Guidance for Industry

Premarketing Risk Assessment

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

March 2005
Clinical Medical

Guidance for Industry

Premarketing Risk Assessment

Additional copies are available from:

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

or

Office of Communication, Training and
Manufacturers Assistance, HFM-40
Center for Biologics Evaluation and Research
Food and Drug Administration
1401 Rockville Pike, Rockville, MD 20852-1448
http://www.fda.gov/cber/guidelines.htm.
(Tel) Voice Information System at 800-835-4709 or 301-827-1800

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

March 2005
Clinical Medical

TABLE OF CONTENTS

Contains Nonbinding Recommendations

Guidance for Industry[1] Premarketing Risk Assessment

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

I. INTRODUCTION

This document provides guidance to industry on good risk assessment practices during the development of prescription drug products, including biological drug products.[2] This is one of three guidances that were developed to address risk management activities. Specifically, this document discusses the generation, acquisition, analysis, and presentation of premarketing safety data.

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

---

[1] This guidance has been prepared by the Center for Drug Evaluation and Research (CDER) and the Center for Biologics Evaluation and Research (CBER) at the Food and Drug Administration.

[2] For ease of reference, this guidance uses the terms *product* and *drug* to refer to all products (excluding blood and blood components) regulated by CDER or CBER, including vaccines. Similarly, for ease of reference, this draft guidance uses the term *approval* to refer to both drug approval and biologic licensure.

Paperwork Reduction Act Public Burden Statement: This guidance contains information collection provisions that are subject to review by the Office of Management and Budget (OMB) under the Paperwork Reduction Act of 1995 (PRA) (44 U.S.C. 3501-3520). The collection(s) of information in this guidance were approved under OMB Control No. 0910-0001 (until March 31, 2005) and 0910-0338 (until August 31, 2005).

II. BACKGROUND

A. PDUFA III's Risk Management Guidance Goal

On June 12, 2002, Congress reauthorized, for the second time, the Prescription Drug User Fee Act (PDUFA III). In the context of PDUFA III, FDA agreed to satisfy certain performance goals. One of those goals was to produce guidance for industry on risk management activities for drug and biological products. As an initial step towards satisfying that goal, FDA sought public comment on risk management. Specifically, FDA issued three concept papers. Each paper focused on one aspect of risk management, including (1) conducting premarketing risk assessment, (2) developing and implementing risk minimization tools, and (3) performing postmarketing pharmacovigilance and pharmacoepidemiologic assessments. In addition to receiving numerous written comments regarding the three concept papers, FDA held a public workshop on April 9-11, 2003, to discuss the concept papers. FDA considered all of the comments received in developing three draft guidance documents on risk management activities. The draft guidance documents were published on May 5, 2004, and the public was provided with an opportunity to comment on them until July 6, 2004. FDA considered all of the comments received in producing the final guidance documents.

- *Premarketing Risk Assessment (Premarketing Guidance)*
- *Development and Use of Risk Minimization Action Plans (RiskMAP Guidance)*
- *Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment (Pharmacovigilance Guidance)*.

B. Overview of the Risk Management Guidances

Like the concept papers and draft guidances that preceded them, each of the three final guidance documents focuses on one aspect of risk management. The *Premarketing Guidance* and the *Pharmacovigilance Guidance* focus on premarketing and postmarketing risk assessment, respectively. The *RiskMAP Guidance* focuses on risk minimization. Together, risk assessment and risk minimization form what FDA calls *risk management*. Specifically, risk management is an iterative process of (1) assessing a product's benefit-risk balance, (2) developing and implementing tools to minimize its risks while preserving its benefits, (3) evaluating tool effectiveness and reassessing the benefit-risk balance, and (4) making adjustments, as appropriate, to the risk minimization tools to further improve the benefit-risk balance. This four-part process should be continuous throughout a product's lifecycle, with the results of risk assessment informing the sponsor's decisions regarding risk minimization.

When reviewing the recommendations provided in this guidance, sponsors and applicants should keep the following points in mind:

- Many recommendations in this guidance are ***not*** intended to be generally applicable to all products.

  Industry already performs risk assessment and risk minimization activities for products during development and marketing. The Federal Food, Drug, and Cosmetic Act (FDCA)

and FDA implementing regulations establish requirements for ***routine*** risk assessment and risk minimization (see e.g., FDA requirements for professional labeling and adverse event monitoring and reporting). As a result, many of the recommendations presented here focus on situations in which a product may pose a clinically important and unusual type or level of risk. To the extent possible, we have specified in the text whether a recommendation is intended for all products or only this subset of products.

- It is of critical importance to protect patients and their privacy during the generation of safety data and the development of risk minimization action plans.

  During all risk assessment and risk minimization activities, sponsors must comply with applicable regulatory requirements involving human subjects research and patient privacy.[3]

- To the extent possible, this guidance reflects FDA's commitment to harmonization of international definitions and standards.

- When planning risk assessment and risk minimization activities, sponsors should consider input from healthcare participants likely to be affected by these activities (e.g., from consumers, pharmacists and pharmacies, physicians, nurses, and third party payers).

- There are points of overlap among the three guidances.

  We have tried to note in the text of each guidance when areas of overlap occur and when referencing one of the other guidances might be useful.

III. THE ROLE OF RISK ASSESSMENT IN RISK MANAGEMENT

Risk management is an iterative process designed to optimize the benefit-risk balance for regulated products. Risk assessment consists of identifying and characterizing the nature, frequency, and severity of the risks associated with the use of a product. Risk assessment occurs throughout a product's lifecycle, from the early identification of a potential product, through the premarketing development process, and after approval during marketing. Premarketing risk assessment represents the first step in this process, and this guidance focuses on risk assessment prior to marketing.

[3] See 45 CFR part 46 and 21 CFR parts 50 and 56. See also the Health Insurance Portability and Accountability Act of 1996 (HIPAA) (Public Law 104-191) and the Standards for Privacy of Individually Identifiable Health Information (the Privacy Rule) (45 CFR part 160 and subparts A and E of part 164). The Privacy Rule specifically permits covered entities to report adverse events and other information related to the quality, effectiveness, and safety of FDA-regulated products both to manufacturers and directly to FDA (45 CFR 164.512(b)(1)(i) and (iii), and 45 CFR 164.512(a)(1)). For additional guidance on patient privacy protection, see http://www.hhs.gov/ocr/hipaa.

It is critical to FDA's decision on product approval that a product's underlying risks and benefits be adequately assessed during the premarketing period. Sponsors seeking approval must provide from the clinical trials a body of evidence that adequately characterizes the product's safety profile.[4]

This guidance provides general recommendations for assessing risk. The adequacy of the assessment of risk is a matter of both quantity (ensuring that enough patients are studied) and quality (the appropriateness of the assessments performed, the appropriateness and breadth of the patient populations studied, and how results are analyzed). Quantity is, in part, considered in other Agency guidances,[5] but it is discussed further here. This guidance also addresses the qualitative aspects of risk assessment.

Although risk assessment continues through all stages of product development, this guidance focuses on risk assessment during the later stages of clinical development, particularly during phase 3 studies. The guidance is not intended to cover basic aspects of preclinical safety assessments (i.e., animal toxicity testing) or routine clinical pharmacology programs. Good clinical risk assessment in the later stages of drug development should be guided by the results of comprehensive preclinical safety assessments and a rigorous, thoughtful clinical pharmacology program (including elucidation of metabolic pathways, identification of possible drug-drug interactions, and determination of any effects from hepatic and/or renal impairment). These issues are addressed in other FDA guidances and guidances developed under the auspices of the International Conference for Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH).

IV. GENERATING RISK INFORMATION DURING CLINICAL TRIALS

Providing detailed guidance on what constitutes an adequate safety database for all products is impossible. The nature and extent of safety data that would provide sufficient information about risk for purposes of approving a product are individualized decisions based on a number of factors (several of which are discussed below). In reaching a final decision on approvability, both existing risk information and any outstanding questions regarding safety are considered in a product's risk assessment and weighed against the product's demonstrated benefits. The fewer a product's demonstrated benefits, the less acceptable may be higher levels of demonstrated risks. Likewise, the fewer the benefits, generally, the less uncertainty may be accepted about a product's risks.

To maximize the information gained from clinical trials, FDA recommends that from the outset of development, sponsors pay careful attention to the overall design of the safety evaluation.

---

[4] Section 505(d)(1) of the Federal Food, Drug, and Cosmetic Act (21 U.S.C. 355(d)(1)) requires the conduct of "adequate tests by all methods reasonably applicable to show whether or not . . . [a] drug is safe for use under the [labeled] conditions. . . ." See also 21 CFR 314.50(d)(5)(vi). Section 351 of the Public Health Service Act (42 U.S.C. 262) requires a demonstration that a biologic is "safe, pure, and potent." See also 21 CFR 601.2.

[5] See the 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*, International Conference on Harmonisation (ICH).

Potential problems that may be suspected because of preclinical data or because of effects of related drugs should be targeted for evaluation. And, because it is impossible to predict every important risk, as experience accrues, sponsors should refine or modify their safety evaluations.

A. Size of the Premarketing Safety Database

Even large clinical development programs cannot reasonably be expected to identify all risks associated with a product. Therefore, it is expected that, even for a product that is rigorously tested preapproval, some risks will become apparent only after approval, when the product is used in tens of thousands or even millions of patients in the general population. Although no preapproval database can possibly be sized to detect all safety issues that might occur with the product once marketed in the full population, the larger and more comprehensive the preapproval database, the more likely it is that serious adverse events will be detected during drug development.

The appropriate size of a safety database supporting a new product will depend on a number of factors specific to that product, including:

- Its novelty (i.e., whether it represents a new treatment or is similar to available treatment)
- The availability of alternative therapies and the relative safety of those alternatives as compared to the new product
- The intended population and condition being treated
- The intended duration of use

Safety databases for products intended to treat life-threatening diseases, especially in circumstances where there are no alternative satisfactory treatments, are usually smaller than for products intended to treat diseases that are neither life-threatening nor associated with major, irreversible morbidity. A larger safety database may be appropriate if a product's preclinical assessment or human clinical pharmacology studies identify signals of risk that warrant additional clinical data to properly define the risk. The appropriate size of the preapproval safety database may warrant specific discussion with the relevant review division. For instance, 21 CFR 312.82(b) (subpart E) provides that for drugs intended to treat life-threatening and seriously debilitating illnesses, end-of-phase 1 meetings can be used to agree on the design of phase 2 trials "with the goal that such testing will be adequate to provide sufficient data on the drug's safety and effectiveness to support a decision on its approvability for marketing."

For products intended for short-term or acute use (e.g., treatments that continue for, or are cumulatively administered for, less than 6 months), FDA believes it is difficult to offer general guidance on the appropriate target size of clinical safety databases. This is because of the wide range of indications and diseases (e.g., acute strokes to mild headaches) that may be targeted by such therapies. Sponsors are therefore encouraged to discuss with the relevant review division the appropriate size of the safety database for such products. Because products intended for life-threatening and severely debilitating diseases are often approved with relatively small safety databases, relatively greater uncertainty remains regarding their adverse effects. Similarly, when

products offer a unique, clinically important benefit to a population or patient group, less certainty in characterizing risk prior to approval may be acceptable.

For products intended for long-term treatment of non-life-threatening conditions, (e.g., continuous treatment for 6 months or more or recurrent intermittent treatment where cumulative treatment equals or exceeds 6 months), the ICH and FDA have generally recommended that 1500 subjects be exposed to the investigational product (with 300 to 600 exposed for 6 months, and 100 exposed for 1 year).[6] For those products characterized as chronic use products in the ICH guidance E1A, FDA recommends that the 1500 subjects include only those who have been exposed to the product in multiple dose studies, because many adverse events of concern (e.g., hepatotoxicity, hematologic events) do not appear with single doses or very short-term exposure. Also, the 300 to 600 subjects exposed for 6 months and 100 subjects exposed for 1 year should have been exposed to relevant doses (i.e., doses generally in the therapeutic range**)**

We note that it is common for well-conducted clinical development programs to explore doses higher than those ultimately proposed for marketing. For example, a dose tested in clinical trials may offer no efficacy advantage and show some dose-related toxicities; therefore, the sponsor does not propose the dose for marketing when the application is submitted. In such cases, data from subjects exposed to doses in excess of those ultimately proposed are highly informative for the safety evaluation and should be counted as contributing to the relevant safety database.

The E1A guidance describes a number of circumstances in which a safety database larger than 1500 patients may be appropriate, including the following:

1. There is concern that the drug would cause late developing adverse events, or cause adverse events that increase in severity or frequency over time. The concern could arise from:

    - Data from animal studies
    - Clinical information from other agents with related chemical structures or from a related pharmacologic class
    - Pharmacokinetic or pharmacodynamic properties known to be associated with such adverse events

2. There is a need to quantitate the occurrence rate of an expected specific low-frequency adverse event. Examples would include situations where a specific serious adverse event has been identified in similar products or where a serious event that could represent an alert event is observed in early clinical trials.

3. A larger database would help make risk-benefit decisions in situations when the benefit from the product:

    - Is small (e.g., symptomatic improvement in less serious medical conditions)

[6] See the 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*, ICH.

- Will be experienced by only a fraction of the treated patients (e.g., certain preventive therapies administered to healthy populations)
- Is of uncertain magnitude (e.g., efficacy determination on a surrogate endpoint)

4. Concern exists that a product may add to an already significant background rate of morbidity or mortality, and clinical trials should be designed with a sufficient number of patients to provide adequate statistical power to detect prespecified increases over the baseline morbidity or mortality.

The determination of whether the above provisions of the ICH E1A guidance are appropriate for a particular product development program and how these considerations would best be addressed by that program calls for evaluation on a case-by-case basis. Therefore, FDA recommends that this issue be discussed with the relevant review division at the end-of-phase 2 meeting, if not earlier.

In addition to the considerations provided in E1A, there are other circumstances in which a larger database may be appropriate.

1. The proposed treatment is for a healthy population (e.g., the product under development is for chemoprevention or is a preventive vaccine).

2. An effective alternative to the investigational product is already available and has been shown to be safe.

FDA is not suggesting that development of a database larger than that described in E1A is required or should be the norm. Rather, the appropriate database size would depend on the circumstances affecting a particular product, including the considerations outlined above. Therefore, FDA recommends that sponsors communicate with the review division responsible for their product early in the development program (e.g., at the pre-IND meeting) on the appropriate size of the safety database. FDA also recommends that sponsors revisit the issue at appropriate regulatory milestones (e.g., end-of-phase 2 and pre-NDA meetings).

B. Considerations for Developing a Premarketing Safety Database

Although the characteristics of an appropriate safety database are product-specific, some general principles can be applied. In general, efforts to ensure the quality and completeness of a safety database should be comparable to those made to support efficacy. Because data from multiple trials are often examined when assessing safety, it is particularly critical to examine terminology, assessment methods, and use of standard terms (e.g., use of the Medical Dictionary for Regulatory Activities (MedDRA)) to be sure that information is not obscured or distorted. Ascertainment and evaluation of the reasons for leaving assigned therapy during study (deaths and dropouts for any reason) are particularly important for a full understanding of a product's safety profile.

The following elements should be considered by sponsors when developing proposals for their clinical programs as these programs pertain to risk assessment.

1. Long-Term Controlled Safety Studies

It is common in many clinical programs for much of subject exposure data and almost all of long-term exposure data to come from single-arm or uncontrolled studies. Although these data can be informative, it may be preferable in some circumstances to develop controlled, long-term safety data. Such data allow for comparisons of event rates and facilitate accurate attribution of adverse events. Control groups may be given an active comparator or a placebo, depending on the disease being treated (i.e., the ethical and medical feasibility of using a placebo versus an active comparator will depend on the disease being treated).

The usefulness of active comparators in long-term safety studies depends on the adverse events of interest.

- Generally, serious events that rarely occur spontaneously (e.g., severe hepatocellular injury or aplastic anemia) would be considered significant and interpretable whenever (1) they are clearly documented and (2) there is no likely alternative explanation, since the expected rate is essentially zero in populations of any feasible size. As a result, the events can usually be appropriately interpreted and regarded as a signal of concern whether or not there is a control group.

- On the other hand, control groups are needed to detect increases in rates of events that are relatively common in the treated population (e.g., sudden death in patients with ischemic cardiac disease). Control groups are particularly important when an adverse event could be considered part of the disease being treated (e.g., asthma exacerbations occurring with inhalation treatments for asthma).

Therefore, FDA decisions as to when long-term comparative safety studies should be conducted for a product should be based on the intended use of the product, the nature of the labeled patient population (e.g., more useful if there is a high rate of serious adverse events), and earlier clinical and preclinical safety assessments. Although it is clear that long-term controlled studies will not usually be conducted, such studies may be particularly useful when a safety issue is identified during earlier development of the drug. In these cases, safety studies designed to test specific safety hypotheses may be appropriate. This would be especially true in situations where the safety issue of concern is more common with cumulative exposure. (See section IV.D below for further discussion of comparative trials.)

2. A Diverse Safety Database

Premarketing safety databases should include, to the extent possible, a population sufficiently diverse to adequately represent the expected target population, particularly in phase 3 studies. FDA has previously addressed this issue in a memorandum,[7] and the recommendations provided

[7] The memorandum from Janet Woodcock, M.D., to Michael Friedman, M.D., dated July 20, 1998, and titled *FDAMA – Women and Minority Guidance Requirements* (with its attached report) discusses the regulations related to diversity. The memorandum can be found on the CDER guidance page under Modernization Act guidance (http://www.fda.gov/cder/guidance/women.pdf).

here are intended to supplement that document. To the extent feasible, only patients with obvious contraindications or other clinical considerations that clearly dictate exclusion should be excluded from study entry. Inclusion of a diverse population allows for the development of safety data in a broad population that includes patients sometimes excluded from clinical trials, such as the elderly (particularly the very old), patients with concomitant diseases, and patients taking concomitant medications. Broadening inclusion criteria in phase 3 enhances the generalizability of the safety (and efficacy) findings. Although some phase 3 efficacy studies may target certain demographic or disease characteristics (and have narrower inclusion and exclusion criteria), overall, the phase 3 studies should include a substantial amount of data from less restricted populations.

3. Exploring Dose Effects Throughout the Clinical Program

Currently, it is common for only one dose, or perhaps a few doses, to be studied during drug development beyond phase 2. Yet, a number of characteristics common to many phase 2 studies limit the ability of these trials to provide definitive data on exposure-response or adequate data for definitive phase 3 dose selection. These characteristics of phase 2 studies (in comparison to phase 3 studies) include the following:

- Shorter durations of exposure
- Common use of pharmacodynamic (PD) endpoints, rather than clinical outcomes
- Smaller numbers of patients exposed
- Narrowly restrictive entry criteria

Although phase 3 trials do not necessarily need to examine a range of doses, such an examination is highly desirable, particularly when phase 2 studies cannot reasonably be considered to have established a single most appropriate dose. When a dose is not established in phase 2, more than one dose level should be examined in phase 3 trials of fixed dose products to better characterize the relationship between product exposure and resulting clinical benefit and risk. Dose-response data from phase 3 trials with multiple dose levels will help to better define the relationship of clinical response to dose for both safety and effectiveness. Furthermore, inadequate exploration of a product's dose-response relationship in clinical trials can raise safety concerns, since recommending doses in labeling that exceed the amount needed for effectiveness may increase risk to patients through dose-related toxicities with no potential for gain. Exposure-response data from phase 3 trials can also provide critical information on whether dose adjustments should be made for special populations. Finally, demonstrating a dose-response relationship in late phase clinical trials with meaningful clinical endpoints may aid the assessment of efficacy, since showing a dose ordering to efficacy can be compelling evidence of effectiveness.[8] When multiple dose levels are examined in phase 3 trials, the appropriate choice of doses to be included in these studies would be based on prior efficacy and safety information, including prior dose-ranging studies. In these circumstances, an end-of-phase 2 meeting with the appropriate review division would be particularly useful.

[8] See FDA's guidance for industry *Exposure-Response Relationships — Study Design, Data Analysis, and Regulatory Applications.*

C. Detecting Unanticipated Interactions as Part of a Safety Assessment

Even a well-conducted and reasonably complete general clinical pharmacology program does not guarantee a full understanding of all possible risks related to product interactions. Therefore, risk assessment programs should examine a number of interactions during controlled safety and effectiveness trials and, where appropriate, in specific, targeted safety trials. This examination for unanticipated interactions should include the potential for the following:

- Drug-drug interactions in addition to those resulting from known metabolic pathways (e.g., the effect of azole antibiotics on a CYP 3A4 dependent drug)

  We recommend that these examinations target a limited number of specific drugs, such as likely concomitant medications (e.g., for a new cholesterol lowering treatment, examining the consequences of concomitant use of HMG CoA reductase inhibitors and/or binding resins). The interactions of interest could be based, for example, on known or expected patterns of use, indications sought, or populations that are likely users of the drug.

- Product-demographic relationships — by ensuring sufficient diversity of the population (including gender, age, and race) to permit some assessments of safety concerns in demographic population subsets of the intended population

- Product-disease interactions — by ensuring sufficient variability in disease state and concomitant diseases

- Product-dietary supplement interactions for commonly used supplements that are likely to be co-administered or for which reasonable concerns exist (e.g., examination of the interactions between a new drug for the treatment of depression and St. John's Wort).

Again, FDA recommends that any such examinations target likely concomitant use based, for example, on indications sought, intended patterns of use, or the population of intended users of the drug and based on a history of drug and dietary supplement use elicited from subjects.

Generally, a sponsor determines its product's intended use and intended population(s) during product development. Decisions as to which interactions to either explore or specifically test in clinical trials could be based on these determinations and/or surveys and epidemiologic analyses.

One important way to detect unexpected relationships is by systematic incorporation of pharmacokinetic (PK) assessments (e.g., universal steady state sampling or population PK analyses) into some or all of the later phase clinical trials, including any specific safety trials. PK assessments can aid in the detection of unexpected PK interactions and, in some cases, could suggest exposure-response relationships for both safety and efficacy. Such data would allow for better assessment of whether pharmacokinetics contribute to any adverse events seen in the clinical trials, particularly rare, serious, and unanticipated events.

When a product has one or more well-established, valid biomarkers pertinent to a known safety concern, the marker should be studied during the PK studies and clinical development (e.g., creatine phosphokinase assessments used in the evaluation of new HMG CoA reductase inhibitors as a marker for rhabdomyolysis, or assessments of QT/QTc effects for new antihistamines).

D. Developing Comparative Safety Data

Depending on the drug and its indication, much of the safety data in an application may be derived from placebo-controlled trials and single-arm safety studies, with little or no comparative safety data. Although comparative safety data from controlled trials comparing the drug to an active control (these could also include placebo group) generally are not necessary, situations in which such data would be desirable include the following:

- The background rate of adverse events is high.

  The new drug may seem to have a high rate of adverse events in a single-arm study when, in fact, the rate is typical of that for other drugs. The additional use of a placebo would help to show whether either drug actually caused the adverse events.

- There is a well-established treatment with an effect on survival or irreversible morbidity.

  In such cases, not only are comparative data important scientifically, but the use of the comparator would likely be required ethically, as a placebo control could not be used and a single-arm trial would generally be uninformative.

- The sponsor hopes to claim superiority for safety or effectiveness.

  If a comparative effectiveness claim were sought, it would be expected that the studies would also address comparative safety, since a gain in effectiveness could be outweighed by or negated by an accompanying safety disadvantage.

In situations where there is a well-established related therapy, a comparative study of the new agent against that well-established therapy would be desirable (e.g., a new NSAID-like drug could be compared to a market-leading NSAID). Such a study could show whether the toxicity profile for the established therapy is generally similar to that for the novel therapy or whether important differences exist.

V. SPECIAL CONSIDERATIONS FOR RISK ASSESSMENT

Although many of the previous comments and recommendations are intended to apply to new product development programs generally, some risk assessment issues would apply only in certain circumstances or to certain types of products. [9]

[9] The *Pharmacovigilance Guidance* discusses additional risk assessment strategies that may be initiated either pre- or postapproval. In particular, the *Pharmacovigilance Guidance* includes a detailed discussion of

A. Risk Assessment During Product Development

The following are examples of how risk assessment strategies could be tailored to suit special situations, where appropriate.

- If a product is intended to be chronically used (particularly when it has a very long half-life) and/or has dose-related toxicities, it can be useful to examine whether a lower or less frequent maintenance dose would be appropriate.

- If a product's proposed dosing includes a proposed titration scheme, the scheme could be based on specific studies to define how titration is best performed and the effects of titration on safety and efficacy.

- Certain kinds of adverse effects are not likely to be detected or readily reported by patients without special attention. When a drug has the potential for such effects, additional testing or specific assessments within existing trials may be appropriate.

  For example, for a new drug with recognized CNS effects (especially sedating effects), sponsors should conduct an assessment of cognitive function, motor skills, and mood. Similarly, since many antidepressants have significant effects on sexual function, new antidepressants should be assessed for these effects. The use of targeted safety questionnaires or specific psychometric or other validated instruments is often important for such assessments, since routine adverse event monitoring and safety assessments tend to underestimate or even entirely miss such effects.

- If a product is to be studied in pediatric patients, special safety issues should be considered (e.g., effects on growth and neurocognitive development if the drug is to be given to very young children/infants; safety of excipients for the very young; universal immunization recommendations and school entry requirements for immunization).

- A sponsor may consider reserving blood samples (or any other bodily fluids/tissues collected during clinical trials) from some or all patients in phase 3 studies for possible assessments at a later time, particularly in circumstances when earlier safety data signal an unusual or important concern. Such later assessments could include pharmacogenomic markers, assessments for immunogenicity, or measurements of other biomarkers that might prove helpful clinically. Having samples available for retrospective analysis of pharmacogenomic markers could help to link the occurrence of serious adverse events to particular genetic markers (e.g., haplotypes).

In unusual circumstances, a large, simple, safety study (LSSS) may be appropriate. An LSSS is usually a randomized clinical study designed to assess limited, specific outcomes in a large number of patients. These outcomes — generally important safety endpoints or safety concerns

---

pharmacoepidemiologic safety studies. Although such studies would principally be initiated after marketing, the *Pharmacovigilance Guidance* discusses certain situations when they could be initiated preapproval.

suggested by earlier studies — should be defined a priori with the study specifically designed to assess them. Although the large simple study model arose in the context of effectiveness assessment, and thus always involved randomized, controlled trials, an LSSS could in some cases be useful even without a control group — for example, to assess the rate of rare events (i.e., events so uncommon that usual safety studies would not be expected to provide good estimates of risk). Although an LSSS would most commonly be performed postapproval, either as a phase 4 commitment to address a lingering safety issue that does not preclude approval or outside of a formal phase 4 commitment in response to a new safety concern that arises after marketing, there are instances where an LSSS may be appropriate prior to approval. This would be the case when, for instance, there is a significant safety signal of concern (e.g., hepatotoxicity, myotoxicity) arising out of the developing clinical trial database that is not sufficiently resolved by the available data or is unlikely to be sufficiently addressed by the remaining ongoing studies. In these circumstances, an LSSS may be appropriate if the safety signal cannot otherwise be better delineated and the safety signal would have an impact on approvability.[10]

In addition, a sponsor seeking to develop a product for preventive use in at-risk, but otherwise healthy, individuals could conduct a large trial to investigate the product's safety. The use of a large trial may increase the chance of showing the product to have an acceptable benefit-risk profile in such cases, because the potential for benefit in the exposed population would generally be small. Such large trials, though not always LSSSs in a strict sense, may in some cases appropriately employ limited, targeted evaluations of both efficacy and safety endpoints, similar to an LSSS.

B. Assessing and Minimizing the Potential for Medication Errors

Sponsors can help minimize the occurrence of medication errors by assessing, prior to marketing, common sources of medication errors. Such errors may arise because of the product's inherent properties or because of the inadvertent contribution of the proposed proprietary name, the established name, the proposed labeling (e.g., container, carton, patient/consumer labeling, or professional package insert), and the proposed packaging.

Some medication errors, especially those involving parenteral products, have been detected in clinical trials prior to marketing. When occurring in clinical trials, events such as improper dilution or improper administration techniques, which may result in non-optimal dosing, should be carefully examined as warning signs that the product could be subject to dosing errors that warrant changes in labeling, packaging, or design. Even if errors are not observed in trials, careful consideration should be given during development to the implications of the design of the product, its packaging, and any device used to administer or deliver the product. For example, when a concentrated product that requires further dilution prior to intravenous administration is being developed, packaging is important. Packaging such a product in a syringe would make it possible to inject the product as a bolus without proper dilution, increasing risks to patients. Similarly, when developing a product that is administered or delivered by a device, the implications of mechanical failure of the device should be examined. Any such occurrences seen

[10] As mentioned in the *RiskMAP Guidance*, an LSSS could also be a method of evaluating the effectiveness of RiskMAP tools in actual practice prior to approval.

or considered during product development should be documented, reported, and analyzed for potential remedial actions (e.g., redesign of the device or modification of instructions for use).

Medication errors arising from confusion because of the similarity of the drug name, when written and spoken, to the name of another drug are less likely to be detected prior to marketing due to the controlled environment of clinical trials. However, the many well-documented cases of medication errors associated with similar proprietary names, confusing labels and labeling, and product packaging suggest it is important that sponsors carefully consider these issues before marketing a product.

Premarketing assessments should focus on:

- Identifying all medication errors that occur during product development
- Identifying the reasons or causes for each identified error (e.g., dosage form, packaging, labeling, or confusion due to trade names when written or spoken)
- Assessing the resultant risk in the context of how and in whom the product will be used
- Identifying the means to minimize, reduce or eliminate the medication errors by ensuring the proper naming, labeling, design, and packaging of the product

Depending on the nature of the product, the indication, how it is administered, who will be receiving it, and the context in which it will be used, one or more of the following techniques may be helpful in assessing and preventing medication errors:

- Conducting a Failure Mode and Effects Analysis[11, 12]
- Use of expert panels
- Use of computer-assisted analysis
- Use of direct observation during clinical trials
- Directed interviews of consumers and medical and pharmacy personnel to better understand comprehension
- Use of focus groups
- Use of simulated prescription and over-the-counter (OTC) use studies

Additional information on the application of these assessment techniques will be published in a future guidance document.

[11] Stamatis, D.H., *Failure Mode and Effect Analysis: FMEA From Theory to Execution*,. Milwaukee: American Society for Quality, Quality Press, 2003.

[12] Cohen, Michael R. ed., *Medication Errors: Causes, Prevention, and Risk Management*, Washington D.C.: American Pharmaceutical Association, 1999.

C. Addressing Safety Aspects During Product Development

FDA recommends addressing the potential for the following serious adverse effects as a part of the new drug application (NDA) for all new small molecule drugs:

- Drug-related QTc prolongation
- Drug-related liver toxicity
- Drug-related nephrotoxicity
- Drug-related bone marrow toxicity
- Drug-drug interactions
- Polymorphic metabolism

Prior experience has shown that these effects can often be identified when properly assessed in clinical development programs. Although FDA believes it is important to address these potential effects in all NDAs, adequately addressing all of these considerations would not necessarily involve the generation of additional data or the conduct of specific trials. (For some issues, such as QTc, specifically conducted preclinical and clinical studies are generally recommended.) For example, a drug that is intended to be topically applied may be shown to have no systemic bioavailability; therefore, systemic toxicities would be of no practical concern.

Some of the above-listed potential effects are relevant to biological products; some are not. In addition, for biological products such as cytokines, antibodies, other recombinant proteins, and cell-, gene-, and tissue-based therapeutics, it may be appropriate to assess other issues. The issues listed here are dependent on the specific nature of the biological product under development.

- Potentially important issues for biological products include assessments of immunogenicity, both the incidence and consequences of neutralizing antibody formation and the potential for adverse events related to binding antibody formation.

- For gene-based biological products, transfection of nontarget cells and transmissibility of infection to close contacts, and the genetic stability of products intended for long-persistence transfections constitute important safety issues.

- For cell-based products, assessments of adverse events related to distribution, migration, and growth beyond the initial intended administration are important, as are adverse events related to cell survival and demise. Such events may not appear for a long time after product administration.

A complete discussion of assessment of safety issues unique to biological products is beyond the scope of this guidance. We recommend that sponsors address the unique safety concerns pertaining to the development of any particular biological product with the relevant product office.

VI. DATA ANALYSIS AND PRESENTATION

Many aspects of data analysis and presentation have been previously addressed in guidance, most notably in FDA's *Guideline for the Format and Content of the Clinical and Statistical Sections of an Application* and the ICH guidances *E3 Structure and Content of Clinical Study Reports* and *M4 Common Technical Document for the Registration of Pharmaceuticals for Human Use*. We do not repeat that guidance here, but offer new guidance on selected issues.

With regard to the guidance offered in this section of the document, it is important to emphasize that the regulatory approach to the evaluation of the safety of a product usually differs substantially from the evaluation of effectiveness. Most studies in the later phases of drug or biologic development are directed toward establishing effectiveness. In such studies, critical efficacy endpoints are identified in advance, and statistical planning is conducted based on being able to make definitive statistical inferences about efficacy. In contrast, these later phase trials are not generally designed to test specified hypotheses about safety or to measure or identify adverse events with any prespecified level of sensitivity. Therefore, the premarket safety evaluation is often, by its nature, exploratory and is intended to identify common adverse events related to the therapy, as well as to help identify signals for serious and/or less common adverse events.

A. Describing Adverse Events to Identify Safety Signals

Because individual investigators may use different terms to describe a particular adverse event, FDA recommends that sponsors ensure that each investigator's verbatim terms are coded to standardized, preferred terms specified in a coding convention or dictionary. Proper coding allows similar events that were reported using different verbatim language to be appropriately grouped. Consistent and accurate coding of adverse events allows large amounts of data regarding these events to be analyzed and summarized and maximizes the likelihood that safety signals will be detected. Inaccurate coding, inconsistent coding of similar verbatim terms, and inappropriate "lumping" of unrelated verbatim terms or "splitting" of related verbatim terms can obscure safety signals.

In general, FDA suggests that sponsors use one coding convention or dictionary (e.g., MedDRA) throughout a clinical program with the understanding that, due to the duration of product development, the coding convention used may undergo revisions. Use of more than one coding convention or dictionary can result in coding differences that prevent adverse event data from being appropriately grouped and analyzed. To the extent possible, sponsors should use a single version of the selected convention or dictionary without revisions. However, if this is not possible, it is important to appropriately group and analyze adverse events taking into account the revisions in subsequent versions. It is not advisable to analyze adverse event data using one version and then base proposed labeling on a different version.

1. Accuracy of Coding

Sponsors should explore the accuracy of the coding process with respect to both investigators and the persons who code adverse events.

- Investigators may sometimes choose verbatim terms that do not accurately communicate the adverse event that occurred.

  — The severity or magnitude of an event may be inappropriately exaggerated (e.g., if an investigator terms a case of isolated elevated transaminases *acute liver failure* despite the absence of evidence of associated hyperbilirubinemia, coagulopathy, or encephalopathy, which are components of the standard definition of acute liver failure).

  — Conversely, the significance or existence of an event may be masked (e.g., if an investigator uses a term that is nonspecific and possibly unimportant to describe a subject's discontinuation from a study when the discontinuation is due to a serious adverse event).

  If an adverse event is mischaracterized, sponsors could consider, in consultation with FDA, recharacterizing the event to make it consistent with accepted case definitions. We recommend that recharacterization be the exception rather than the rule and, when done, be well documented with an audit trail.

- We recommend that in addition to ensuring that investigators have accurately characterized adverse events, sponsors confirm that verbatim terms used by investigators have been appropriately coded.

  Sponsors should strive to identify obvious coding mistakes as well as any instances when a potentially serious verbatim term may have been inappropriately mapped to a more benign coding term, thus minimizing the potential severity of an adverse event. One example is coding the verbatim term *facial edema* (suggesting an allergic reaction) as the nonspecific term *edema*; another is coding the verbatim term *suicidal ideation* as the more benign term *emotional lability*.

- Prior to analyzing a product's safety database, sponsors should ensure that adverse events were coded with minimal variability across studies and individual coders.

  Consistency is important because adverse event coding may be performed over time, as studies are completed, and by many different individuals. Both of these factors are potential sources of variability in the coding process. FDA recommends that to examine the extent of variability in the coding process, sponsors focus on a subset of preferred terms, particularly terms that are vague and commonly coded differently by different people. For example, a sponsor might evaluate the consistency of coding verbatim terms such as *weakness* and *asthenia* or *dizziness* and *vertigo*. NOS (not otherwise specified)-type codes, such as *ECG abnormality NOS*, are also coding terms to which a variety of verbatim terms may often be mapped. These should be examined for consistency as well. Sponsors should pay special attention to terms that could represent serious or otherwise important adverse reactions.

In addition to considering an adverse event independently and as it is initially coded, sponsors should also consider a coded event in conjunction with other coded events in some

circumstances. Certain adverse events or toxicities (particularly those with a constellation of symptoms, signs, or laboratory findings) may be defined as an amalgamation of multiple preferred coding terms. Sponsors should identify these events (e.g., acute liver failure) based on recognized definitions.

2. Coding Considerations During Adverse Event Analysis

When analyzing an adverse event, sponsors should consider the following:

- Combining related coding terms can either amplify weak safety signals or obscure important toxicities.

  For example, the combination of dyspnea, cough, wheezing, or pleuritis might provide a more sensitive, although less specific, appraisal of pulmonary toxicity than any single term. Conversely, by combining terms for serious, unusual events with terms for more common, less serious events (e.g., constipation might include cases of toxic megacolon), the more important events could be obscured.

- Coding methods can divide the same event into many terms. Dividing adverse event terms can decrease the apparent incidence of an adverse event (e.g., including pedal edema, generalized edema, and peripheral edema as separate terms could obscure the overall finding of fluid retention).

Although potentially important safety events cannot always be anticipated in a clinical development program, sponsors, in consultation with the Agency, should prospectively group adverse event terms and develop case definitions or use accepted standardized definitions whenever possible.

- A prospective grouping approach is particularly important for syndromes such as serotonin syndrome, Parkinsonism, and drug withdrawal, which are not well characterized by a single term.

- Some groupings can be constructed only after safety data are obtained, at which time consultation with FDA might be considered.

- Sponsors should explain such groupings explicitly in their applications so that FDA reviewers have a clear understanding of what terms were grouped and the rationale for the groupings.

- For safety signals that are identified toward the end of a development program, the pre-NDA meeting would be a reasonable time to confer with FDA regarding such groupings or case definitions.

B. Analyzing Temporal or Other Associations

For individual safety reports, the temporal relationship between product exposure and adverse event is a critical consideration in the assessment of potential causality. However, temporal factors, including the duration of the event itself, are often overlooked during the assessment of aggregate safety data. Simple comparisons of adverse event frequencies between (or among) treatment groups, which are commonly included in product applications and reproduced in tabular format in labeling, generally do not take into account the time dependency of adverse events. Temporal associations can help further understand causality, adaptation, and tolerance, but may be obscured when only frequencies of adverse events are compared.

Temporal analyses may be warranted for important adverse events whether they arise from controlled clinical trial data or treatment cohorts. In both cases, analyzing changes over time may be important for assessing risk and potential causality. Analyses of temporal associations are particularly worth conducting in situations where prior experience (e.g., experience from similar products) has shown that a temporal relationship between product exposure and ensuing adverse events is likely to exist. In addition, in the context of controlled clinical trials, temporal analyses may provide insight into the relative importance of differences in adverse event frequencies between study groups.

Descriptions of risk as a function of subjects' duration of exposure to a product, or as a function of time since initial exposure, can contribute to the understanding of the product's safety profile. Assessments of risk within discrete time intervals over the observation period (i.e., a hazard rate curve) can be used to illustrate changes in risk over time (e.g., flu-like symptoms with interferons that tend to occur at the initiation of treatment but diminish in frequency over time). It may be useful for sponsors to consider event rates (events per unit of time) in reconciling apparent differences in the frequencies of events between studies when there are disparities in subjects' time of exposure or time at risk.

For important events that do not occur at a constant rate with respect to time and for events in studies where the size of the population at risk (denominator) changes over time, a life-table or Kaplan-Meier approach may be of value for evaluating risks of adverse events. Clinically important events (e.g., those events for which the occurrence of even a few cases in a database may be significant) are of particular interest. Examples of such events include the development of restenosis following coronary angioplasty, cardiac toxicity, and seizures.

Temporal associations identified in previous experience with related products can help focus sponsor analyses of potential temporal associations for a product under study, but sponsors should balance this approach with an attempt to detect unanticipated events and associations as well. Knowledge of a product's pharmacokinetic and pharmacodynamic profiles, as well as an appreciation of physiologic, metabolic, and host immune responses, may be important in understanding the possible timing of treatment-related adverse events.

It is important to consider study and concomitant treatment regimens (i.e., single treatment; short course of treatment; continuous, intermittent, titrated, or symptom-based treatment) in temporal analyses. Other important factors to consider in planning and interpreting temporal analyses are

(1) the initiation or withdrawal of therapies and (2) changes in the severity or frequency of subjects' preexisting conditions over time.

For events that decrease in frequency over time and are found to be associated with the initiation of treatment, supplemental analyses may be of value to discriminate the relative contributions of adaptation, tolerance, dose reduction, symptomatic treatment, decreases in reporting, depletion of susceptibles, and subject dropout.

C. Analyzing Dose Effect as a Contribution to Risk Assessment

Sponsors should analyze event rates by dose for clinically important adverse events that may be product related and events that might be expected based on a product's pharmacologic class or preclinical data.

For studies involving the evaluation of a range of doses, dose response is most commonly assessed by analyzing adverse event frequencies by administered dose. In such studies, it may also be useful to consider event frequencies by weight-adjusted or body surface area-adjusted dose, especially if most patients are given the same dose regardless of body weight or size. It should be recognized, however, that when doses are adjusted by a subject's weight or body surface area, women are commonly overrepresented on the upper end of the range of adjusted doses, and men are commonly overrepresented on the lower end of this range. For products administered over prolonged periods, it may be useful to analyze event rates based on cumulative dose. In addition, when specific demographic or baseline disease-related subgroups may be at particular risk of incurring adverse events, exploration of dose-response relationships by subgroup is important. Subgroup analyses have the potential to provide a more reliable and relevant estimate of risk for important subgroups of the target population. Alternatively, multiplicity issues could result in an apparent signal that does not represent a real finding (i.e., a false positive).

Although the most reliable information on dose response comes from randomized fixed dose studies, potentially useful information may emerge from titration studies and from associations between adverse events and plasma drug concentrations.

For dose titration or flexible dose studies, it would generally be useful to assess the relationship between adverse event frequencies and the actual doses subjects received preceding the adverse events or the cumulative dose they received at the onset of the events. The choice is a function of the mode of action, pharmacokinetics, and pharmacodynamics of the product.

For products with a stepped dosing algorithm (i.e., incremental dosing based on age or weight), the actual cut points of the paradigm are often selected relatively early in product development. Although the cut points may be based on the best knowledge available at the time, it is useful in such cases to make a specific effort to explore safety (and efficacy) just above and below these points. For example, if the dose of a product is to be 100 mg for patients weighing less than 80 kg and 150 mg for patients weighing 80 kg or more, an assessment of the comparative safety profiles of patients weighing from 75 to 79.9 kg versus patients weighing from 80 to 84.9 kg would be valuable.

As is typical of most safety evaluations, the likelihood of observing false positive signals increases with the number of analyses conducted. Positive associations between adverse events and dose, as well as signals that emerge from subgroup analysis, should be considered with this in mind. Such associations should be examined for consistency across studies, if possible.

D. Role of Data Pooling in Risk Assessment

Data pooling is the integration of patient-level outcome data from several clinical studies to assess a safety outcome of interest. Generally, data pooling is performed to achieve larger sample sizes and data sets because individual clinical studies are not designed with sufficient sample size to estimate the frequency of low incidence events or to compare differences in rates or relative rates between the test drug (exposed group) and the control (unexposed group). Use of pooled data does not imply that individual study results should not be examined and considered. When pooling data, sponsors should consider the possibility that various sources of systematic differences can interfere with interpretation of a pooled result. To ensure that pooling is appropriate, sponsors should confirm that study designs, as well as ascertainment and measurement strategies employed in the studies that are pooled, are reasonably similar.

Used appropriately, pooled analyses can enhance the power to detect an association between product use and an event and provide more reliable estimates of the magnitude of risk over time. Pooled analyses can also provide insight into a positive signal observed in a single study by allowing a broader comparison. This can protect against undue weight being given to chance findings in individual studies. However, a finding from a single study should not be automatically dismissed because of the results of a pooled analysis, especially if it is detected in a study of superior design or in a different population. Any pooled analysis resulting in a reduced statistical association between a product and an observed risk or magnitude of risk, as compared to the original safety signal obtained from one or more of the contributing studies, should be carefully examined.

Some issues for consideration in deciding whether pooling is appropriate include possible differences in the duration of studies, heterogeneity of patient populations, and case ascertainment differences across studies (i.e., different methods for detecting the safety outcomes of interest, such as differences in the intensities of patient follow-up). When there is clinical heterogeneity among trials with regard to the safety outcome of interest (e.g., major disparity in findings for particular safety endpoints), sponsors should present risk information that details the range of results observed in the individual studies, rather than producing a summary value from a pooled analysis.

E. Using Pooled Data for Risk Assessment

All placebo-controlled studies in a clinical development program should be considered and evaluated for appropriateness for inclusion in a pooled analysis. Decisions to exclude certain placebo-controlled studies from, or to add other types of studies (such as active-controlled studies or open-label studies) to, a pooled analysis would depend on the objectives of the analysis. Such analyses should be conducted in a manner that is consistent with the following guiding principles:

- Generally, phase 1 pharmacokinetic and pharmacodynamic studies should be excluded.

  These are usually single- or multiple-dose trials of a short duration conducted in healthy subjects or in patients with refractory or incurable end-stage disease who have confounding symptoms. Unless a risk were limited to a short period immediately after the first dose, inclusion of these studies in a pooled analysis would not increase the statistical power or contribute to the precision of the risk estimates. However, inclusion of these studies could (1) diminish the magnitude of apparent risk by including a population with little or no possibility of having had the adverse reaction or (2) increase the apparent magnitude of risk because of significant baseline symptoms unrelated to the drug.

- The risk of the safety outcome of interest should be expressed in reference to total person-time (exposure time) or be evaluated using a time-to-event analysis.

  When the duration of drug exposure for the individual subjects included in a pooled analysis varies, sponsors should not express the risk merely in terms of *event frequency* (that is, using persons as the denominator). Use of the person-time approach relies on the assumption that the risk is constant over the period of the studies. Whenever there is concern regarding a non-constant nature of a risk, a time-to-event log-rank type analysis may be helpful, as it is a robust approach even when risk is not constant over time.

- The patient population in the pooled analysis should be relatively homogeneous with respect to factors that may affect the safety outcome of interest (e.g., dose received, duration of therapy).

  The pooled analysis should be of a size sufficient to allow analyses of demographic subgroups (gender, age, race, geographic locations).

- The studies included in a pooled analysis should have used similar methods of adverse event ascertainment, including ascertainment of the cause of dropouts.

  Study-specific incidence rate should be calculated and compared for any signs of case ascertainment differences. Since study-to-study variation is to be expected, it is a challenge to distinguish between possible case ascertainment differences and study-to-study variation.

There are some situations in which pooling may be relatively straightforward. For example, a pooled analysis of similarly designed phase 3 studies could readily be used to create a table of common adverse events. This type of analysis is typically less subject to the problems discussed above because (1) the studies are similar in study design and patient population and (2) the intent of such an analysis is often more descriptive than quantitative. However, if a specific safety concern is raised during the clinical development program, the guiding principles discussed above should be closely followed when conducting a prespecified pooled analysis.

F. Rigorous Ascertainment of Reasons for Withdrawals from Studies

Subjects may drop out or withdraw from clinical trials for many reasons, including perceived lack of efficacy, side effects, serious adverse events, or an unwillingness to expend the effort necessary to continue. The reasons for dropout are not always clear. This lack of information may be largely irrelevant (e.g., discontinuation due to moving from the area) or indicative of an important safety problem (e.g., stroke). Therefore, regardless of the reason for withdrawal, sponsors should attempt to account for all dropouts.

- Sponsors should try to ascertain what precipitated dropout or withdrawal in all cases, particularly if a safety issue was a part of the reason for withdrawal.

  It is not helpful to simply record vague explanations such as "withdrew consent," "failed to return," "administratively withdrawn," or "lost to follow-up."

- Participants who leave a study because of serious or significant safety issues should be followed closely until the adverse events are fully and permanently resolved or stabilized (if complete resolution is not anticipated), with follow-up data recorded in the case report forms.

- Follow-up information should be pursued on patients withdrawn from the study (for reasons other than withdrawing consent in the absence of an adverse event).

  If this information is not obtainable, FDA recommends that the measures taken to obtain follow-up information be reflected on the case report forms and the resultant failure to obtain the information should be discussed in the clinical discussion of safety.

- Patients considering withdrawing consent should be encouraged to provide the reason, and patients who withdraw should be encouraged to provide information as to whether the withdrawal of consent resulted from a serious or significant safety issue.

- Some patients withdraw due to abnormal laboratory values, vital signs, or ECG findings that are not characterized as adverse events. Sponsors should include information on these types of discontinuations in addition to information on discontinuations due to adverse events.

G. Long-term Follow-up

In some cases, it is recommended that all subjects be followed to the end of the study or even after the formal end of the study (e.g., where the drug has a very long half-life, is deposited in an organ such as bone or brain, or has the potential for causing irreversible effects, such as cancer). The concern over adequate follow-up for ascertaining important safety events in such cases is particularly critical in long-term treatment and clinical outcome studies. In such cases, FDA recommends the follow-up for late safety events, even for subjects off therapy, include those subjects who drop out of the trial or who finish the study early due to meeting a primary outcome of interest. The duration of follow-up, however, would be dependent on the circumstances of the

product development and therefore should be discussed with the appropriate review division (e.g., during end-of-phase 2 meetings).

H. Important Aspects of Data Presentation

We recommend that once a product's safety data have been analyzed, comprehensive risk assessment information be presented succinctly. FDA and ICH have provided extensive guidance regarding the presentation of safety data,[13,14, 15] and we offer these additional recommendations, which have not been addressed previously.

- For selected adverse events, adverse event rates using a range of more restrictive to less restrictive definitions (e.g., myocardial infarction versus myocardial ischemia) should be summarized.

  The events chosen for such a summary might be limited to more serious events and events that are recognized to be associated with the relevant class of drugs;

- For a drug that is a new member of an established class of drugs, the adverse events that are important for the class of drug should be fully characterized in the NDA's integrated summary of safety.

  That characterization should include an analysis of the incidence of the pertinent adverse events, as well as any associated laboratory, vital sign, or ECG data. For example, the characterization of a drug joining a class that is associated with orthostatic hypotension would include analyses of orthostatic blood pressure changes as well as the incidence of syncope, dizziness, falls, or other events. We recommend that when sponsors are establishing case definitions for particular adverse events, they consider definitions previously used for the other drugs in the class or, if available, standard definitions.

- The distribution of important variables across the pooled data, such as gender, age, extent of exposure, concomitant medical conditions, and concomitant medications (especially those that are used commonly to treat the indication being studied), should be included in the integrated summary of safety.

- The effect of differential discontinuation rates by treatment on adverse event occurrence should be characterized (e.g., when placebo-treated patients drop out of a trial earlier than patients being treated with an active drug). This differential discontinuation can lead to misleading adverse event incidences unless patient exposure is used as the denominator for risk calculations.

- Case report forms (CRFs) submitted for patients who died or discontinued a study prematurely due to an adverse event should include copies of relevant hospital records,

[13] See *Guideline for the Format and Content of the Clinical and Statistical Section of an Application*.
[14] See the guidance for industry *E3 Structure and Content of Clinical Study Reports*, ICH.
[15] See the guidance for industry *M4 Common Technical Document for the Registration of Pharmaceuticals for Human Use*, ICH.

autopsy reports, biopsy reports, and radiological reports, when feasible. The possibility that such information may be reported to FDA should be stated in the informed consent document with a notation that the patient would not be identified in such reports.

These source documents should become a formal part of the official CRF and be properly referenced.

- Narrative summaries (as previously described in guidance[16]) of important adverse events (e.g., deaths, events leading to discontinuation, other serious adverse events) should provide the detail necessary to permit an adequate understanding of the nature of the adverse event experienced by the study subject. (This level of detail may be unnecessary for events expected in the population (e.g., late deaths in a cancer trial). This issue should be discussed with the appropriate review division.)

  Narrative summaries should not merely provide, in text format, the data that are already presented in the case report tabulation, as this adds little value. A valuable narrative summary would provide a complete synthesis of all available clinical data and an informed discussion of the case, allowing a better understanding of what the patient experienced. The following is a list of components that would be found in a useful narrative summary:

  – Patient age and gender
  – Signs and symptoms related to the adverse event being discussed
  – An assessment of the relationship of exposure duration to the development of the adverse event
  – Pertinent medical history
  – Concomitant medications with start dates relative to the adverse event
  – Pertinent physical exam findings
  – Pertinent test results (e.g., lab data, ECG data, biopsy data)
  – Discussion of the diagnosis as supported by available clinical data
  – For events without a definitive diagnosis, a list of the differential diagnoses
  – Treatment provided
  – Re-challenge results (if performed)
  – Outcomes and follow-up information

[16] See the guidance for industry *E3 Structure and Content of Clinical Study Reports*, ICH.