causality is when an individual is rechallenged with the 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 subject's whole course may be helpful (i.e., both challenge periods and dechallenge periods). 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. Such factors to consider are 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 clinical trial (e.g., what is in the investigator's brochure), and are not informed by awareness of the entire safety database. Generally, these analyses are not expected to provide much useful information in assessing causality and should be disregarded.

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 clinical trial drug group to that in the placebo (or other control) group. 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 that does not provide unique efficacy or some other advantage over available treatments.

## 7. REVIEW OF SAFETY

The following sections bear the same names and numbers of section 7, Review of Safety, in the clinical review template, and organize the safety review into eight main sections:

- (1) Safety Summary (unnumbered section located below the main section 7 heading)
- (2) Section 7.1, Methods
- (3) Section 7.2, Adequacy of Safety Assessments
- (4) Section 7.3, Major Safety Results
- (5) Section 7.4, Supportive Safety Results
- (6) Section 7.5, Other Safety Explorations
- (7) Section 7.6, Additional Safety Explorations
- (8) Section 7.7, Additional Submissions/Safety Issues

**When presenting analyses in the safety review, reviewers should distinguish between their own analyses and conclusions and those of the applicant's.** Reviewers should avoid cutting and pasting information such as tables and figures directly from the application into the body of the review.

**The annotated outline of the review begins here.**

### Safety Summary

The safety summary is written as a narrative with reference to the appropriate numbered subsections. The summary is not intended to be exhaustive. It is considered a comprehensive discussion of safety findings prioritized to cover the **major safety issues and critical concerns**. The summary is also a place to note if any subsections were excluded from the review. The 120-day safety update data are expected to be integrated within these subsections, particularly sections 7.3 through 7.6. This summary is intended to be flexible and should contain at a minimum the following topics:

- A discussion of the adequacy of exposure with regard to the size of the safety database and the duration of exposure.
- A **safety issue problem list**. The purpose of this problem list is to convey the key data about safety issues that may affect approval, result in a risk evaluation and mitigation strategy (REMS), be discussed in labeling, need additional data collected in the postmarketing period such as a postmarketing requirement, or require monitoring after approval through a Document Archiving, Reporting and Regulatory Tracking System safety application. Each safety issue on the problem list should have the key adverse event and measurement data summarized (e.g., laboratory, vital sign, or ECG), including references to the relevant subsections of the review.

- A list of *pertinent negatives* (e.g., no cases of blood dyscrasias, serious skin rashes, Torsades de Pointes (TdP), drug-induced liver injury).
- Safety concerns suggested by the safety database but not worrisome enough to warrant inclusion on the problem list. These concerns are areas of uncertainty that may need to be resolved by acquisition of additional data, performing additional analyses, and/or monitoring in the postmarketing period.
- A discussion of any subsections excluded from the safety review and reasons for exclusion.
- Overall conclusions about the safety of the drug:
  - Overall assessment of the adequacy of the available safety information
  - The limitations of the available data, including analyses that would be important if the data existed
  - Additional information needed, including both additional analyses and additional studies/clinical trials
  - Comparison, to the extent possible, of the safety of the drug under review to the safety of other available drugs, and the basis for that comparison (direct comparative data versus inference)
  - Whether a REMS is needed and why

The following are several sample formats for a safety issue problem list entry.

#### **Example Safety Issue #1: QT Prolongation**

*Dose-related QT prolongation compared to control was seen in all controlled clinical 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 subjects 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 4.3, Nonclinical Pharmacology/Toxicology, for discussion of the animal models used to evaluate effects on K channels, and QT prolongation*
- *See section 7.3.1, Deaths, for discussion of deaths that may be related to QT prolongation and detailed discussion of the finding*

- See section 7.4.4, *Electrocardiograms (ECGs)*, for discussion of ECG changes
- See section 7.4.5, *Special Safety Studies/Clinical Trials*, for dose-response clinical trial of QT prolongation (doses of 10, 40, and 100 mg)
- See section 7.2.5, *Metabolic, Clearance, and Interaction Workup*, 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 9, *Appendices*, at page \_\_\_ for published articles about similar drugs and methodological suggestions
- Subject ID numbers for possibly relevant deaths: \_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_

As the example safety issue #1 QT prolongation shows, it is useful to identify the various sections of the rest of the clinical review that can be referenced for additional details about an identified adverse event. If the review is converted into a portable document format (PDF) file, internal HTML electronic links can be used in the problem list to link to earlier sections of the review.

As an alternative and to streamline the safety summary if there are many issues to highlight, the use of brief narrative text can be used as shown in the following examples.

### **Example Safety Issue #2: Application Site Reactions**

*Drug X-related application site reactions are unique to the class of dopamine agonists used to treat Parkinson's disease. Application site reactions were common, but led to discontinuation of a small percentage of users and rarely were serious adverse events. Most subjects' reactions resolved following discontinuation of drug X. Data from a clinical pharmacology clinical trial suggest that sensitization to drug X can occur. These application site reactions may limit the ability of patients to continue treatment and have led to recommendations that increase the complexity of use (recommendations to rotate patch site and not reapply to the same site for 14 days). The applicant provided literature references supporting that application site reactions are seen commonly with other drugs administered by patch.*

### **Example Safety Issue #3: Sleep Attacks**

*One of the most concerning safety issues with drug X is the increased risk of sleep attacks. Sleep attacks or sudden onset of sleep are somewhat unique in that they are potentially harmful not only to the treated patient but, depending on the circumstances, to the general public as well. Sleep attacks were reported for 1.3 percent of the drug X-*

*treated population and in controlled clinical trials, and 1.4 percent of drug X subjects. No placebo subjects experienced sleep attacks. This risk seems high but active comparator data from the controlled clinical trials found a risk of sleep attacks of 1.8 percent for drug Y. Historical comparisons to NDA data for recently approved dopamine agonists are not useful because drug X sleep attacks were recognized as related to dopamine agonist treatment and prospectively designated events of special concern. At the time when drug Y and drug Z were being developed, sleep attacks were not yet recognized as related to treatment and were not prospectively identified events of concern.*

## **7.1 Methods**

This section should describe the relevant data sources that were in the safety pool (i.e., all primary safety data sources), and the applicant's general methods used to evaluate safety. The reviewer should comment on whether the data sources used in the safety assessment were adequate, and whether the applicant's methods were appropriate. This section should discuss in detail the pooled studies/clinical trials used for the safety analyses, and the applicant's logic in pooling various groups of studies/clinical trials, if applicable. The applicant's categorization of data (i.e., whether or not the safety data were appropriately coded) should be discussed as well.

### **7.1.1 Studies/Clinical Trials 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 sources are described elsewhere in the review, such as in section 5, Sources of Clinical Data, this section can reference those sections or be omitted to avoid repetition. (This section is especially useful for reviewers who are completing only a safety review and did not complete section 5. For those reviewers, a list of studies and clinical trials that make up the safety database is helpful for the reader to have when discussing the studies/clinical trials further in the review.)

Generally, the primary safety data source is the database derived from the applicant's development program. Studies/clinical trials in this program generally will have full study/clinical trial reports related to safety, or studies/clinical trials that are grouped for analysis of safety in an ISS; CRFs will be available. These studies/clinical trials usually will have been closely monitored. Secondary sources also may 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).

Secondary source data are: (1) data derived from studies/clinical trials not conducted under the applicant's investigational new drug application (IND) and for which CRFs and full final reports are not available,<sup>7</sup> or studies/clinical trials so poorly designed or

---

<sup>7</sup> If CRFs are available from any such clinical trials and the data quality is comparable to that of data from clinical trials conducted under the applicant's IND, these data ordinarily would be included in the primary source database.

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/clinical trials 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.

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.

Reviewers should be able to determine from the NDA/BLA exactly what other studies/clinical trials provided data and what the basis was for not integrating such data with the primary source data (e.g., no CRFs, no final reports, not adequately monitored).

#### Study/Clinical Trial Type and Design/Subject Enumeration

When expanding on safety populations and considering the differences between the safety pool and the efficacy pool, the reviewer should include a table, such as that illustrated in Table 1: Enumeration of Subjects for New Drug Development Program (see Sample Tables). This is a critical table that identifies the important subject pools and denominators for subsequent analyses and incidence estimates. An NDA/BLA generally includes data from subject samples that are at different levels of completeness in terms of data entry, information collected, and validation. Table 1 should include subject counts (or estimates) from all studies/clinical trials contributing data, regardless of these factors.

The reviewer should also include a table that provides brief descriptive information for all individual studies/clinical trials, including study design (e.g., case-control, cross-sectional) and/or clinical trial design (e.g., fixed dose versus flexible dose, parallel versus crossover), dosing schedule, study/clinical trial location (e.g., foreign versus domestic), treatment groups and doses, sample sizes, and patient population (e.g., elderly). Studies/clinical trials that were designed to assess a particular aspect of safety should be noted (e.g., ECG, ophthalmic).

Applicants sometimes segregate certain studies/clinical trials from their primary source data, especially foreign data. This separation 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 were excluded from the primary source data.

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/clinical trials might be more distant, whereas the cutoff date for submitting serious adverse events from all studies/clinical trials may be more recent. These dates may likely need updating during the course of NDA/BLA review as more data become available.

### 7.1.2 Categorization of Adverse Events

Although investigator adverse reaction terms are provided as part of final reports and are listed in case report tabulations, the integrated analysis of the ISS requests the applicant to use some way of grouping closely related events to obtain an overall rate for a category of events. This grouping is accomplished by using a *dictionary* of preferred terms, such as Coding Symbols for Thesaurus of Adverse Reaction Terms (COSTART),<sup>8</sup> or MedDRA,<sup>9</sup> the latter a more granular listing developed under the auspices of ICH. These dictionaries are in fact lists of preferred terms and leave 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. 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.

When assessing the applicant's coding of events, the reviewer should compare the applicant's preferred terms to the verbatim terms used by investigators and subjects, for at least a selected subset of events. This assessment should focus on the events leading to dropouts, other changes in treatment, or serious adverse events. The applicant will usually provide the following tables and listings for assistance with this assessment (ideally this will have been agreed to at the pre-NDA/pre-BLA meeting, but if not, it should be sought early in the review); they should be provided in a form the reviewer can manipulate, such as a SAS transport file, not just in PDF:

- Adverse event tables in individual study/clinical trial reports based on investigator terms for events.
- A comprehensive line listing of all adverse events in phase 2 and phase 3 studies/clinical trials with a column containing investigator terms coded under a preferred term (see Table 2: Treatment Emergent Adverse Event Listing). This table is for reference only; it would not be included in the review.
- A listing of preferred terms and the investigator and subject 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.

---

<sup>8</sup> Food and Drug Administration, 1995, Coding Symbols for Thesaurus of Adverse Reaction Terms, 5<sup>th</sup> ed., FDA, Rockville, MD.

<sup>9</sup> See <http://www.meddrasso.com>.

The reviewer should consider the following:

- Whether the 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 the terms lack a commonly understood meaning (e.g., mouth disorder, tooth disorder, gastrointestinal 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 the terms exaggerate a finding (e.g., acute liver failure for a transaminase elevation) or minimize the importance of an event (e.g., 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.

Generally, a subset of adverse event terms should be evaluated in detail. Focus on instances in which important events may be concealed when a preferred term subsumes verbatim terms. 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 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.

### 7.1.3 Pooling of Data Across Studies/Clinical Trials to Estimate and Compare Incidence

#### Pooled Data vs. Individual Study/Clinical Trial Data

Before estimating the incidence of adverse events, the reviewer must select the subject sample of interest. Pooling data from different studies/clinical trials 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/clinical trials. 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/clinical trials. 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/clinical trials that are of similar design, that is, similar in dose, duration, choice of control, methods of ascertainment, and population (checklist versus general inquiries versus no prompt at all). (In psychiatric drug trials it is typical for obsessive compulsive subjects to spontaneously report adverse events more frequently than schizophrenic subjects. It is also possible that different populations may have different vulnerabilities to a drug, and therefore, different risk profiles.) When the studies/clinical trials 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/clinical trials being pooled. For a specific adverse event, if the incidence differs substantially across the individual studies/clinical trials in a pool, the pooled value should not be used because it is probably not meaningful and, in some cases, can obscure important information about predictors for that event. In one case, for example, several studies/clinical trials were combined and a reassuringly low estimate of phototoxicity was obtained. Subsequent examination of individual study/clinical trial results found one clinical trial with a substantial rate of phototoxicity. The clinical trial was the only outpatient clinical trial done (i.e., the only one in which subjects had an opportunity to be exposed to sunlight). In some situations, the incidence may be best described by the range in the various studies/clinical trials. For the phototoxicity example, however, the most relevant data are those from the outpatient clinical trial, the only clinical trial that was conducted under conditions pertinent to intended use.
- In some cases, observed differences in rates in various studies/clinical trials can be explained (e.g., better ascertainment, different populations), so that a consistent rate can be determined from a subset of studies/clinical trials.
- Formal tests for extreme values may be useful when assessing appropriateness of assay pooled data (e.g., test of heterogeneity such as the Breslow-Day chi square test can 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/clinical trials, or use a graphic display of incidence by study/clinical trial to informally consider the extent of variability and to identify

outliers; outliers may be important in identifying subgroups of subjects who are at particular risk for certain adverse reactions.

### Combining Data

When pooling data, a common practice is to simply combine the numerator events and denominators for the selected studies/clinical trials. However, statistical expertise should be sought to ascertain whether this is appropriate for the particular studies/clinical trials under consideration. Other more formal weighting methods may be required (e.g., weighting studies/clinical trials on the basis of study/clinical trial 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.

## **7.2 Adequacy of Safety Assessments**

This section 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 subjects and in appropriate demographic subsets of subjects? Were doses and durations of exposure appropriate? Were all (or not all) appropriate tests performed in the exposed subjects? 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 (e.g., to what extent was psychomotor impairment specifically assessed in a drug that is sedating)?

Important data that are missing can 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 after approval. 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 describe the studies/clinical trials 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.

Reviewers may find the following guidances to be useful for this part of the review: the ICH guidance for industry *E1A The Extent of Population Exposure to Assess Clinical Safety: For Drugs Intended for Long-Term Treatment of Non-Life-Threatening Conditions* and the guidance for industry *Premarketing Risk Assessment*.<sup>10</sup>

---

<sup>10</sup> See <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm>.

## 7.2.1 Overall Exposure at Appropriate Doses/Durations and Demographics of Target Populations

### Adequacy of Overall Clinical Experience

When 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 as well as the guidance for industry *Premarketing Risk Assessment*.<sup>11,12</sup> 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/clinical trials (e.g., open, active-controlled, placebo-controlled) was adequate to answer critical questions.
- Whether potential class effects were evaluated (e.g., for antiarrhythmic effects, evaluation of the potential for proarrhythmic effects) and whether problems suggested by nonclinical data were assessed.
- Whether subjects excluded from the study/clinical trial limit the relevance of safety assessments (e.g., diabetics, people older than 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 subjects who were studied.

### Demographics

The reviewer should include appendix tables in a format similar to that illustrated in Table 3: Demographic Profile, and Table 4: Number (Percent) of Subjects Receiving New Drug.

---

<sup>11</sup> The ICH guidance for industry *E1A The Extent of Population Exposure to Assess Clinical Safety: For Drugs Intended for Long-Term Treatment of Non-Life-Threatening Conditions* 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 healthy populations, where only a small fraction of patients will benefit).  
(<http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm065004.htm>)

<sup>12</sup> See <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm>.

Tables that provide overall demographic information for phase 1 (when appropriate, omitting healthy volunteers, single-dose clinical trials, and others) and phases 2 and 3 clinical trial pools separately can be included (or referenced from section 5.3, Discussion of Individual Studies/Clinical Trials, or section 6, Review of Efficacy, of the clinical review template). It may be appropriate to provide demographic displays for subsets within these larger pools at other points in the review (e.g., a discussion of geriatric exposures). The reviewer should distinguish the efficacy from the safety populations when there are notable differences.

### 7.2.2 Explorations for Dose Response

There are many ways to summarize the dose and duration experience with a new drug. Either can be expressed as mean, median, or 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 1. If the clinical trial used a titration design, the modal dose may be the more useful summary statistic (if two different doses were used for the same duration, the larger, or maximal modal dose). It is particularly important to examine the subgroup of subjects who received a dose at least as large as the dose intended for marketing.

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

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

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

### 7.2.3 Special Animal and/or In Vitro Testing

The clinical reviewer should not attempt a general assessment of the nonclinical program, but rather, comment on whether nonclinical testing was adequate to explore certain potential adverse reactions, using nonclinical 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 clinical trials, there are in vitro models to evaluate this potential. The reviewer should note whether such studies were done. If such phase 1 clinical trials were conducted in humans, those results should be summarized in the clinical pharmacology review, and any associated nonclinical studies in animals should be discussed in section 4.3, Nonclinical Pharmacology/Toxicology.

#### 7.2.4 Routine Clinical Testing

The reviewer should comment on the adequacy of routine clinical testing of clinical trial subjects, including efforts to elicit adverse event data and monitor laboratory parameters, vital signs, and ECGs. When assessing the adequacy of clinical testing, the reviewer should consider the adequacy of the methods and tests used and the frequency of testing.

The reviewer should be alert to the absence of data in an NDA laboratory database for analyses that are typically included in routine laboratory monitoring. For example, it was discovered after approval that the NDA laboratory database for the antiepileptic 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.

#### 7.2.5 Metabolic, Clearance, and Interaction Workup

Knowledge of how a drug is metabolized, transported, 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 or enhancement of 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 studies 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 transporters (e.g., P-glycoprotein)
- The effect of the drug on CYP450 enzymes and/or transporters (inhibition, induction) and the effects of the drug on the pharmacokinetics (PK) of model compounds
- The major potential safety consequences of drug-drug interactions (these results should be included in section 7.5.5, Drug-Drug Interactions)

Details of these assessments will be found in the clinical pharmacology review and may be summarized earlier in section 4.4, Clinical Pharmacology. If so, the reviewer can refer the reader to that section for more details.

#### 7.2.6 Evaluation for Potential Adverse Events for Similar Drugs in Drug Class

The reviewer should discuss the adequacy of the applicant's efforts to detect specific adverse events 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 an effort 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 anticholinergics, QT prolongation with type III antiarrhythmics, extrapyramidal effects with antipsychotics, muscle pain with statins), whereas experience with other members of the class would suggest others (e.g., hepatotoxicity with thiazolidinedione peroxisome proliferator-activated receptors (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):

- Suicidal ideation (antidepressants, antipsychotics)
- 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)
- Central nervous system stimulation
- Anticholinergic activity
- Allergic reactions
- Sexual dysfunction (any antidepressant, sedating drug)
- Elevated intraocular pressure
- Cataracts
- Retinopathy
- Worsening glucose tolerance/diabetes (diuretics, atypical antipsychotics)
- Proarrhythmic effects and increased mortality (most nonbeta blocker antiarrhythmics)
- Increased congestive heart failure (CHF) and standard deviation 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 clinical trials, the reviewer should ascertain whether the applicant made efforts, beyond routine ECG testing, in phases 2 and 3 to explore the consequences in subjects 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 on subjects who experienced clinical events that may be

manifestations of TdP (e.g., syncope, dizziness, or palpitations)? Holter monitoring, for example, might have been appropriate in such subjects.

### **7.3 Major Safety Results**

This section details the specific major safety results that are required by FDA regulations, considered by ICH guidance, and/or would have a significant effect on the ability to approve the drug or result in a warning, precaution, or other major labeling change. This section consists of five subsections (e.g., death, nonfatal serious adverse events, dropouts and/or discontinuations, significant adverse events, submission-specific primary safety concerns). Each of these subsections is organized somewhat differently, depending on the content.

When discussing deaths, serious adverse events, and dropouts and discontinuations, it is critical that the reviewer identify individual subjects in a blinded way that enables subsequent readers to readily access data and supporting information if needed (e.g., study/clinical trial number, investigator number, subject ID number).

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.

#### Considerations for Uncommon but Serious Adverse Events

Causality judgments are difficult for uncommon serious events (e.g., fewer than 1/1,000) where there are, in most cases, no useful comparisons to control groups. The reviewer must form a judgment as to the plausibility of drug-relatedness for the individual cases.

- The following questions should be considered:
  - Was the subject in fact exposed to the drug and did the adverse event occur after drug exposure?
  - Did the subject have a clinical experience that meets the criteria for the adverse event of interest? (Establishing a standard case definition may be helpful here.)
  - 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 subject's underlying illness.)

- Is the adverse event of a type commonly associated with drug exposure, such as hematologic, hepatic, renal, dermatologic, or proarrhythmic events? (Also see the following 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., myocardial infarction, 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 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:
  - Whether the drug is a member of a class of drugs known to be causally associated with the event of interest
  - 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)
  - Positive rechallenge with the drug (although it would be unusual to deliberately rechallenge for a serious event, there occasionally may be inadvertent re-exposures that are informative)
- The reviewer should be cautious about dismissing uncommon serious events *that don't seem plausibly drug-related* and should consider differences in common with 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 methysergide
  - Pulmonary hypertension with aminorex (a European weight loss drug) and various other drugs

- Suicidal ideation with interferons, isotretinoin
- Intussusception with rotovirus vaccine
- Pulmonary fibrosis with amiodarone

### 7.3.1 Deaths

#### Identifying Deaths Relevant to the Safety Review

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

- During participation in any clinical trial, or during any other period of drug exposure
- After a subject leaves a clinical trial, or otherwise discontinues the drug, whether or not the subject completes the clinical trial to the nominal endpoint, if the death:
  - Is the result of a process initiated during the clinical trial or other drug exposure, regardless of when it actually occurs.
  - Occurs within a time period that might reflect drug toxicity for a subject leaving a clinical trial or otherwise discontinuing the 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 5: Deaths Listing), unless they are an effectiveness outcome.

Applicants must provide line listings and CRFs of all subjects who died in clinical trials (21 CFR 314.50(d)(5)(vi)(b) and 314.50(f)(2)). Narratives are expected for each death but are not required by the NDA regulations.

#### Distinguishing Expected Deaths From Unexpected Deaths

Certain causes of death are sufficiently unusual in the absence of drug therapy, even in large databases, that they are usually considered unexpected (e.g., aplastic anemia, progressive multifocal leukoencephalopathy, 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.<sup>13</sup> 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 clinical trials in which mortality is an endpoint and the cause of death is expected for the disease or condition.
- Deaths in clinical trials 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 clinical trials 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 clinical trial (e.g., a subject who dies from progression of cancer or Alzheimer's disease or an acute myocardial infarction attributed to underlying coronary artery disease present before clinical trial entry).
- Deaths from intercurrent long-term illness. These deaths 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 caused by 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 is by comparison with a control group (a single clinical 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 enrolled patients were chosen because they were not expected to die soon; hematologic deaths are unexpected in a postinfarction clinical trial). For unexpected deaths, the individual medical events associated with the death should be 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

---

<sup>13</sup> Note that the term *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.

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).

### Pooling of Relevant Data

Before performing any mortality analyses, the reviewer must consider the poolability of the data pertinent to deaths. If data are not poolable, analyses should be performed for separate databases and then examined together. See section 7.1.3, Pooling of Data Across Studies/Clinical Trials to Estimate and Compare Incidence.

### Overall Mortality Analysis

The review should include an analysis of overall mortality for all phase 2 and phase 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 (e.g., cardiovascular (broad) versus acute myocardial infarction, sudden death, CHF (more specific)), recognizing that assessing cause-specific mortality is difficult even in the best circumstances, such as a clinical trial in which there is an attempt to describe such endpoints prospectively.<sup>14</sup> Death from an acute myocardial infarction can be indistinguishable from death resulting from an arrhythmia.

Analyses should be corrected for differences in drug exposure using person-time in the denominator to calculate mortality rates.<sup>15</sup> 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 6: Mortality by Treatment Group). Life table approaches may be helpful in cases when there are more than a few deaths, and when the direction of different clinical trials 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)

---

<sup>14</sup> Temple, R and G Pledger, 1980, The FDA's Critique of the Anturane Reinfarction Trial, N Engl J Med, 303:1488-1492.

<sup>15</sup> Since placebo- and active-controlled subjects generally have shorter durations of exposures than subjects given the new drug, they may have less opportunity for serious events to occur.

- 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

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

- A listing of information upon which reviewer assessment is based (e.g., CRFs, narrative summaries, consultant reports, autopsy reports).
- A tabular summary of deaths. Deaths should be summarized in an appendix table, as illustrated in Table 5: Deaths Listing. 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 an identifier so that subsequent reviewers can identify the particular subject.
- An analysis of overall mortality for phase 2 and phase 3 drug exposures across treatment groups (see Table 6: Mortality by Treatment Group).
- An analysis of cause-specific mortality across treatment groups (a table similar to Table 6 can be used).
- 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 investigational 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 (e.g., 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 7.5, Other Safety Explorations).
- When deaths occur in uncontrolled clinical trials, best available estimates of mortality in the population studied, in the absence of the treatment. Compare to background incidence in the population data to compare incidence.
- When deaths are relatively frequent, the reviewer should consider some of the approaches described for common adverse events (see section 7.4.1, Common Adverse Events).

### 7.3.2 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.<sup>16</sup>

Applicants generally provide a line listing of all subjects in phase 2 and phase 3 of the development program who had an event meeting the FDA's criteria for a serious adverse event. For each such event, the applicant should also provide a brief narrative (see the ICH guidance for industry *E3 Structure and Content of Clinical Study Reports*, section 12.3.2).<sup>17</sup> 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.

This section of the review should contain the following:

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

---

<sup>16</sup> See 21 CFR 312.32(a), 314.80(a), and 600.80(a).

<sup>17</sup> See <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm065004.htm>.



- A tabular summary of serious adverse events (see sample listing, Table 7: Serious Adverse Event Listing).
- An analysis of the overall rate of serious events and the 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 subjects in the denominator, similar to the presentation for deaths in Table 6: Mortality by Treatment Group.
- The reviewer's overall assessment of which serious adverse events were probably explained by factors other than the investigational drug (e.g., another drug, underlying illness, another illness common in the population) and which could not reasonably be explained by such factors, including any pertinent information from *nonserious* events that may be related to the serious event (e.g., seizures not leading to hospitalization, all syncope).
- 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 a tabular summary, illustrated in Table 5: Deaths Listing).

If serious nonfatal adverse events are relatively frequent, the reviewer should consider some of the approaches described for common adverse events (see section 7.4.1, Common Adverse Events).

### 7.3.3 Dropouts and/or Discontinuations

FDA regulations require that the CRFs from subjects who discontinue treatment in association with an adverse event (adverse dropout) be submitted with the application (21 CFR 314.50(f)(2)) and their analysis constitutes a critical part of the safety evaluation. It is imperative that the reviewer considers how the applicant defined the terms *dropout* and *discontinuation* in the ISS and how these data were analyzed. Dropouts without adequate follow-up by the applicant may be caused by unreported adverse events. Reviewers should consider those dropouts or discontinuations of treatments by subjects to be related to possible adverse events. Any dropout or discontinuation of drug for which there is no adequate reason, poor narrative, or CRF should be pursued with the applicant for more detailed information.

### 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 phase 3 clinical trial 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 versus dropouts in other clinical trials; 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 subjects are counted only once.<sup>18</sup> Ordinarily, the dropouts should be categorized in a table or tables appended to the safety review (see Table 8: Dropout Profile). It can be useful to display, graphically or in tables, the cumulative dropout rates for each treatment group within each clinical trial, especially for cause-specific reasons, when this information is available and to assess subject baseline risk factors that contribute to differential cumulative dropout patterns. When pooling data, consideration of dropout patterns over all clinical trials may reveal information that is useful to the overall safety evaluation.

When classifying dropouts, the reviewer should 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

As mentioned, 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 the decisions.

Ordinarily, the reviewer should combine subjects categorized as dropping out for intercurrent illness and subjects 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

---

<sup>18</sup> *Mutually exclusive* refers to the reason for dropping out. Subjects should be identified with only one of the reasons. However, subjects may be represented in more than one column (treatment group) of a table (e.g., subjects in a crossover study/clinical trial may have survived several treatment arms and then dropped out).

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 clinical trial conduct or analyses (e.g., a substantial number of dropouts caused by *lost to follow-up* and sites with disproportionately high dropout rates should be a sign of concern). For example, early dropouts generally, and differential early dropouts in particular (drug group versus placebo group), 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.

#### 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 subjects 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 analysis 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 caused by adverse events are provided automatically to reviewers, these adverse dropouts may provide a clue to unexpected, but important, adverse reactions that can be easily dismissed as intercurrent illness (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). The frequency of these events is likely to be low, and the review should contain an analysis of each such adverse event that resulted in withdrawal from the clinical trial, 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 caused by a known effect of the drug might reflect an unexpected effect of the drug. The applicant usually provides a line listing of adverse dropouts (similar to Table 9: Adverse Event Dropout Listing), 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 he or she 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 9). Tables displaying incidence of events should include each event that led to a dropout even if a single subject had more than one such event.
- Whether the event can be reasonably 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 (see section 7.4.1, Common Adverse Events) and the known pharmacology of the drug.
- The dose response and time dependency of the dropouts and drug-demographic, drug-disease, and drug-drug interactions (see section 7.5, Other Safety Explorations).

For the rarer events that can 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 subject to withdraw, in which case it would represent the actual incidence of specific adverse events that led to drop out. This approach is preferred. Alternatively, a table may list the adverse events that a subject experienced at the time of drop out and not identify any event (or events) as causing the drop out. 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.

If a submission separates out information on adverse events that led to dose reduction or significant additional concomitant therapy,<sup>19</sup> but not to discontinuation of treatment, those findings should be described using an approach similar to that proposed above for adverse dropouts.

#### 7.3.4 Significant Adverse Events

This section should contain a discussion of adverse events consistent with the definition of **significant adverse events** in the ICH guidance for industry *E3 Structure and Content of Clinical Study Reports*.<sup>20</sup>

---

<sup>19</sup> This is recommended in ICH E3.

<sup>20</sup> See <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm065004.htm>.

ICH E3 defines an additional category of *other significant adverse events*. It includes:

- Marked hematological or other lab abnormalities not meeting the definition of serious. This adverse event will need to be an individual judgment, probably depending on the drug (e.g., creatinine phosphokinase elevation can 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), the guideline *Format and Content of the Clinical and Statistical Sections of an Application*, and the ICH guidance for industry *M4E: The CTD — Efficacy*).<sup>21</sup>
- Potentially important abnormalities not meeting the above definition of serious and not leading to death or modification of therapy (e.g., a single seizure, syncope episode, orthostatic symptoms).

Those adverse events that did not lead to discontinuation but otherwise meet the definition above should be described in this section. In addition, an analysis of *severe* adverse events should be included in this section. These are events that are characterized as severe in intensity, but may not reach the regulatory definition of a serious adverse event.

### 7.3.5 Submission-Specific Primary Safety Concerns

Reviews often identify a specific safety issue related to the drug that does not meet the regulatory definition of either a serious adverse event or the ICH criteria for other significant adverse events. For example, a biologic product may reveal an immunogenicity issue that might affect long-term administration. Hepatotoxicity or a QT issue might not meet the regulatory definition of serious or significant adverse event, but be notable enough at this point in the review to affect labeling or regulatory decision making and therefore not be appropriate for a lower-priority section of the safety review. Reviewers can use this section 7.3.5 to elevate and highlight a concern from a lower-numbered section in the safety review.

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. These elevated transaminase levels across the safety database may be a good subject to discuss in this section rather than later in the review under 7.4.2, Laboratory Findings. This idea is expanded upon below.

---

<sup>21</sup> See <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm>.

Example: Hepatotoxicity/Drug-Induced Liver Injury

Reviewers will find the guidance for industry *Drug-Induced Liver Injury: Premarketing Clinical Evaluation* valuable when assessing the potential for a drug to cause a severe liver injury.<sup>22</sup>

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 (e.g., iproniazid) to the present (e.g., ticrynafen, benoxaprofen, troglitzone, bromfenac) and has led to important limitations on the use of many more drugs (e.g., 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 (DILI), that a pure hepatocellular injury leading to jaundice had serious implications, a 10 to 50 percent mortality. Any Hy's Law cases should be identified in the treatment group (e.g., subjects with any elevated aminotransferase (AT) of >3x upper limit of normal (ULN), alkaline phosphatase (ALP) >2xULN, and associated with an increase in bilirubin  $\geq 2xULN$ ). Narratives and CRFs should be expected as well for these particular cases as they would be consistent with unexpected serious adverse events associated with the use of the drug.

Finding one Hy's Law case in the clinical trial database is worrisome; finding two is considered highly predictive that the drug has the potential to cause severe DILI when given to a larger population. In a drug development database, a potential for severe DILI is signaled by the following set of findings (as reproduced from the guidance):

- **An excess of AT elevations to >3xULN compared to a control group**

AT elevations to >3xULN are relatively common and may be seen in all groups, but an excess of these elevations compared to a control group is nearly always seen for drugs that ultimately prove severely hepatotoxic at relatively high rates (1/10,000). Therefore, the sensitivity of a significantly increased incidence compared to control (e.g., of >3xULN AT elevations) as an indicator of a potential for liver injury is high. But many drugs show this signal without conferring a risk of severe injury (e.g., tacrine, statins, aspirin, heparin), indicating low specificity for an excess of AT elevations alone. There are no good data to predict how great this excess incidence of AT elevations should be compared to controls to suggest an increased risk of DILI. Such an excess may not be apparent for drugs with a potential to cause idiosyncratic DILI that are used for short treatment courses, such as many antibiotics.

---

<sup>22</sup> See <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm064993.htm>.

- **Marked elevations of AT to 5x-, 10x-, or 20xULN in modest numbers of subjects in the test drug group and not seen (or seen much less frequently) in the control group**

Many, but not all, severely hepatotoxic drugs show such elevations, indicating high sensitivity for predicting severe DILI; again, however, some drugs, such as tacrine and others that are not severely hepatotoxic, also can cause AT elevations to this degree, so that specificity of this finding is suboptimal.

- **One or more cases of newly elevated total serum bilirubin to >2xULN in a setting of pure hepatocellular injury (no evidence of obstruction, such as elevated ALP typical of gall bladder or bile duct disease, or malignancy, or impaired glucuronidation capacity caused by genetic (Gilbert syndrome) or pharmacologic (treatment with atazanavir or other drugs) factors), with no other explanation (viral hepatitis, alcoholic or autoimmune hepatitis, other hepatotoxic drugs), accompanied by an overall increased incidence of AT elevations >3xULN in the test drug group compared to placebo**

## 7.4 Supportive Safety Results

For the prioritized safety review, the following sections are considered supportive to the rest of the safety results. However, there may be major issues related to laboratory findings or ECG that might be considered or already discussed as major rather than supportive. If so, the reviewer can choose to use section 7.3.5, Submission-Specific Primary Safety Concerns, for those discussions. Otherwise, this section consists of six subsections (e.g., common adverse events (including less common adverse events), laboratory findings, vital signs, ECGs, special safety studies/clinical trials, and immunogenicity). Each of these subsections is organized somewhat differently, depending on the content.

### 7.4.1 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, including less common adverse events by incidence and body system in order of decreasing frequency. NDAs typically contain numerous tables and analyses of adverse event incidence (e.g., by study/clinical trial, by various pools of studies/clinical trials, and for the overall database). In general, what are included are treatment emergent signs and symptoms (i.e., signs and symptoms not present at baseline, or not present at the severity seen on treatment). To approach these data, the reviewer should generally go through the steps outlined below.

#### 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 clinical trials (it is not usually 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 subjects were assessed, and whether the approaches differed among clinical trials. Identification of signs (abnormal findings observed by a health care provider) would seem to be less of a problem, as these are elicited by physical examination, but use of a physician exam checklist can lead to a different result from a more general requirement for physical exam. If different approaches were used (e.g., checklists in clinical trials conducted in the United States, open-ended inquiries in European clinical trials), the reviewer should consider and discuss in the review the effect, if any, on the adequacy of adverse event information collected.

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

Where events are common and occur in multiple subjects 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 subjects treated with the test drug and in control groups. If an event is more frequent with the test drug than the control, it can be attributed to treatment with the test drug.

#### Incidence of Common Adverse Events — Assessment of Various Databases

Applicants typically prepare a wide variety of tables of adverse event rates for individual clinical trials and pools of various clinical trials. Those tables generally include investigator causality assessments and severity ratings. Incidence rates for common adverse events may be estimated from the relatively small portion of the overall database that is contained in the controlled trials (especially placebo-controlled). For these more common events, the ability to compare rates experienced while taking the investigational drug with those reported during treatment 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 phase 3 database that will provide the best estimate of rates and develop tables of event rates based on that judgment. It is important to ensure that similar events are not divided into categories that dilute their effect. For example, although there may be subtle differences between *azotemia*, *BUN elevated*, and *acute renal insufficiency*, they largely reflect the same process and probably should be considered together. The data will be less informative if the events are considered separately.



- If possible, the reviewer should rely on pooled data from clinical trials 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 10: Treatment Emergent Adverse Event Incidence). 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 clinical trials (drug at a particular dose and placebo), but the others (active controls, individual dose groups) may also be useful (see section 7.1.3, Pooling of Data Across Studies/Clinical Trials to Estimate and Compare Incidence, for a broader discussion of pooling).
- If there are not adequate numbers of subjects in such trials to give meaningful rate estimates, the reviewer should consider pooling placebo-controlled trials, active-controlled 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 clinical trials, or even individual clinical trials, 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 subjects experiencing an adverse event in a clinical trial (or the absolute number of adverse events experienced in a group), without regard to the duration of treatment received. This approach is often satisfactory for relatively short-term clinical trials. If clinical trials 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 (number of subjects with adverse event total person-time exposure), rather than the risk (number of subjects with adverse event total number of subjects). This approach 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 subjects, 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 risk-benefit conclusion for the drug. However, for events with high background rates (e.g., headache, fatigue, and other events that occur frequently independent of drug therapy), display of all reported events can result in a high event rate that obscures drug-relatedness. This result can be a particular problem when time on drug is prolonged. For example, it is common for clinical trials of 4 to 6 weeks duration to report headache at a high rate (20 to 25 percent). 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/clinical trials intended to identify these events, they generally should be given more credence than nontargeted studies/clinical trials, which tend to substantially underestimate rates (see section 7.4.5, Special Safety Studies/Clinical Trials). Incidence rates should be based on findings from the targeted studies/clinical trials.

### 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 large databases as with less common adverse events; see below). The table (or tables) will form the basis for the adverse reaction table in labeling. The table may use a higher cutoff 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 cutoff for inclusion of adverse events in the table is inherently arbitrary (e.g., greater than 1 percent). If a cutoff is used, the review should explain how the threshold was determined. It also may 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 also may be useful.

### Identifying Common and Drug-Related Adverse Events

For common adverse events, the reviewer should attempt to identify those events that can be reasonably considered drug-related. Although it is tempting to use hypothesis-testing methods, any reasonable correction for multiplicity would make a *finding* untenable. The most persuasive evidence for causality is a consistent difference from control across

clinical trials, and evidence of dose response (population-based assessment rather than case-based assessment). The argument for causality is strengthened when there is a similar safety issue in a related drug, a related finding in nonclinical studies, and/or biological plausibility.

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.

#### Less Common Adverse Events

In general, a 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 phase 3 database, including data for which there are no useful concurrent controls. 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 1/1,000
- Adverse events estimated to occur at rates less than 1/1,000

The reviewer should then develop a condensed list of reactions to be included in the ADVERSE REACTIONS section of labeling. 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 7.1.2, Categorization of Adverse Events).

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, and therefore included in the WARNINGS AND PRECAUTIONS section of labeling. In that case, it is useful to notify the safety evaluator in the Office of Surveillance and Epidemiology who will be monitoring the drug after marketing.

## 7.4.2 Laboratory Findings

The approach to review of laboratory findings (e.g., 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 trials, describe the dataset from which laboratory findings information is obtained, describe the 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 7.3.2, Nonfatal Serious Adverse Events) need not be discussed again in detail in this section. This section should refer to the more detailed discussions of such findings elsewhere in the review.

The following lower-level subsections should be included in the review, but **the reviewer may choose to combine sections or highlight sections where necessary to focus the discussion on significant results**. Also, many of the subsections below may not apply or adequately express significant results; therefore, **the reviewer should choose the best way to present the data and avoid cutting and pasting of results to fit these suggested subsections**. Reviews should not be padded with extraneous tables filled with normal results for the sake of filling out these sections. Reviewers may optionally label these lower-numbered sections (7.4.2.1, 7.4.2.2) when necessary, but should avoid fragmenting discussion and should not add these subsections to the table of contents. If needed, unnumbered headings (such as the underlined headings below) may be created and are preferred.

### Overview of Laboratory Testing in the Development Program

The review should provide an overview of what laboratory testing was carried out (e.g., chemistry, hematology, and urinalysis). It is preferable to summarize the overall approach, rather than provide detailed comments about laboratory testing for each clinical trial. 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 subjects were followed until their values normalized, whether any subjects were rechallenged, the procedures used for sample analysis (i.e., central or local labs, *windows* of time in which lab values were considered<sup>23</sup>))
- A summary table identifying the number of subjects exposed to the test drug who had baseline laboratory values and follow-up assessments

---

<sup>23</sup> 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.

- 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 subject 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., blood urea nitrogen (BUN)/creatinine). In such cases, the laboratory data of interest should be requested from the applicant.

#### Selection of Clinical Trials and 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 and unsuitable for assessing late-developing abnormalities; therefore, 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). When identifying the sample population for comparison of laboratory values, the reviewer should pool relevant clinical trials. The review should explain how the clinical trials to be pooled were selected. When comparing laboratory values, there are additional considerations when using pooled data including:

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

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

#### Standard Analyses and Explorations of Laboratory Data

In general, this review should include three standard approaches to the analysis of laboratory data, noted as: (1) Analyses Focused on Measures of Central Tendency; (2) Analyses Focused on Outliers or Shifts From Normal to Abnormal; and (3) Marked Outliers and Dropouts for Laboratory Abnormalities. The first two analyses are based on comparative trial data. The third analysis should focus on all subjects in the phase 2 to phase 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 data should be thought of as descriptive. Generally, the magnitude of change is more important than the p-value for the difference. The reviewer should recognize that there are key differences between laboratory values, and these differences affect how they are best analyzed. For example, elevation of hepatic transaminases to 50 percent above the upper limit of normal may have some clinical significance, whereas elevation of serum sodium to this level is universally fatal.

**It may be best to prioritize the analyses that best illustrate the data.** For example, if measures of central tendency tend to be *normal* but shift from normal to abnormal or marked outliers reveal more significant results, these analyses should be presented first.

#### 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 in inadequately hydrated patients (e.g., ticrynafen, suprofen). 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 11: Mean Change From Baseline for Serum Chemistry Parameters). The reviewer should note and discuss signals that emerge from these tables and indicate those for which additional clinical trials are needed, if any.

#### Analyses Focused on Outliers or Shifts From Normal to Abnormal

The review should focus on subjects 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 subjects across treatment groups who had a potentially clinically important deviation from normal on one or more laboratory parameters while on treatment (see Table 12: Incidence of Potentially Clinically Significant Changes in Serum Chemistry Parameters). When analyzing outliers, the reviewer should be aware of the following:

- Regression to the mean (and an apparent upward shift) can be expected if subjects 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 taken 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 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 nonhepatotoxic drugs; 3-fold and higher elevations appear to be more discriminating).

Decisions about what criteria to use to identify outliers should have been made at the pre-NDA meeting, if possible. Because it is not possible to know in advance what criteria will be optimal for detecting between-group differences, it may be useful to perform 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 subjects meeting some arbitrary criterion
- Subjects with large shifts within the normal reference range
- Subjects who meet outlier criteria for more than 1 variable simultaneously (e.g., transaminase and bilirubin)
- Subjects 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 the review as appropriate.

#### Marked Outliers and Dropouts for Laboratory Abnormalities

The reviewer should analyze individual subjects 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 subjects with extreme changes, usually specified in advance. Individual subject data displays

should be available to the reviewer for all such subjects. 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 white blood count, may signal major problems), all such dropouts in the phase 2 to phase 3 population should be identified. The reviewer should generally analyze and comment on each individual subject 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 nondrug-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).

#### 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 7.5, Other Safety Explorations). The review should discuss the rationale for additional explorations, the methods used, and the results and interpretations.

#### Special Assessments

Certain laboratory assessments are so critical that they deserve special attention in any review. For example, hepatotoxicity has been an important cause of market withdrawal since the 1950s and may deserve a special assessment in this section, or if the safety issues are important, these particular issues can be elevated to section 7.3.5, Submission-Specific Primary Safety Concerns.

#### 7.4.3 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. As in section 7.4.2, Laboratory Findings, the reviewer should choose the best way to present the data and avoid copying of results to fit these suggested subsections.

Analyses can be combined or reordered if they help the review flow better, but overall the reviewer should include all the necessary information similar to laboratory testing above to express significant data to the reader.



Overview of Vital Signs Testing in the Development ProgramSelection of Clinical Trials and Analyses for Overall Drug-Control ComparisonsStandard Analyses and Explorations of Vital Signs DataAnalyses Focused on Measures of Central TendencyAnalyses Focused on Outliers or Shifts From Normal to AbnormalMarked Outliers and Dropouts for Vital Signs AbnormalitiesAdditional Analyses and Explorations

## 7.4.4 Electrocardiograms (ECGs)

ECG data can be analyzed and reported using an essentially identical approach to that taken for laboratory data, with the reviewer choosing the best way to present the data. The adequacy of this safety assessment may be especially important in this case (see section 7.2, Adequacy of Safety Assessments), given recent experience with drugs that prolong the QT interval and cause the ventricular tachycardia known as TdP. Reviewers are referred to ICH guidance for industry *E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs* as a good resource.<sup>24</sup> 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 clinical trials designed specifically to assess the effects of the drug on the QT interval. Thorough QT clinical trials may be discussed in section 7.4.5, Special Safety Studies/Clinical Trials.

Overview of ECG Testing in the Development Program, Including Brief Review of Nonclinical Results

This section should describe the number of baseline and on-clinical trial ECGs obtained, who read the ECGs, and what methodology was used (e.g., automatic, blinded cardiologists). Abnormalities in cardiac rhythm and conduction should be described, if detected. Dropouts for ECG abnormalities should be highlighted.

Selection of Clinical Trials and Analyses for Overall Drug-Control ComparisonsStandard Analyses and Explorations of ECG DataAnalyses Focused on Measures of Central TendencyAnalyses Focused on Outliers or Shifts From Normal to AbnormalMarked Outliers and Dropouts for ECG AbnormalitiesAdditional Analyses and Explorations

---

<sup>24</sup> See <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm065004.htm>.

#### 7.4.5 Special Safety Studies/Clinical Trials

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

- Clinical trials to assess whether a drug has safety concerns common to its pharmacologic class (e.g., a clinical trial to assess effects of a benzodiazepine hypnotic on driving, respiration, memory, or next-day psychomotor functioning).
- Clinical trials in topical products (including systemic products delivered by a patch) to assess cumulative irritancy, contact sensitizing potential, photosensitivity, and photoallergenicity.
- Clinical trials to characterize a drug's effect on QT interval; the thorough QT clinical trial should be described. Reviewers should consult the ICH guidance for industry *E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs*.<sup>25</sup>
- Clinical trials intended to demonstrate a safety advantage over therapeutic alternatives (e.g., less extrapyramidal effect for an antipsychotic, less sedation for an anti-histamine, less cough from an angiotensin II blocker than from an ACE inhibitor). Such clinical trials should include the comparator drug (a failure to see the side effect in a placebo-controlled clinical trial is usually not informative without the active control to demonstrate assay sensitivity).
- Studies/clinical trials in special populations thought to be at increased risk and likely to use the drug.

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

#### 7.4.6 Immunogenicity

Data on the effect 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 can result in one or more of the following outcomes:

---

<sup>25</sup> See <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm065004.htm>.

- For a drug that is intended as a replacement for a missing endogenous substance, antibodies can neutralize the replacement drug and generate a clinical deficiency syndrome
- Neutralization of a protein drug by *blocking* antibodies can reduce the efficacy of a life-saving drug
- Antibody development can 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 is administered via the subcutaneous route (more likely if it is), and whether the protein is intended for chronic use (more likely if it is). This section of the review should assess the adequacy of the immunogenicity data provided to address these issues.

## 7.5 Other Safety Explorations

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, and concomitant illnesses, are considered below. In general, these explorations are meaningful only for adverse events that appear to be drug-related.

### 7.5.1 Dose Dependency for Adverse Events

If data from randomized, parallel, fixed-dose clinical trials (or data from clinical trials in which subjects 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<sup>2</sup>, to decrease the effect of size or weight differences on drug exposure. 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 clinical trial 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 Clinical Trials

Although it is tempting to try to extract dose-response or plasma level-response data from flexible dose (titration) clinical trials, and the ICH dose-response guideline encourages this,<sup>26</sup> there are many potential problems with such analyses. In particular, many adverse

---

<sup>26</sup> See the ICH guidance for industry *E4 Dose-Response Information to Support Drug Registration* (<http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm065004.htm>).

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 anticholinergic or sedating reactions, the drug is dose-adjusted to toxicity, which will often obscure any dose-response relationship. In addition, if only subjects who have not had adverse effects are given increased doses (i.e., subjects who are resistant to adverse effects), 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 (erythropoiesis), analysis of adverse events by dose or cumulative dose before 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.

#### 7.5.2 Time Dependency for Adverse Events

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

#### Time of Onset

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

subjects (e.g., two subjects each exposed for 6 months = 1 subject-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 subjects though it is still occurring at the same rate, and reduced dose or dropping out in subjects 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 subjects 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 can be compared to a similar cohort of placebo recipients who experienced the same event at baseline. The same approach can be used for adverse events occurring later in treatment. It is also important to evaluate not only duration of the adverse event, but degree of resolution of the adverse event — particularly important for toxic cyclic therapies, such as chemotherapy.

### 7.5.3 Drug-Demographic Interactions

Numerous methods can be used to analyze age, sex, and race implications for safety, and applicants should present analyses of safety information for these population subsets. In most cases, there will be PK 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 drugs 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 sex-weight subgroups can identify the group at greatest risk of bleeding (thin women).

It may also be useful to consider age-sex or race-sex 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 greater rate on drug than placebo. In small clinical trials or for low frequency events, usually there is insufficient power to detect differences between groups, so that these analyses will 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 versus women, old versus young, black versus white); because 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 (cumulative risk on drug/cumulative risk on comparison

drug or placebo); and (2) evaluation of attributable risk (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 for both sexes are 3 and the relative risk for men and women is 1 (no difference), yet the attributable risk is much greater for women than men (6 percent versus 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 sex difference, even though the risk ratios are the same.

#### 7.5.4 Drug-Disease Interactions

The reviewer should be aware of the possibility that comorbidity will affect the adverse reaction profile of the drug (i.e., a drug-disease interaction). Such interactions can arise from abnormalities of excretory or metabolic function (e.g., renal or hepatic disease), and typically, the applicant will have carried out formal PK clinical trials in subjects with hepatic or 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 (PD) differences). In general, the same methods described for exploring drug-demographic interactions can be applied here.

#### 7.5.5 Drug-Drug Interactions

The reviewer should be aware of the potential of drug-drug interactions to affect the safety profile of the drug. Again, these interactions can be either pharmacokinetic (affecting the absorption, distribution, and/or 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 clinical trials to evaluate potential PK effects of concomitant therapy on drugs metabolized by CYP450 enzymes, but PK interactions can also occur through effects on oral absorption, renal excretion and transport (e.g., P-glycoprotein) proteins. True PD 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.

### 7.6 Additional Safety Explorations

NDA's may or may not have particular issues related to the following subsections, or may have greatly expanded discussions. For subsections that have been combined, but require separation because of lengthy discussion (such as section 7.6.4, Overdose, Drug Abuse

Potential, Withdrawal and Rebound), the reviewer may split each main section out with additional subsection numbers, but should not add these subsections or create any lower-numbered subsections to the table of contents of the review (i.e., 7.6.4.1, Overdose, 7.6.4.2, Drug Abuse Potential).

### 7.6.1 Human Carcinogenicity

Although formal clinical trials in humans of the carcinogenic effects of drugs and biologics are uncommon, because of the expectation that induction of cancer would occur over a 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/clinical trials 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.

### 7.6.2 Human Reproduction and Pregnancy Data

Although formal clinical trials 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.<sup>27</sup>

### 7.6.3 Pediatrics and Assessment of Effects on Growth

Increasingly, reviewers are presented with analyses of height and weight data collected during clinical trials 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 the drug on growth requires accurate measurements, particularly for height, and in most clinical trials, height is not measured accurately. Growth is a process that occurs over a long period 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 clinical trials 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 in the laboratory data section. For example, analysis of changes in central tendency and outlier analysis apply to the evaluation of the effect of a drug on growth. There are, however, some distinctive issues that must be considered.

---

<sup>27</sup> See the draft guidances for industry *Pharmacokinetics in Pregnancy — Study Design, Data Analysis, and Impact on Dosing and Labeling* and *Clinical Lactation Studies — Study Design, Data Analysis, and Recommendations for Labeling* (<http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm064982.htm>). When final, these guidances will represent the FDA's current thinking on their respective topics.

First, a description of how the height and weight were measured should be available. 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 staff were trained in measurement will result in more reliable data than a development program that did not standardize procedures. The review should 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 height and weight 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 against 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 clinical trials are less easily interpreted because these scores can result from the effect of the drug or can be caused by the disease for which the treatment is being studied.

Applicants 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 clinical trials of the effect of drug on growth in children. The review team should have requested such analyses at the pre-NDA meeting.

#### 7.6.4 Overdose, Drug Abuse Potential, Withdrawal and Rebound

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 clinical trial phases. In addition, subjects with certain physiological differences that would compromise their ability to clear a drug (e.g., renal impairment, hepatic impairment, limited CYP450 2D6 activity for a drug cleared by this isozyme) may provide data relevant to the clinical implications of overdose.

The review should contain a discussion of abuse potential and any apparent withdrawal symptoms. Usually, clinical trials are conducted to assess these issues for drug classes with a history of abuse potential and withdrawal phenomena (e.g., sedatives/hypnotics and anxiolytics). The review should comment on the adequacy and findings of these clinical trials. 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.

### **7.7 Additional Submissions/Safety Issues**

The initial NDA/BLA submission may not contain all information pertinent to the safety evaluation. Additional 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.

As a general rule, **the 120-day safety update should be integrated with the rest of the safety results above** but may be included here separately for a priority review where integrating the information may be too time consuming.

This section should also describe additional safety submissions or other safety issues, noting whether the results have been incorporated into the rest of the review or are considered in this section instead. This information may include negative or positive clinical trials that may have been conducted on a different population and could not, for appropriate reasons, be pooled. For those safety matters not incorporated into the rest of the review, discuss any data with important implications for safety. In general, this discussion will involve deaths, serious adverse events, and dropouts caused by serious adverse events. Even so, these data should be considered in the higher numbered sections under section 7.3, Major Safety Results, and 7.4, Supportive Safety Results, 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 be needed (e.g., deaths, liver injury).

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

## SAMPLE TABLES

Some of these sample tables may work better if they are developed in landscape page orientation, as that would allow more room for data presentation.

### **Table 1: Enumeration of Subjects for New Drug Development Program**

This table provides a count by clinical trial type of the subjects exposed to new drug, active control, and placebo across the entire set of clinical trials in the development program that contributed safety and efficacy data for the new drug. It should include all subjects known or assumed to have received even a single dose of assigned treatment. It should exclude subjects 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 subjects should be provided. (Note: If this list includes more than a few subjects, this may indicate a potentially important problem in the conduct of clinical trials.)

When creating this table, it is necessary to classify and group clinical trials on the basis of several characteristics. For the purposes of this table, the following characteristics and distinctions are deemed important:

- Phase 1 versus phases 2 to 3
- Completed versus ongoing and blinded
- Single dose versus multiple dose
- Controlled versus uncontrolled
- Short term versus long term
- Placebo-controlled versus active-controlled
- Fixed dose versus flexible dose

Other important features could lead to additional breakdowns within the table or to separate tables (e.g., different indications, inpatient versus outpatient status, differences in the quality and completeness of data collected across different clinical trials, foreign versus domestic). The characteristics to be used in classifying clinical trials for the purpose of this table should be decided in consultation with the designated review division.

In addition to this table that enumerates subjects by category of clinical trials, it would be useful to have a table that enumerates subjects by each individual clinical trial in the development program. This table would be an expanded version of Table 1 that enumerates subjects for each clinical trial (i.e., each of the categories in Table 1 would identify and provide data for the individual clinical trials comprising that category). Applicants ordinarily provide such a table.

Subjects participating in crossover trials should be counted in each of the pertinent columns of the table (e.g., a subject receiving treatment in each of the three arms of a

three-way crossover clinical trial comparing new drug, active control, and placebo would be counted in all three columns).

Footnotes to this table should identify by clinical trial number all those clinical trials comprising the various clinical trial 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.

We prefer that the applicant provide this table in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the applicant and the review division.

Additional table instructions are included as footnotes in the sample table.

**Table 1: Enumeration of Subjects for New Drug Development Program**

<b>Table 1                      Enumeration of Subjects for                      New Drug Development Program                      Cutoff Date: xx,xx,xx<sup>1</sup></b>			
<b>Clinical Trial                      Groups</b>	<b>Treatment Groups</b>		
	<b>New Drug</b>	<b>Active Control<sup>2</sup></b>	<b>Placebo</b>
<b>Completed Phase 1 (Clinical Pharmacology)</b>			
<b>Single Dose</b>	<b>120</b>	<b>30</b>	<b>30</b>
<b>Multiple Dose</b>	<b>60</b>	<b>30</b>	<b>30</b>
<b>Ph 1 Subtotal</b>	<b>180</b>	<b>60</b>	<b>60</b>
<b>Completed Phase 2-3 (Clinical Trials of Proposed Indication)</b>			
<b>Placebo Control<sup>3</sup></b>			
<b>Fixed Dose</b>	<b>500</b>	<b>150</b>	<b>150</b>
<b>Flexible Dose</b>	<b>100</b>	<b>100</b>	<b>100</b>
<b>Active Control</b>			
<b>Fixed Dose</b>	<b>200</b>	<b>100</b>	<b>0</b>
<b>Flexible Dose</b>	<b>100</b>	<b>100</b>	<b>0</b>
<b>Uncontrolled</b>			
<b>Short Term</b>	<b>100</b>	<b>0</b>	<b>0</b>

*continued*

Table 1, continued

Clinical Trial Groups	Treatment Groups		
	New Drug	Active Control <sup>2</sup>	Placebo
Long Term	700	0	0
Ph 2-3 Subtotal	1,200 <sup>4</sup>	450	250
<b>Ongoing Phase 2-3 (Clinical Trials of Proposed Indication)</b>			
Placebo Control			
Flexible Dose	150 <sup>5</sup>	0	150 <sup>5</sup>
Single Dose Subtotal	120	30	30
Multiple Dose Subtotal	1,410	480	430
<b>Grand Total</b>	<b>1,530</b>	<b>510</b>	<b>460</b>

<sup>1</sup> This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed subjects were not available for entry). Generally, this date should be no more than several months before the submission date for an NDA. This date, as well as this table, will likely need to be updated during the course of NDA review as more data become available.

<sup>2</sup> In the sample table, only one column is provided for an *active control* group. One such category may suffice for certain NDAs, but not for others, and the decision regarding how to categorize active-controlled subjects should be made in consultation with the review division.

<sup>3</sup> In this table, a decision was made to pool all clinical trials having a placebo arm, whether or not an active-controlled arm was also included. Thus, the active control category includes only those active-controlled clinical trials that did not have a placebo-controlled arm. Other approaches to grouping clinical trials may be equally appropriate.

<sup>4</sup> The intent of this table is to provide a count of unique subjects exposed to the new drug in the development program. Because subjects often participate in more than one clinical trial in a development program, it is necessary to have an approach to avoid counting subjects 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 subjects in that cell total who have already been counted by virtue of having participated in a previous clinical trial (e.g., a subject 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 the new drug is 1,700 less the 500 subjects already counted in short-term controlled trials, or 1,200).

<sup>5</sup> Frequently, some clinical trials may be ongoing and blinded at the time of NDA submission, even though some individual subjects who experienced serious adverse events may have been unblinded. In these instances, the table should include estimates of the number of subjects exposed to the new drug from these clinical trials, since exact counts may not be available. Footnotes should indicate when the table entries are based on estimates rather than exact counts.

**Table 2: Treatment Emergent Adverse Event Listing**

This table is a line listing of all reported treatment emergent adverse events, regardless of whether or not they were considered drug related, for all subjects participating in trials identified as sources for this listing. A footnote should identify all clinical trials contributing to this pool.

The variables included in this listing include:

- Trial number
- Center number
- Subject number (a unique number that identifies this subject 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 subject
- An indication of whether or not the event met the 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<sup>2</sup>, 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 (e.g., mild, moderate, severe)
- Action taken (e.g., none, decrease dose, discontinue treatment)
- Outcome
- Causality assessment by investigator (e.g., definitely, probably, possibly, or unlikely related)
- Location in NDA of CRF (e.g., subject narrative summary)

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 applicant and the review division.

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

Additional table instructions are included as footnotes in the sample table.

**Table 2: Treatment Emergent Adverse Event Listing**

Table 2 Treatment Emergent Adverse Event Listing New Drug Clinical Trials Source: Phase 2-3 Database Sorting A: Randomized Treatment, Trial #, Investigator/Center #, Subject # <sup>1</sup> Treatment = New Drug <sup>2</sup> Cutoff Date: xx,xx,xx <sup>3</sup>											
Trial	Center	Subject	Age (yrs)	Sex	Dose <sup>4</sup> (mg)	Time <sup>5</sup> (days)	Body System	Preferred Term	Adverse Event <sup>6</sup>	Serious <sup>7</sup>	W/D <sup>8</sup>

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

<sup>2</sup> This sample listing is for new drug subjects (i.e., for all subjects exposed to the new drug in the phase 2 to phase 3 clinical trials that are part of the integrated primary database). Similar listings should be provided for active control and placebo subjects.

<sup>3</sup> This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed subjects were not available for entry). Generally, this date should be no more than several months before the submission date for an NDA. This date, as well as this table, will likely need to be updated during the course of NDA review as more data become available.

<sup>4</sup> This column should include the dose being administered (in mg/day) at the time the event occurred.

<sup>5</sup> This column should include the time (i.e., duration of exposure (in days)) that the event occurred.

<sup>6</sup> This column should include the adverse event in the language reported by the investigator and/or subject (i.e., before coding).

<sup>7</sup> 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.

<sup>8</sup>This column should include an indication of whether or not the adverse event led to discontinuation of the assigned treatment.

### **Table 3: Demographic Profile**

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.

Subjects participating in crossover trials should be included in the calculations for each of the pertinent columns of the table (e.g., a subject receiving treatment in each of the three arms of a three-way crossover clinical trial comparing new drug, active control, and placebo would be included in the calculations for all three columns).

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

This sample table includes four 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 sex and age (e.g., women younger than 50).

We prefer that the applicant provide this table in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the applicant and the review division.

Additional table instructions are included as footnotes in the sample table.

**Table 3: Demographic Profile**

<b>Table 3 Demographic Profile for Phase 2-3 Clinical Trials With New Drug Cutoff Date: xx,xx,xx<sup>1</sup></b>			
<b>Demographic Parameters</b>	<b>Treatment Groups<sup>2,3</sup></b>		
	<b>New Drug N =</b>	<b>Placebo N =</b>	<b>Active Control N =</b>
<b>Age (years)</b>			
<b>Mean</b>			
<b>Range</b>			
<b>Groups<sup>4</sup></b>			
<b>&lt; 40</b>	%	%	%
<b>40-64</b>	%	%	%
<b>≥ 65</b>	%	%	%
<b>Sex</b>			
<b>Female</b>	%	%	%
<b>Male</b>	%	%	%
<b>Race<sup>5</sup></b>			
<b>Caucasian</b>	%	%	%
<b>Noncaucasion</b>	%	%	%
<b>Weight (kg)</b>			
<b>Mean</b>			
<b>Range</b>			

<sup>1</sup> This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed subjects were not available for entry). Generally, this date should be no more than several months before the submission date for an NDA. This date, as well as this table, will likely need to be updated during the course of NDA review as more data become available.

<sup>2</sup> In the sample table, only one column is provided for an *active control* group. One such category may suffice for certain NDAs, but not for others, and the decision regarding how to categorize active-controlled subjects should be made in consultation with the review division. Similarly, for this table, only one column is provided for the new drug, with the implication that all new drug subjects, regardless of dose, should be included in the calculations for that column. Other approaches (e.g., distinguishing subjects on the basis of dose) may be equally appropriate.

<sup>3</sup> 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.

<sup>4</sup> If there are pediatric exposures, these should be broken out as well.

<sup>5</sup> Other approaches to racial categorization may be substituted for that proposed in this sample table.



**Table 4: Number (Percent) of Subjects Receiving New Drug**

This table is calculated by first categorizing subjects on the basis of the interval of exposure for each (e.g., a subject exposed for 6 weeks would be counted in the 4<Dur<12 row). The mean daily dose is then calculated for each subject for dose categorization (e.g., a 6-week subject with a mean daily dose of 15 mg would be counted in the 10<Dos<20 column). Subjects are enumerated in only one cell of the matrix (i.e., this display is mutually exclusive). 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 is useful.

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

The same table can be generated for any individual clinical trial or for any pool of clinical trials.

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).

Similar tables should be provided for active-controlled drugs and placebo.

We prefer that the applicant provide this table in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the applicant and the review division.

Additional table instructions are included as footnotes in the sample table.

**Table 4: Number (Percent) of Subjects Receiving New Drug**

Table 4 Number (Percent) of Subjects Receiving New Drug According to Mean Daily Dose and Duration of Therapy in Phase 2-3 Clinical Trials (N=2,500) Cutoff Date: xx,xx,xx <sup>1</sup>								
Duration (Weeks)	Dose (mg) <sup>2</sup>						Total (AnyDos)	(%)
	0<Dos ≤5	5<Dos ≤10	10<Dos ≤20	20<Dos ≤30	30<Dos ≤50	50<Dos		
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%)
<b>Total (AnyDur)</b>	<b>125</b>	<b>375</b>	<b>623</b>	<b>625</b>	<b>500</b>	<b>252</b>	<b>2,500</b>	<b>(100%)</b>
<b>(%)</b>	<b>(5%)</b>	<b>(15%)</b>	<b>(25%)</b>	<b>(25%)</b>	<b>(20%)</b>	<b>(10%)</b>	<b>(100%)</b>	

<sup>1</sup> This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed subjects were not available for entry). Generally, this date should be no more than several months before the submission date for an NDA. This date, as well as this table, will likely need to be updated during the course of NDA review as more data become available.

<sup>2</sup> Dose may also be expressed as mg/kg, mg/m<sup>2</sup>, or in terms of plasma concentration if such data are available.

**Table 5: Deaths Listing**

A footnote should describe the rule for including deaths in this 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 deaths 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.

Deaths occurring outside the time window for this table should be listed elsewhere.

We prefer that the applicant provide this table in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the applicant and the review division.

Similar lists should be provided for subjects exposed to placebo and active-controlled drugs.

Additional table instructions are included as footnotes in the sample table.

**Table 5: Deaths Listing**

<p style="text-align: center;"><b>Table 5 Deaths Listing Treatment = New Drug Cutoff Date: xx,xx,xx<sup>1</sup></b></p>									
<b>Trial</b>	<b>Center</b>	<b>Subject</b>	<b>Age (yrs)</b>	<b>Sex</b>	<b>Dose (mg)<sup>2</sup></b>	<b>Time (Days)<sup>3</sup></b>	<b>Source<sup>4</sup></b>	<b>Person Time<sup>5</sup></b>	<b>Description<sup>6</sup></b>

<sup>1</sup> This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed subjects were not available for entry). Generally, this date should be no more than several months before the submission date for an NDA. This date, as well as this table, will likely need to be updated during the course of NDA review as more data become available.

<sup>2</sup> Dose at time of death; or if death occurred after discontinuation, note that dose, as well as last dose before discontinuation.

<sup>3</sup> 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.

<sup>4</sup> 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<sup>0</sup> for deaths arising from primary source clinical trials and 2<sup>0</sup> for deaths arising from secondary sources).

<sup>5</sup> This column should identify subjects (yes/no) for whom person-time data are available, so the reviewer can know which subjects were included in the mortality rate calculations.

<sup>6</sup> Because narrative summaries should be available for all deaths, the description can be brief (e.g., myocardial infarction, stroke, pancreatic cancer, suicide by drowning).

### **Table 6: Mortality by Treatment Group**

This table provides data comparing overall mortality across treatment groups for the pool of all phase 2 to 3 clinical trials 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 applicant's judgment about causality, including:

- Any deaths occurring during participation in any of the clinical trials in the target pool
- Any deaths occurring after a subject leaves any of the targeted clinical trials, whether prematurely or after completion to the nominal endpoint, if the death:
  - Is the result of a process initiated during the clinical trial, regardless of when it actually occurs
  - Occurs within 4 weeks of a subject leaving a clinical trial, 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.

Subjects participating in crossover trials should be included in the calculations for each of the pertinent columns of the table (e.g., a subject receiving treatment in each of the three arms of a three-way crossover clinical trial comparing the new drug, active control, and placebo would be included in the calculations for all three columns).

We prefer that the applicant provide this table in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the applicant and the review division.

Additional table instructions are included as footnotes in the sample table.

**Table 6: Mortality by Treatment Group**

<p align="center"><b>Table 6</b>  <b>Mortality by Treatment Group for</b>  <b>Pool of Phase 2-3 Clinical Trials with New Drug</b>  <b>Cutoff Date: xx,xx,xx<sup>1</sup></b></p>						
<b>Treatment Group<sup>2</sup></b>	<b>Total Number of Subjects</b>	<b>Total Number of Deaths<sup>3</sup></b>	<b>Crude Mortality<sup>4</sup></b>	<b>Patient Exposure Years (PEY)<sup>5</sup></b>	<b>Total Deaths With Person-Time<sup>6</sup></b>	<b>Mortality per 100 PEY<sup>7</sup></b>
New Drug						
Active Control						
Placebo						

<sup>1</sup> This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed subjects were not available for entry). Generally, this date should be no more than several months before the submission date for an NDA. This date, as well as this table, will likely need to be updated during the course of NDA review as more data become available.

<sup>2</sup> In the sample table, only one row is provided for an *active control* group. One such category may suffice for certain NDAs, but not for others, and the decision regarding how to categorize active-controlled subjects should be made in consultation with the review division. Similarly, for this table, only one row is provided for the new drug, with the implication that all new drug subjects, regardless of dose, should be included in the calculations for that column. Other approaches (e.g., distinguishing subjects on the basis of dose) may be equally appropriate.

<sup>3</sup> This column is the total number of deaths for each group.

<sup>4</sup> This column is simply the total number of deaths divided by the total number of subjects exposed in each group.

<sup>5</sup> 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.

<sup>6</sup> This column is the subset of total deaths for which person-time is available.

<sup>7</sup> This column is the number of deaths for whom person-time is available divided by PEY for each group, and multiplied by 100.

**Table 7: Serious Adverse Event Listing**

This table is a line listing of all reported adverse events that met the applicant's definition of being a *serious* adverse event, regardless of whether or not they were considered drug related, for all subjects participating in the phase 2 to phase 3 trials in the development program. This listing is a critical component of the ISS.

The following variables are included in this listing:

- Trial number
- Center number
- Subject number (a unique number that identifies this subject 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 subject
- 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/m<sup>2</sup>, or even plasma concentration, if available
- Other drug treatment
- Severity of adverse event (e.g., mild, moderate, severe)
- Action taken (e.g., none, decrease dose, discontinue treatment)
- Outcome
- Causality assessment by investigator (e.g., definitely, probably, possibly, or unlikely related)
- Location in NDA of CRF (e.g., subject narrative summary)

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 applicant and the review division.

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

Additional table instructions are included as footnotes in the sample table.

**Table 7: Serious Adverse Event Listing**

<p style="text-align: center;"><b>Table 7</b>  <b>Serious Adverse Event Listing</b>  <b>New Drug Clinical Trials</b>  <b>Source: Phase 2-3 Trials</b>  <b>Sorting A: Randomized Treatment, Trial #, Investigator/Center #, Subject #<sup>1</sup></b>  <b>Treatment = New Drug<sup>2</sup></b>  <b>Cutoff Date: xx,xx,xx<sup>3</sup></b></p>										
Trial	Center	Subject	Age (yrs)	Sex	Dose (mg) <sup>4</sup>	Time (days) <sup>5</sup>	Body System	Preferred Term	Adverse Event <sup>6</sup>	W/D <sup>7</sup>

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

<sup>2</sup> This sample listing is for all new drug subjects across all clinical trials in the phase 2 to 3 development program. Similar listings should be provided for active-controlled and placebo subjects.

<sup>3</sup> This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed subjects were not available for entry). Generally, this date should be no more than several months before the submission date for an NDA. This date, as well as this table, will likely need to be updated during the course of NDA review as more data become available.

<sup>4</sup> This column should include the dose being administered (in mg/day) at the time the event occurred.

<sup>5</sup> This column should include the time (i.e., duration of exposure (in days)) that the event occurred. If the event occurred after discontinuation of drug, a footnote should note how long after discontinuation.

<sup>6</sup> This column should include the adverse event in the language reported by the investigator and/or subject (i.e., before coding).

<sup>7</sup> This column should include an indication of whether or not the adverse event led to discontinuation of the assigned treatment.

**Table 8: Dropout Profile**

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.

Subjects participating in crossover trials should be included in the calculations for each of the pertinent columns of the table (e.g., a subject receiving treatment in each of the three arms of a three-way crossover clinical trial comparing the new drug, active control, and placebo would be included in the calculations for all three columns).

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

We prefer that the applicant provide this table in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the applicant and the review division.

Additional table instructions are included as footnotes in the sample table.

**Table 8: Dropout Profile**

<p align="center"><b>Table 8</b>  <b>Dropout Profile: Incidence of Dropout by Treatment Group and</b>  <b>Reason for Phase 2 to 3 Clinical Trials with New Drug</b>  <b>Cutoff Date: xx,xx,xx<sup>1</sup></b></p>			
<p align="center"><b>Reasons for Dropout<sup>2</sup></b></p>	<p align="center"><b>Treatment Groups<sup>3</sup></b></p>		
	<p align="center"><b>New Drug</b> N =</p>	<p align="center"><b>Placebo</b> N =</p>	<p align="center"><b>Active Control</b> N =</p>
<b>Lack of Efficacy</b>	%	%	%
<b>Adverse Event</b>	%	%	%
<b>Lost to Follow-Up</b>	%	%	%
<b>Other</b>	%	%	%
<b>Total Dropouts</b>	%	%	%

<sup>1</sup> This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed subjects were not available for entry). Generally, this date should be no more than several months before the submission date for an NDA. This date, as well as this table, will likely need to be updated during the course of NDA review as more data become available.

<sup>2</sup> This sample table includes four categories for dropout, but a more detailed breakdown may be of interest as well.



- The *adverse event* category would include all subjects identified as dropping out for adverse events, regardless of whether or not the events were judged by the investigator or applicant to be drug-related and regardless of what other reasons may have been identified in association with drop out. Subjects identified as dropping out for intercurrent illness would ordinarily be included under this adverse event category. Similarly, a subject 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, because it reflects on the overall conduct of the clinical trials.
- The *other* category is intended to include all other reasons that generally may be considered nontreatment related. This category is often identified as *administrative*, and includes such reasons as subject refused further participation, subject moved away, subject improved, subject not eligible, protocol violation, and unknown.

Decisions about what categories to include should be made in consultation with the review division.

<sup>3</sup> In the sample table, only one column is provided for an *active control* group. One such category may suffice for certain NDAs, but not for others, and the decision regarding how to categorize active-controlled subjects should be made in consultation with the review division. Similarly, for this table, only one column is provided for the new drug, with the implication that all new drug subjects, regardless of dose, should be included in the calculations for that column. Other approaches (e.g., distinguishing subjects on the basis of dose) may be equally appropriate.

### **Table 9: Adverse Event Dropout Listing**

This table 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 subjects 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 ISS.

The following variables are included in this listing:

- Trial number
- Center number
- Subject number (a unique number that identifies this subject 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 subject
- An indication of whether or not the event led to withdrawal
- Outcome

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

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

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 applicant and the review division.

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

Additional table instructions are included as footnotes in the sample table.

**Table 9: Adverse Event Dropout Listing**

Table 9 Adverse Event Dropout Listing New Drug Clinical Trials Source: Phase 2-3 Database Sorting A: Randomized Treatment, Trial #, Investigator/Center #, Subject # <sup>1</sup> Treatment = New Drug <sup>2</sup> Cutoff Date: xx,xx,xx <sup>3</sup>											
Trial	Center	Subject	Age (yrs)	Sex	Dose (mg) <sup>4</sup>	Time (days) <sup>5</sup>	Body System	Preferred Term	Adverse Event <sup>6</sup>	Serious <sup>7</sup>	Outcome <sup>8</sup>

*continued*

Table 9, continued

Trial	Center	Subject	Age (yrs)	Sex	Dose (mg) <sup>4</sup>	Time (days) <sup>5</sup>	Body System	Preferred Term	Adverse Event <sup>6</sup>	Serious <sup>7</sup>	Outcome <sup>8</sup>

<sup>1</sup> It is essential to provide this listing in two different forms (i.e., sorting A (by subject) and sorting B (by adverse event)). This listing is for sorting A (by subject) and permits the reviewer to explore all the adverse events reported as leading to discontinuation for each individual subject. Sorting B (by adverse event) should be as follows: Randomized Treatment, Body System, Preferred Term, Adverse Event, Trial, Center, Subject #, 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.

<sup>2</sup> This sample listing is for all new drug subjects across all clinical trials in the phase 2 to 3 development program. Similar listings should be provided for active-controlled and placebo subjects.

<sup>3</sup> This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed subjects were not available for entry). Generally, this date should be no more than several months before the submission date for an NDA. This date, as well as this table, will likely need to be updated during the course of NDA review as more data become available.

<sup>4</sup> This column should include the dose being administered (in mg/day) at the time the event occurred.

<sup>5</sup> This column should include the time (i.e., duration of exposure (in days)) that the event occurred.

<sup>6</sup> This column should include the adverse event in the language reported by the investigator and/or subject (i.e., before coding).

<sup>7</sup> 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.

<sup>8</sup> 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 10: Treatment Emergent Adverse Event Incidence**

This table compares the incidence of treatment emergent adverse events across treatment groups for a pool of similarly designed placebo-controlled trials of the new drug. Generally, an arbitrary threshold incidence for new drug subjects is used as a criterion for selecting which adverse events to include; greater than 1 percent for the 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.

Clinical trial 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.

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 investigator's judgments about drug-relatedness. Either approach is acceptable, but it is critical that a footnote indicate when adverse events are not included because of the investigator's judgments that they were not drug-related, since this approach may reduce the adverse event rates that appear in the table.

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.

Not uncommonly, a new drug is developed for more than one indication. If adverse event rates appear 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.

Adverse events that occur at a rate for placebo that is greater than the rate for new drug should be removed from the table and noted only as a footnote.

Subjects participating in crossover trials should be included in the calculations for each of the pertinent columns of the table (e.g., a subject receiving treatment in each of the three arms of a three-way crossover clinical trial comparing the new drug, active control, and placebo would be included in the calculations for all three columns).

We prefer that the applicant provide this table in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the applicant and the review division.

Additional table instructions are included as footnotes in the sample table.

**Table 10: Treatment Emergent Adverse Event Incidence**

<b>Table 10</b> <b>Treatment Emergent Adverse Event Incidence</b> <b>for Pool of 6-Week Placebo-Controlled Trials</b> <b>Cutoff Date: xx,xx,xx<sup>1</sup></b>			
<b>Body System/ Adverse Event<sup>2</sup></b>	<b>Percentage of Subjects Reporting Event<sup>3</sup></b>		
	<b>New Drug<sup>4</sup> N=<sup>6</sup></b>	<b>Active Control<sup>5</sup> N=</b>	<b>Placebo N=</b>
<b>Body as a Whole</b>			
<b>Headache</b>			
<b>Etc.</b>			
<b>Cardiovascular System</b>			
<b>Postural Hypotension</b>			
<b>Etc.</b>			
<b>Gastrointestinal System</b>			
<b>Constipation</b>			
<b>Etc.</b>			
<b>Urogenital System Impotence<sup>7</sup></b>			
<b>Etc.</b>			

<sup>1</sup> This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed subjects were not available for entry). Generally, this date should be no more than several months before the submission date for an NDA. This date, as well as this table, will likely need to be updated during the course of NDA review as more data become available.

<sup>2</sup> Adverse events should be organized under body system categories. Within each body system category, adverse events should be ordered according to decreasing frequency. 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.

<sup>3</sup> Percentages should be rounded to the nearest integer. Although not strictly hypothesis testing, p-values give some idea 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 percent.

<sup>4</sup> For this table, only one column is provided for the new drug, with the implication that all new drug subjects, regardless of dose, should be included in the calculations for that column. Other approaches (e.g., dividing subjects on the basis of dose) may be equally appropriate. If the clinical trials used were fixed-dose clinical trials, 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 be reasonably omitted from this table. It is generally not useful to try to artificially construct dose categories from dose titration clinical trials, since there is often confounding of dose and time.

<sup>5</sup> For this table, only one column is provided for an *active control* group. One such category may suffice for certain NDAs, but not for others, and the decision regarding how to categorize active-controlled subjects should be made in consultation with the review division.

<sup>6</sup> The N for each column should be provided at the column heading, so that only the percentage of subjects having that adverse event need be included in the table, and not the actual number.

<sup>7</sup> The rates for sex-specific adverse events (e.g., impotence) should be determined using the appropriate sex-specific denominator, and this fact should be indicated with a footnote.

### **Table 11: Mean Change From Baseline for Serum Chemistry Parameters**

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.

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.

Subjects participating in crossover trials should be included in the calculations for each of the pertinent columns of the table (e.g., a subject receiving treatment in each of the three arms of a three-way crossover clinical trial comparing the new drug, active control, and placebo would be included in the calculations for all three columns).

We prefer that the applicant provide this table in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the applicant and the review division.

Additional table instructions are included as footnotes in the sample table.

**Table 11: Mean Change From Baseline for Serum Chemistry Parameters**

Table 11 Mean Change From Baseline for Serum Chemistry Parameters in Pool of Placebo-Controlled Clinical Trials Cutoff Date: xx,xx,xx <sup>1</sup>									
Serum Chemistry Parameters and Units of Measure <sup>2</sup>	Treatment Groups								
	New Drug <sup>3</sup>			Placebo			Active Control <sup>4</sup>		
	N <sup>5</sup>	BL <sup>6</sup>	Change From BL <sup>7</sup>	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)									

<sup>1</sup> This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed subjects were not available for entry). Generally, this date should be no more than several months before the submission date for an NDA. This date, as well as this table, will likely need to be updated during the course of NDA review as more data become available.

<sup>2</sup> 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 clinical trials 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 review division.

<sup>3</sup> For this table, only one column is provided for the new drug, with the implication that all new drug subjects, regardless of dose, should be included in the calculations for that column. Other approaches (e.g., dividing subjects on the basis of dose) may be equally appropriate. If the clinical trials used were fixed-dose clinical trials, 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 be reasonably omitted from this table. It is generally not useful to try to artificially construct dose categories from dose titration clinical trials, since there is often confounding of dose and time.

<sup>4</sup> For this table, only one column is provided for an *active control* group. One such category may suffice for certain NDAs, but not for others, and the decision regarding how to categorize active-controlled subjects should be made in consultation with the review division.

<sup>5</sup> N represents the number of subjects who had the serum chemistry parameter of interest assessed at baseline and at least one follow-up time.

<sup>6</sup> This column should provide the baseline means for all the serum chemistry parameters of interest.

<sup>7</sup> This column should provide the mean change from baseline to subject's worst on drug value for each of the serum chemistry parameters of interest. Although 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 12: Incidence of Potentially Clinically Significant Changes in Serum Chemistry Parameters**

This table provides data comparing the incidence across treatment groups of subjects 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 subjects who were abnormal at baseline and met these PCS criteria.

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.

Subjects participating in crossover trials should be included in the calculations for each of the pertinent columns of the table (e.g., a subject receiving treatment in each of the three arms of a three-way crossover clinical trial comparing the new drug, active control, and placebo would be included in the calculations for all three columns).

We prefer that the applicant provide this table in electronic format. The exact design of the table and the preferred electronic format should be established in discussions between the applicant and the review division.

Additional table instructions are included as footnotes in the sample table.



**Table 12: Incidence of Potentially Clinically Significant Changes in Serum Chemistry Parameters**

<b>Table 12</b> <b>Incidence of Potentially Clinically Significant Changes in Serum Chemistry Parameters for Pool of Placebo-Controlled Clinical Trials for New Drug</b> <b>Cutoff Date: xx,xx,xx<sup>1</sup></b>									
<b>Serum Chemistry Parameters and PCS Criteria<sup>2</sup></b>  L=Low; H=High; ULN=Upper Limits of Normal	<b>Treatment Groups<sup>3</sup></b>								
	<b>New Drug</b>			<b>Placebo</b>			<b>Active Control</b>		
	<b>Total Pts<sup>4</sup></b>	<b>Abnormal</b>		<b>Total Pts</b>	<b>Abnormal</b>		<b>Total Pts</b>	<b>Abnormal</b>	
		<b>Nbr<sup>5</sup></b>	<b>%<sup>6</sup></b>		<b>Nbr</b>	<b>%</b>		<b>Nbr</b>	<b>%</b>
<b>Albumin-L (&lt; 2.5 g/dl)</b>									
<b>Alkaline P'tase-H (&gt; 400 U/L)</b>									
<b>Bilirubin, total-H (&gt; 2 mg/dl)</b>									
<b>BUN-H (&gt; 30 mg/dl)</b>									
<b>CK-H (&gt; 3XULN)</b>									
<b>Calcium-L (&lt; 7 mg/dl)</b>									
<b>Calcium-H (&gt; 12 mg/dl)</b>									
<b>Cholesterol-H (&gt; 300 mg/dl)</b>									
<b>Creatinine-H (&gt; 2 mg/dl)</b>									
<b>GGT-H (&gt; 3XULN)</b>									
<b>Glucose-L (&lt; 50 mg/dl)</b>									

*continued*

Table 12, continued

<b>Serum Chemistry Parameters and PCS Criteria<sup>2</sup></b>  L=Low; H=High; ULN=Upper Limits of Normal	<b>Treatment Groups<sup>3</sup></b>								
	<b>New Drug</b>			<b>Placebo</b>			<b>Active Control</b>		
	<b>Total Pts<sup>4</sup></b>	<b>Abnormal</b>		<b>Total Pts</b>	<b>Abnormal</b>		<b>Total Pts</b>	<b>Abnormal</b>	
		<b>Nbr<sup>5</sup></b>	<b>%<sup>6</sup></b>		<b>Nbr</b>	<b>%</b>		<b>Nbr</b>	<b>%</b>
<b>Glucose-H</b> (> 250 mg/dl)									
<b>LDH-H</b> (> 3XULN)									
<b>Phosphorus-L</b> (< 2.0 mg/dl)									
<b>Phosphorus-H</b> (> 5.0 mg/dl)									
<b>Potassium-L</b> (< 3.0 mmol/L)									
<b>Potassium-H</b> (> 5.5 mmol/L)									
<b>AST-H</b> (> 3XULN)									
<b>ALT-H</b> (> 3XULN)									
<b>Sodium-L</b> (< 130 mmol/L)									
<b>Sodium-H</b> (> 150 mmol/L)									
<b>Triglycerides-H</b> (> 300 mg/dl)									
<b>Uric Acid (F)-H</b> (> 8.0 mg/dl)									
<b>Uric Acid (M)-H</b> (> 10.0 mg/dl)									

<sup>1</sup> This is the data lock date for entering data into this table (i.e., the date beyond which additional exposed subjects were not available for entry). Generally, this date should be no more than several months before the submission date for an NDA. This date, as well as this table, will likely need to be updated during the course of NDA review as more data become available.

<sup>2</sup> 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 clinical trials 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 review division.

<sup>3</sup> In the sample table, only one column is provided for an *active control* group. One such category may suffice for certain NDAs, but not for others, and the decision regarding how to categorize active-controlled subjects should be made in consultation with the review division. Similarly, for this table, only one column is provided for the new drug, with the implication that all new drug subjects, regardless of dose, should be included in the calculations for that column. Other approaches (e.g., distinguishing subjects on the basis of dose) may be equally appropriate.

<sup>4</sup> The total number of subjects for each parameter should represent the number of subjects 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.

<sup>5</sup> The number in this column represents the subset of the total number of subjects who met the criterion in question at least once during treatment. A separate listing should provide subject identification for those subjects meeting the criterion.

<sup>6</sup> Percentage of the total number of subjects meeting the criterion should be rounded to the nearest integer. Although 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 percent.

**GLOSSARY OF ACRONYMS**

ALP	alkaline phosphatase
ALT	alanine aminotransferase
AT	aminotransferase
BLA	biologics license application
BUN	blood urea nitrogen
CHF	congestive heart failure
COSTART	Coding Symbols for Thesaurus of Adverse Reaction Terms
CRF	case report form
CTD	common technical document
DILI	drug-induced liver injury
ECG	electrocardiogram
GRP	good review practice
IND	investigational new drug application
ISS	integrated summary of safety
MedDRA	Medical Dictionary for Regulatory Activities
NDA	new drug application
NSAID	nonsteroidal anti-inflammatory drug
PCS	potential clinical significance
PD	pharmacodynamics
PDF	portable document format
PEY	patient exposure years
PK	pharmacokinetics
PPAR	peroxisome proliferator-activated receptors
REMS	risk evaluation and mitigation strategy
SSRI	selective serotonin reuptake inhibitor
TdP	Torsades de Pointes
ULN	upper limit of normal