
Good Review Practice: Clinical Review of Investigational New Drug Applications

This document has been prepared by the Office of New Drugs in the Center for Drug Evaluation and Research at the Food and Drug Administration.

December 2013

Table of Contents

1.	INTRODUCTION	1
2.	GENERAL CONSIDERATIONS	2
2.1	Pre-IND Meeting/IND Original Submission	2
2.2	Phase 1 Clinical Trial Protocol	4
2.3	End-of-Phase 2/Phase 3 Planning	5
2.4	Controlled Clinical Trial Protocol Review (including Special Protocol Assessments)	8
2.5	Fast Track or Breakthrough Designation	11
2.6	IND Safety Reports (21 CFR 312.32(c))	12
3.	DOSING AND CLINICAL PHARMACOLOGY.....	13
3.1	Phase 1 Tolerability Trials	13
3.1.1	Choosing a Starting Dose for Phase 1	13
3.1.2	Dose-Escalation and Maximum Dose and Duration in Phase 1 ...	15
3.1.3	Toxicity-Induced Modifications in Enrollment or Dosing (Safety Rules)	17
3.2	Pharmacokinetic and Pharmacodynamic Trials.....	17
3.2.1	Effect of Intrinsic and Extrinsic Factors on PK and PD	18
3.2.2	Classic PK Clinical Trials (Frequent Sampling).....	19
3.2.3	Population PK Clinical Trials	19
3.2.4	Bioavailability and Bioequivalence Trials.....	19
3.2.5	Drug-Drug Interactions.....	20
3.3	Choice of Dosing Interval.....	21
4.	ASSESSING DOSE-RESPONSE.....	22
4.1	Fixed-Dose Clinical Trials	24
4.2	Titration Clinical Trials.....	25
4.3	Crossover Dose-Response Trials	26
5.	CONTROLS, TRUTH STANDARDS, AND COMPLIANCE	26
5.1	Types of Controls.....	26
5.1.1	Placebo Control.....	27
5.1.2	No-Treatment Control.....	28
5.1.3	Dose-Comparison Control	28
5.1.4	Active-Treatment Control.....	29
5.1.5	External (Historical) Control	31
5.2	Trial Design Features.....	32
5.2.1	Randomized Withdrawal Trials	32
5.2.2	Adaptive Designs	33
5.2.3	Enrichment.....	33
5.2.4	Crossover and Multiple Treatment Trials	34
5.2.5	Trials in Nonresponders or Intolerants	34
5.3	Truth Standards	34
5.4	Assessing Treatment Compliance.....	35
5.5	Background Care and Standard of Care.....	35
6.	RANDOMIZATION AND BLINDING.....	36

6.1	Randomization	36
6.1.1	Fixed-Randomization Schemes	37
6.1.1.1	Blocked randomization	37
6.1.1.2	Stratified randomization	38
6.1.2	Adaptive-Randomization Schemes	39
6.1.3	Allocation Ratio	39
6.1.4	Review of Randomization.....	40
6.2	Blinding.....	40
6.2.1	Optimizing and Maintaining the Blinding	41
6.2.1.1	Character of the placebo and trial drug	41
6.2.1.2	Unblinding drug effects	41
6.2.1.3	Dealing with imperfect blinding	42
6.2.1.4	Assessing unblinding	42
6.2.1.5	Blinded evaluators or evaluation committees	42
7.	PATIENT POPULATIONS, SPECIAL POPULATIONS.....	43
7.1	Trial Population	43
7.1.1	Homogeneity vs. Heterogeneity; Individualization of Treatment	44
7.1.2	Factors Influencing the Nature of the Trial Population	44
7.2	Special Populations, Demographic Subgroups	45
7.2.1	Subgroup Analyses vs. Special Population Trials	46
7.2.2	Pediatric Populations	46
7.2.2.1	Timing of studies in pediatric populations	49
7.2.2.2	Types of studies in pediatric populations	51
7.2.3	Women.....	52
7.2.3.1	PK issues regarding women	53
7.2.3.2	Interactions with oral contraceptives	53
7.2.3.3	Studying pregnant women	54
7.2.4	Elderly Subjects	54
7.2.5	Racial Groups.....	55
7.2.6	Other Subpopulations of Interest: Genetic, Proteomic, and Concomitant Illness	56
7.3	Patient Population Size	57
7.3.1	Sample Size in Phase 1 Clinical Trials	57
7.3.2	Sample Size in Phase 2 and Phase 3 Clinical Trials	57
7.3.3	Total Population Exposure.....	59
8.	STATISTICAL ANALYSIS PLANS	60
8.1	Planned Analyses	60
8.1.1	Adequacy of the Statistical Analysis Plan	61
8.1.2	Reviewing Changes to the Statistical Analysis Plan.....	63
8.1.3	Interim Analysis Plans	64
8.1.3.1	Confidentiality of interim data	64
8.1.3.2	Stopping rules for an early finding of efficacy or toxicity	65
8.1.3.3	Risks of early stopping for efficacy	66
8.1.3.4	Unplanned interim analyses	66
8.1.3.5	Reviewer's role during the trial	67

	8.1.4	Intent-to-Treat Analysis	67
8.2		Endpoints	68
	8.2.1	Primary Endpoints	69
		8.2.1.1 Composite endpoints	72
	8.2.2	Secondary Endpoints	73
		8.2.2.1 Descriptive analyses	74
	8.2.3	Surrogate Endpoints.....	74
	8.2.4	Patient-Reported Outcome Measures	76
8.3		Discussions With the Sponsor	77
	8.3.1	Target Product Profile.....	77
	8.3.2	Continuous Involvement.....	77
9.		GOOD CLINICAL PRACTICES	78
	9.1	The Institutional Review Board	79
	9.2	Informed Consent.....	80
		9.2.1 Consent in Pediatric or Other Vulnerable Populations	81
		9.2.2 Waiver of Informed Consent	82
	9.3	Investigator’s Brochure.....	82
	9.4	Investigator Qualifications and Responsibilities	83
	9.5	Trial Monitoring and Auditing.....	83
10.		ADVERSE DRUG EXPERIENCES AND REPORTS.....	84
	10.1	Safety Monitoring	84
	10.2	Reporting Requirements for Sponsors	86
	10.3	IND Safety Reports — Written Reports	87
		10.3.1 Definition: Adverse Event	88
		10.3.2 Definition: Adverse Reaction	88
		10.3.3 Definition: Suspected Adverse Reaction	89
		10.3.4 Definition: Unexpected.....	89
		10.3.5 Definition: Serious	89
		10.3.6 Definition: Life-Threatening.....	90
	10.4	Assessment of Adverse Drug Reports	90
	10.5	Actions After Reviewing Adverse Drug Reports	91
	10.6	Annual Reports	92
	10.7	Other Safety Data.....	93
11.		EXPEDITED DRUG DEVELOPMENT PROGRAMS	93
	11.1	Serious or Life-Threatening Condition.....	94
	11.2	Available Therapy.....	95
	11.3	Demonstrating the Potential to Address Unmet Medical Need	96
	11.4	Qualifying Criteria for Expedited Drug Development Designations.....	97
		11.4.1 Fast Track	97
		11.4.2 Breakthrough Therapy	98
	11.5	Features of Expedited Drug Development Designations....	100
		11.5.1 Features of Fast Track.....	100
		11.5.2 Features of Breakthrough.....	100
12.		REFERENCES	101

13. GLOSSARY OF ACRONYMS107

Good Review Practice: Clinical Review of Investigational New Drug Applications

1. INTRODUCTION

This good review practice (GRP) document was prepared to assist FDA clinical review staff in reviewing clinical submissions to an investigational new drug application (IND) from the pre-IND phase to the time of the pre-new drug application/biologics license application meeting.¹ Although the primary focus is on the clinical review of INDs for new molecular entities, many of the principles can be applied to all INDs and to new drug applications/biologics license applications (NDAs/BLAs) and their supplements.

This document identifies and describes issues that should be prospectively considered during IND development to facilitate development of a complete, high-quality database that could be submitted in an NDA/BLA. It also identifies important areas that may warrant additional consideration and discussion.²

The extent and type of clinical review and communication with the sponsor for each submission to the IND will vary, depending upon the drug's novelty, FDA familiarity with other drugs in the same class, potential safety concerns, the stage of development, and the disease the drug is intended to diagnose, treat, or prevent. Because CDER has limited resources, CDER review staff will prioritize submission review based on: (1) relative importance to patient safety; and (2) the context of the sponsor's development plan.

The term *review* refers to the formal process of review and what should be considered, not to the reviewer's report documenting that review. The written report will reflect many of the considerations discussed in this document, but the length of the document is not necessarily reflective of the thoroughness of the review.

This GRP document is organized as follows:

- Section 2, General Considerations: Organized by major stages of drug development; lists questions for reviewers to consider
- Sections 3 to 11: Presents detailed discussions of critical aspects of overall development and trial design; expands on components of section 2

¹ An investigational drug is defined as any drug or biologic that is used in a study or clinical trial. For the purposes of this document, all references to drugs include both human drugs and therapeutic biological products regulated by CDER unless otherwise specified.

² The Federal Food, Drug, and Cosmetic Act (FD&C Act), the regulations, and finalized guidances take priority over this document.

- Sections 12 to 13: Glossary and references to other relevant documents

Hypertext links are provided for ease of navigation and cross-reference throughout this document.

2. GENERAL CONSIDERATIONS

The following links contain lists of questions to help the reviewer during various stages of drug development to ensure review completeness and consistency. The lists can be particularly helpful when a sponsor has not raised critical issues for discussion and for reminding reviewers of the most important issues that they should communicate to sponsors. They are arranged by major stages in drug development.

Select the appropriate quicklink for the submission type:

- Section 2.1 [Pre-IND Meeting/IND Original Submission](#)
- Section 2.2 [Phase 1 Clinical Trial Protocol](#)
- Section 2.3 [End-of-Phase 2/Phase 3 Planning](#)
- Section 2.4 [Controlled Clinical Trial Protocol Review \(including Special Protocol Assessments\)](#)
- Section 2.5 [Fast Track or Breakthrough Designation](#)
- Section 2.6 [IND Safety Reports \(21 CFR 312.32\(c\)\)](#)

2.1 Pre-IND Meeting/IND Original Submission

When reviewing an original IND submission or planning for a pre-IND meeting, the FDA review team's primary concern is the safety of the subjects who will receive the drug during the proposed clinical trial. A secondary consideration is evaluation of the initial drug development plan and the role of the proposed study or clinical trial in that plan. The following list includes questions that should be considered during the pre-IND/IND period.

- Are the purity, potency, stability, and sterility (if applicable) of the drug adequate to support the proposed development phase? This item is the primary responsibility of the review chemists/product quality reviewers and microbiology review staff with whom this should be discussed.³

³ See the guidance for industry *IND Meetings for Human Drugs and Biologics; Chemistry, Manufacturing, and Controls Information*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM070568.pdf>)

- Do animal studies provide sufficient safety support for the starting dose and schedule in the proposed or planned clinical trials? This item is the primary responsibility of the pharmacology and toxicology review staff with whom this should be discussed.
 - Choice or relevance of animal species and model used as a basis for dose selection? (See [section 3.1.1](#), Choosing a Starting Dose for Phase 1.)
 - Identification of potential target organs of toxicity and potential means for monitoring those toxicities?
- Is the plan for human pharmacokinetic (PK) and pharmacodynamic (PD) trials sufficient and does it appropriately reflect nonclinical findings? This item is the primary responsibility of the clinical pharmacology review staff with whom this should be discussed.
 - Duration, dose, schedule, and route of exposure? (See [section 3.1.1](#), Choosing a Starting Dose for Phase 1; and [section 3.1.2](#), Dose-Escalation and Maximum Dose and Duration in Phase 1.)
 - PK and PD assessments? (See [section 3.2](#), Pharmacokinetic and Pharmacodynamic Trials.)
 - Identification of dose-response? (See [section 3.1.2](#), Dose-Escalation and Maximum Dose and Duration in Phase 1.)
 - Safety in special populations to be studied (e.g., neonates, pregnant women, renal and hepatic impaired patients)? (See [section 7.2](#), Special Populations, Demographic Subgroups.)
- Clinical/regulatory issues
 - Is the overall drug development plan described in detail and appropriate? If so, are the indications sought clear? (See [section 8.3.1](#), Target Product Profile.)
 - Are the phase 1 clinical trial protocols appropriately designed to ensure safety and meet objectives, including criteria for dose escalation (e.g., schema, number of subjects per dose level, observation period, and number of subjects exposed before dose escalation)? (See [section 7.3.1](#), Sample Size in Phase 1 Clinical Trials.)
 - Does the overall drug development plan include consideration of or plans to study appropriate pediatric populations? (See [section 7.2.2](#), Pediatric Populations.)
 - Is the drug being developed under the animal efficacy rule (21 CFR part 314, subpart I, Approval of New Drugs When Human Efficacy Studies Are Not Ethical or Feasible, or 21 CFR part 601, subpart H, Approval of Biological Products When Human Efficacy Studies Are Not Ethical or Feasible)? CDER's Office of Counter-Terrorism and Emergency Coordination should be consulted regarding any protocol or meeting regarding animal models.

- For indications with a history of high trial failure rate and/or for submissions with poor dose selection processes, has the sponsor considered requesting an end-of-phase 2A (EOP2A) meeting?⁴

2.2 Phase 1 Clinical Trial Protocol

- Are the purity, potency, stability, and sterility (if applicable) of the drug adequate to support phase 1? This item is the primary responsibility of the review chemists/product quality reviewers and microbiology review staff with whom this should be discussed.⁵
- Is subject selection appropriate? Are healthy volunteers acceptable, given drug toxicity or mechanism of action?
- Is the proposed starting dose appropriate with an acceptable safety margin? Are nonclinical data properly taken into account? (See [section 3.1.1](#), Choosing a Starting Dose for Phase 1.)
- Is the dose escalation scheme appropriate? Are nonclinical data and concerns addressed? (See [section 3.1.2](#), Dose-Escalation and Maximum Dose and Duration in Phase 1; [section 3.3](#), Choice of Dosing Interval; and [section 4](#), Assessing Dose-Response.)
- Is the amount of information available (e.g., duration of observation, laboratory, and clinical observations) and the plan to consider additional accrued data before each escalation appropriate? (See [section 3.1.2](#), Dose-Escalation and Maximum Dose and Duration in Phase 1; [section 3.3](#), Choice of Dosing Interval; and [section 4](#), Assessing Dose-Response.)
- Is the number of subjects treated at each dose appropriate, and is there an appropriate duration of observation before treating the next subject in each cohort and before subsequent dose escalation? (See [section 3.1.2](#), Dose-Escalation and Maximum Dose and Duration in Phase 1; [section 3.3](#), Choice of Dosing Interval; and [section 4](#), Assessing Dose-Response.)
- Is the size of each dose increment appropriate? (See [section 3.1.2](#), Dose-Escalation and Maximum Dose and Duration in Phase 1.)
- Is the monitoring scheme for toxicity and pharmacologic activity appropriate? Is the case report form (CRF) adequate, if applicable? (See [section 10.1](#), Safety Monitoring.)

⁴ See the guidance for industry *End-of-Phase 2A Meetings*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm079690.pdf>)

⁵ See the guidance for industry *Content and Format of Investigational New Drug Applications (INDs) for Phase 1 Studies of Drugs, Including Well-Characterized, Therapeutic, Biotechnology-derived Products*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM071597.pdf>)

- For multicenter trials, is the plan adequate for sharing safety data among clinical investigators as well as with the sponsor?
- Are the rules for stopping the administration of the investigational drug, stopping enrollment, and stopping dose escalation clear and appropriate? (See [section 3.1.3](#), Toxicity-Induced Modifications in Enrollment or Dosing (Safety Rules).)
- Are PK and PD data appropriately collected? Are immunogenicity data for biologics appropriately collected? (See [section 3](#), Dosing and Clinical Pharmacology.)
- Does the consent form (when submitted or requested by the FDA for review) contain all the necessary elements and is it inaccurate or misleading? (See [section 9.2](#), Informed Consent.)
- Is the investigator's brochure complete, not misleading, and up to date? (An investigator's brochure is not required for single center trials.) (See [section 9.3](#), Investigator's Brochure.)
- If the investigator has a particular interest in the success of the trial, such as financial (development of a patent or stake in a drug as an employee of clinical research organization, or being paid with stock options) or intellectual (evaluation of pet theory), what arrangements have been made for involvement, accompaniment, and assessment by a disinterested party?
- For new phase 1 trials submitted to an IND, were data from previously conducted trials adequately taken into account when designing the new trial, including the choice of dose and dose-escalation schema?
- Is the product being developed under the animal efficacy rule, 21 CFR 314.600 (subpart I, Approval of New Drugs When Human Efficacy Studies Are Not Ethical or Feasible) or 21 CFR 601.90 (subpart H, Approval of Biological Products When Human Efficacy Studies Are Not Ethical or Feasible)?⁶ CDER's Office of Counter-Terrorism and Emergency Coordination should be consulted regarding any protocol or meeting regarding animal models.

2.3 End-of-Phase 2/Phase 3 Planning

Before initiation of phase 3 trials, both development to date and planned development should be reviewed to ensure that the development program will address the relevant regulatory requirements and issues. This review is typically done at an end-of-phase 2 (EOP2) meeting with the sponsor.

For drugs developed under the animal efficacy rule, human trials before approval are required to assess safety and PK/PD. The PK/PD results are used to extrapolate animal findings to humans to select an appropriate dose. These trials will be conducted in healthy human volunteers, and the sample size will depend upon factors such as whether

⁶ See the draft guidance for industry *Animal Models — Essential Elements to Address Efficacy Under the Animal Rule*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm078923.pdf>)

the product is a new molecular entity (versus an approved drug seeking a new indication where human efficacy trials are not feasible or ethical), or if the indication is for treatment versus prophylaxis. Confirmatory efficacy trials, for ethical reasons, may only be conducted after approval in a setting when the approval medical countermeasure is required.

A review of an EOP2 meeting package (or other meeting of similar intent) should take into account the following.

- Are the purity, potency, stability, and sterility (if applicable) of the drug adequate to support phase 2 and/or phase 3? This item is the primary responsibility of the review chemists/product quality reviewers and microbiology review staff with whom this should be discussed.⁷
- Are sufficient data available to plan a phase 3 program?
- Are the assessments of general safety and safety in specific subpopulations pertinent to the planned use? (See [section 7](#), Patient Populations, Special Populations.)
- Is the target population well defined and appropriate? (See [section 7.1](#), Trial Population.) Does the sponsor include an assessment of prior therapy in the target population for use in selecting patients, stratifying patients, or other purposes?
- Are the proposed endpoints well defined and appropriate? (See [section 8.2](#), Endpoints.)
- Is the choice of dosing regimen, including dose and dose interval, appropriate? (See [section 3](#), Dosing and Clinical Pharmacology.)
- Are there appropriate choices of control and for diagnostic tests, truth standards? (See [section 5](#), Controls, Truth Standards, and Compliance)⁸
- Has the use of concomitant therapies been sufficiently addressed? (See [section 5.5](#), Background Care and Standard of Care.)
- Will the total planned population exposure (e.g., subject numbers, duration of dosing, exposures at relevant dose levels) be adequate to assess safety? (See [section 7](#), Patient Populations, Special Populations.)⁹
- Will the planned development, together with ongoing and completed trials, provide adequate data regarding drug effects in a broad population including subpopulations

⁷ See the guidance for industry *INDs for Phase 2 and Phase 3 Studies Chemistry, Manufacturing, and Controls Information*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070567.pdf>)

⁸ See the ICH guidance for industry *E10 Choice of Control Group and Related Issues in Clinical Trials*. (<http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm129460.pdf>)

⁹ See 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*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073083.pdf>)

of interest (see [section 7.2](#), Special Populations, Demographic Subgroups), particularly including those defined by:

- Sex? (See [section 7.2.3](#), Women.)
- Race? (See [section 7.2.5](#), Racial Groups.)
- Age (various pediatric groups and geriatric groups)? (See [section 7.2.2](#), Pediatric Populations; and [section 7.2.4](#), Elderly Subjects.)
- Body weight/body surface area?
- Genetic difference in metabolism, risk factors? (See [section 7.2.6](#), Other Subpopulations of Interest: Genetic, Proteomic, and Concomitant Illness.)
- Disease severity?
- Patients with single or multiple concomitant illnesses? (See [section 7.2.6](#), Other Subpopulations of Interest: Genetic, Proteomic, and Concomitant Illness.)
- Immunodeficiency (where appropriate)?
- Pregnancy (if there is reason to expect use in pregnancy)? (See [section 7.2.3.3](#), Studying pregnant women.)
- Concomitant medications? (See [section 3.2.5](#), Drug-Drug Interactions.)
- Renal/hepatic/excretory organ impairment? (See [section 3.2.1](#), Effect of Intrinsic and Extrinsic Factors on PK and PD.)
- Are there exclusions for demographic factors (e.g., older than 75 years of age) or concomitant illness that are not needed but will decrease the breadth of the population studied? (See [section 7.2.4](#), Elderly Subjects; and [section 7.2.6](#), Other Subpopulations of Interest: Genetic, Proteomic, and Concomitant Illness.)
- Will there be adequate assessment of drug-drug interactions of likely importance? (See [section 3.2.5](#), Drug-Drug Interactions.)
- Will there be adequate assessment of a broad enough range of doses to provide informative labeling regarding dosing? (See [section 4](#), Assessing Dose-Response.)¹⁰
- Are the submitted clinical trial protocols adequately designed and powered (i.e., adequate sample size) to meet their stated objectives and provide the basis for evaluating the safety and effectiveness of the drug?
- Has the sponsor submitted an initial pediatric study plan (PSP), including a request for waiver or deferral if needed, and been informed of the requirement under the Food and Drug Administration Safety and Innovation Act of 2012 (FDASIA) to do so no later than 60 days after the EOP2 meeting?

¹⁰ See the ICH guidance for industry *E4 Dose-Response Information to Support Drug Registration*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073115.pdf>)

- Will there be adequate assessments of specific potential safety problems (e.g., QT prolongation, hepatotoxicity, immunogenicity) either suggested by nonclinical data or always needed?¹¹
- Have any diagnostic tests necessary for use of the drug been validated (e.g., to identify patients based on genomic characteristics) and is there a plan to ensure availability of the test if the drug is marketed?

2.4 Controlled Clinical Trial Protocol Review (including Special Protocol Assessments)

- Is there a statement describing the trial hypothesis and trial type (e.g., superiority, noninferiority)?
- Is the choice of control (placebo or active), and trial design (e.g., withdrawal, crossover) appropriate and well supported? (See [section 5.1](#), Types of Controls, and the ICH guidance for industry *E10 Choice of Control Group and Related Issues in Clinical Trials*.¹²)
- If there is a placebo-controlled trial, is there also an active control to assess assay sensitivity, if appropriate?
- If an active-controlled noninferiority trial, is the noninferiority margin described and supported? (See [section 5.1.4](#), Active-Treatment Control.)
- If this is a randomized fixed-dose dose-response trial, are the doses reasonable, based on phase 2 experience? Is the dosage spread sufficient? If a titration design is used, is there any plan for additional dose finding? (See [section 4](#), Assessing Dose-Response.)
- Is the choice of primary endpoints clear and acceptable? (See [section 8.2.1](#), Primary Endpoints.)
- If the endpoint is a surrogate, is the surrogate adequately supported? (See [section 8.2.3](#), Surrogate Endpoints.)
- Are the methods used to assess the endpoints well validated and appropriate? (See [section 8.2.1](#), Primary Endpoints.)

¹¹ See the ICH guidance for industry *E2A Clinical Safety Data Management: Definitions and Standards for Expedited Reporting* (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073087.pdf>), the ICH guidance for industry *E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs* (<http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm129357.pdf>), and the guidance for industry *Drug-Induced Liver Injury: Premarketing Clinical Evaluation*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM174090.pdf>)

¹² <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm129460.pdf>

- Are the secondary endpoints clearly and appropriately defined and is there a plan for their analysis (e.g., shared, alpha, step-down)? (See [section 8.2.2](#), Secondary Endpoints.)
- Are the endpoints assessed at the appropriate time points? (See [section 8.2.1](#), Primary Endpoints.)
- Are subject inclusion and exclusion criteria appropriate? (See [section 7.1](#), Trial Population.)
- Is there a screening period for eligibility for enrollment and is it well described?
- Are there enrichment features of the trial, including strategies to decrease heterogeneity (e.g., excluding patients whose disease or symptoms improve spontaneously), prognostic enrichment (e.g., choosing patients with a greater likelihood of having a disease-related endpoint), or predictive enrichment (e.g., choosing patients more likely to respond to treatment, based on a disease characteristic related to the drug's mechanism of action)? (See [section 6.1](#), Randomization, and the draft guidance for industry *Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products*.¹³)
- Are subjects stratified as part of randomization? If not, should they be? (See [section 6.1](#), Randomization.)
- Is appropriate demographic and important baseline information collected? (See [section 7.2](#), Special Populations, Demographic Subgroups.)
- Is the drug allocation scheme clear and appropriate? Are subjects randomized within trial sites? (See [section 6.1](#), Randomization.)
- Is the plan for blinding appropriate? (See [section 6.2](#), Blinding.)
- Is the trial duration appropriate? (See [section 4](#), Assessing Dose-Response.)
- Is a prospective statistical analysis plan (SAP) submitted as part of the protocol? (See [section 8.1](#), Planned Analyses.)
- Is the method of analysis of the primary endpoints clear and reasonable, with appropriate accounting for any interim analyses and for other types of multiplicity? Does the analytic plan describe how missing data will be dealt with? If an analysis of covariance is planned, is the method described? Is the plan for dealing with dropouts clear? (See [section 8.1](#), Planned Analyses.)
- Is the use of demographic and baseline information in the outcome analysis (e.g., for performing adjusted analyses) appropriately prespecified? (See [section 8.1.1](#), Adequacy of the Statistical Analysis Plan.)
- Is any planned interim analysis well described and appropriate? Is there adequate assurance that the integrity of the trial caused by interim unblinding is not compromised? (See [section 8.1.3](#), Interim Analysis Plans.)

13

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM332181.pdf>

- Is the planned trial size appropriate? Does the trial appear to have adequate power to assess the primary endpoint(s)? (See [section 7.3.2](#), Sample Size in Phase 2 and Phase 3 Clinical Trials.)
- If this is the only planned well-controlled trial, what level of evidence needs to be achieved for the trial to be considered a success (e.g., is the nominal p-value appropriate) and what other supportive/confirmatory evidence will be available to support approval?¹⁴ (See [section 4](#), Assessing Dose-Response)
- Is there a data monitoring committee (DMC)? If not, should there be a DMC?¹⁵ (See [section 9.5](#), Trial Monitoring and Auditing)
- Is subject monitoring for adverse drug events (AEs) adequate, including assessment tools (e.g., CRFs), frequency of assessment, and duration of follow-up? (See [section 10.1](#), Safety Monitoring.)
- Are the planned handling and analysis of safety data appropriate? (See [section 8.1.1](#), Adequacy of the Statistical Analysis Plan.)
- Is information being collected for which the use is not specified in the protocol? (See [section 8.1.1](#), Adequacy of the Statistical Analysis Plan.)
- Are uses of concomitant therapies and standard care and any requirements for prior therapies appropriately specified and controlled by protocol? (See [section 5.5](#), Background Care and Standard of Care.)
- Are the procedures for encouraging and assessing drug compliance appropriate? (See [section 5.4](#), Assessing Treatment Compliance.)
- Are data quality assurance approaches (e.g., investigator qualifications and training, monitoring, and auditing) described and appropriate? (See [section 9.4](#), Investigator Qualifications and Responsibilities; and [section 9.5](#), Trial Monitoring and Auditing.)
- Does the informed consent document (ICD), if submitted or requested by the FDA, comply with the informed consent elements required by 21 CFR 50.25? Is the information in the ICD inaccurate or misleading? Is there a need for a consult to the Human Subject Protection Branch in the Office of Scientific Investigations (OSI) and/or an ethics consult? (See [section 9.2](#), Informed Consent, and MAPP 6030.2 *INDs: Review of Informed Consent Documents*.¹⁶)

¹⁴ See the guidance for industry *Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products*.
(<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM078749.pdf>)

¹⁵ See the guidance for clinical trial sponsors *Establishment and Operation of Clinical Trial Data Monitoring Committees*.
(<http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm127073.pdf>)

¹⁶
<http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/MannualofPoliciesProcedures/ucm082024.pdf>

- Is the investigator's brochure, or other information for investigators, adequate? Has it been updated (e.g., with data from earlier phases)? (See [section 9.3](#), Investigator's Brochure.)
- Has a special protocol assessment (SPA) been requested?¹⁷ Should the sponsor be encouraged to submit the protocol for SPA review?

2.5 Fast Track or Breakthrough Designation

- Does the investigational drug address a serious aspect of a serious or life-threatening disease or condition? (See [section 11.1](#), Serious or Life-Threatening Condition.)
- Does the drug show potential to treat this serious aspect of the condition? (See [section 11.3](#), Demonstrating the Potential to Address Unmet Medical Need.)
- Is the drug development program designed to determine whether the drug will affect a serious aspect of the condition and is the degree of specificity appropriate to the development stage?
- Is there any accepted or approved treatment for the same serious or life-threatening aspect of the condition being studied? If so, will the drug development program assess the ability of the drug to address an unmet medical need (e.g., by studying the new drug in addition to the standard or as an alternative therapy in those who do not respond to or cannot tolerate the alternative)? (See [section 11.3](#), Demonstrating the Potential to Address Unmet Medical Need.)
- Is the drug designated as a qualified infectious disease product under section 505E(d) of the Federal Food, Drug, and Cosmetic Act (FD&C Act) and thus designated as fast track under FDASIA?¹⁸
- Is there preliminary clinical evidence that demonstrates the potential for substantial improvement over available therapies for the treatment of a serious aspect of a serious disease that would warrant consideration for breakthrough designation? If so, this designation requires discussion and decision-making from high-level CDER managers.

¹⁷ See the guidance for industry *Special Protocol Assessment*. (<http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm080571.pdf>)

¹⁸ Title VIII of FDASIA, Generating Antibiotic Incentives Now (GAIN), provides incentives for the development of antibacterial and antifungal drugs intended to treat serious and life-threatening infections. Under GAIN, a sponsor may be granted a *qualified infectious disease product (QIDP)* designation for a drug that meets the criteria outlined in the statute. A drug that receives a QIDP designation is eligible for fast track designation and, upon submission of an NDA or supplement for that designated use, will receive a priority review. Upon approval of an application for a QIDP, a 5-year extension will be added to any exclusivity granted with that approval.

2.6 IND Safety Reports (21 CFR 312.32(c))

Under 21 CFR 312.32(c), sponsors “must notify FDA and all participating investigators in a written IND safety report of . . . any adverse experience associated with the use of the drug that is both serious and unexpected.” The IND safety reporting rule has been revised to clarify when the adverse event is likely enough to have been caused by the drug for the experience to meet the definition of a *suspected adverse reaction*. The following list should be used to help assess the adequacy and significance of the adverse drug report (ADR). This list can also be used for reviewing safety information in annual reports.

- Is the information in the report adequate? (See [section 10.2](#), Reporting Requirements for Sponsors, and [section 10.3](#), IND Safety Reports — Written Reports.)
- Does the report contain information about:
 - Trial drug, if appropriate? (Note: For blinded trials, the FDA would need to request unblinding of the subject if knowledge of the treatment group is essential.)
 - Concomitant therapies and illnesses that might be relevant?
 - Subject’s medical condition at time of reaction?
 - Nature, severity, and duration of reaction (including response to dechallenge and rechallenge, if applicable)?
 - Time of onset in relation to dosing?
 - Outcome of reaction?
- Does it meet the definition of serious or unexpected? (See [section 10.3.5](#), Definition: Serious; and [section 10.3.4](#), Definition: Unexpected.)
- Have similar reactions been reported for other subjects to the same or related therapies?
 - From trials in other countries?
 - From postmarketing studies or trials in other indications (potentially reviewed in other divisions)?
 - From trials of other formulations, delivery systems, or routes of administration?
 - From related drugs or metabolites?
 - From nonclinical studies?
- Is there a causality assessment?
 - From the investigator?
 - From the sponsor?
- Should the trial be placed on clinical hold?
 - Until it is modified to provide an acceptable risk profile?
 - While awaiting more data?

- Should the consent form be modified? (See [section 9.2](#), Informed Consent.)
- Does the investigator's brochure need to be modified? (See [section 9.3](#), Investigator's Brochure.)
 - To list new types of serious unexpected adverse events?
 - To list new frequencies of serious unexpected adverse events?
- Does the protocol need to be modified?
 - Entry criteria to exclude subjects at risk?
 - Dose or regimen adjustments?
 - Concomitant medications to exclude?
 - New toxicity monitoring?
 - New stopping rules?
- Have investigators (as required by 21 CFR 312.32(c)) been notified?
- Should other trials of the drug, or related drugs, be placed on clinical hold?
 - Until they are modified or collect more frequent information, alter dose, or take other action to minimize risk in order to provide an acceptable risk profile?
 - While awaiting more data?

3. DOSING AND CLINICAL PHARMACOLOGY

3.1 Phase 1 Tolerability Trials

3.1.1 Choosing a Starting Dose for Phase 1

Nonclinical evaluation of a new drug should provide information that guides the choice of a safe and appropriate starting dose for the initial trial of the new drug in humans.¹⁹ Information frequently used for this purpose includes:

- The *no observed effect* dose or plasma drug level.
- The *no observed adverse effect* (NOAEL) dose or plasma drug level.
- The ED₅₀ (dose inducing 50 percent of a specified response in an animal population), using the most appropriate animal species.

¹⁹ See the guidance for industry *Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers*.
<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm078932.pdf>

- The EC₅₀ (plasma concentration inducing 50 percent of the group response), in an in vitro system or in a whole animal.
- PK data obtained in animals, including information about metabolism and the enzymes responsible.
- Immunogenicity data obtained in animals for biologics.
- PK data (or immunogenicity data) obtained in trials conducted outside the United States.
- The severely toxic dose to 10 percent of rodents (STD₁₀) in oncology studies.
- The highest nonseverely toxic dose (HNSTD) in nonrodents in oncology studies.
- Data obtained from exploratory IND trials, used to evaluate potential mechanism of action, preliminary PK data, or biodistribution characteristics²⁰

When selecting an appropriate starting dose, it is critical to consider how the data would be extrapolated across species and how well the animal data are likely to predict responses in humans. In addition, in vitro studies of cellular responses and of relative receptor binding affinity in human versus animal cells may be informative in selecting the appropriate starting dose. Comparisons of an investigational drug's potency and activity with that of other well-characterized drugs of the same class may also be useful in determining the starting dose.

Usually, the initial human dose is a small fraction of the NOAEL in the most sensitive animal species, often about 1/10 to 1/100, or a dose projected to provide 1/10 or 1/100 of the exposure at the NOAEL in the most sensitive animal species. This starting dose could be higher in some cases, depending on the familiarity of the drug class and the nature of the adverse effect on animals. For anticancer drugs investigated in patients with metastatic or locally advanced solid tumors or serious and life-threatening hematologic malignancies, the usual approach to setting a clinical starting dose is 1/10th the STD₁₀ from a rodent study or 1/6th the HNSTD from a nonrodent study, using the most appropriate animal species.²¹

Factors that suggest use of the lower starting dose include:

- Steepness of the toxicity dose-response

²⁰ See the guidance for industry, investigators, and reviewers *Exploratory IND Studies*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm078933.pdf>)

²¹ See the ICH guidance for industry *S9 Nonclinical Evaluation for Anticancer Pharmaceuticals*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm085389.pdf>)

- Severity of the toxicity
- Monitorability of the toxicity in the human subject
- Reversibility of the effect
- Wide variability between species in doses or exposures eliciting toxicity
- Novel therapeutic targets
- Immune-stimulatory compounds (e.g., drugs that stimulate cytokine release)
- Animal models of limited utility
- New excipients or adjuvants
- Poorly understood PK and PD

Reviewers should determine whether sponsors have given consideration to all relevant nonclinical data and foreign human data, if available, before selecting a dose range for an initial phase 1 clinical trial in the United States.

3.1.2 Dose-Escalation and Maximum Dose and Duration in Phase 1

For new investigational drugs given to humans for the first time, the time course of any potential AE is unknown. When toxicity is delayed, repeated administration of the drug may place the subject at added risk and multiple dosing of the new drug could lead to accumulated toxicity before the toxicity profile of a single dose can be defined.

Therefore, the most conservative and commonly used approach is to initiate phase 1 clinical trials with single dose exposures, with adequate follow-up to evaluate PK and potential AEs, before proceeding to multiple dose trials. Once the initial PK and safety profile of the investigational drug has been better elucidated, multiple dose trials can be undertaken. Repeat-dose initial phase 1 trials may be more appropriate for certain drug types intended for use in particular patient populations (e.g., cancer or HIV patients).

Most phase 1 clinical trials are characterized by progressive dose escalation with sponsors collecting safety information on each dose experience. The safety information obtained at each level is then used to allow dose escalation in subsequent cohorts of subjects, where again, one dose level per cohort is administered. Often all members of a cohort are dosed simultaneously, but for a first-in-human trial, particularly for drugs with a novel mechanism of action, a more cautious approach is to dose sequentially (i.e., one member is dosed and a subsequent member of a cohort is not dosed until safety has been demonstrated in the previous member). Such an approach may be appropriate depending on the specifics of the drug and the available nonclinical data.

Occasionally, a dose may be escalated within the cohort (i.e., intrasubject escalation: subjects in a low-dose group may later receive a higher dose), particularly when the initial dose is deliberately chosen below the expected active dose and unlikely to give a response. If this approach is taken, there is the possibility that prior receipt of the lower dose may affect the response to higher doses (e.g., a subject who receives progressively higher doses of a drug over time may better acclimate to these higher doses) alternatively, there could be cumulative toxicity.

Reviewers should encourage sponsors to base the rate of dose escalation between dosing cohorts upon nonclinical findings and expectations as well as the half-life of the drug.

Many of the same factors that lead to the selection of a low starting dose should also lead to cautious dose escalation. These factors can include a small therapeutic window (e.g., low ratio of toxic dose to therapeutic dose) in nonclinical data, poor animal models, and concerns about toxicity.

Usually, dose increments in phase 1 dose-escalation trials are on a linear or logarithmic scale. Common practices include increasing drug dose by fixed intervals either on a linear scale (e.g., 25 mg, 50 mg, 75 mg, 100 mg) or on a logarithmic scale (e.g., 25 mg, 50 mg, 100 mg, 200 mg). The latter approach is sometimes modified to give a so-called modified Fibonacci sequence (e.g., 50 mg, 75 mg, 125 mg, 200 mg). In certain disease entities, more aggressive dose-escalation schemes may be appropriate. In some designs, the rate of escalation is slowed after the first pharmacological activity is observed (especially if adverse).

The choice of the highest dose to be used in a phase 1 trial is a complex decision. In some phase 1 clinical trials, the highest dose to be used will be predetermined and will be a dose anticipated to be well tolerated and sufficiently high to yield desired blood levels and effects based on animal data or related drugs. This design has limitations, however, as it offers no information about the consequences of higher exposures (e.g., those that might arise clinically in patients with poor excretory function, impaired renal or hepatic metabolic polymorphisms, drug-drug interactions, or from overdose, or those that might be needed to explore the full dose-response relationship). Therefore, doses well above those expected to be needed should be studied (when safety allows) in phase 1 trials, to identify toxicities and to identify the highest dose that is reasonably well tolerated.

With this design type, dose escalation is terminated when some or all subjects no longer tolerate the investigational drug or develop a laboratory finding that is potentially dangerous (e.g., QT or QRS prolongation, first degree heart block, transaminase elevations). This strategy should be employed with caution and close monitoring, and may be inappropriate, depending upon the nature and severity of the toxicities that are expected with higher doses.

Reviewers should ensure that safety data will be thoroughly monitored during phase 1 trials. Trials should ensure collection of AE data, appropriate laboratory testing (e.g., blood tests, electrocardiogram (ECG) with QT/QTc intervals), PD data, blood levels of the parent and metabolites to allow PK measurements, and immunogenicity data, when appropriate. Data collection should occur at appropriate time points.

Clinical trials should be designed to assess toxicities, including (but not limited to) those that may be anticipated from nonclinical studies. Toxicity grading scales may be useful in some phase 1 protocols (some well-known grading scales, such as the Common Toxicity Criteria developed for cancer therapies, may not be appropriate for other disease states). There should be dose-modification and dose-escalation stopping rules in place to deal with serious unexpected toxicity (see [section 3.3](#), Choice of Dosing Interval). If nonclinical observations suggest that a syndrome may occur, it should be prospectively

defined with a relevant grading system. Any serious adverse event should be considered potentially drug-induced.

3.1.3 Toxicity-Induced Modifications in Enrollment or Dosing (Safety Rules)

Planned modifications to the dosing schedule that will be implemented during a trial based on observed toxicities (expected or unexpected) are often referred to as safety stopping rules. They are most often used for phase 1 trials.

To generate such rules, sponsors should develop both: (1) a list of acceptable toxicities (i.e., toxicities that, if observed within specified parameters, will not result in changes to subject enrollment and dosing); and (2) a procedure for dealing with the occurrence of other toxicities (i.e., not on the list of acceptable toxicities). These procedures should be specified in writing. Most procedures specify one of the following: (1) a halt to subject dosing or trial enrollment until toxicity data can be further studied; (2) evaluation of additional subjects in a particular dose cohort or in each dose cohort without exposing subjects to a higher dose to make the trial more sensitive to characterizing AE data; (3) implementation of smaller dose increases between dose cohorts; and (4) exclusion of certain subjects thought to be more at-risk for a particular AE. Sponsors should be encouraged to devise and implement stopping rules for phase 1 trials.

3.2 Pharmacokinetic and Pharmacodynamic Trials

At any stage of IND development that is being considered, the reviewer should assess the state of characterization of the PK of the drug and of the relationship between blood levels and response (e.g., PK/PD relationship, the totality of PK information, and plans for further characterization of PK and PD).

Understanding of PK (C_{max} (peak concentration), T_{max} (time to C_{max}), C_{min} (trough concentration), $T_{1/2}$ (elimination half-life), accumulation ratio (the relationship of dose to concentration), and the metabolism of the parent drug and active metabolites is critical to designing trials, choosing doses and dose intervals, and anticipating which subjects might accumulate the drug and what concomitant drugs might lead to interactions. Having these data early can inform a wide range of critical decisions and should be encouraged.

When there is a readily measurable, rapid pharmacological response to a drug that is thought to be related to a clinical effect, it is generally useful to evaluate the relationship of that response to dose and plasma concentration. Such PK/PD information can provide useful guidance regarding dosing, dosing regimens, target population, concomitant medications, and trial characteristics. It can also give insight into how high doses in phase 1 trials should be escalated. Although these data can suggest clinical responses, all critical findings (i.e., dose-response, duration of action) need to be confirmed using clinical data, because the relationship between PD measures and clinical outcomes is not completely predictable.

3.2.1 Effect of Intrinsic and Extrinsic Factors on PK and PD

The PK characteristics of a drug (i.e., the rate and pattern of its absorption, distribution, metabolism, or excretion, including deactivation) can differ in patient populations. In addition, patient populations can also vary in their PD responses to a drug, and the consequences of these variations can be neutral, beneficial, or adverse. Differing PK or PD characteristics can result in dissimilar profiles of drug efficacy or safety. Therefore, it is important that differences in PK and PD in various populations be investigated during drug development.²²

Factors that could result in differences in PK or PD responses have been categorized in the ICH guidance for industry *E5 Ethnic Factors in the Acceptability of Foreign Clinical Data* as either intrinsic ethnic factors or extrinsic ethnic factors.²³ Generally, the intrinsic factors that should be evaluated for effect on PK or PD during drug development include age (e.g., elderly, adult, pediatric groups), sex, and race. PK differences can be easily studied because blood levels are readily measured. PD differences have historically been studied less, except where mechanisms of difference are well understood and anticipated (e.g., differences between races in response to angiotensin converting enzyme (ACE) inhibitors, beta-blockers, and angiotensin II antagonists), and they are usually assessed principally by the demographic analyses of safety and effectiveness data called for in regulations 21 CFR 314.50(c)(2)(viii).

Genetic variations among patients are increasingly recognized as important determinants of PK and/or PD variability. There is a growing interest in studying the effect of genetic factors and well-recognized genetic influences on PK and PD. With respect to PK, there is interest in the activity of a number of metabolic enzymes (e.g., CYP2D6, CYP2C19, CYP2C9, UGT1A1, VKORC1) which can markedly affect the blood levels of a drug in patients who have low enzyme activity or patients in whom an interacting drug can decrease enzyme activity. Where an active metabolite is the source of a drug's effect, poor metabolism can render the drug less effective (e.g., low CYP2D6 for codeine, low CYP2C19 for clopidogrel).

There is also increasing interest in genetic PD differences, notably differences in receptors on cancer cells that predict drug response or predict outcome, and genetic markers that predict certain AEs. Approaches to some of these genetic factors, principally related to PK differences, are discussed in more detail below and in [section 7.2](#), Special Populations, Demographic Subgroups. Extrinsic factors of particular interest include concomitant medications, diet, and the use of alcohol (see [section 3.2.5](#), Drug-Drug Interactions).

²² See the draft guidance for industry *Clinical Pharmacogenomics: Premarketing Evaluation in Early Phase Clinical Studies*.
<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM337169.pdf>

²³
<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073117.pdf>

3.2.2 Classic PK Clinical Trials (Frequent Sampling)

Classic PK clinical trials are conducted by taking multiple-timed samples following a particular dose, either given as a single dose or following a multidose period. Such trials can be conducted in healthy volunteers (for most drugs except when risks preclude this), in patients, or in special groups (e.g., elderly, renally impaired, racial groups). The results in normal volunteers may not fully reflect PK in the target patient population because volunteers for such trials tend to be healthier on average than the patient population that will receive the drug.

3.2.3 Population PK Clinical Trials

Exposure-response provides confirmatory and/or supportive evidence of effectiveness and allows deriving rational dosing recommendations. Sponsors should be encouraged to collect sparse, but informative, PK data from phase 2 and 3 trials. Population PK clinical trials are trials in which relatively few samples are taken per subject but are collected from a larger sample of subjects, usually within a larger randomized clinical trial being otherwise conducted. Population PK trials can provide estimates of variability in the serum concentration for a wide range of subjects and examine the PK effects and interactions of a wide variety of intrinsic and extrinsic factors. Analysis can be performed with linear regression or mixed effects modeling.

In many cases, evidence of extreme deviation in the PK data occurring within a particular subgroup would require a more detailed study of that subgroup. Adequate number of patients with renal or hepatic dysfunction is needed to reliably use population PK.

3.2.4 Bioavailability and Bioequivalence Trials

Measures of the degree to which an active drug substance gets into the circulation (i.e., bioavailability (BA) for systemically active drugs) are important for nonintravenous preparations. When changes are made in the drug formulation that may affect bioavailability or when substantive changes are made in manufacturing complex substances, trials measuring the relative bioavailability (or, in the case of intravenous preparations, the PK) should be performed and compared to data from the previous formulation.²⁴ Bioequivalence (BE) documentation can be useful during the IND or NDA period to establish links between: (1) early and late clinical trial formulations; (2) formulations used in clinical trial and stability studies, if different; (3) clinical trial formulations and to-be-marketed drug product; and (4) other comparisons, as appropriate. It is also important to know whether ingestion of a drug with food affects its bioavailability.

²⁴ See the guidance for industry *Bioavailability and Bioequivalence Studies for Orally Administered Drug Products — General Considerations* (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM070124.pdf>) and the guidance for industry *Food-Effect Bioavailability and Fed Bioequivalence Studies* (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070241.pdf>).

For trials of a controlled release dosage form of an already approved immediate release counterpart, relevant comparisons include a number of PK parameters, including C_{min} , area under the curve (AUC), and C_{max} . Information from immediate release dosage forms relating efficacy/PD to drug exposure levels (such as C_{min}) can facilitate the evaluation of a controlled release dosage form. Ordinarily, a controlled release drug's C_{min} should be the same as or greater than that demonstrated for the immediate release dosage form. If that is not the case, particular attention will need to be paid to the duration of drug effect and during the entire dosing interval.

3.2.5 Drug-Drug Interactions

Drug-drug interactions can cause important AEs or can interfere with a drug's effectiveness. Therefore, potentially important drug-drug interactions should be evaluated in nonclinical studies or in specifically designed clinical trials and in observation of overall clinical data.

When a drug is marketed, its use with a wide range of other drugs is often inevitable. Drug-drug interactions can alter PK in various ways, affecting both the C_{max} , AUC, T_{max} , $T_{1/2}$ of the parent molecule, the predominant route of clearance, formation of active metabolites, and formation of toxic metabolites, and can lead to increased or decreased exposure to the parent or to its metabolites. PK interactions usually are of two broad types. The first type is when the investigational drug is the substrate on which another drug acts; the second is when the investigational drug alters the metabolism or transport of another drug. The role of a drug as substrate and inhibitor/inducer can be critically important.

Less recognized but potentially important, PD interactions of co-administered drugs may affect important clinical outcomes (e.g., additive hypotensive effects of organic nitrates and sildenafil or ibuprofen's interference with aspirin's platelet inhibition).

It is critical to understand the metabolic pathway and excretory mechanisms of the parent drug and its active metabolites to anticipate drug-drug interactions. Metabolism occurs primarily by the cytochrome P450 family (CYP) of enzymes, but may also occur by non-P450 enzyme systems, such as glucuronosyl transferases (UGTs). Recently, membrane transporters were found to have important effects on differences in exposure. For example, a polymorphism in organic anion transporting polypeptide (OATP) 1B1 was found to play a role in the exposure of several statin drugs, including pravastatin, simvastatin acid, and rosuvastatin.

In vitro techniques are available to identify the specific cytochrome p450 (CYP450) isozymes involved in the metabolic oxidation of a drug and indicate which drugs might inhibit or induce these processes. At present, although in vitro methods can indicate the potential for some CYP-related drug-drug interactions, they cannot yet define the magnitude of the interaction, nor do they predict other important kinds of interactions, such as competition for renal excretory sites that is mediated by transporters (e.g., effect of probenecid on penicillin excretion). In vitro techniques are becoming available to help identify drugs that are substrates or inhibitors for transporters. Use of in vitro tools to

evaluate a drug's potential to be a substrate, inhibitor, or inducer of metabolizing enzymes or transporters, followed by in vivo interaction studies to assess potential interactions, has become an integral part of drug development and regulatory review. In vitro criteria are being developed so that in vitro data may be sufficient to ensure lack of interaction.

Many drugs can be substrates, inhibitors, or inducers for both metabolizing enzymes and transporters that have overlapping selectivity. Interplay between enzymes and transporters can make the prediction of in vivo interactions based on in vitro assessment challenging. Where there are multiple active metabolites, or several metabolic enzymes and transporters involved in metabolism and excretion, modeling and simulation of drug interactions can be helpful in the design of clinical trials to inform drug interaction potential. The evaluation of potential drug-drug interactions requires the development of sensitive and specific assays for a drug and its important metabolites. Major metabolizing enzymes and transporters should be evaluated during drug development.

Useful designs for human trials, including the use of standard inducers, inhibitors, and substrates, are described in the draft guidance for industry *Drug Interaction Studies — Study Design, Data Analysis, Implications for Dosing, and Labeling Recommendations*.²⁵ Such trials can assess the effects of the investigational drug on concomitant drugs (both inhibiting or inducing metabolic enzymes or transporters) and the effects of concomitant drugs on the investigational drug, again both by inhibiting or inducing its metabolic enzymes. Drug labels should include appropriate instructions for dosing based on results of these drug-drug interaction trials.

3.3 Choice of Dosing Interval

The frequency of dosing can be an important factor in determining the relationship of benefit to risk. Dosing more often than twice per day appears to lead to poorer patient compliance and is unattractive to patients, so sponsors usually try to avoid more frequent dosing even with relatively short half-life drugs. However, less frequent dosing than is indicated by the pharmacokinetics of the drug can lead to diminished drug efficacy near the end of the dosing interval. In some cases, the dose is increased to overcome this loss of effect. When the dosing interval is longer than the half-life of the drug, particular attention should be paid to the sponsor's basis for choosing this specific dosing interval, and to: (1) whether the desired effect persists to the end of the dosing interval; and (2) whether, to achieve a reasonable C_{\min} , dosing has been increased to the point where it causes undesirable effects at C_{\max} .

Where the dosing interval appears long compared to the half-life of the drug, sponsors should be encouraged to compare a longer dosing-interval regimen with the same total daily dose given as more divided doses (to examine the persistence of desired drug effect throughout the dosing interval (e.g., assessing efficacy at likely C_{\min}) and the incidence of

25

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM292362.pdf>

AEs (particularly those related to the pharmacologic effect of the drug) associated with the increased C_{\max} compared with the incidence observed with less frequent dosing). These issues are important for drugs with relatively short-term effects (both favorable and unfavorable).

In the past, for example, short half-life dihydropyridines proved to be poorly tolerated when attempts to use twice-daily dosing with maintenance of good trough (C_{\min}) blood pressure (BP) effect led to use of excessive doses. These drugs were instead developed as controlled-release preparations. For some drugs, particularly biologics and some antibiotics (e.g., gentamicin), pharmacologic effects persist well beyond the duration of adequate blood drug levels. Other drugs, such as anticancer agents that are intentionally cytotoxic, may have a long lasting effect that does not need to be maintained by frequent dosing.

In some cases, these concerns may be adequately addressed by short-term PD trials, but in others, clinical trials may be needed to evaluate the safety and tolerability of alternative regimens.

4. ASSESSING DOSE-RESPONSE

Dose-response data are critical to good drug development and close attention should be paid both to the adequacy of plans to generate such data and its interpretation and use. The sponsor should obtain dose-response data from well-controlled, rigorous trials. In addition to relating dose to effect size for both desired and undesired effects of the drug, trials should identify titration steps (if titration is needed) and optimal dose interval. It is particularly important to identify a dose beyond which up-titration should not be attempted because of the low likelihood for further benefit or potential for unacceptable toxicity.

Dose-response and concentration-response information for both effectiveness and adverse effects of drugs are critical components of the evaluation, safety, and effectiveness of drugs and the findings are important components of drug labeling.²⁶ Specific dose adjustments for subject size, sex, age, concomitant illness, and concomitant therapy should also be defined (21 CFR 214.50) and findings should appear in prescribing information. In assessing dose-response, it is important to allow adequate duration at a given drug dose to allow full effect of the specific drug to be manifested, a potential problem for titration designs. Because the subject's response during the early dosing period may not be the same as in the subsequent maintenance dosing period, it is desirable to study dose-response during maintenance treatment. In rare cases, responses to a drug have been related to cumulative rather than daily dose, to duration of exposure, and to the time of day (e.g., morning versus evening dosing).

²⁶ See ICH E4

(<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073115.pdf>) and the guidance for industry *Exposure-Response Relationships — Study Design, Data Analysis, and Regulatory Applications* (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072109.pdf>).

Reviewers should recognize that many dose-response trials can have two purposes. First, demonstration of dose-response is evidence of effectiveness, and the dose-response trial is one kind of control described in 21 CFR 314.126 as suitable for an adequate and well-controlled trial. Second, dose-response trials provide critical information about how to use a drug to optimize the beneficial and minimize the adverse effects. Some trial designs serve both purposes equally but others may not. For example, a randomized, fixed-dose, dose-response trial without placebo that shows positive slope demonstrates effectiveness well and may identify optimal dose and maximal effective dose, but it usually does not reveal effect size because there is no placebo. A randomized, fixed-dose, dose-response trial with a placebo control is the best design, not only to study multiple doses, but to evaluate the relation of dose to toxicity and effectiveness.

Reviewers should encourage sponsors to conduct dose-ranging or concentration-response trials early in drug development to reduce the possibility of: (1) failed phase 3 clinical trials; or (2) too much data that represents subject experience with ineffective doses that does not represent useful safety exposure. Use of excessive doses may yield misleading side effect profiles and cause excessive dropout.

In early drug development, rapidly available quantitative information from a PD endpoint considered likely to correspond to clinical effect is often useful in choosing a range of possible active and tolerable doses. Subsequently, in controlled phase 2 and phase 3 trials with clinical endpoints, sponsors can study a more limited set of doses. It is highly desirable that sponsors continue examining a range of doses in phase 3, except where sample size would make this unrealistic (e.g., in large outcome trials), because phase 2 trials give only limited information about less common adverse effects and may not fully elucidate how dose affects efficacy. For example, a drug for heart failure can use surrogate endpoints such as cardiac output or wedge pressure early in drug development for dose-response trials. Later in development, a narrower range of doses can be studied using endpoints such as exercise tolerance, mortality, or irreversible morbidity.

A common error in drug development is to use a single dose or dose regimen in phase 3 trials. This design can either lead to disappointing efficacy results, or evaluation of an effective dose with undesirable safety issues, where another dose might have shown more acceptable risk-benefit characteristics.

To support the choice of dosing range for phase 2 and phase 3 clinical trials, sponsors should be encouraged to attempt to determine the shape of the population average dose-response curve for both desirable and undesirable effects in advance of conducting the trials. Selection of drug doses should be based upon such information, in conjunction with an evaluation of the relative importance of both the desirable and undesirable effects of the drug. For example, a high starting dose of a drug, one that is near the plateau of the effectiveness dose-response curve, might be indicated in the following situations: (1) for a drug with a large demonstrated separation between its useful and undesirable dose ranges; or (2) when a life-threatening disease requires rapid intervention at the fully effective dose.

A lower starting dose can be used when the drug has significant first dose effects (e.g., alpha blockers for hypertension), serious dose-related adverse effects, and adverse effects that decrease with continued use or that are decreased by titration. There is value in examining not only the mean (population average) dose-response, but also individual responses, looking for subsets of patients also responding to lower doses. The need for lower starting doses could be suggested by intersubject variability in PD response to a given concentration of drug in the blood or by intersubject PK differences that could arise from nonlinear kinetics, metabolic polymorphism, or drug-drug interactions. In certain cases, adequate knowledge of exposure-response can allow assessment of the appropriateness of proposed doses and/or dosing regimens. In some cases, such analysis leads to approval of doses not directly studied in clinical trials to improve benefit-to-risk ratio.

The choice of design for dose-response trials depends on many factors, including the development phase, the therapeutic indication, the ability to rapidly ascertain pharmacological or clinical effects, the time to equilibration and manifestation of drug effects, and the severity of the disease in the population of interest. Many potential trial designs can be used to assess dose-response; those designs used most frequently are discussed below. Sponsors should be encouraged to support the choice of dose and/or dosing regimen using clinical trial simulations. The goal of these simulations would be to explore competing doses, based on earlier trials and information from other drugs and the likelihood of identifying a dose or exposure-response. In addition to the specific designs described below, there are many approaches to derive concentration-response relationships from available data when there is a combination of frequent measurements of drug plasma concentration and a PD measurement.²⁷

For drugs developed under the animal efficacy rule (21 CFR part 314, subpart I, Approval of New Drugs When Human Efficacy Studies Are Not Ethical or Feasible, or 21 CFR part 601, subpart H, Approval of Biological Products When Human Efficacy Studies Are Not Ethical or Feasible), the data or information on the pharmacokinetics and pharmacodynamics of the drug or other relevant data or information in animals or humans must be sufficiently well understood to allow selection of an effective dose in humans. Therefore, it is reasonable to expect the effectiveness of the drug in animals to be a reliable indicator of its effectiveness in humans.

4.1 Fixed-Dose Clinical Trials

A widely used trial design generally preferred for its ability to provide the clearest dose-response data is the randomized, parallel group, fixed-dose, dose-response trial.²⁸ Subjects are randomized to receive one fixed dose throughout the trial with several doses examined (e.g., placebo, 50 mg, 100 mg, 500 mg). Dosing may be initiated at the final

²⁷ See ICH E4.

(<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073115.pdf>)

²⁸ Ibid.

fixed dose or can be reached after dose-titration up to the target, but the critical comparisons are assessed at the target dose for each group. Fixed doses should be maintained for a sufficient time to allow for adequate dose-response comparisons. To measure the absolute size of the drug effect, a placebo or other control with negligible effect on the endpoint of interest is usually needed (see [section 5.1](#), Types of Controls), although evidence of effectiveness can be based on a positive dose-response slope. In this design, there is no potential confounding of dose and duration of exposure.

Inclusion of a placebo group in a dose-response trial can provide critical information in interpreting trials in which all doses tested resulted in indistinguishable outcomes, usually because the doses are all above the minimum effective dose (on the *plateau*) or because the doses are too close together. Without the presence of a placebo group, it may be impossible to tell whether any of the doses were effective at all in the trial. In such a case, the trial provides no evidence of effectiveness and no useful dose-response information. With a placebo group, the trial can provide evidence of effectiveness and, if efficacy is seen, may be able to identify where on the dose-response curve of the examined doses fall.

A common failing in dose-response trials is insufficient spread between the doses. Ideally, the range should be at least an order of magnitude unless earlier trials give a basis for narrowing the range.

4.2 Titration Clinical Trials

In titration trials, subjects receive several doses so that sample size is far smaller than the fixed-dose parallel-dose-response trial, because a trial of many doses needs only a single treated group and a placebo group. This design contrasts with an n-dose trial, which will need n+1 groups. In addition, individual dose-response information is available in such a trial, which is not the case in the randomized, parallel fixed-dose, dose-response design. Individual dose-response is clearest for the forced titration design but also exists in optional titration designs.

In a *forced titration trial*, which generally is placebo-controlled, subjects move through a series of escalating drug doses so that every subject is exposed to every dose. One significant disadvantage of titration designs is that a response to an increased dose cannot be distinguished from a response to increased duration of drug therapy, a cumulative drug dosage effect, or a spontaneous change in the disease state. This design may also generate inadequate data on the relationship of drug dose to adverse effects because these effects are often time-dependent. Despite these deficiencies, the forced titration trial can provide a reasonable first approximation of both the population average dose-response curve and the distribution of individual dose-response curves. Favorable data are more likely to be generated if drug effects develop rapidly in subjects, the cumulative dose effect is minimal, and the number of drug withdrawals in the trial is not excessive.

In an *optional titration trial*, which is also generally placebo-controlled, subjects move to the next drug dose only if they fail to meet a specified clinical response. Since only poor responders receive the higher doses, crude analyses of response by dose received in these

trials can be misleading, often showing an *umbrella-shaped* dose-response curve with smaller effects in subjects titrated to the highest doses. More sophisticated approaches to determining population dose-response from such trials, generally involving individual patient modeling, are available and may at least partially overcome these defects.

Titration designs with only two groups (e.g., a titrated group and placebo) require fewer subjects than a randomized, parallel dose-response trial of the same number of doses, and, with a concurrent placebo group included in the trial design, can provide evidence of effectiveness and an early estimate of dose-response. If the effect of the drug and of dose changes is fairly rapid (e.g., as it is for many antihypertensives), such a trial can explore many doses effectively. This kind of trial can be particularly valuable early in the drug development process and can be used to narrow the dose range for a later, more definitive randomized parallel group dose-response trial.

4.3 Crossover Dose-Response Trials

A *randomized, multiple crossover trial* of different doses can provide useful information if a drug effect develops rapidly and subjects return to baseline conditions quickly after cessation of therapy. This design can give individual patient and population dose-response information. Unlike titration trials, which are a type of crossover trial but without random order of doses, this design allows for better discrimination between dose effect and time effect. The crossover design is problematic if there are many withdrawals from the trial. In addition, the trial duration can be quite long, depending on the pharmacokinetics and dynamics of the drug and there can be uncertainty about carryover effects. Trials of by-period interactions can help evaluate the presence of such effects. Similar to the titration designs, fewer subjects are needed compared to a parallel design, but the increased trial duration may present significant problems. For drugs with a prolonged duration of effect, crossover trials are not realistic.

5. CONTROLS, TRUTH STANDARDS, AND COMPLIANCE

5.1 Types of Controls

The control group is a group of subjects whose characteristics are similar to those of the investigational group, but who do not receive the investigational drug. By comparing results between the two groups, the effect of the test drug can be distinguished from other influences that could affect subjects' clinical status (e.g., subject characteristics, spontaneous change, regression to the mean, and investigator expectations). The two groups should be treated identically, with the exception of treatment assignment (investigational treatment or control treatment).²⁹ Treatment assignment is determined by randomization. To help ensure that differences seen between the groups are real and not the result of distorted perceptions (i.e., bias), the assignment to a particular group, drug, or control is usually blinded (i.e., what treatment a patient receives is not known to the patient, the investigator, or anyone assessing response or analyzing the data).

²⁹ For detailed guidance about the choice of control, see ICH E10. (<http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm129460.pdf>)

Controls are either *concurrent* as described above (i.e., dividing a population into two or more groups, usually by randomization) or *historical* (i.e., comparing a group treated with a test treatment with a group treated in some other way at some other time). A concurrent control group may be defined as one of four major categories: (1) placebo; (2) no treatment; (3) different doses of the trial drug; and (4) a different active drug. Note that while a placebo does not provide treatment, it is not the same as a *no treatment* control. To minimize bias, a placebo-controlled trial administers an inert drug designed to look like the investigational drug. A no treatment control is, by definition, unblinded and potentially open to bias that may affect the conduct or interpretation of the trial. An external or historical control can be an untreated or actively treated group, but is a population external to the trial.

Note that a trial is defined by the treatments the patients are randomized to receive, not by other treatments being given. Thus a trial in which a new treatment is compared to placebo (with both added to one or more standard *background* treatments that all patients receive), is still deemed a placebo-controlled trial (sometimes called an *add-on trial*), as the control itself is the placebo, not the background therapy. Conversely, a trial with an active control (A versus B) may use placebos for A and B (a *double dummy* trial) to mask the treatments, but the comparison is between treatment A and B and therefore is an active-controlled design.

5.1.1 Placebo Control

A trial with a placebo concurrent control is a blinded controlled trial. It generally allows for the optimal assessment of a drug's absolute clinical effect by having a group that receives no treatment (although they are unaware that they are receiving no treatment) and a blinding arrangement that helps to ensure an unbiased assessment of the results. The use of a concurrent placebo controls not only for an actual placebo effect (e.g., a patient's response to medical treatment itself, unrelated to pharmacological activity), but also for a wide variety of factors that can lead to improvement or apparent improvement in a patient, such as spontaneous improvement of the disease, regression toward the mean, a medically supportive environment, ancillary care, and better compliance with other treatment.

The principal concern with placebo-controlled (or any no-treatment control) is that one group of patients may be denied an existing effective treatment. This concern does not apply when there is no effective treatment. Placebo-controlled trials also can be ethically conducted even when there is an available effective treatment, when assignment to placebo will not harm the patient.³⁰ When the trial has an add-on design (all patients receive standard therapy to which either drug or placebo is added by random allocation), there is no ethical issue, as all patients receive the existing effective treatment. Similarly, a placebo-controlled trial in a symptomatic condition, conducted with fully informed and noncoerced patients, is generally acceptable.

³⁰ See ICH E10. (<http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm129460.pdf>)

Placebo-controlled trials are not generally ethical where available therapy is known to prevent mortality or irreversible serious morbidity (e.g., long-term treatments of congestive heart failure or hypertension). It is often possible, however, to use trial designs that avoid such problems. As noted, add-on trials do not raise this issue and randomized withdrawal trials can be used to establish long-term effects where use of a long-term placebo might be unacceptable. In these trials, patients are treated with the test drug, then randomized to drug or placebo. Patients randomized to placebo are removed from the trial when symptoms reoccur. These designs are commonly used to show long-term effectiveness of antihypertensives and antidepressants.

It is common in certain settings to include an active-controlled arm in placebo-controlled trials, in addition to the investigational drug, to assess *assay sensitivity*, the ability of the trial to distinguish effective from ineffective treatments. Clinical trials of antidepressant drugs, for example, often fail to show effectiveness of active agents, and it is therefore common to include an active control in addition to the test drug. If the trial cannot distinguish the established treatment or the test drug from placebo, these results show the trial lacked assay sensitivity and would not suggest ineffectiveness of the test drug. If in contrast, the established treatment can be distinguished from placebo but the test drug could not, the trial would suggest that the test drug was indeed ineffective.

5.1.2 No-Treatment Control

Considerations regarding use of no-treatment controls are largely the same as considerations for use of placebo controls, except that blinding is considered neither feasible nor necessary, generally in the setting of endpoints that are completely objective (e.g., mortality). Because blinding is an important tool for reducing bias, reviewers should encourage sponsors to use placebo controls rather than no-treatment controls whenever the use of placebo would be feasible. Even if the overall trial cannot be blinded, it is often possible to have blinded evaluation of endpoints (e.g., using blinded adjudication committees that review cases for events without knowledge of treatment assignment).

5.1.3 Dose-Comparison Control

Dose-comparison concurrent control trials compare several doses of an experimental regimen. These trials can provide important information regarding dose-response, as well as compelling evidence of drug effectiveness. If a trial demonstrates a positive dose-response on an outcome measuring clinical benefit (e.g., a higher dose is significantly better than the lowest dose), it can provide evidence of drug effectiveness, even without a placebo group. However, if the doses are indistinguishable, the trial usually will be uninterpretable absent a placebo group.

If a dose-response control trial will be used to provide evidence of efficacy, reviewers should ensure that the proposed analysis for determining efficacy was prospectively defined. Many options exist for determining efficacy, although their success can depend on the (not yet known) actual shape of the dose-response relationship. For example, the highest dose can be compared to the lowest dose (and/or placebo), or the two highest doses can be combined and compared to the lowest dose (and/or placebo). All doses can

be evaluated to determine whether the slope of the dose-response curve is greater than zero. The relative power of these approaches depends on the true shape of the dose-response curve. The protocol should include a prospectively designed SAP that includes the method of analysis. If more than one approach is contemplated, the SAP should account for multiplicity or use a sequential approach (e.g., test for positive slope, then high dose versus placebo). The choice of primary analysis should be clear.

Comparison of a drug to a lower dose of the same drug can sometimes be a more attractive alternative to the use of a placebo in situations in which investigators or subjects are reluctant to use a placebo. However, if the lower dose selected is known to be minimally active or inactive, the ethical considerations are no different from those concerning placebos. Subjects must be appropriately informed of that information in the informed consent (see [section 9.2](#), Informed Consent).

Although a positive slope provides evidence of drug effect, even without a placebo group, the addition of a placebo control to a dose comparison trial allows for the assessment of absolute effect size, increases the chances that the trial will be able to identify a minimal effective dose, and greatly increases the statistical power to determine a treatment effect. Sponsors should be encouraged to include a placebo group when possible. Data from the placebo group also should be used in the slope analysis, instead of simply subtracting the effect in the placebo group from each of the active dose groups. Failure to do so has resulted in uninterpretable inferences.

Many of the principles regarding dose-response and clinical trial design are also discussed in guidances.³¹

5.1.4 Active-Treatment Control

Clinical trials using active-treatment concurrent controls compare the effect of an experimental drug to the effect of an active treatment. Such trials can have various objectives, including demonstration of effectiveness by showing noninferiority to an active drug, demonstration of effectiveness by showing superiority to an active control, and assessment of the relative effectiveness of two drugs.

To demonstrate effectiveness by showing noninferiority to an active drug in a trial without an additional placebo control, the control drug should be of established effectiveness and its effect in the current trial must be able to be reliably estimated by past experience with the drug.³² This design depends on an estimate of the effect that is not actually measured, giving the trial some similarity to a historically controlled trial.

³¹ See ICH E4

(<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073115.pdf>) and the guidance for industry *Exposure-Response Relationships — Study Design, Data Analysis, and Regulatory Applications*.

(<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072109.pdf>).

³² See 21 CFR 314.126.

The design is rarely useful in symptomatic conditions where placebo-controlled trials of effective drugs often fail to show an effect (e.g., studies of depression, anxiety, pain, allergic rhinitis), making it difficult to be reasonably sure that a noninferiority study will have assay sensitivity. The difference between the test and control drug that should be ruled out (i.e., the noninferiority margin) can be no larger than the known effects of the control drug in the new trial, and usually is some fraction thereof (e.g., rule out a more than a 50 percent decrease in effect size). The fraction of the effectiveness of the control drug to be preserved depends on the seriousness of the disease and other factors.

To set a noninferiority margin based on the past performance of the control drug, the design of the historical trials and the characteristics of the populations used in those historical trials should be similar to those proposed for the investigational trial. Also, there should be reasonable assurance that the treatment effect of the active control (versus placebo) has remained constant over time (e.g., that additional treatments that would make its effects smaller have not become standard). Moreover, the trial should be well conducted because poor performance can lead to a *bias toward the null*, decreasing the effect size of the control and any difference between treatments, undermining the fundamental premise of the noninferiority trial. Also, for some situations, the sample size needed to demonstrate noninferiority of a new therapy to a highly effective standard therapy may preclude the practical conduct of the trial.

The difficulties of using active-controlled, noninferiority trials to show effectiveness in many situations, in particular the problem of assuring the assay sensitivity of the trial, are discussed in ICH E10,³³ the guidance for industry *Antibacterial Drug Products: Use of Noninferiority Trials to Support Approval*,³⁴ and the draft guidance for industry *Non-Inferiority Clinical Trials*.³⁵

When the trial's objective is to establish efficacy of the new drug by showing superiority to the control, the control should have demonstrated efficacy, but its effect size is not critical. When the trial's purpose is to compare the relative effectiveness of two drugs, the comparison should be *fair*, with optimal use (e.g., dose, timing of measurements) of the control drug.

Designing a *fair comparison* of an experimental drug to an active-treatment concurrent control involves consideration of all variables that might affect the safety or efficacy of the drugs being compared. The target patient population (i.e., demographic and baseline disease state), concomitant therapies, and endpoints should be examined for their effect on expected drug activity overall, and for their differential effect on the activity of the

³³ <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm129460.pdf>

³⁴

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM070951.pdf>

³⁵

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM202140.pdf>

two drugs. Consideration should be paid to selection of the dose and regimen of both drugs. In general, the active control should be studied at an approved and well-studied dose; sometimes, use of multiple doses of one or both drugs may be optimal to allow comparison of efficacy at equally tolerated doses or safety at equally effective doses.

5.1.5 External (Historical) Control

External controls, whether truly historical or concurrent from another trial, arise from a population different from the trial population given the drug. There is no randomization to ensure comparability of the populations, nor blinding to ensure comparability of treatment and assessment. In some cases, there may not be an explicit control but rather a general knowledge of the progression of untreated disease. Reviewers should be fully aware of the limitations associated with the use of external controls and should notify sponsors who propose using external controls of the potential problems with this clinical trial design. Concerns regarding comparability between subjects and external controls include the following:

- Assessment of baseline disease (e.g., new diagnostic techniques)
- Baseline disease severity (e.g., patients now present earlier or survive longer compared to the past), or shifts in staging classification criteria (e.g., altered criteria that define *class 2* severity differently in an older and newer trial)
- Subject demographics
- Treatments received before the trial
- Ancillary treatments on trial
- Assessment of response to therapy (e.g., more frequent or sensitive assessment techniques)
- Differential placebo effects
- Lack of critical covariate and outcome data measured in the same manner and at the same time as in the clinical trial

The use of historical controls is inappropriate for most circumstances, given the usually modest treatment effects of most therapeutic drugs. In limited cases, historical controls can be acceptable for establishing drug efficacy, primarily in cases where natural history of the disease is well established and highly predictable and the treatment effect is so large that bias introduced from the problems previously listed is unlikely to lead to incorrect decisions. For example, a tumor response rate is interpretable as a treatment effect without a concurrent control group, as tumors do not ordinarily shrink by themselves. Similarly, cures of a tumor rarely occur spontaneously (e.g., acute leukemia, metastatic testicular cancer). On the other hand, modest drug effects on either survival or

time to progression of cancer cannot be evaluated without a concurrent control, because these parameters cannot be accurately predicted from past experience.

When reviewing a protocol proposing to use historical controls, the reviewer should assess the trial's assumptions and consider the feasibility of alternative trial designs. If historical controls are of potential use, the sponsor should be encouraged to define the control group before obtaining data on the new treatment rather than choosing the control with the new drug data in hand.³⁶ The sponsor should be encouraged to examine the prospectively identified historical control group with respect to the elements previously listed. In addition, the face validity of the data for the historical controls should be examined by comparing the results with similar databases. Valid historical databases should have subjects similar to those being studied with an investigational new drug, and should also represent the common experience with a particular therapy. When widely differing subject outcomes occur in otherwise similar historical databases, it is difficult to use historical controls for establishing effectiveness.

5.2 Trial Design Features

5.2.1 Randomized Withdrawal Trials

A trial design that is particularly useful to establish long term effectiveness, examine optimal duration of treatment, and provide rapid confirmation of effectiveness is the randomized withdrawal design. In this design, treated patients who appear to be responding to treatment are randomized to continued treatment (possibly with several different doses) or placebo. Endpoints can include the rate of a specified reoccurrence of signs or symptoms (BP or Ham-D Depression Scale reaching a specified level) or a measurement over a defined period (e.g., average BP over weeks 2 to 4, at week 4).

As noted, this design has been used to show how long a treatment continues to provide benefit (e.g., the NSABP B-14 trial, in which women were randomized to receive tamoxifen or placebo for 5 years; women on tamoxifen were then re-randomized to receive either tamoxifen or placebo for an additional 5 years; the FIT and FLEX trials of alendronate, in which women who completed 3 years of therapy were randomized to continue it for 5 years or to take placebo) to document maintenance of effects of antidepressants, antipsychotics, and antihypertensives; and to provide rapid means of conducting a confirmatory trial in a population maintained indefinitely on treatment (e.g., nifedipine in vasospastic angina, GHB in cataplexy, tetrabenazine for treatment of choreiform movement in patients with Huntington's Chorea). In the last example, often used to study treatment of rare diseases, the ability to conduct a confirmatory trial in an already identified population can save years of time that would be needed to recruit new patients. These confirmatory trials are also enriched with apparent responders, increasing statistical power. (See [section 5.2.3](#), Enrichment.)

³⁶ See ICH E10. (<http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm129460.pdf>)

5.2.2 Adaptive Designs

There is a growing interest in the use of adaptive designs, broadly defined as any design that uses accumulating information in the course of the trial.³⁷ Reviewers who encounter an adaptive trial design should contact their statistical team members so that appropriate statistical advice may be provided. Although many of these designs raise complex analytic issues related to multiplicity, control of Type I error, and potential unblinding, several adaptive features are well established and should be considered in many trials and, in some cases, suggested if not proposed by the sponsor:

- In almost any outcome trial, a critical determinant of statistical power is the event rate. A straightforward adaptive feature is a blinded assessment of total events, with a plan to increase enrollment or extend the trial duration to attain the needed number of events. Indeed, many trials are designed as event-driven trials, planning for a specified number of events rather than subjects.
- In most trials with a significant endpoint (stroke, death), there will be a plan for an interim examination of results by a DMC, including prospectively planned interim assessment(s), using one of several alpha-error conserving approaches (e.g., O'Brien-Fleming, Peto, Lan-DeMets). It is important to consider in these trials how blinding of investigators not doing the interim analysis is maintained.

5.2.3 Enrichment

In one way or another, most trials are enriched to improve the chance of detecting an effect.³⁸ Enrichment can include identifying a population likely to be fully compliant who still has a sign or symptom after a placebo lead-in period, who has a prognostic marker that will result in many of the events the drug is intended to prevent (e.g., the use of high C-reactive protein patients in the JUPITER Trial,³⁹ or class III, IV heart failure patients in early trials of ACE inhibitors (CONSENSUS)⁴⁰), or who has a marker that predicts response to the test drug (of increasing interest in cancer trials).

All such approaches need to be clearly described in the protocol, and any implications for analysis and interpretation of the trial fully considered in the protocol and other documents.

³⁷ See the draft guidance for industry *Adaptive Design Clinical Trials for Drugs and Biologics*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM201790.pdf>)

³⁸ See the draft guidance for industry *Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM332181.pdf>)

³⁹ Ridker, P et al., 2008, Rosuvastatin to Prevent Vascular Events in Men and Women With Elevated C Reactive Protein, *New England Journal of Medicine*, 359:2195–2207.

⁴⁰ The Consensus Trial Study Group, 1987, Effects of Enalapril on Mortality in Severe Congestive Heart Failure, *New England Journal of Medicine*, 316:1429–1435.

5.2.4 Crossover and Multiple Treatment Trials

As noted in [section 4.3](#), Crossover Dose-Response Trials, crossover trials can be used to assess dose-response more efficiently than parallel design trials. More generally, crossover designs can be more efficient because patient characteristics are the same for treatment and control groups (unlike a parallel trial where there are covariates) and because, other things being equal, they will need half as many subjects for a two-group trial. The problem is their potential for time-dependent changes and carry-over effects, which can make the two periods noncomparable.

A potential approach in studying rare diseases with reversible or intermittent manifestations is the *n of 1* multiple crossover trial design. In one variant, patients are assigned to a random sequence of drug and placebo periods, moving to the next treatment when an endpoint is reached. Such a trial, enrolling just 9 patients but encompassing 47 treatment periods, was the basis for approval of danazol for treatment of hereditary angioedema.⁴¹

5.2.5 Trials in Nonresponders or Intolerants

It is of great value to know whether a new treatment is effective in patients whose disease fails to respond to another therapy. To show this vigorously, it is critical to randomize patients to the new and failed drug, a design rarely used. Merely studying the new drug versus placebo in a nonresponder population does not provide information on the performance of the failed drug in the new trial. The failed drug cannot be severely toxic, which would make re-randomization unethical.

This design was used to approve clozapine (compared to typical antipsychotics), despite a greater than 1 percent rate of agranulocytosis; bepridil (compared to diltiazem) for angina, despite Torsades De Pointes potential; and captopril (compared to triple antihypertensive therapy of hydrochlorothiazide, hydralazine, and reserpine), despite what was thought to be a risk of agranulocytosis. In each case, the new drug demonstrated markedly superior efficacy to the control drugs that were ineffective in the enrolled population, positively influencing an assessment of risks and benefits. It is a potential trial design to develop drugs whose toxicity could make them otherwise unacceptable. Similar approaches can be used to evaluate a new drug in patients who experienced unacceptable (but not dangerous) adverse effects on a prior treatment.

5.3 Truth Standards

The evaluation of a diagnostic test judges how well it can assess the true state of what it is designed to measure (e.g., presence or absence of a lesion). A truth standard is the measure that is used to describe the true state of a subject or a true value of a measurement. Truth standards provide a means of evaluating diagnostic tests, such as

⁴¹ Gelfand, JA, RJ Sherins, DW Alling, and MM Frank, 1976, Treatment of Hereditary Angioedema With Danazol – Reversal of Clinical and Biochemical Abnormalities, *New England Journal of Medicine*, 295:1444–1448.

medical imaging, including the critical effectiveness parameters of a diagnostic test (sensitivity, specificity, positive and negative predictive values).⁴²

5.4 Assessing Treatment Compliance

The extent of treatment compliance by patients is often an important factor in being able to show drug effects, but removing poorly compliant patients from the evaluation after randomization is generally unacceptable, because poor compliance, even compliance while taking a placebo, can result from other factors that predict or are associated with poor outcome, as was seen in the Coronary Drug Project.⁴³ Therefore, sponsors should be encouraged to ensure good compliance insofar as this is possible.

When reviewing protocols for trials in which compliance is particularly critical, reviewers should assess the adequacy of procedures in place to promote and assess compliance. Such procedures can include patient education, reminders, pill counts, drug levels, and interviews.

Some trials, including the *Physician's Health Study* and the *Veteran's Administration Cooperative Study on Antihypertensive Agents*,^{44,45} have tested compliance before randomization. In some cases, *medication event monitoring systems* are used and poorly compliant patients can be encouraged to do better. Examples of such systems include vials or blister packs with a medication monitoring microprocessor. In general, however, poor compliance tends to bias the results toward the null, weakening the ability of a trial to show superiority to a control. In active-controlled noninferiority trials, poor compliance tends to minimize differences between treatments and might allow a finding of noninferiority for an inferior drug.

5.5 Background Care and Standard of Care

Many drugs are studied against a background of standard-of-care practices, including concomitant medications, procedures, diet, diagnostic evaluations, and other interventions. Sponsors should be encouraged to describe expected care in the clinical protocol and ensure that it falls within accepted norms. The standard-of-care practices

⁴² Additional discussions of how to use truth standards in imaging trials can be found in the guidance for industry *Developing Medical Imaging Drug and Biological Products, Part 3: Design, Analysis, and Interpretation of Clinical Studies*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM071604.pdf>)

⁴³ Influence of Adherence to Treatment and Response of Cholesterol on Mortality in the Coronary Drug Project, 1980, *New England Journal of Medicine*, 303:1038–1041.

⁴⁴ Steering Committee of the Physician's Health Study Research Group, 1989, Final Report on the Aspirin Component of the Ongoing Physician's Health Study, *New England Journal of Medicine*, 321(3):129–35.

⁴⁵ The Veterans Administration Cooperative Study Group on Antihypertensive Agents, 1967, Effects of Treatment on Morbidity in Hypertension: Results in Patients with Diastolic Blood Pressure Averaging 115-129 mmHg, *Journal of the American Medical Association*, 202:1028–1034.

can be left to the investigator or can be very fully specified. Sponsors should be strongly encouraged to consider how other modalities of care may interact with the trial drug. When other treatment may affect critical outcome measures of the trial, it is advisable to minimize possible differences in use of those treatments between the investigational drug and control groups and to assess the effect of any differences. It is also important to ensure that the type of background care provided to the patients is consistent with care used by the medical community in the United States, so that trial results are applicable to the potential postmarketing setting.

Even if background care is left to the investigator, the reviewer should nonetheless ensure that the protocol is fairly detailed regarding the types of patient care permitted and what rules should be used for determining what care is provided. It is also important to ensure that the trial captures data regarding the types of relevant care actually administered to the patients. In general, more restrictive approaches to the type of care given may decrease variability and provide greater statistical power, but more flexible approaches with broader entry criteria may facilitate enrollment and generalization of results to the target population after approval, giving a more *real world* result. If the trial is successful, potential differences in response among differently treated subsets can be explored.

6. RANDOMIZATION AND BLINDING

Randomization and blinding are the two principal means of reducing bias and ensuring validity of trial conclusions. Randomization helps protect against the possibility that differences between groups at baseline will lead to outcome differences that might mistakenly be attributed to drug effect. Blinding protects against the possibility that differences in the on-trial treatment or assessment of subjects will lead to spurious outcome differences that are mistakenly attributed to a drug effect.

6.1 Randomization

In the context of clinical trial design, randomization is defined as the allocation of patients to the investigational drug and control arms *by chance*. Randomization is intended to prevent any systematic difference between patients assigned to the treatments being compared and is a critical assumption for valid statistical comparisons. It is also intended to produce groups that are comparable (statistically balanced) with respect to both known and unknown factors.

Although covariate adjustment may be able to deal with imbalances in certain identified factors, one cannot *adjust* for unknown prognostic factors. Randomization provides reasonable assurance that the unknown factors will be randomly distributed and that the overall prognosis is equivalent across trial groups. If important covariates (i.e., variables that predict risk or prognosis) are identified before trial initiation, stratification for such factors (i.e., randomizing such groups separately), should be considered, especially if the factors are relatively uncommon. More common factors should be evenly distributed by the randomization process.

Sometimes, sponsors propose systematic allocation schemes that would appear to produce randomized groups. For example, some trials have been conducted with groups assigned according to the day of the week the patients were present in clinic, or on their birthdays, or by the first letter of their last names. These procedures, if followed without any other influence, might produce random assignment, but they should be discouraged because they may not ensure blinding and assignments could therefore be biased. Some of these methods could produce other imbalances. For example, the letter Z is a common first letter for surnames of Chinese/Chinese-Americans, but not for other ethnic groups, and there could be genetic or other baseline differences between Chinese and other ethnic groups. Therefore, assignment of all patients with surnames beginning with the letter Z to treatment A only, and none to treatment B, can create an imbalance between treatments A and B with respect to ethnicity.

The implementation of an appropriate randomization procedure depends upon the trial design. There are two general types of randomization schemes: fixed randomization (the most common type) and adaptive (or dynamic) randomization.

6.1.1 Fixed-Randomization Schemes

Fixed-randomization schemes have probabilities for being assigned to the investigational or control drug that remain fixed during the entire course of the trial, usually 1:1, but possibly 2:1, 3:1, and so on. The allocation ratio, any stratification, and block sizes are defined before the trial and remain unchanged following trial initiation. Fixed-randomization schemes are generally easier to manage than adaptive schemes.

6.1.1.1 Blocked randomization

The simplest form of fixed randomization, unrestricted randomization, can be implemented by tossing an unbiased coin for each eligible patient or generating a single sequence of random numbers. The major difficulty with this procedure is that large imbalances between the sizes of the assigned groups can arise by chance at any point during the clinical trial. This is an especially important consideration for small trials, or for trials in which an interim analysis is performed that could lead to early termination.

Blocked randomization procedures are frequently implemented to avoid this potential problem and ensure balance between arms throughout the trial. In such procedures, balance of group assignments is ensured within each of a series of patient entries of a prespecified size (e.g., if the block size is six and randomization is 1:1 in a two-arm trial, three patients will be assigned to each treatment in that entry block). Thus, if the first three patients are randomized to get the investigational drug, the next three will get the control drug. If small and constant block sizes are used (e.g., size of two), treatment assignments may come to be predicted by the investigators if the trial is not fully blinded. Therefore, the block sizes generally should be greater than two and varied and not divulged until trial completion. For trials in which blinding is likely to be at least partially compromised, reviewers should advise sponsors to use a variable block size.

Sponsors also should be encouraged to use blocking if: (1) patient enrollment is likely to continue over an extended period of time, especially if demographic or clinical

characteristics of the trial population can be expected to change over the course of enrollment; or (2) there are practical or statistical reasons why it is important to satisfy the specified allocation ratio at various points during the enrollment process.

6.1.1.2 Stratified randomization

In most trials, it is important to achieve comparability between the treatment arms with respect to known prognostic factors measured at the start of the trial (e.g., demographic, disease stage, trial site). Stratified randomization schemes are often used to accomplish this balance. In stratified randomization, individuals are randomized within strata depending on their baseline characteristics. For example, if the stratification factors are sex, age (e.g., younger or older than 60), and disease stage (early versus late), there will be 8 strata, 1 for each possible combination of categories (e.g., younger than 60, female and early would be one stratum; older than 60, male, late would be another).

Stratification provides additional assurance, over and above that provided by randomization, that important factors will be equally represented in the two (or more) trial groups. Use of stratification can be especially important in small trials, where randomization alone will less assuredly lead to balance of important covariates, although small trials may be able to support only a few strata (e.g., in a trial of 60 individuals, if there were 8 strata and a block size of 6, it is easy to see that most of the strata will not be completed; therefore, the assurance of balance will be largely lost). In large trials there is little or no need to stratify for common variables at randomization (e.g., age, sex, disease duration), because these variables generally will be similar in the two groups even without stratification. In multicenter trials, stratification by center helps to ensure balanced treatment assignment within each center, thereby controlling for center-specific factors (e.g., differences in patient population, patient care, and management policies).

Stratified randomization is also used when the magnitude of the treatment effect is expected to vary among important groups. Analyses that reflect this type of stratification are more likely to detect a treatment effect, if one exists, than analyses that ignore the stratification.

The larger the number of stratification variables and strata, the smaller the number of subjects per stratum. Stratification for more than two or three factors often leads to strata that are so small relative to block size that balance is not ensured. The following is a list of some basic tenets of stratification:

- Only variables that can be observed before randomization can be a basis for stratification.
- It is impractical to control for more than a few sources of variation by stratification.
- Variables that are subject to major sources of error because of differing interpretations will not be helpful stratification variables.

- It is unreasonable to expect that all important sources of baseline variation will be controlled by stratified randomization.
- Multicenter trials with plans to enroll substantial numbers of subjects at each center generally should stratify by center (i.e., randomized within centers). Stratification for this variable controls for differences in the trial population because of environmental, social, demographic, and other factors and differences in management. In multicenter trials in which only small numbers of subjects are entered at each center, stratification by center generally is not practical.
- Stratification should be limited to those variables that are considered potentially correlated with treatment effect.

6.1.2 Adaptive-Randomization Schemes

Adaptive-randomization schemes assign experimental and control therapies with probabilities that are modified during the trial as a function of prerandomization characteristics, such as age or sex of subjects entered to that point. The most common type of adaptive randomization is sometimes referred to as *minimization*, and uses an algorithm that assigns treatment for an individual by using that individual's characteristics to calculate which assignment would result in the best overall balance with respect to all factors of interest, and then making that assignment. A preferred variation includes an element of randomization, as well as minimization; the *optimal* treatment assignment is made with a specified probability, such as 80 percent, rather than 100 percent.

Although this approach may be useful when there are many important factors on which balance is desirable, adaptive randomization, like stratified randomization, can become ineffective as the number of factors increases and the number of subjects in each block becomes too small to ensure balance.

Adaptive randomization, intended to ensure similarity of randomized populations, should not be confused with adaptive trial designs that attempt to alter assignment based on interim outcomes, which have a different purpose. For example, the *Play the Winner* allocation algorithm assigns new individuals (often with some element of randomization) to the treatment that appears to be more successful at that point in the trial. The purpose of this approach is to minimize the number of patients assigned to the inferior therapy. These design types, which are used rarely in the regulatory environment, are not applicable to trials in which the primary endpoint does not occur soon after the initiation of treatment, and are subject to difficulties if real-time data on outcome are not highly reliable. It is important not to alter assignments so markedly that one group is assigned very few patients.

6.1.3 Allocation Ratio

Allocation of patients to treatment and control arms can be uniform or nonuniform. Uniform allocation (i.e., equal numbers allocated to each arm) is the usual practice and provides the most statistical power for a given total sample size. Nonuniform allocation

may lower costs (if one arm is substantially more expensive) and improve recruitment (if one arm is generally preferred) and may increase the size of the exposed patient safety database. In general, the loss of statistical power in seeking to detect a difference between treatments going from uniform allocation to 2:1, or even 3:1, is fairly small; however, as more imbalanced allocation occurs, power drops off more rapidly. A special case is where a trial seeks both to show effectiveness versus placebo and to compare the test drug with an active control. In that case, it usually is necessary for the active treatment groups to be substantially larger to examine the smaller differences between the active treatments.

6.1.4 Review of Randomization

Reviewers should ensure that the protocol prospectively defines and explains both the randomization scheme and the proposed mechanism for its implementation. In addition, the appropriateness of allocation ratio, block size, and stratification variables should be evaluated. The seed used in the random number generator should be kept on file so that, if needed, the implemented assignment can be reproduced and audited at the completion of the trial. Centralized randomization, using computers or telephone, is usually preferable to envelope randomization with envelopes sent in advance to each site, because it is less vulnerable to manipulation.

6.2 Blinding

Patient or investigator knowledge of the treatment being administered can introduce bias into a trial through influences on patient management or assessment or by affecting expectations on outcomes. Whenever possible, investigators should be blinded to the drug they are administering and evaluating. Blinding is especially important for those trials in which endpoints contain elements of subjectivity, whether patient-reported or physician-reported outcomes (e.g., pain or depression), outcomes (or reporting of an outcome) that can be influenced by expectations (e.g., hospitalization for pneumonia), or outcomes that can depend on patient effort (e.g., exercise test). Many *objective* outcomes (such as occurrence of heart attack or stroke), also depend critically on interpretation of patient data reported or collected (e.g., symptoms, ECG, enzymes), and therefore also have subjective elements. Even completely objective outcomes, such as time to death, can be affected by factors such as use of do-not-resuscitate orders, concomitant therapies, or intensity of follow-up efforts.

In *single-blind trials*, only the patients are unaware of the treatment arm to which they have been assigned. In *double-blind trials*, the patients and the clinical investigator are unaware of the assigned treatment. The term *triple-blind* is sometimes used to refer to double-blind trials in which additional parties are also blinded (e.g., pharmacists, assessors of activity, data analysts, or other professionals). However, these people also should be blinded in a double-blind trial. Exactly what sponsors mean by single-, double-, or triple-blind trials should be specified, as definitions may vary. In general, all parties to a trial who will be making any judgments about patient management, outcome, or consideration of the design features and the analytical approaches to be taken (e.g., inclusion, exclusion, covariates, choice of analyses) should be blinded to patient

assignments. The exception is members of a DMC and the statistician who prepares reports for the DMC.⁴⁶

Since most trials incorporate some subjective assessments (e.g., tolerability), sponsors generally should be encouraged to use a high degree of blinding when feasible. In addition to the extent of blinding planned (e.g., single-, double-), reviewers should consider the specific procedures used to maintain and assess the blinding and approaches used to deal with imperfect blinding as described by the sponsor. Even in an unblinded trial, endpoint assessment is often made by a blinded evaluation group and overall results can be presented blindly to analysts of a multicenter trial; there still can be a concern as to whether the events sent to the evaluation group are identified in an unbiased way.

6.2.1 Optimizing and Maintaining the Blinding

Reviewers should consider the following factors in assessing and improving the likely quality of the blinding.

6.2.1.1 Character of the placebo and trial drug

A placebo should be indistinguishable from the trial drug, with respect to color, smell, taste (if orally administered), clarity (if liquid), size, or volume, and any other characteristic that might allow it to be identified. In addition, a placebo should be administered identically to the trial drug, including the time, duration, and mode of administration.

In some cases, characteristics of a trial drug that can be difficult to mimic in the control can be masked or altered. For parenteral medications, an opaque wrap around the syringe or intravenous bottle may serve to mask color and clarity. For oral medications, recognizable pills are sometimes crushed and placed in capsules. In such cases, sponsors should ensure that changes to the drug are not pharmacologically significant.

6.2.1.2 Unblinding drug effects

Reviewers should be aware of possible unblinding effects of treatments and should emphasize to sponsors the importance of anticipating possible unblinding effects and managing them insofar as this is possible (e.g., by having blinded assessments of particular outcomes). Patients and investigators can become unblinded when a trial drug causes potentially recognizable effects, whether intended effects or side effects. There is no remedy for this, but it may be useful for sponsors to ask patients what treatment they think they received and pose similar questions to investigators. An exploratory analysis could consider results in patients who were and who were not unblinded. Other possible interventions that can help maintain the blinding of investigators include clothing that covers injection sites and instructions to patients and investigators not to discuss injection site reactions or other potential drug side effects.

⁴⁶ See the guidance for clinical trial sponsors *Establishment and Operation of Clinical Trial Data Monitoring Committees*.
(<http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm127073.pdf>)

In some cases, a drug can cause potentially unblinding effects related to laboratory measurements. For example, erythropoietin treatment results in higher reticulocyte counts; various drugs raise uric acid levels or alter electrolytes. When possible, and if it can be done without compromising patient care, investigators of therapies that alter laboratory functions should be blinded to the laboratory analyses. In some cases, where needed for patient care, another party can manage the laboratory abnormality.

When drug levels are measured as part of the trial, the investigator should remain blinded to the results. If, as a result of PK results, someone other than the investigator makes dose and regimen adjustments, similar adjustments should be made in the control arm to preserve the blinding. Reviewers should be aware that in some cases of serious and chronic diseases, patients have unblinded themselves by having blood levels tested at nontrial laboratories.

Drugs highly effective in alleviating symptoms can unblind patients and investigators as a result of their efficacy. Such unblinding should have little effect on the evaluation of effects of the outcome that caused the unblinding because of the noticeable efficacy, but can bias evaluation of longer term effects, adverse effects, and other efficacy measures.

6.2.1.3 Dealing with imperfect blinding

In many cases, the blinding will not be perfect. In addition to optimizing the blinding, sponsors should be encouraged to try to minimize the effect of unblinding on the trial and to assess its potential affect by relying on objective endpoints with prespecified evaluation criteria and by using assessment by blinded evaluators for endpoints that entail subjectivity (e.g., medical evaluations or interpretation of radiographs).

6.2.1.4 Assessing unblinding

It may be useful to capture data regarding the extent of investigator unblinding. In most trials, clinical investigators have the ability to unblind a patient if they deem it important for the patient's management. When physicians are provided the opportunity to break the blind, it should be done in a manner in which there is a record of when and by whom the blind was broken. Drug containers that contain the identity of the drug (e.g., under an opaque patch) should provide a means to assess whether they have been unblinded. In other types of unblinding, such as that caused by drug effects, the extent of unblinding may be more difficult to quantify, but questionnaires administered to patients might be helpful.

6.2.1.5 Blinded evaluators or evaluation committees

If the clinical investigator cannot be completely blinded, it is often helpful to use independent blinded evaluators to assess important endpoints. Such evaluators can be kept blinded to knowledge of any factors that might cause unblinding. For example, in trials of interferon in multiple sclerosis, both injection site reactions and the occurrence of flu-like syndrome can unblind some patients and investigators. In such trials, determination of exacerbations and of progression of disability can be performed by a neurologist other than the one treating the patient.

A closely related approach is to establish a clinical endpoints committee to adjudicate the occurrence of an endpoint while blinded. In blinded trials, there is debate as to whether such committees represent an improvement over the on-site assessments, but in trials without blinding, such committees can be a means of ensuring that endpoint assessments are not biased (see [section 8.2](#), Endpoints).

7. PATIENT POPULATIONS, SPECIAL POPULATIONS

7.1 Trial Population

In general, the choice of the trial population in a phase 2 or phase 3 clinical trial should reflect the intended use of the drug. This principle should not be interpreted to preclude use of selection criteria that improve the power and practicality of the trial. It is common, for example, to require persistence of disease over a run-in period; stability of baseline measures such as BP, exercise tests, or pulmonary function tests; or factors that improve the likelihood of compliance. In outcome trials, it is common to choose patients who are expected to have a high rate of primary endpoint events (prognostic enrichment) on the basis of clinical history, pathophysiologic observations, disease severity, or genetic or proteomic predictors. Any differences (e.g., disease stage or severity, risk factors, demographics) between the intended population and the population in which efficacy and safety are to be studied should be identified and their effect on generalizability of results and the applicability of results to a specific population and labeling examined. It should be recognized, however, that study of a broad population raises many of the same issues, even if differences in response among population subsets are well studied.

In an oncology trial, for example, there might be a marker that predicts response, such as overexpression of HER2 on breast cancers treated with Herceptin. In a broad population not selected for that marker, about one-third would have HER2 overexpression. If a trial conducted in an unselected population of breast cancer patients demonstrated an improved survival of 2 months compared with control, this finding could represent a 6-month improvement in survival in patients with HER2-overexpressing tumors and no effect at all in other patients.

If predictors of response can be identified prospectively, they are used to selectively enroll or treat patients who will potentially benefit from therapy. It also may be important to evaluate response in the nonenriched population (normal HER2 expression in the above example), because other factors may potentially affect response. Decisions on whether to require evaluation in both the enriched and nonenriched population involve many complex factors and must be carefully considered before providing advice to sponsors. CDER policy in this area is evolving in response to the rapidly growing science of genomics and other biomarkers that can support development of drugs targeted to be effective only in the marker-positive population. Reviewers should consult with their supervisors and current CDER policy statements before making recommendations to sponsors regarding preapproval data requirements for evaluation of the nonenriched patient population.

Trials with broad populations are likely to have subsets in which the drug effect is greater or smaller than the average. If anticipated, randomization can be stratified to facilitate examination of such subjects, but even without stratification, subjects can be examined. The regulations require data to be examined for differences in effectiveness, safety, and dose-response among age, sex, and racial subgroups as well as other pertinent subsets (21 CFR 314.50(d)(5)(v) and (vi)(a)). There also is a growing interest in genomic and other predictors of response. Sponsors should be encouraged to consider collection of DNA samples in a substantial proportion of their trial population to allow examination of data in relevant pharmacogenetic subsets. See the draft guidance for industry *Qualification Process for Drug Development Tools*.⁴⁷

7.1.1 Homogeneity vs. Heterogeneity; Individualization of Treatment

There are powerful reasons to include heterogeneous populations in a trial. Study of a population substantially more restricted than the broader patient population with an indication can limit the generalizability of the results. Important differences in safety, efficacy, or optimal dosing in subpopulations, which should be evaluated, might be missed. At the same time, as noted above, there may be valid reasons for focusing at least initially on population subsets with more severe disease or particular disease characteristics, who may be more likely to have the endpoints of interest (prognostic enrollment) or more likely to respond (predictive enrichment).

7.1.2 Factors Influencing the Nature of the Trial Population

The inclusion and exclusion criteria determine the nature of the trial population and it is important to pay close attention to them.

There are a variety of maneuvers, generally referred to as *enrichments*, that can greatly influence the power of the trial.⁴⁸ Some are obvious, such as excluding subjects already taking the test drug (or one related to it) or taking a drug that would interfere with the test drug. It is common to screen possible subjects during a single-blind placebo period before randomization to see whether symptoms persist (eliminate placebo responders), whether subjects have consistent measurements (BP and exercise test), and whether they take their medications (smart bottles). These selections do not affect generalizability.

There are a wide variety of other possible enrichment maneuvers. It may be useful to screen subjects for ability to tolerate the drug. For example, some beta blocker heart failure trials enrolled only patients who demonstrated that they could tolerate beta blockers. Trials of various cardiovascular (CV) interventions have enrolled people with severe disease (e.g., severe heart failure, high cholesterol) or risk factors (e.g., diabetes,

⁴⁷

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM230597.pdf>

⁴⁸ See the draft guidance for industry *Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM332181.pdf>)

recent acute myocardial infarctions (AMIs) or stroke), all of which yield a population likely to have many events, critical to being able to show an effect.

A short-term response such as tumor glucose uptake could be used to identify people with the potential to have a long-term favorable response to treatment. For example, patients have been selected for trials of hormonally active breast cancer drugs on the basis of estrogen and progesterone receptor content in their tumors. As noted above, Herceptin has been studied in patients whose tumors overexpressed HER2. In some cases, a history of past response to the drug class (or past failure on other drugs) has been used to find a population of interest. Reviewers can expect to see subjects chosen (or excluded) on the basis of genomic or proteomic characteristics. These choices are intended to increase the power of a trial, but the selection criteria must be considered to understand the applicability of the results.

7.2 Special Populations, Demographic Subgroups

The identification of differences and similarities in response among populations is an important part of drug development. In 1998, the regulations were modified to require analyses of effectiveness and safety by age, sex, and race. Assessment of drugs in demographic subgroups is addressed in various regulations and guidances.

- 21 CFR 314.50(d)(5)(v) and (vi)(a)
- Guidance for industry *Guideline for the Study and Evaluation of Gender Differences in the Clinical Evaluation of Drugs*⁴⁹
- Guidance for industry *Guideline for the Study of Drugs Likely to be Used in the Elderly*⁵⁰
- ICH guidance for industry *E7 Studies in Support of Special Populations: Geriatrics*⁵¹
- ICH guidance for industry *E11 Clinical Investigation of Medicinal Products in the Pediatric Population*⁵²

⁴⁹

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072044.pdf>

⁵⁰

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072048.pdf>

⁵¹ <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM129519.pdf>

⁵² <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm129477.pdf>

Regulations require *IND annual reports* to tabulate subjects entered into trials by age group, sex, and race (21 CFR 312.33(a)(2)). Reviewers should consider the distribution of subjects and ensure that there are no unjustified subject exclusions (e.g., subjects over 75 years of age), that PK differences among different subpopulations (including age, gender, race and organ dysfunction) are examined in specific trials or by population PK to determine the need for dosage adjustment in these subpopulations, and that the integrated analyses of safety and effectiveness will look for potentially important differences in dose response. Where a condition is particularly important in a demographic subgroup, it may be appropriate to enrich the population for that subgroup. It has been recognized, for example, that many drugs for chronic illnesses are heavily used in a very elderly population (older than 75 years of age).

7.2.1 Subgroup Analyses vs. Special Population Trials

In general, the size of a clinical trial is established to demonstrate an overall treatment effect, not to allow assessment of effects in particular patient subgroups. Therefore, the analyses of specific subgroups will be possible principally in the integrated analyses of safety and effectiveness. A possible exception is large outcome trials in which it is usual to show effects in a variety of cohorts defined by demographics, severity of illness, use of concomitant drugs, and other factors in so-called *forest plots*. These have at times been included in labeling. Care must be taken to avoid specific efficacy claims based on such a subset analysis.

As an alternative to assessing effects in special populations through subgroup analysis or as an approach to confirming a differential effect noted on such an analysis, sponsors can choose to evaluate specific groups of patients in small trials. For example, a sponsor can conduct specific trials of the drug's effects in elderly patients, commonly done for sedative-hypnotic drugs and in patients with varying degrees of renal function. In general, apart from enrichment attempts, sponsors should be encouraged to conduct major efficacy trials in demographically heterogeneous patient populations and in patients with a wide range of concurrent illnesses and treatments to ensure that the results are reasonably generalizable. Within those trials, subset analysis can help identify important differential treatment effects.

In particular, reviewers should closely examine exclusions in phase 3 trials to consider whether they are really needed. It has been common, for example, to exclude patients older than 75, but there is no good reason to do this. Similarly, exclusions of patients with a history of psychiatric or cardiovascular illness, unless dictated by the drug's pharmacology, decrease the opportunity to detect important drug-drug interactions and should be discouraged.

7.2.2 Pediatric Populations

In recent years, new initiatives, including laws, regulations, and guidances, have sought to ensure that drug testing in children occurs at an appropriately early and safe stage in drug development so that drugs are properly labeled for use in pediatric age groups. Reviewers should be aware of the following:

- ICH E11⁵³
- Draft guidance for industry *Pediatric Oncology Studies in Response to a Written Request*⁵⁴
- Guidance for industry *Nonclinical Safety Evaluation of Pediatric Drug Products*⁵⁵
- Draft guidance for industry *General Considerations for Pediatric Pharmacokinetic Studies for Drugs and Biological Products*⁵⁶
- Title V of FDASIA (permanent reauthorization of the Pediatric Research Equity Act and the Best Pharmaceuticals for Children Act)⁵⁷

NDA and BLA submissions for a new active ingredient, new indication, new dosage form, new dosing regimen, or new route of administration must contain a pediatric assessment unless the applicant has obtained either a *waiver* or *deferral*, or the requirement is inapplicable. The pediatric assessment consists of studies to determine safety and effectiveness of a drug or biological product for the approved indication(s) in all relevant pediatric subpopulations. It must also support dosing and administration for each pediatric subpopulation for which the drug or biological product has been assessed to be safe and effective. These data can come from pediatric populations, or may be extrapolated from adult trials when appropriate.

A waiver (or partial waiver if it only applies to a specific age subgroup of the pediatric population) can be granted releasing the sponsor or applicant from the requirement to conduct studies in children. Waivers are only granted if studies are not feasible (e.g., the disease does not occur in children, the number of patients in this population is so small that studies would not be practical), the drug would be unsafe or ineffective in children, the drug does not represent a meaningful therapeutic benefit over existing pediatric therapies, or the applicant has demonstrated that an age-appropriate formulation cannot be made. If studies are waived because there is evidence that the drug would be unsafe or would be ineffective in the pediatric population, this information must be included in labeling. A deferral can be granted to delay the initiation and completion of required

⁵³ <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm129477.pdf>

⁵⁴

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071976.pdf>

⁵⁵

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm079247.pdf>

⁵⁶

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072114.pdf>

⁵⁷ <http://www.gpo.gov/fdsys/pkg/BILLS-112s3187enr/pdf/BILLS-112s3187enr.pdf>

studies until after approval of a submission. Deferrals generally are granted when additional safety or effectiveness data must be collected before pediatric studies can be safely initiated or because the drug is ready to be approved in adults.

The Pediatric Review Committee (PeRC) is an internal FDA committee comprised of experts in pediatrics and pediatric research. The PeRC is required to review all pediatric assessments, PSPs, pediatric waiver requests, deferral requests, and deferral extension requests, as well as all written requests issued under the BPCA. Reviewers are encouraged to consult the pediatric experts in the Pediatric and Maternal Health Staff for assistance with any pediatric issues when necessary.

Under 21 CFR 201.57(c)(9)(iv), the pediatric population is defined as the age group from birth to 16 years, including age groups called neonates, infants, children, and adolescents. Any classification of the pediatric population into age categories is to some extent arbitrary. One classification is shown below, but others should be used if they are more appropriate to the issue under investigation (e.g., age defined by Tanner stage for safety or efficacy issues in adolescent girls). If the clearance pathways of a drug are well established and age-related changes are understood, age categories for PK evaluation might be chosen based on any breaking point where clearance is likely to change significantly. Sometimes it may be more appropriate to collect data over broad age ranges and examine the effect of age as a continuous covariate. For evaluation of effectiveness, different endpoints can be established for pediatric patients of different ages, and the age groups might not correspond to the following categories.

The following list is one possible pediatric categorization. There is, however, considerable overlap in developmental issues across the age categories. Further discussion of each age category can be found in ICH E11.⁵⁸

- Preterm newborn infants
- Term newborn infants: 0 to 27 days
- Infants and toddlers: 28 days to 23 months
- Children: 2 to 11 years
- Adolescents: 12 to 16-18 years (dependent on region)

Drugs or biologics administered to pediatric patients can differ in their PK (e.g., because of immaturity of hepatic microsomal enzymes), PD, immunogenicity (biologics), efficacy, or safety profiles. Furthermore, these profiles can differ among pediatric age groups (e.g., neonates, infants, children, or adolescents), leading to a need for dosage adjustments.

To assess the need for pediatric studies, the potential health benefit and use or potential use of the drug in the pediatric population should be determined. If this assessment indicates usefulness of the drug in one or more pediatric age groups, development of the drug in the pediatric population should be discussed with the applicant.

⁵⁸ <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm129477.pdf>

It is also important to assess the applicant's progress in developing appropriate pediatric formulations. Formulation development can become the rate-limiting step in making a drug available for younger pediatric patients. Where appropriate, applicants should begin development of a pediatric formulation for the purposes of submission and approval before initiation of pediatric studies.

Finally, the FDA has the authority to require holders of applications of previously approved marketed drugs to submit a pediatric assessment under certain circumstances (see section 505(b) of PREA). Because this process is complex and may involve issuance of a written request, a review process that differs from standard review processes, and the possibility of a pediatric advisory committee meeting, experts within the FDA should be consulted on questions related to studying already-marketed drugs.

7.2.2.1 Timing of studies in pediatric populations

For drugs with potentially important uses in pediatric populations, studies in pediatric patients should be carried out relatively early in drug development. Applicants should be encouraged to consider the following factors in deciding when to initiate such studies.

- The prevalence of the condition to be treated in the pediatric population
- The seriousness of the condition to be studied
- The availability and suitability of alternative treatments
- Whether the drug is novel
- What is known or suspected about the safety of the drug in adults
- The need for the development of pediatric-specific endpoints
- The age ranges of pediatric patients likely to be treated
- Unique pediatric (developmental) safety concerns with the drug, including any nonclinical safety issues
- Potential need for pediatric formulation development

Of these factors, the most important is the presence of a serious or life-threatening disease for which the drug represents a potentially important advance in therapy. This situation suggests relatively urgent and early initiation of pediatric studies. Applicants should be encouraged to provide a plan for pediatric studies as early as possible in the drug development process. The plan should be considered at meetings throughout development. Although decisions on timing of pediatric studies are complex, several specific situations should be noted.

- **Drugs for diseases predominantly or exclusively affecting pediatric populations.** The entire development program will be conducted in the pediatric population except for initial safety and tolerability data, which usually will be obtained in adults. Some drugs may be reasonably studied only in the pediatric population even in the initial phases (e.g., when trials in adults would yield little useful information or expose them to inappropriate risk). Examples include surfactant for respiratory distress syndrome in preterm infants and therapies targeted at metabolic or genetic diseases unique to the pediatric population.

- **Drugs intended to treat serious or life-threatening diseases in both adults and pediatric patients for which there are currently no or limited therapeutic options.** The presence of a serious or life-threatening disease for which the drug represents a potentially important advance in therapy suggests the need for relatively urgent and early initiation of pediatric studies. In such cases, drug development should begin early in the pediatric population, following assessment of initial safety data and reasonable evidence of potential benefit. Pediatric study results should be part of the marketing application database.
- **Drugs intended for other diseases and conditions.** In these cases, although the drug ultimately may be used in pediatric patients, there is less urgency than in the previous cases, and studies usually would begin at later phases of clinical development or, if a safety concern exists, even after substantial postmarketing experience in adults. Testing of these drugs in the pediatric population usually would not begin until phase 3 (when applicants have acquired at least some effectiveness data in adults), and often begins after marketing.

ICH E11 contains more specific guidance as to when pediatric studies should or must be initiated.⁵⁹

Additionally, with the passage of FDASIA in 2012,⁶⁰ sponsors developing a drug for a new active ingredient, new indication, new dosage form, new dosing regimen, or new route of administration must submit a PSP within 60 days of the EOP2 meeting. A PSP must contain information regarding specific plans to study the drug in children:

- Outline of the pediatric study or studies planned (including to the extent practicable study objectives and design, age groups, relevant endpoints, and statistical approach)
- Plans and justification for request of a deferral, full waiver, or partial waiver

Additionally, a PSP should contain the following information:

- Overview of the disease in the pediatric population, and the drug under development
- Potential plans and justification for use of extrapolation
- Plans for pediatric-specific formulation development

⁵⁹ <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm129477.pdf>

⁶⁰

<http://www.fda.gov/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticActFDCA/SignificantAmendmentstotheFDCA/FDASIA/ucm20027187.htm>

- Nonclinical data, complete or planned, to support studies in children
- Timeline for completion of the studies contained within the plan
- Any agreements with other health authorities (e.g., Pediatric Investigation Plan for European Medicines Agency)

All PSPs submitted to the FDA must be reviewed by the PeRC (see [section 7.2.2](#), Pediatric Populations). If an EOP2 meeting is not held, then there is some flexibility in the timing of the submission of a PSP. See the draft guidance for industry *Pediatric Study Plans: Content of and Process for Submitting Initial Pediatric Study Plans and Amended Pediatric Study Plans*.⁶¹

7.2.2.2 Types of studies in pediatric populations

Whether clinical effectiveness studies are needed in the pediatric population or in all pediatric populations depends on what is known about the disease, drug concentration response relationships, and other factors.

When a drug will be used in the pediatric population for the same indications as those approved in adults and the disease process is similar in adults and pediatric patients, the outcome of therapy is likely to be similar, and extrapolation from adult efficacy data may be appropriate (see 21 CFR 314.55(a) and 21 CFR 601.27(a)). In such cases, PK trials (and/or immunogenicity trials for biologics) in all age ranges of pediatric patients likely to receive the drug, together with safety trials, may provide adequate information for use by allowing selection of pediatric doses that will produce drug levels in blood similar to those observed in adults. If this approach is taken, adult PK data (and/or immunogenicity data for biologics) should be available to plan the pediatric studies. A similar approach may allow extrapolation of results from older to younger pediatric patients. In such cases, PK trials (or immunogenicity trials for biologics) in the relevant age groups of pediatric patients likely to receive the drug, together with safety trials, can be sufficient to provide adequate information for the younger patients.

When the outcome of therapy in pediatric patients is expected to be similar to the outcome in adults, but the appropriate blood levels in pediatric patients are not known because there is concern that concentration response relationships may differ between the adult and pediatric populations, an approach based solely on attaining similar blood levels will not be sufficient. In that case, it may be possible to use measurements of a PD effect related to clinical effectiveness to confirm the expectations of effectiveness and to define the dose and concentration needed to attain PD effect. Thus, a PK/PD approach combined with safety and other relevant trials could obviate the need for clinical efficacy trials.

61

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM360507.pdf>

In situations where a PK approach is not applicable (such as for topically active drugs), extrapolation of efficacy from one population to another can be based on trials that include PD endpoints or appropriate alternative assessments. Local tolerability assessment (not defined in phase 1 trials) may be needed. It may be important to determine blood levels and systemic effects to assess safety.

When novel indications are being sought for the drug in pediatric populations, or when the disease course and outcome of therapy are likely to be different in adults and pediatric patients, clinical efficacy trials in the pediatric population will be needed.

A complete discussion of PK and efficacy trials can be found in ICH E11.⁶² The principles of clinical trial design, statistical considerations, and choice of control groups detailed in the ICH guidance for industry *E9 Statistical Principles for Clinical Trials*⁶³ and ICH E10⁶⁴ generally apply to pediatric efficacy studies. ICH E11 also discusses ethical issues in pediatric studies, including institutional review board/independent ethics committee (IRB/IEC), recruitment of patients, consent and assent, minimizing risk, and minimizing distress.

7.2.3 Women

Historically (in part as a result of a 1977 guideline), women of childbearing potential were excluded from early clinical trials in an attempt to protect any potential fetus from unanticipated exposure to potentially harmful drugs, particularly in early drug development. Excluding women of childbearing potential rather than informing them of the need to avoid pregnancy while participating in the trial implied to many a lack of respect for autonomy and responsibility and carried the potential of denying women with serious illnesses access to potentially lifesaving experimental therapies. Such exclusions are now strongly discouraged and indeed, if present, are now a basis for a clinical hold when they involve drugs for a life-threatening disease or condition (21 CFR 312.42(b)(v)).

Generally, we encourage the inclusion of women of all age groups throughout drug development, but protection of an existing or potential fetus remains critically important. Therefore, restrictions on inclusion of pregnant women are still justifiable; and women of childbearing potential, when included, are instructed to avoid pregnancy at least until animal reproduction studies are complete. In short-term trials, it is possible to limit potential fetal exposure by giving a drug immediately following a woman's menstrual period after a negative result from a pregnancy test. A requirement that women take appropriate measures to avoid pregnancy while participating in the trial is a common and appropriate inclusion criterion. Current FDA guidance and regulations in this area are summarized as follows:

⁶² <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm129477.pdf>

⁶³ <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073137.pdf>

⁶⁴ <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm129460.pdf>

- Women should be appropriately represented in clinical trials during all phases of drug development. We expect sufficient representation of both sexes in clinical trials to permit detection of important differences.⁶⁵
- Marketing applications must include assessment of potential differences in drug effectiveness, safety, and dose-response between sexes (see 21 CFR 314.50(d)(5)(v) and (vi)).
- Women (or men) with reproductive potential who have a life-threatening disease or condition cannot be excluded from eligibility for a trial (even an early trial) of an investigational drug intended to treat the disease or condition because of a risk or potential risk of reproductive toxicity (i.e., affecting reproductive organs) or developmental toxicity (i.e., affecting potential offspring) from use of the investigational drug (see 21 CFR 312.42(b)(v)).
- Efforts to prevent pregnancy should persist until reproductive and developmental toxicity studies in animals demonstrate at most a low risk.

7.2.3.1 PK issues regarding women

Possible PK differences between women and men should be assessed, either by formal trials or using population PK methods. PK trials of female subjects should evaluate the following issues, as appropriate:

- Effects of the menstrual cycle (and circulating hormones such as estradiol and progesterone)
- Differences between premenopausal and postmenopausal women, including the effects of hormone replacement therapy and use of systemic contraceptive agents
- Effects of body size (e.g., smaller size), body composition (e.g., higher fat content), and endogenous hormones
- Effects of pregnancy

7.2.3.2 Interactions with oral contraceptives

The influence of a drug on the effectiveness of oral contraceptives generally should be studied by assessing drug blood levels unless interaction data or other information allow the potential for interaction to be ruled out.

⁶⁵ See the guidance for industry *Guideline for the Study and Evaluation of Gender Differences in the Clinical Evaluation of Drugs*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072044.pdf>)

7.2.3.3 Studying pregnant women

The physiological changes that occur in pregnancy can alter the PK profile of many drugs.⁶⁶ Changes include plasma volume expansion, increased cardiac output, increased renal blood flow with increased glomerular filtration rate, decreased gastrointestinal motility, and altered serum protein profiles. The effect of those changes cannot be assessed without PK trials in various stages of pregnancy.

Moreover, there is often a need to use drugs during pregnancy (e.g., for ongoing treatment of a woman who becomes pregnant or for treatment of a significant problem that emerges during pregnancy). Historically, pregnant women have not been included in clinical trials premarketing, but this practice deserves reconsideration. It is almost inevitable that drugs will be used by pregnant women. Drugs directed at diseases present in women of reproductive age should at a minimum be the subject of trials of their PK profile in pregnancy, particularly if the drug is directed at serious or life-threatening diseases.

The timing of such trials in the course of drug development should be based on evaluation of many factors, including (but not limited to):

- The disease that is the target of therapy
- Availability of good alternative therapies for the pregnant patient
- Frequency of the disease or condition in women
- Effect of pregnancy on the course of the disease
- Effect of the disease on pregnancy
- Reproductive toxicity profile of the drug

Whether and when pregnant women should be included in or studied separately in trials of effectiveness and safety also deserves consideration. Questions regarding trials in pregnant women should lead to appropriate clinical consultations with the Maternal Health Team within the Pediatric and Maternal Health Staff. In some cases, consultations from experts outside of the FDA also should be considered.

7.2.4 Elderly Subjects

Although elderly patients comprise a significant portion of the consumer population for drugs, they are often underrepresented in clinical trials. Data collected from elderly subjects can be of particular value because elderly subjects are more likely to have organ impairment, take a larger number of concomitant medications, and be susceptible to certain drug-related toxicities. In general, subjects over 65 years of age are considered elderly for the purposes of data evaluation, but it is particularly important to have data on subjects 75 years of age and older. For drugs intended for a population that includes the elderly, substantial numbers of elderly subjects should be included in trials by the time of

⁶⁶ See the draft guidance for industry *Pharmacokinetics in Pregnancy—Study Design, Data Analysis, and Impact on Dosing and Labeling*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM072133.pdf>)

the marketing application. ICH E7 should be consulted about the inclusion of elderly subjects in clinical trials.⁶⁷ A recent amendment to this guidance strongly emphasizes the need for exposure of patients above 75 years of age. As noted earlier, arbitrary upper age limits for trial entry are almost never justified and should be discouraged.

7.2.5 Racial Groups

The database submitted in a marketing application should reflect usage in a diverse racial population, one reflective of the likely patient mix postmarketing, for potential differences in response to become apparent. Although racial differences in drug effects have not commonly been reported, some have been noted. For example, ACE inhibitors in general have a smaller BP effect in self-identified Black patients than in White patients. Similarly, in a comparison of losartan with atenolol in hypertensives, the Losartan Intervention for Endpoint reduction in hypertension (LIFE) trial showed an overall advantage of losartan in stroke reduction but no suggestion of such an effect on Black patients, who appeared to do better on atenolol.⁶⁸

There also are race-associated differences in CYP450 (e.g., CYP2C6, 2C9, 2C19) metabolic activity. When a sponsor is planning to conduct the majority of the clinical trials overseas, or limit the number of centers, the database may not include sufficient representation of races to assess even large differences in drug effects. Different outcomes in racial groups could reflect racially associated genetic differences but could also reflect a variety of other differences between the populations in intrinsic and extrinsic factors. ICH E5 offers guidance relevant to the wide variety of factors that can influence drug effects, particularly when populations are located in different regions.⁶⁹

The EOP2 meeting with the sponsor is an appropriate time to discuss approaches for obtaining adequate information regarding diverse racial groups. The guidance for industry *Collection of Race and Ethnicity Data in Clinical Trials* recommends a standardized approach for collecting and reporting race and ethnicity information in clinical trials.⁷⁰ Interpretation of race differences in response observed in clinical trials can be greatly enhanced if genetic information is available. DNA sample collection should be encouraged if trials will be conducted in diverse populations or in regions outside of the United States. It should be noted that a diverse trial population may be limited in certain settings, such as a trial conducted primarily in Eastern Europe.

⁶⁷

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073131.pdf>

⁶⁸ See current losartan labeling for further details.

⁶⁹

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM073117.pdf>

⁷⁰

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071596.pdf>

7.2.6 Other Subpopulations of Interest: Genetic, Proteomic, and Concomitant Illness

Although special populations of interest traditionally have been defined on the basis of demographics or in some cases physiological features (e.g., plasma renin activity, systolic versus diastolic function in heart failure, organ dysfunction including renal and hepatic impairment), the rapidly evolving technologies of pharmacogenomics and growing recognition of drug-disease interactions and various risk factors for outcome under the broad heading of *individualization* of therapy is expected to increase efforts to assess the effect of such factors on drug effects. Genetic factors are known to determine how drugs are absorbed, distributed, metabolized, and eliminated and can significantly affect PK, dosing, and drug interactions. Genetic factors can also affect the pharmacologic actions and PD effects of a drug, most strikingly seen up until now in cancer treatments but also for some adverse effects in non-oncologic drugs. We anticipate seeing increasing amounts of data regarding the effect of genetic, proteomic, and other factors on drug effects.

When evaluating such data, reviewers should scrutinize the validation of the submitted test methodologies and the multiplicity of hypotheses that arise when many genes are analyzed. Interpretation of pharmacogenetic data requires evaluation of results in the context of what is known with respect to the clinical pharmacology of the drug, disease biology, and genetic variability. Reviewers are encouraged to consult with the clinical pharmacology reviewers in the Genomics Group, in the Office of Clinical Pharmacology for submissions containing pharmacogenetic or biomarker trial objectives and/or data.

If the reviewer or the sponsor thinks a label might indicate that genetic (or other) testing is essential for the safe and effective use of a drug (e.g., to define the indication, to determine dosing, to identify high-risk patients or to identify likely responders), then the sponsor should be reminded that both the drug and the companion test must, in most circumstances, be approved at the same time. When such testing may be appropriate, reviewers should contact the Office of Combination Products, with a consult to the Center for Devices and Radiological Health, as early as feasible in the drug development and review process. Additional information is available in the draft guidance for industry and Food and Drug Administration staff *In Vitro Companion Diagnostic Devices*.⁷¹

The growing interest in individualization of treatment is also reflected in the increasing examination of effects in population subsets defined by concomitant illness or disease severity in addition to the traditional demographic groups, especially in outcome trials, with resulting so-called *forest plots* appearing in published reports and in drug labeling.

⁷¹ <http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm262292.htm>

7.3 Patient Population Size

7.3.1 Sample Size in Phase 1 Clinical Trials

Most phase 1 clinical trials involve a small number of healthy subjects ($n < 20$). A small cohort is treated at each drug dose and is closely observed for adverse effects.⁷² Doses are generally subsequently escalated according to published, well-known escalation plans. The phase 1 trial should be large enough to identify important dose-limiting toxicities and to identify doses suitable for additional clinical trials. It also should be designed to minimize the likelihood that subjects will be exposed to toxic doses. For drugs in which there is a risk of serious toxicity (e.g., novel compounds, immune stimulating agents, and compounds for which animals are poor predictors of human results), it may be advisable for a single subject to receive a given dose of the test drug followed by a lengthened observation period before additional subjects are exposed or before initiating phase 1 trials.

7.3.2 Sample Size in Phase 2 and Phase 3 Clinical Trials

The sample size needed to meet clinical objectives of both phase 2 and phase 3 clinical trials is generally estimated using statistical methods based on the power of the trial. Power is the ability to demonstrate a specified effect on the primary endpoint with a specified alpha (Type I) error, calculated by making assumptions about event rate (in an outcome trial) and extent of spontaneous improvement (in trials of symptomatic patients). The power of a trial to succeed (i.e., reject the null hypothesis) is typically chosen to be 80 percent (i.e., the planned Type II error is 20 percent). The principal factors that influence the sample size requirements are thus related to the following parameters:

- The selection of the primary endpoint and any secondary or safety endpoint for which a definitive answer is critically important
- Event rate in the control group, for a trial that measures an effect on events
- Treatment effect size to be detected (posited difference between the control and the investigational drug), with respect to primary and important secondary endpoints
- Type I error (alpha error, acceptable rate of false positives)
- Type II error (power, acceptable rate of false negatives)
- Trial drug/control allocation ratio
- Expected rates of dropouts and other protocol deviations

⁷² See the guidance for industry *Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers*.
(<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm078932.pdf>)

- Measurement and other variability (e.g., variance of the outcome measurements)

Many of these variables are fixed. Type I error (alpha error) for most trials is set at 5 percent, two-sided, but it generally would be lower in cases where a single trial is being proposed as providing the evidence for effectiveness. As noted, trials conducted for establishing effectiveness generally are designed with a maximum Type II error of 20 percent (i.e., a minimum power of 80 percent). Allocation ratios generally fall between 1:2 and 2:1 and, if within this range, have little effect on trial size. Given these conventionally accepted limits on Type I and Type II errors, the sample size is driven primarily by the anticipated outcomes in the control arm (e.g., event rate, when events are an endpoint), and by the anticipated magnitude of the difference between the investigational drug and control groups.

Many trials fail to show significant evidence of efficacy despite some positive trends. The positive trends may represent chance or they may reflect a true finding from an underpowered trial. A number of published articles illustrate how clinical trials are frequently undersized because of unrealistic (overoptimistic) assumptions by sponsors about event rates and magnitude of the effect of the investigational drug. In some cases, limited resources on the part of the sponsor can influence these assumptions and lead to an undersized trial.

Reviewers should examine the reasonableness of the proposed sample size in effectiveness trials. In general, we do not accept proposals to allow an unusually high Type I error (e.g., 5 percent one-sided or 10 percent two-sided) in a trial intended to support effectiveness. If there appears to be an inadequate sample size likely to result in an unusually high Type II error rate (perhaps because the effect size used in the power calculations was overly optimistic), this should be discussed with the sponsor. It should be appreciated that sponsors may not insist on as rigorous a response in a phase 2 trial as would be needed in a more definitive trial, a risk a sponsor can decide to take. Sample size is generally calculated based on the primary endpoint. Reviewers should assess whether other critical objectives, such as subgroup hypotheses and safety endpoints or a reasonable estimate of an important secondary endpoint (such as mortality), suggest a larger sample size.

The uncertainty surrounding event rates (e.g., mortality) in an event-driven trial can be dealt with by careful planning. The most straightforward approach is an *event-driven* trial where the enrollment and trial duration are not fixed, rather, enrollment continues or the trial is carried on until the predetermined number of endpoint events is observed. The uncertainty of the ultimately enrolled population size can make this choice a problem for sponsors (e.g., uncertainty about costs, resources, and powering for secondary and safety endpoints). Another approach is to allow sample size adjustment based on observed event rates as data accrue in the trial. This approach may be accomplished in different ways with different implications for trial conduct, and sponsors should plan and specify the procedures to be used.

The simplest situation is where the pooled (overall) event rate is used to determine the final sample size without any unblinding, an approach that does not involve statistical adjustment. Adjustment of sample size by using the event rate observed only in the control arm also can be done with no statistical adjustment, but only if the overall blinding of the trial is unequivocally maintained. However, this is a difficult task because the patients in the control group would need to be identified and this would mean some entity involved with the trial would need to be unblinded.

There are also approaches that adjust for effect size (e.g., where effect size was overestimated so that the trial is underpowered). These approaches are far more complex and involve some alpha (Type 1 error) adjustment. Experienced clinical and statistical review staff should scrutinize such approaches to adjusting sample size based on effect size. For any of these methods, the protocol should specify the relationship between the findings on interim analysis of event rates and the re-calculated sample size, as well as any adjustment to the alpha error caused by the re-estimation of the appropriate sample size.

It is also possible to assess the variance of a trial and adjust sample size upward if it is greater than expected. If the analysis is blinded for treatment assignment, there should be no need for adjustment of the alpha error.

Apart from the potential modifications mentioned, sponsors should be encouraged to specify and justify the trial size before trial initiation. Sponsors may have valid reasons for changing the trial size after a trial has begun (e.g., a new trial showing that assumptions in the original calculation were incorrect, changes in medical practice), but they should be encouraged to submit any changes as a protocol amendment that includes close attention to the potential for the introduction of bias. These amendments should be reviewed by both the clinical and statistical review staff. If the sponsor or any other individuals involved in planning and proposing the change in sample size had access to interim results, such change can create bias. There should be sufficient detail in the protocol to allow the statistical reviewer to reconstruct the computations supporting the proposed sample size.

7.3.3 Total Population Exposure

The precision of estimates of common AE rates and the likelihood of detecting uncommon events associated with a trial drug will be functions of the total number of subjects studied. When reviewing a sponsor's drug development plan, the reviewer needs to assess whether the size of the safety database will be sufficient at the time a marketing application is planned. Of importance during this review process is the total number of subjects who: (1) received the investigational drug; (2) received the drug at or above specific dose levels; and (3) received the drug at or beyond specific durations (e.g., 6 and 12 months). The numbers treated in important subpopulations also should be tabulated.

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*⁷³ and the guidance for industry *Premarketing Risk Assessment*⁷⁴ provide general guidance regarding adequacy of population exposure. Since publication of the ICH guidance, there has been growing consensus at the FDA that the number of subjects suggested in the guidance may be insufficient to ensure safety in some clinical settings. The exposed population needed can be drug-class and disease-specific and this issue needs to be discussed with team leaders and division and office directors. There is also a growing recognition of the importance of studying a demographically diverse population and of avoiding unnecessary exclusion criteria.

8. STATISTICAL ANALYSIS PLANS

Sponsors should be encouraged to include the SAP as part of the protocol, rather than providing it in a separate document, even if the SAP has not been finalized. If the SAP is changed late in the trial, particularly after the data may be available, it is critical for the sponsor to assure the FDA that anyone making such changes has been unaware of the results. Sponsors should be encouraged to describe the methods used to ensure compliance. Additional information on the principles of statistical analyses of clinical trials is available in ICH E9.⁷⁵ The review of the SAP requires close collaboration with the biostatistical reviewer.

8.1 Planned Analyses

Analyses intended to support a marketing application (generally analyses for the phase 3 efficacy trials) should be prospectively identified in the protocol and described in adequate detail. An incomplete description of the proposed analyses in the protocol can leave ambiguity after trial completion in how the trial will be analyzed.

Nonprospectively defined analyses pose problems because they leave the possibility that various statistical methods were tried and only the most favorable analysis was reported. In such cases, the estimates of drug effect may be biased by the selection of the analysis, and the proper correction for such bias can be impossible to determine. Preplanning of analyses reduces the potential for bias and often reduces disputes between sponsors and the FDA on the interpretation of results. The same principles apply to supportive and/or sensitivity analyses. These analyses should be prospectively specified, despite the fact that the results of such analyses cannot be used as a substitute for the primary analysis. If the protocol pertains to a multinational trial, it is important that an analysis of the regional

⁷³

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073083.pdf>

⁷⁴

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072002.pdf>

⁷⁵

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073137.pdf>

differences be prespecified. Clinical reviewers should review these considerations for planned analyses in collaboration with statistical reviewers.

Although detailed prespecification is essential for the primary efficacy analysis, the ability to interpret findings on other outcomes, such as important secondary efficacy endpoints for which a claim might be sought, is also dependent on the presence of a prospectively described analysis plan. Observations of potential interest, termed *descriptive endpoints* because the trial will almost always be underpowered in their respect, may be considered in a trial that is successful on its primary endpoint to further explore consistency in demographic subgroups (e.g., sex, age, and race) or evaluate regional differences in multinational trials. Safety outcomes are also important and should be specified prospectively. They will often not be part of the primary analysis unless the trial was designed to assess such an endpoint. Analyses not prospectively defined will in most cases be considered exploratory; see [section 8.2.2.1](#), Descriptive Analysis, for potential use of such descriptive analyses.

Interim analyses may play an important role in trial design. They present complex issues, including preservation of overall Type I error (alpha spending function), re-estimation of sample size, and stopping guidelines. Plans for interim analyses should be prospectively determined and reviewers should discuss these plans with the statistician. See [section 8.1.3](#), Interim Analysis Plans, for further discussion of these plans.

8.1.1 Adequacy of the Statistical Analysis Plan

When reviewing the SAP, it is critical to consider whether there is ambiguity about the planned analyses. Particular attention should be paid to the primary endpoint and how it will be analyzed. If there are multiple primary endpoints or analyses, the Type 1 error rate should be controlled appropriately. If there is a single primary endpoint, details of the analysis are important. For example, an SAP that defines the primary analysis as a comparison of the time to event between treatment arms leaves open many possibilities, such as the specific analytical approach (e.g., Cox regression, log rank test), whether the analyses will be adjusted for covariates (and which covariates would be included), and the method for this adjustment. Censoring for subjects who drop out of the trial or who are lost to follow-up should be discussed, particularly since dropout may not be random. Post dropout follow-up may have different implications for superiority and noninferiority trials.

Consideration also should be paid to other preplanned analyses, such as secondary endpoint analysis, population subset analysis, regional analysis, and interim analysis. Both clinical and statistical reviewers should collaborate in order to make appropriate recommendations.

When there are possible secondary efficacy endpoints (e.g., different time points, population subsets, different statistical tests, different outcome measures), it is critical to determine how they will be analyzed and their role in the efficacy assessment. In general, secondary analyses are not considered in regulatory decision-making unless there is an effect on the primary endpoint, so that no Type 1 error adjustment is needed

for the primary endpoint. A secondary endpoint intended to represent a trial finding (and thus a possible claim) after success on the primary endpoint should be considered as part of the overall SAP and, if there is more than one of these, a multiplicity adjustment or *gatekeeper* approach may be necessary to protect the Type 1 error rate at a desired level ($\alpha = 0.05$) for such analyses. Positive results in a secondary analysis when the primary endpoint did not demonstrate a statistically significant difference generally will not be considered evidence of effectiveness.

Protection of the overall (family-wide) Type 1 error rate at a desired level ($\alpha = 0.05$) is essential when the protocol has designated multiple hypotheses testing. Examples include efficacy comparisons among multiple doses with respect to primary and secondary endpoints, subpopulation analysis, and regional analysis. Various commonly used statistical procedures can be used for this multiplicity adjustment (e.g., Bonferroni, Dunnett, Hochberg, Holm, Hommel, and gatekeeping procedures), and these procedures will be considered in a multiplicity guidance under development. The proper use of each procedure depends on the priority of the hypotheses to be tested and the definition of a successful trial outcome. The following two examples are illustrative:

- **Example 1.** A placebo-controlled trial with one primary endpoint and three treatment doses (low, medium, and high) is planned. To assess the efficacy of the three doses as compared to placebo, a commonly used hierarchical procedure tests sequentially from high dose to low against placebo, each at $\alpha = 0.05$, until a p-value ≤ 0.05 is not attained for a dose. Significance is then declared for all doses that achieved a p-value ≤ 0.05 .

The Bonferroni correction approach also can be used to share $\alpha = 0.05$ among the three doses and test each one at $\alpha = 0.05/3 = 0.017$. This method will be less efficient than the sequential method, if the effect is likely to be positively associated with dose. The primary analysis could also evaluate all three doses pooled versus placebo (less efficient if the low doses are not effective) or of the two highest doses versus placebo.

- **Example 2.** A placebo-controlled trial of two endpoints, A and B, and three treatment doses (low, medium, and high) is planned. Suppose endpoint A is thought to be more indicative of the true effect than B and so is placed higher in the hierarchy than B. Also, suppose the medium and high doses are hypothesized to be equally effective while the low dose is considered less likely to exhibit significance. Multiple clinical decision rules are designated in hierarchical order to demonstrate the efficacy:
 - Show benefit for each of two higher doses individually compared to placebo with respect to endpoint A and endpoint B
 - Show benefit for the two higher doses pooled compared to placebo with respect to endpoint A or B

- Show benefit for the low dose with respect to endpoint A
- Show benefit for the low dose with respect to endpoint B

Although the Bonferroni, Holm, or Hommel procedure can be applied to test these four hypotheses, a gatekeeper procedure that sequentially tests the four sets of hypotheses in hierarchical order, each at $\alpha = 0.05$, is likely to be more efficient.

How missing data (particularly from dropouts) are handled can profoundly affect trial outcomes. Sponsors should be encouraged to detail how they plan to minimize dropouts and to specify particular methods for assessing data from dropouts. It is not credible to design these analyses once unblinded data are available. Reviewers should consider the best approach for the particular situation, recognizing that such classic methods as last observation carried forward (known as LOCF) can bias a trial for or against a drug, depending on the reasons for attrition, the time course of the disease, and response to treatment. Dropout may not be random, as subjects may drop out of either the new drug or the control therapy for toxicity or for lack of efficacy. Depending on the cause of the dropout, the use of modeling approaches might have advantages.

8.1.2 Reviewing Changes to the Statistical Analysis Plan

Changes to critical elements of the analysis (e.g., the primary endpoint, handling of dropouts) during a trial can raise concerns regarding bias, specifically whether the changes could reflect knowledge of unblinded data. Concerns are inevitably greatest when the change is made late and has an important effect on outcome. In theory, if such changes are unequivocally made blindly (e.g., because of data from other trials or careful reconsideration) they should not pose problems, but the assurance of blinding can be hard to provide. For obvious reasons, changes made with data in hand (but purportedly still blinded) pose the greatest difficulties and are hard to support.

When changes to the original SAP are proposed during the course of conducting the trial, it is critical to determine exactly what information, if any, regarding trial outcomes was available to those involved in proposing the change. Changes made with knowledge of results can introduce bias that can be substantial and impossible to measure. Note that such biases can occur subtly (e.g., the likelihood of adoption of a proposal made by an individual with no knowledge of data can be influenced by the comments or nonverbal communication of an individual who does have such knowledge). Therefore, major protocol changes are not credible if knowledge of interim outcome data is available to any individual who is involved with those planning the change. If there is any potential for such changes, sponsors should be encouraged to describe fully who has had access to data and how the firewalls were maintained, among other information.

After trial data collection is completed, and before unblinding, there is often a blinded data cleanup phase. During that phase, previously unaddressed specific concerns about the data may be identified (e.g., types and amounts of missing data, concomitant therapies), and decisions are often made by the sponsor as to how to address those

concerns. Typically, any changes made during this data cleanup phase should be minor clarifications of the SAP. If more than minor clarifications are made to the SAP, sponsors should be encouraged to submit these changes to the FDA for review as protocol amendments.

8.1.3 Interim Analysis Plans

Interim analysis is a systematic approach to assessing clinical data during the course of a trial. Sponsors should monitor all clinical trials (phases 1, 2, and 3) for subject safety. Interim monitoring of overall trial data efficacy results is needed in only a small subset of randomized, controlled trials. These include trials with mortality or major morbidity as primary endpoints, or trials in high-risk populations in which major safety concerns may not be identified without interim comparisons of important safety outcomes. In such trials, early termination can be considered if interim results demonstrate definitively the superiority of one of the trial groups.

Interim effectiveness analyses should be performed and evaluated by independent experts with no interest in the trial results (i.e., not the sponsor or its steering committee). For trials requiring such interim reviews, DMCs are typically established.⁷⁶ Review staff should be familiar with the issues discussed in this guidance when evaluating sponsor proposals for managing interim review of data. If no DMC is named in an outcome trial, the reviewer should suggest that the sponsor consider a DMC to perform interim analyses.

Adaptive designs, which can lead to a change in sample size, treatment arms, and trial endpoints, should be given special attention and consideration. These changes are usually made based on the knowledge and analysis of the treatment effect in an interim analysis. The rationale for an adaptive design and the proposed plan should be prespecified and reviewed by FDA staff before implementation.⁷⁷ The plan should describe who will be doing the analysis of the data and who will be recommending changes to be made. Moreover, the procedures for ensuring the confidentiality of the data and preserving the integrity of the rest of the trial should be described in detail. Statistical methods that will ensure the preservation of the overall Type I error rate also should be addressed.

8.1.3.1 Confidentiality of interim data

In general, plans for interim analyses should be defined prospectively and include a set of operating procedures to protect the confidentiality of the interim results. Disclosure of

⁷⁶ See the guidance for clinical trial sponsors *Establishment and Operation of Clinical Trial Data Monitoring Committees*.
(<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm124936.pdf>)

⁷⁷ See the draft guidance for industry *Adaptive Design of Clinical Trials for Drugs and Biologics*.
(<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm201790.pdf>)

the interim results to investigators and other trial personnel can be problematic in several ways, such as:

- Knowledge of early trends by investigators and subjects can inhibit continued willingness to enroll, affect the types of subjects subsequently enrolled, or affect use of concomitant therapies
- Knowledge of interim results by the steering committee or sponsor can affect their ability to recommend or implement protocol changes in ways that are free of bias

Therefore, if interim analyses are planned, they generally should be conducted by a DMC that has no potential responsibility for modifying the protocol or for conduct of the trial. Such committees function to enhance subject safety and protect trial integrity.

Information on results of DMC analyses available to the sponsor, those conducting the trial, and/or the FDA should be kept to a minimum (e.g., the data do not suggest that any modifications are needed) and documented. The preferred approach is to not have sponsor statisticians prepare interim results for a DMC because even the best within-sponsor firewalls may fail. If sponsor statisticians do prepare interim results, they should be kept separate from anyone with trial responsibilities and there should be a description of who will have information, what information they will have, what their role will be in the trial, and the procedures that will protect against dissemination of information to others who also have roles in the trial. As noted earlier, knowledge of interim results will lead to an inability to make modifications to the trial without concern about introduction of bias.

8.1.3.2 Stopping rules for an early finding of efficacy or toxicity

If accumulating data are analyzed repeatedly at various time points, with the possibility of stopping the trial, then the likelihood of a false positive result (alpha error) increases with the number of analyses. To preserve the overall Type I error for the primary hypothesis test, appropriate significance levels are calculated for each interim analysis (usually with extremely stringent levels for early analyses, and gradually increasing so that the final analysis is performed at a level close to the nominal level) so that the overall alpha does not exceed a given fixed level (generally 5 percent). Sponsors generally plan the number and timing of interim analyses in advance and should describe them in the SAP. Clinical reviewers should consult with statistical reviewers regarding the SAP. Statistical reviewers should be actively involved in reviewing the sponsor's plans regarding stopping rules.

Sponsors should be encouraged to prospectively describe in the protocol the specific approach to be taken to sequential analysis along with the early termination guidelines to be used at interim analysis (commonly called *stopping rules*). Usually, stopping rules are based on the analysis of the primary efficacy variable, along with a formal rule for alpha spending. In some circumstances, however, decisions about early stopping may be appropriately based on an endpoint that is not the primary endpoint. For example, when the primary efficacy endpoint is a composite of components including death (see [section](#)

[8.2.1.1](#), Composite endpoints) and lesser endpoints (such as the need for revascularization from an AMI), it may be appropriate to stop early for a clear effect on death but not on the primary composite endpoint that may be driven by other outcomes of lesser importance. In such cases, there should be a clear and consistent understanding among the sponsor, the DMC, and the FDA regarding the stopping rules.

Early stopping also can be based on safety outcomes in cases when toxic effects emerge at unexpectedly high or severe levels. In all situations in which early stopping is a possibility, preservation of the Type I error and the protocol for ensuring that the monitoring approach is conveyed in the informed consent process are complex issues requiring consideration.

It is important to consider whether a trial with the proposed stopping rules will be able to provide sufficient information for regulatory decision-making should the trial be stopped early.

8.1.3.3 Risks of early stopping for efficacy

When reviewing plans for interim analyses, reviewers should remember that clinical trials are designed to provide information on both the safety and efficacy of a drug. A drug or biologic can show evidence of efficacy at the interim analysis but not show convincing evidence of safety. A short duration of follow-up and relatively small number of subjects accrued at the time of interim analysis may not allow for adequate evaluation of its safety profile. Other problems can arise when a trial is stopped early. The number of subjects can be inadequate to perform important subgroup analyses that support the primary endpoint. This is especially important when a marketing application might be filed on the basis of a single trial. Additionally, data on important secondary efficacy endpoints, particularly long-term efficacy, can be limited. Furthermore, stopping one trial can have an effect on other ongoing trials.

All these concerns should be balanced against the type of benefit being demonstrated. For example, when a trial shows an early strong survival benefit, the value of clarifying issues regarding other safety factors or effects in subgroups will not outweigh the value of making a lifesaving therapy available as early as possible and the questionable ethical acceptability of continuing to randomize to a treatment leading to a decreased survival. An effect on other endpoints, in contrast, may not represent so clear a choice.

It should also be noted that while the trial is ongoing, the DMC may see data that have not been fully verified or not completely up to date. Sponsors should be advised that before early stopping of a trial, they should move rapidly to assess all current database information and to assess the likelihood that further data evaluation could have a substantial effect on the findings. The risk that the findings will become less impressive as additional data are accumulated and validated is always a concern.

8.1.3.4 Unplanned interim analyses

Unplanned analyses of accumulating data from an ongoing trial to evaluate treatment efficacy can cause the overall Type I error (alpha level), generally fixed at the initial

design stage, to become inflated. Nonetheless, analyses which were not originally planned may be undertaken, such as those recommended by the DMC during routine monitoring.

Recent developments in group sequential methods during the course of a trial allow changes to be made in the number of interim analyses by spending the nominal level of alpha according to a prespecified *use function*. Sponsors should be encouraged to use sequential designs that incorporate this flexibility. However, if the plans for handling an additional analysis are not prespecified, once the analysis has occurred, it may be difficult to determine what result would have led to stopping for efficacy and, therefore, how much alpha was *spent*. This tends to lead to relatively conservative assumptions about what correction to use and can affect interpretability of the trial results. Therefore, plans for potential *unplanned analyses* should be addressed before the trial, if possible, or at least before the analysis. Note that so-called administrative review of the data may need to be considered interim review. Such a practice might be acceptable when there is no possibility for stopping the trial based on the result of the administrative review of the data, but providing this assurance is often an impossible task.

8.1.3.5 Reviewer's role during the trial

Occasionally, a sponsor may contact a division to note that a DMC has proposed early stopping and to ask for FDA advice. The FDA does not determine whether trials should be stopped, but can discuss relevant considerations with the sponsor. Such cases can be complex, with potentially major effect on drug approval, and they should always be considered by division- and office-level supervisors. If the proposal is to stop the trial early for efficacy, the sponsor should be advised of the potential ways in which the early stopping might prevent the results from being sufficient for a marketing application (see [section 8.1.3.3](#), Risks of early stopping for efficacy). Review of the interim results with respect to whether they support efficacy should be avoided; such a review can compromise the FDA's ability to both perform an objective review in the future and to provide further advice later in the trial should the trial not be terminated early.

Sometimes events external to a trial (e.g., failure of a similar trial with the sponsor's drug or a related drug) will cause the sponsor to ask for advice regarding continuing an ongoing trial based on inspection of the interim results. Viewing unblinded data from an ongoing trial should be avoided, because such a practice may impede the ability to evaluate future protocol amendments without bias. The existence of a well-constituted DMC can be useful in such cases, in that the sponsor may ask the DMC to review both the external information and interim results and make recommendations on how to proceed. CDER has an internal DMC advisory group that provides consultation to review staff on all issues related to DMCs.

8.1.4 Intent-to-Treat Analysis

In most trials, some patients do not receive the treatment assigned by randomization or may withdraw early because of poor response, toxicity, improvement or worsening of disease, and other reasons. Early withdrawals raise concern that dropping out may be *nonrandom* (if it is random, there is no concern) or *informative*, such that the results will

be distorted. This is a potential concern in outcome trials, where, for example, a person might drop out because of worsening illness so that his or her impending event is lost to the trial.

Nonrandom or informative dropout may be of concern even if it occurs before the initiation of treatment. The SAP for all outcome trials should include an intent-to-treat (ITT) analysis or an all-patients-with-data analysis.⁷⁸ In an ITT and an all-patients-with-data analysis, patients are not dropped as *unevaluable*. These analyses ensure that the comparability of populations created by randomization is maintained and reduces the risk that bias will be introduced during the trial (dropping patients with an apparent poor prognosis in one arm) or during the analysis (excluding patients from analysis because they are *not evaluable*). These analyses can dilute the observed treatment effect, but such dilution generally is preferable to the potential bias that can result from alternative approaches.

However, there may be circumstances in which an analysis of evaluable patients provides useful secondary information, particularly in determining the magnitude of an effect in patients actively taking a drug. Such analyses are generally of interest in a superiority trial only if the ITT or all-patients-with-data analysis is successful in meeting the prespecified level of statistical significance. For example, it might be important to know the magnitude of the effect of an antihypertensive drug on the BP of patients who had at least 2 weeks of treatment. An analysis of outcomes in patients with known good compliance versus poor compliance may seem desirable. However, a compliance analysis in the Coronary Drug Project demonstrated a profound effect for good versus poor adherence to placebo. Such analyses are rarely performed or credible. The incidence of an adverse effect is best calculated in the population that has taken the drug, and events occurring long after discontinuation can minimize the drug effect through a bias towards the null.

Similarly, in trials designed to establish efficacy of a drug through demonstrating *noninferiority* to the control drug, an ITT analysis that includes all randomized patients who did not take the drug, or who crossed over onto the other therapy, can lead to an incorrect conclusion of no difference because such occurrences decrease the observed treatment effect. For noninferiority trials, the primary efficacy analysis is usually the evaluable patient analysis, with an ITT analysis performed as a secondary analysis.

Sponsors should prospectively define all populations for each analysis wherever possible.

8.2 Endpoints

It is important that endpoint definitions and assessments be used consistently by and across all investigators and all centers of a multicenter clinical trial. The reviewer should

⁷⁸ In an ITT analysis, all patients randomized are included in the group to which they were randomized regardless of therapy received and followed up even if they leave the trial. For symptomatic conditions, this is usually modified to an all-patients-with-data analysis so that all patients who receive any drug are counted and generally data after drop out are not used.

ensure that each endpoint is prospectively defined in the clinical protocol, and that its definition includes the nature, timing, and method of assessment. If the endpoint uses laboratory measurements, a delineation of the assay type and number of samples to be taken should be specified. Clinical endpoint determinations should be done by qualified and trained individuals only. If there is to be a review of endpoints by an external endpoint evaluation committee, the protocol should state what data the investigator needs to collect and define the referral criteria.

8.2.1 Primary Endpoints

The protocol should define the primary endpoint or endpoints. There is considerable confusion about the distinction between primary and secondary endpoints. In general, the primary endpoint is the trial outcome that will be used to show whether the drug has clinically significant beneficial effect. If it does not, in all but the most unusual cases, the trial will not be considered to demonstrate a drug effect and will not support marketing approval. A secondary endpoint, in most cases, cannot be used as evidence of efficacy if the primary endpoint analysis was not clinically and statistically significant. In a trial designed to establish efficacy, a primary endpoint should measure a clinically meaningful therapeutic effect or should have demonstrated ability to predict clinical benefit (e.g., BP as a measure of major cardiac benefit). If subparts H and E (accelerated approval) will be sought, the primary endpoint should be a surrogate endpoint that is *reasonably likely* to predict clinical benefit or an effect on a clinical endpoint other than survival or irreversible morbidity.

Generally, the primary endpoint is used to determine a trial's sample size. Although there is usually only one primary efficacy endpoint specified in a protocol, there are important exceptions, including the following:

- **Winning on any endpoint:** Sometimes the effect of drug on any of two or more endpoints (A, B, or more) may be considered clinically important. In an example where two endpoints, A and B, are described in the protocol, the trial can win on A, on B, or on both. An appropriate adjustment for multiplicity is needed for the win on either endpoint. Various procedures (e.g., Bonferroni, sequential testing, and others), can be used to adjust for multiplicity. Both clinical and statistical reviewers should review the use of multiple primary endpoints to ensure that they are clinically justified and that the statistical procedures used to account for multiplicity are appropriate.
- **Composite endpoints:** In clinical trials with time-to-event endpoints, an effect on either of two endpoints (e.g., death or myocardial infarction (MI)), whichever comes first, may be considered clinically important. In such a case, the primary endpoint is defined as a composite endpoint (see [section 8.2.1.1](#), Composite endpoints). The associated analyses are survival or Cox Regression analyses. If the individual components of the composite endpoint will be tested as well, appropriate adjustments for multiplicity should be made. The type of multiplicity adjustment depends on how closely the endpoints are correlated and how the testing of hypotheses is specified. Choosing the multiplicity correction can be

difficult, because many procedures can be used. Both clinical and statistical reviewers should review the use of multiple primary endpoints to ensure that they are clinically justified and that the statistical procedures used to account for multiplicity are appropriate.

- Co-primary endpoints: To establish effectiveness in some disease areas, the sponsor will have to demonstrate an effect on more than one endpoint (i.e. co-primary endpoints). The effects should be highly correlated. For example, for an indication for the treatment of psychosis or cognitive impairment in Alzheimer's disease, it has been considered necessary to show an effect on both a standard disease-related endpoint as well as a measure of global function. Similarly, it is recognized that an antihypertensive drug should have effects on both systolic and diastolic pressure. These requirements inflate beta or Type II error, although not to any large degree.

A long history of use of a primary endpoint in clinical trials generally supports use of the endpoint for drug approval and precedents applicable to the planned trial should be considered by the reviewer. However, if new information indicates that prior decisions may have been flawed, or that new alternatives better measure clinical benefit, use of a previously used endpoint should be reconsidered. Also, reviewers should remain open to the use of a novel primary endpoint proposed by a sponsor. This situation generally requires discussion at the division and office levels.

An important role of an advisory committee is to consider and comment upon the appropriateness of an endpoint. In the absence of advice from an advisory committee, and if there is doubt about the acceptability of an endpoint (particularly if there are no regulatory precedents), it may be appropriate, preferably before the trial has begun, to consult individual experts for their advice or seek internal discussion.

The following list includes some useful questions about the choice of the primary endpoint:

- Does the endpoint measure a meaningful clinical benefit or (for subpart H and E approval) is it reasonably likely to predict clinical benefit?
- Has the endpoint been used before in other similar trials? Were problems identified in those trials or applications?
- Has an advisory committee provided advice on the endpoint?
- Do other endpoints better measure clinical benefit?
- Why has the sponsor chosen this endpoint if viable alternatives exist?
- Would a composite endpoint be more appropriate?

- Is more than one effect needed to establish effectiveness (i.e., two co-primary endpoints such as those noted above in psychosis and Alzheimer’s disease)?
- Should some safety measures be part of the primary endpoint (e.g., hemorrhagic strokes included in the endpoint for a platelet active drug intended to decrease thrombotic endpoints)?
- Is the SAP appropriate for the chosen endpoint?
- For drugs developed under the animal efficacy rule, is the animal study endpoint related to the desired benefit in humans, which is generally the enhancement of survival or prevention of major morbidity?⁷⁹

In general, it is desirable for the primary endpoint to directly and completely measure the most important clinical benefit that is anticipated from investigational drug therapy. The endpoint also should be practical and there may be a tension between these goals. For example, the ultimate benefit of most interventions for cardiovascular disease (e.g., hypertension, coronary artery disease, heart failure) is improved survival, but demonstrating that benefit may take an unacceptably long time or require impracticably large trials, especially as trials move from the sickest patients (worst heart disease, worst coronary disease history) to less affected patients. Therefore, we may rely on other meaningful clinical endpoints (such as heart attack, stroke, or hospitalization for heart failure) or composites of these endpoints.

In some cases, there can be a credible surrogate endpoint. Thus, we have accepted an effect on BP as a basis for approving antihypertensive drugs because numerous outcome trials with antihypertensive drugs representing diverse pharmacological classes repeatedly have shown beneficial effects on stroke, heart attacks, death, and other outcomes. BP is an established surrogate endpoint that is much more practical to study than is prevention of stroke (see [section 8.2.3](#), Surrogate Endpoints).

In life-threatening diseases such as congestive heart failure (CHF) and cancer, survival can be objectively measured, is obviously clinically meaningful, and is a highly useful endpoint that has been used as a primary efficacy endpoint. When a survival benefit is impractical to demonstrate (e.g., because of the number of patients and time required), other measurements of clinical benefit can be proposed as primary endpoints. For example, for heart failure patients classified as class III-IV by the New York Heart Association’s functional classification system, mortality is so high that survival trials are practical and the first outcome trial of an ACE inhibitor (the previously mentioned CONSENSUS trial), used a survival endpoint successfully in a trial with only 253 patients. In less-ill patients, such trials would be difficult or impossible, so trials have used death plus CHF hospitalizations. In cases where less-direct measures of clinical benefit are used in the presence of high patient morbidity or mortality, it may be

⁷⁹ 21 CFR part 314, subpart I, Approval of New Drugs When Human Efficacy Studies Are Not Ethical or Feasible, or 21 CFR part 601, subpart H, Approval of Biological Products When Human Efficacy Studies Are Not Ethical or Feasible.

especially important to determine whether the beneficial effects measured are counterbalanced by adverse effects on survival or other outcomes.

The primary endpoint can be either a continuous variable (e.g., BP) or a dichotomized variable (e.g., number/proportion of patients reaching goal BP, or number/proportion of patients with a given size fail). In some cases (notably drugs for Alzheimer's disease), labeling has described both the primary endpoint (mean effect on a cognitive scale) and the distribution of individual results without any statistical multiplicity adjustment. Even if the primary endpoint is the continuous variable, it is often clinically informative to include the distribution of results in labeling.

8.2.1.1 Composite endpoints

As previously noted, composite endpoints combine multiple events (e.g., death, hospitalization, MI) into one endpoint. When there are multiple outcomes of importance and the event rate for individual endpoints is low, a composite endpoint can be the best primary endpoint rather than any of the individual endpoints of interest. For example, in drug trials for HIV infection, a variety of opportunistic infections and malignancies have been combined as the composite *AIDS defining event*. Composite endpoints are widely used for a number of reasons. In many outcome trials, there may be too few of the events of greatest interest (e.g., deaths) to allow a trial of reasonable size to be conducted. For example, because deaths are uncommon in patients receiving percutaneous coronary interventions, assessing effects of antiplatelet drugs on survival would require extremely large trials. Therefore, a composite endpoint has been developed (e.g., death plus heart attack plus need for repeat procedure).

In general, the components of a composite endpoint should be of reasonably comparable clinical significance, although it is obvious that other endpoints are almost never as significant as death. Usually, the effect of the intervention on all components is expected to trend in the same direction. An exception is the evaluation of strokes (hemorrhagic and thrombotic) in aspirin trials, where aspirin could affect each type of stroke differently. In this case, adverse and therapeutic effects should be distinguished from each other and should trend in opposite directions. Composite endpoints may be a poor choice when the following conditions apply:

- The endpoint components are of different clinical importance (death and angina episodes).
- The endpoint components occur with different frequencies (in this case, the observed effect is attributable to its effect on the more common endpoint). Note that death is often included, even if it is less frequent, because of its unquestioned clinical importance.

Weighted composite outcomes, in which more significant events count proportionally more, have been suggested and can address some of these problems. Assigning weights is an inherently arbitrary process and, therefore, should be considered cautiously. Ranked analyses, in which outcomes are ranked from best to worst (e.g., full recovery,

disability, death), are not often used, because their meaning can be difficult to explain or justify.

Findings on composite endpoints may be more difficult to communicate than findings on single endpoints. Labeling should show results for each component of the composite result.⁸⁰ Generally, when composite endpoints are used or when scoring systems composed of several components are used as primary endpoints, individual components should be specified as secondary endpoints. Where benefit has unequivocally been confined to a single component, only that component has been the basis for an effectiveness claim. For example, in the LIFE trial comparing losartan to atenolol, a composite endpoint of death, AMI, and stroke showed a benefit of losartan, but all of the benefit was attributable to the stroke component. Thus, only that claim was described in labeling.

8.2.2 Secondary Endpoints

It is well recognized that the primary outcome of a trial should be prespecified and that if there is more than one *primary* outcome (i.e., the hypothesis that the trial is designed to support or reject), the SAP should account for this by *sharing alpha error* between the two or more hypotheses. However, investigators carrying out trials usually have many hypotheses of interest, aside from the primary outcome, and how to perform appropriate analyses and interpret the results of the additional hypotheses is not as well established or recognized. The areas of interest for secondary endpoints include:

- Components of a composite endpoint
- Effects in demographic subsets of the population
- Effects in subsets of different disease severity, different concomitant illness, different background drug therapy
- Long-term effects when the primary endpoint is measured early (e.g., 30-day mortality versus 7-day mortality)
- Safety endpoint in an effectiveness trial
- Additional effects beyond those in the primary endpoint, such as patient-reported outcomes (PROs), additional symptomatic benefits
- Individual doses versus placebo where the primary analysis considers the overall effect or a pooled analysis of more than one dose (e.g., the two highest doses versus placebo)

⁸⁰ See the guidance for industry *Clinical Studies Section of Labeling for Human Prescription Drug and Biological Products — Content and Format*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm075059.pdf>)

In some cases, the secondary endpoints are *exploratory* (i.e., not intended to lead to a conclusion in the current trial), but in others it is hoped that they will support additional conclusions and claims. If this is intended, careful planning is needed. A secondary endpoint generally cannot be used to *salvage* a trial that fails on its primary endpoint. If the primary endpoint is successful, the secondary endpoints as a group generally can share an overall alpha at 0.05, divided as desired among the endpoints. With more than a few such endpoints, the alpha for each will be small and not likely to be achieved. An alternative is a sequential approach, with secondary endpoint #1 tested at $p=0.05$, then if #1 is successful, testing endpoint #2 at $p=0.05$, and so on. In this case, testing stops once a secondary endpoint in the order fails. The order of secondary endpoints for testing should be prespecified in the protocol. Sponsors should be advised that it is important to choose the order based on power and likely correlation with the primary endpoint.

8.2.2.1 Descriptive analyses

Formal secondary analyses are not the only way to explore subgroups of interest. Analyses beyond the primary endpoint are increasingly common; most large trials are presented in publications with so-called *forest plots* showing outcomes in demographic, disease severity, concomitant drug, and concomitant illness subsets. Findings in such subgroups cannot be given the same weight as primary analyses (and cannot overcome the failure of the primary endpoint to show a statistically significant result), but they are used to suggest consistency or lack of it. Such forest plots have appeared in FDA labeling. Demographic analyses of individual trial data and the pooled data in the integrated summary of safety (ISS) and integrated summary of effectiveness (ISE) are explicitly called for in regulations (21 CFR 314.50) and guidance.⁸¹ Other areas of interest are (as noted in [section 8.2.1](#), Primary Endpoints) presentation of the distribution of results (even if a continuous variable was the primary endpoint), and time-to-event displays, commonly displayed as Kaplan-Meier curves, for results of outcome trials and symptomatic conditions.

The distinction between true secondary endpoints and *descriptive* approaches to data is an area of continued discussion.

8.2.3 Surrogate Endpoints

Established surrogate endpoints such as BP, LDL cholesterol, and HbA1c have long been the basis for approval. Surrogate endpoints are attractive because trials using these endpoints can assess effectiveness in much less time and in far fewer patients than trials using actual clinical outcomes.

⁸¹ See the guidances for industry *Guideline for the Study and Evaluation of Gender Differences in the Clinical Evaluation of Drugs* (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072044.pdf>) and *Guideline for the Study of Drugs Likely to be Used in the Elderly* (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072048.pdf>).

A potential problem with even a well-supported surrogate is that drugs may have positive and negative effects in addition to their effect on the surrogate. For example, some antihypertensives have favorable effects apart from BP effects on reducing the risk of heart failure; they could prove superior to drugs lacking this property in some populations. A particular concern would be an adverse effect on the clinical endpoint for which the surrogate is supposed to be predictive.

When a surrogate primary efficacy endpoint is used, sponsors should provide the evidence for its validity as a surrogate for clinical benefit. Validation means that there exists evidence that changes in the surrogate after therapy lead to clinical benefit. The correlation of the surrogate to clinical benefit should be reasonably certain before it is relied upon as the sole basis of approval.

A reduction in BP or serum cholesterol to the normal range has long been accepted as evidence of a clinical benefit. There has been clear evidence since the late 1960s and early 1970s that lowering BP with a wide range of drugs reduces stroke and, to a lesser degree, CV death. Thus, the epidemiological evidence showing a relation of BP elevation to rates of stroke and CV death was supported by clinical intervention data. The benefits of lowering cholesterol, in contrast, although strongly supported by epidemiological and pathophysiological evidence, were not well supported until the 1990s when trials of some of *statins* clearly showed improved survival and decreased rates of AMI.

Epidemiological data and pathophysiological data can sometimes be reasonably persuasive to the acceptance of a surrogate, but the actual benefits of effects on surrogates has been shown for a relatively limited range of therapies, including BP (reduction of stroke, heart attack, and CV death), cholesterol (reduction of death and AMI), CD4 counts and viral loads (long-term survival in AIDS), and complete responses in oncology (long-term disease-free survival).

It is important to remain aware of the possibility that an epidemiologic association of one outcome with a marker does not always predict a benefit from altering the marker, either because the mechanistic connection was misinterpreted or because the treatment had an adverse effect, as well as its effect on the positive surrogate.⁸² For both reasons, the use of surrogate markers is approached with caution and a marker that has proven prognostic value may not predict clinical benefit. A classic example of this is the results of the Cardiac Arrhythmia Suppression Trial in which highly successful suppression of premature ventricular beats (a strong predictor of death postinfarction), by encainide and flecainide did not lead to improved survival and led to a more than two-fold increase in mortality.⁸³ Similarly, the cholesterol drug torcetrapib had a highly unfavorable effect on CV outcome (deaths) despite substantially raising HDL levels. Similar problems can arise with short-term symptomatic benefits. For example, several inotropic agents that

⁸² Katz, R, 2004, Biomarkers and Surrogate Markers: An FDA Perspective, *NeuroRx*, 1, 189–195.

⁸³ Echt, DS et al., 1991, Mortality and Morbidity in Patients Receiving Encainide, Flecainide or Placebo: The Cardiac Arrhythmia Suppression Trial, *New England Journal of Medicine*, 324:781–788.

improved hemodynamic signs and symptoms in CHF proved to increase mortality, not reduce it.^{84,85}

For drugs to treat serious or life-threatening diseases, effects on surrogate markers that are not yet validated and that would not be a basis for traditional approval but are *reasonably likely* to predict clinical benefit can serve as trial endpoints to support accelerated approval under subparts H and E. On occasion, even when an applicant did not request accelerated approval, the submitted data may be inadequate for traditional approval, but we may consider it potentially adequate for accelerated approval. Sponsors should be advised that they must directly measure and confirm clinical benefit in postmarketing trials, after accelerated approval, if the drug is to remain on the market (21 CFR part 314, subpart H, for drugs and 21 CFR part 601, subpart E, for biological products).

8.2.4 Patient-Reported Outcome Measures

PRO instruments are included as endpoints in clinical trials because: (1) some treatment effects are known only to the patient; (2) the patient provides a distinct, important perspective about the effectiveness of a treatment; and (3) systematic assessment of the patient's perspective on (1) and (2) using a PRO instrument often is preferable to obtaining the information filtered through an investigator's evaluation of the patient's response to clinical interview questions.

Endpoints measured by PRO instruments are most often used to assess a patient's symptoms or ability to function. Many PRO instruments are specifically designed not only to assess symptoms but also to examine other possible consequences of treatment (i.e., effects on activities of daily living or psychological state). PRO instruments serve as primary endpoint measures for trials in many conditions where the patient is the only source of information on the effect of a treatment (e.g., pain therapies, treatments targeted at symptom relief). PRO instruments also can augment what is known about a treatment based on the investigator's perspective or physiological measures. Improvements in clinical measures do not always correspond to improvements in how the patient functions or feels (e.g., incomplete correlation of spirometric assessments of lung function with asthma symptoms).

The amount and type of evidence expected to support a claim measured by a PRO instrument is similar to that needed for any other measure of effectiveness. The determination of whether the PRO instrument is an adequate endpoint to establish effectiveness is based on an assessment of the ability of the PRO instrument to measure the claimed treatment benefit and is specific to the intended population and to the characteristics of the condition or disease treated.

⁸⁴ Packer, M et al., 1991, Effect of Oral Milrinone on Mortality in Severe Chronic Heart Failure. The PROMISE Study Research Group, *New England Journal of Medicine*, 325 (21):1468–75.

⁸⁵ Packer, M et al., 1993, Effect of Flosequinan on Survival in Chronic Heart Failure: Preliminary Results of the PROFILE Study (abstract), *Circulation*, 88(suppl 1).

The guidance for industry *Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims* describes the FDA's recommendations for PRO endpoints used in clinical trials,⁸⁶ including recommended processes for development and validation of PRO instruments, issues to consider when incorporating PRO instruments in clinical protocols, and considerations for statistical analysis of PRO endpoints. PRO instruments raise unique challenges for:

- Multinational trials
- Research in pediatric populations
- Research in cognitively impaired populations

8.3 Discussions With the Sponsor

8.3.1 Target Product Profile

The target product profile is a tool currently used in CDER to provide a format for discussions between a sponsor and the FDA of the critical aspects of clinical trial design and analysis. By identifying the effectiveness claims the sponsor hopes to support, the dosing recommendations the sponsor hopes to make, and any comparative statement with respect to other drugs the sponsor hopes to support, the sponsor can design trials to potentially support these statements and identify the trials that do not support these statements.⁸⁷

8.3.2 Continuous Involvement

If possible, the acceptability of the design, endpoints, and analysis of any proposed trial identified by the sponsor as critical to the demonstration of safety and efficacy should be discussed with the sponsor before trial initiation.

If a sponsor does not request review of critical efficacy protocols at an EOP2 meeting, at an ad hoc meeting regarding the protocol, or through an SPA,⁸⁸ the acceptability of the endpoints and other critical features of the protocol should nonetheless be considered and discussed with the sponsor by telephone during the protocol review with other staff participating as appropriate (e.g., team leader, statistics consult).

⁸⁶

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071975.pdf>

⁸⁷ See the draft guidance for industry and review staff *Target Product Profile — A Strategic Development Process Tool*.
(<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm080593.pdf>)

⁸⁸ See the guidance for industry *Special Protocol Assessment*.
(<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm080571.pdf>)

The sponsor should be advised of issues regarding endpoints or other design features that would make a trial potentially not acceptable for the intended regulatory use. If possible, these issues should be resolved before trial initiation. If no agreement can be reached, reviewers should consider whether the design is inadequate to meet stated objectives. However, the adequacy of a protocol and appropriateness of an endpoint can be a matter of judgment. All discussions with the sponsor should be documented.

9. GOOD CLINICAL PRACTICES

Good clinical practice (GCP) is an international ethical and scientific quality standard for designing, conducting, recording, and reporting trials that involve the participation of human subjects. Compliance with this standard provides assurance that the rights, safety, and welfare of trial subjects are protected and that the clinical trial data are credible.

Several sections of FDA regulations, notably 21 CFR parts 50, 54, and 56, and 21 CFR part 312, subpart D, address various aspects of clinical practice. The ICH guidance for industry *E6 Good Clinical Practice: Consolidated Guidance* provides a well-organized discussion of the standards for GCP.⁸⁹ Clinical trials should be conducted in accordance with the ethical principles that are consistent with GCP and applicable regulatory requirements. To help ensure GCP, sponsors should be encouraged to consider the following issues regarding the conduct of a trial:

- Before a trial is initiated, foreseeable risks and inconveniences should be weighed against the anticipated benefit for the individual trial subject and society. A trial should be initiated and continued only if the anticipated benefits justify the risks.
- The rights, safety, and welfare of trial subjects are the most important considerations and should prevail over interests of science and society.
- The available nonclinical and clinical information on an investigational drug should be adequate to support the proposed clinical trial.
- A trial should be conducted in compliance with the protocol that has received prior IRB/IEC approval or favorable opinion.
- The medical care given to, and medical decisions made on behalf of, subjects should be the responsibility of a qualified physician or, when appropriate, of a qualified dentist.
- Each individual involved in conducting a trial should be qualified by education, training, and experience to perform his or her respective tasks.

89

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073122.pdf>

- Freely given informed consent should be obtained from every subject before clinical trial participation, including screening assessments.
- All clinical trial information should be recorded, handled, and stored in a way that allows its accurate reporting, interpretation, and verification.
- The confidentiality of records including name, date of birth, and other personal data that could identify subjects should be protected, respecting the privacy and confidentiality rules in accordance with applicable regulatory requirements, including the Health Insurance Portability and Accountability Act of 1996.
- Investigational drugs should be manufactured, handled, and stored in accordance with applicable current good manufacturing practice. They should be used in accordance with the approved protocol.
- Systems with procedures that ensure the quality of every aspect of the trial should be implemented.

Foreign trials conducted under an IND should comply with U.S. regulations unless a requirement is waived. Because foreign IECs differ in some respects from IRBs, the requirement of an IRB meeting U.S. regulatory specifications is often waived. Trials conducted entirely outside the United States need not be conducted under an IND; but if they are to be relied upon by the United States for regulatory purposes, they should still follow standards for GCPs.⁹⁰

9.1 The Institutional Review Board

The IRB or, in Europe, the IEC, helps to safeguard the rights, safety, and well-being of subjects participating in trials by reviewing the protocol, investigator's brochure, written consent forms, and informed consent process. The IRB is also responsible for continuing review of the progress of the trial, including reviewing protocol amendments and investigator submissions of unexpected problems. For multicenter trials, a centralized IRB review process can be used to improve efficiency and to minimize delays and duplication of efforts.⁹¹

FDA review of a new IND submission may precede IRB approval but a trial cannot proceed until IRB approval is obtained, except in unusual emergency situations (generally involving a single patient) (21 CFR 312.36).

⁹⁰ See 21 CFR 312.120 and the guidance for industry *Acceptance of Foreign Clinical Studies*.
<http://www.fda.gov/downloads/regulatoryinformation/guidances/ucm124939.pdf>

⁹¹ See the guidance for industry *Using a Centralized IRB Review Process in Multicenter Clinical Trials*.
<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm080606.pdf>

Usually, reviewers will not have direct contact with an IRB, although occasionally an IRB will seek information about the investigational drug or about trial design issues. The IRB may be given any information that is not proprietary or confidential and can be referred to the sponsor if additional information is required.

9.2 Informed Consent

All clinical trials conducted under an IND must be conducted in compliance with the informed consent requirements in 21 CFR part 50 and the IRB requirements in 21 CFR part 56 unless waived to allow use of an appropriate IEC for foreign trials. Additional discussion of informed consent can be found in ICH E6.⁹² MAPP 6030.2 *INDs: Review of Informed Consent Documents* describes when an ICD should be reviewed, when CDER should request that an ICD be submitted for review, and the procedures for reviewing an ICD.⁹³ In some circumstances, a consult to the Human Subject Protection Branch of OSI and/or an ethics consultation may be necessary as part of the ICD review process. Adequacy of informed consent must be reviewed and approved by an IRB (21 CFR part 56).

There are situations in which review of the consent form by the FDA, in addition to IRB review, is particularly important to determine whether a clinical investigation may safely proceed under investigational regulations. These situations include the following:

- Unusual toxicity is associated with the investigational drug or the drug class
- The trial population is particularly vulnerable⁹⁴
- The trial design is unusual for the therapeutic class
- CDER has unique knowledge about a particular concern related to a drug area and may be in a better position than the IRB to assess whether the ICD addresses the concern
- The clinical investigation has significant potential for serious risk to human subjects

92

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073122.pdf>

93

<http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/MannualofPoliciesProcedures/ucm082024.pdf>

⁹⁴ Examples of vulnerable categories of subjects include children, prisoners, pregnant women, handicapped or mentally disabled persons, or economically disadvantaged persons (see, for example, 21 CFR 56.111(a)(3)). The regulations do not define what *vulnerable* means specifically. In general, subjects may be considered vulnerable when they have impaired decision-making capacity, an increased susceptibility to undue influence or coercion, or an increased susceptibility to the risks associated with a particular clinical investigation.

- The clinical investigation is a postmarketing safety trial conducted to better characterize the safety profile of a drug
- The clinical investigation involves the first administration of a drug in humans
- The clinical investigation will ask patients to forego or delay effective treatment that is known to decrease long-term mortality or irreversible morbidity⁹⁵

In these cases, the following areas should receive particular attention in the review:

- Description of potential adverse effects
- Discussion of potential benefits
- Discussion of alternative therapies

In the following two circumstances, the ICD should be reviewed by CDER:

- Treatment INDs and treatment protocols (21 CFR part 312, subpart I; see MAPP 6030.6 *INDs: Processing Treatment INDs and Treatment Protocols*⁹⁶)
- Exception from informed consent requirements for emergency research (21 CFR 50.24; see MAPP 6030.8 *INDs: Exception From Informed Consent Requirements for Emergency Research*⁹⁷)

The FDA's role in the review of consent forms is intended to complement the role of the IRB, not replace it, because FDA review staff may have a unique perspective not always available to the IRB. For practical reasons, the general consent form being developed, rather than the form approved at each site, is reviewed and each modification made to a consent form generally is not reviewed. The division director should communicate his or her expectations regarding informed consent review.

9.2.1 Consent in Pediatric or Other Vulnerable Populations

Informed consent in pediatric trials includes additional elements (21 CFR part 50, subpart D). Informed consent must be provided by the minor's legally authorized representative. However, to the extent that a minor child is capable of understanding the trial, the minor

⁹⁵ See ICH E10.

(<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073139.pdf>)

⁹⁶

<http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/MannualofPoliciesProcedures/UCM283082.pdf>

⁹⁷

<http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDER/MannualofPoliciesProcedures/ucm082027.pdf>

should be asked to give assent and should sign and personally date the consent form if capable of doing so.

Similarly, for adult subjects with mental or physical conditions that prevent them from giving valid consent, consent must be obtained from the legally authorized representative. If the subject is capable of understanding the trial, assent should be obtained from the subject.

Note should also be made of subjects who may be considered vulnerable by virtue of their membership in a group with a hierarchical structure such as students in biomedical fields, employees in the pharmaceutical industry, members of the armed forces, and persons kept in detention. Participation in a clinical trial may be influenced by the perception that it will result in some benefit and that refusal to participate may be met with retaliation by senior members of the hierarchy. Additional vulnerable subjects include those with incurable disease, and those who are homeless or economically disadvantaged.⁹⁸

9.2.2 Waiver of Informed Consent

The FDA has regulations and guidance defining conditions and specific procedures under which the requirement for informed consent may be waived.⁹⁹ The regulations permit access to experimental therapies and conduct of acute care research in situations in which the patient's condition is such that he or she cannot provide legally effective informed consent and there is no legally authorized representative to give informed consent before the treatment must be initiated.

Because the absence of informed consent is such a critical matter, there are a variety of additional protections such as community discussion (a potential benefit for patients), public disclosure of the proposed trial before its initiation, and adequate dissemination of the results upon completion of the trial associated with waiver under 21 CFR 50.24. When faced with a proposal to use the waiver provision, the relevant policy and guidance should be reviewed and the issue should be discussed with a supervisor.

9.3 Investigator's Brochure

The investigator's brochure is an important tool ensuring that investigators are familiar with the use of the investigational drug as required by GCP and regulations (21 CFR 312.23). The investigator's brochure should be assessed to ensure that it is presented in an objective, balanced, and nonpromotional format that provides complete information relevant to the safe and appropriate use of the drug. Sponsors should be encouraged to

⁹⁸ See the ICH guidance for industry *E6 Good Clinical Practice: Consolidated Guidance* (section 1.61, Vulnerable Subjects).
(<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073122.pdf>)

⁹⁹ See 21 CFR 50.24 and the guidance for institutional review boards, clinical investigators, and sponsors *Exception from Informed Consent Requirements for Emergency Research*.
(<http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM249673.pdf>)

update the investigator's brochure on a regular basis with pertinent information obtained from ongoing trials, such as safety data, PK and PD properties (especially for subpopulations that are outliers), changes in drug formulations, and addenda to the protocol.

9.4 Investigator Qualifications and Responsibilities

In general, the investigator's qualifications should be examined. The investigator should be qualified by education, training, and experience to assume responsibility for the proper conduct of the trial,¹⁰⁰ and should provide evidence of such qualifications through up-to-date curriculum vitae or other relevant documentation. Typically the clinical investigator should be a physician or dentist with appropriate expertise.

9.5 Trial Monitoring and Auditing

Both clinical trial monitoring and auditing are quality assurance efforts and are described in ICH E6.¹⁰¹

Monitoring refers to a set of oversight procedures conducted by the sponsor or its agent (e.g., a contract research organization), intended to ensure that the trial is properly conducted and documented in accordance with the protocol, GCP, good laboratory practice, and applicable regulatory requirements. Monitoring activities largely take place during a trial and commonly are performed both centrally and at the clinical sites. When noncompliance with the protocol, GCP, or the regulations is found, the monitor and sponsor should promptly secure compliance, discontinue shipments of the test article, or end the clinical investigator's participation in the trial (21 CFR 312.56).

Sponsors are responsible for ensuring that a trial is adequately monitored. In many traditional pharmaceutical trials, monitors visit each site before, during (on a regular schedule or after a predetermined number of subjects are enrolled), and after the trial, to assess protocol adherence and other procedures and check clinical data entries against the original source data. The amount of monitoring is flexible and can depend on the extent and nature of investigator training, how much is known about the drug, and the nature of the trial. It is not realistic, for example, to expect monthly monitoring of every site in a 5,000 subject trial nor does it appear to be necessary for the endpoints in those trials.

It is often valuable to consider and discuss in advance with sponsors the adequacy of the proposed scheme of monitoring for important trials. This is particularly true when the monitoring scheme or trial setting is *not* one that has been used extensively or with proven success.

¹⁰⁰ See

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073122.pdf> and 21 CFR 312.53.

¹⁰¹

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073122.pdf>

Factors to be considered in assessing the adequacy of the nature and extent of monitoring include the objective, design, complexity, blinding, size, and endpoints of the trial; the training and experience of investigators; and the ability of central monitoring to ensure quality. In general, trials that are less susceptible to bias (e.g., trials with simple entry criteria and procedures), have objective outcomes (particularly binary results such as death versus survival), and a high level of blinding may require less monitoring. Additionally, trials conducted by a group with established standard operating procedures (SOPs) and a track record of generating strong data may require less monitoring. Note that intensive monitoring may be more critical in noninferiority trials (where obscured differences, carelessness, and poor performance can lead to a false conclusion of equivalence) than in superiority trials, where it is in a sponsor's interest to ensure quality.

The source records should contain documentation of the original investigator assessments as well as any changes made by the medical monitor to these assessments. Corrective action may be warranted when a pattern of repeated errors is observed, when an observed error is likely to be repeated, or when the risk to subjects is significantly increased.

Auditing should be carried out after the trial is completed and performed by agents of the sponsor. The purpose of auditing, which should be independent of trial conduct and monitoring, is to evaluate the trial conduct; evaluate compliance with the protocol, SOPs, GCPs, and regulatory requirements; and verify the reliability of the data. See ICH E6 for additional guidance on monitoring and auditing.¹⁰² See the guidance for industry *Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products* for guidance about when less extensive monitoring may be acceptable.¹⁰³

10. ADVERSE DRUG EXPERIENCES AND REPORTS

10.1 Safety Monitoring

Adequate safety monitoring is critical in all clinical trials. Subjects in clinical trials should be adequately observed for AEs, and any such events should be appropriately collected, analyzed, and reported.

When there are limited data on the safety of the investigational drug (i.e., early in development), it is especially important that an extensive array of clinical and laboratory assessments be performed frequently. As safety data accumulate, the nature and extent of safety monitoring should be adjusted accordingly.

¹⁰²

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073122.pdf>

¹⁰³

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072008.pdf>

The critical components of safety monitoring are appropriate trial design, monitoring of AEs by investigators and the sponsor, and review of potentially important findings by the sponsor and the FDA. The critical elements of safety monitoring are thus partly the design of clinical trials and partly what is done with emerging data. Reviewers should consider the following list of important elements when assessing a safety monitoring plan:

- Is the duration of clinical observation adequate with respect to the stated objectives and endpoints, the anticipated response to drug, and the health-related conditions being studied? For a chronically used drug, ICH E1A needs to be considered,¹⁰⁴ but that guidance notes circumstances suggesting a need for larger databases.
- Is there a need for prolonged observation of the subject in a hospital or other closely monitored setting following initial dosing with the drug?
- Is the interval between clinic or hospital visits appropriate?
 - Is there a need for more frequent observation during the first week following initial dosing?
 - Are there provisions for more frequent clinic visits for subjects found to have developed AEs or laboratory abnormalities?
 - Are intervals in later stages of development realistic (i.e., less frequent) with respect to clinical use?
 - Is follow-up of subjects with an AE long enough to detect the course of the event?
- Are the laboratory test data to be collected appropriate and adequate?
 - Do they include routine assessment of all organ systems?
 - Are they sufficiently detailed and complete for organs more likely or known to be affected by the drug?
 - Do they include all relevant tests for the drug class under study (e.g., bicarbonate and chloride for all drugs known to inhibit carbonic anhydrase)?
 - Are there stopping rules for subjects whose laboratory test abnormalities reach a certain threshold?

104

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073083.pdf>

- Are clinical evaluations appropriate and adequate?
 - Are subjects to be seen regularly by trained medical personnel?
 - Are evaluations sufficiently detailed and complete for systems more likely to be affected by the drug (e.g., drowsiness for drugs with antihistaminic effects) and for adverse effects of particular concern such as suicidality with antidepressants and orthostatic hypotension with vasoactive drugs?
 - Do evaluations include all relevant tests for the drug class under study (e.g., assessment of sexual dysfunction for selective serotonin reuptake inhibitors)?
 - Are phone calls to subjects to be made between clinic or hospital visits? If so, are scripted interviews available for such calls?
- Are clinical signs or symptoms of AEs likely to be associated with the drug outlined in sufficient detail in the patient monitoring plan?
- Are the CRFs clear and complete? Do they allow for collection of information on AEs of particular interest, as well as on unexpected serious AEs? Is there room for further description (narratives) for events of particular interest, such as those leading to change in therapy (e.g., discontinuation, dose reduction, or new treatment) or death?
- Does the sponsor have a plan for assessing the accumulating data? In large outcome trials, this generally involves analysis of important events by a DMC, but even in smaller trials or groups of trials, a systematic review of important AE reports is useful. This review can be performed by either the sponsor or an outside expert.

10.2 Reporting Requirements for Sponsors

The investigator and sponsor have primary responsibility for monitoring the safety of subjects given investigational drugs, but certain potentially important AEs must be reported promptly to the FDA for evaluation and the sponsor must report more broadly on safety annually.

These reports allow reviewers to bring a different perspective as well as broad experience with related drugs to bear on consideration of the safety of the investigational drug. The regulations in 21 CFR 312.32 specify a sponsor's obligations for evaluating and reporting AE data for INDs and in 20 CFR 320.31 for sponsors conducting BA and BE trials that are exempt from IND requirements (see also [sections 10.3 through 10.7](#), IND Safety Reports — Written Reports through Other Safety Data).

The FDA's final rule for safety reporting requirements clarifies the type of safety information drug sponsors must report to the FDA for INDs. For more in-depth details on reporting requirements and expected IND safety reports, reviewers should consult the draft guidance for industry and investigators *Safety Reporting Requirements for INDs and BA/BE Studies*.¹⁰⁵ It is important to note that the sponsor must promptly review all information relevant to the safety of the drug obtained or otherwise received by the sponsor from any source, foreign or domestic, including information derived from clinical investigations, animal investigations, commercial marketing experience, reports in the scientific literature, and unpublished scientific papers.

A sponsor's principal reporting requirements for AEs are as follows:

- Serious and unexpected AEs associated with the use of the investigational drug must be reported in writing as an IND safety report to the FDA no later than 15 calendar days after the sponsor's initial receipt of this information ([see section 10.3](#), IND Safety Reports — Written Reports).
- Fatal or life-threatening unexpected experiences for which there is a possibility that the experience may have been caused by the drug must be reported by the sponsor to the FDA by telephone or facsimile transmission no later than 7 calendar days after receipt of this information.
- A summary of all IND safety reports, including a discussion of the most frequent and most serious AEs, must be submitted to the FDA each year in the annual report. This report must be submitted within 60 days of the anniversary of the date the IND went into effect. [Section 2.6](#), IND Safety Reports (21 CFR 312.32(c)), should be consulted after receiving AE data to ensure that the sponsor has submitted all relevant information and has taken appropriate action (e.g., has modified the dose or instituted new monitoring). If the sponsor voluntarily suspends a trial based on new safety data, the division should consider placing the IND on clinical hold to prevent further development until safety issues are adequately addressed.

10.3 IND Safety Reports — Written Reports

Under 21 CFR 312.32, all AEs associated with the use of the drug that are both serious and unexpected require an IND safety report. Findings from tests in laboratory animals that suggest a significant risk for human subjects also require an IND safety report.

Individual reports of *disease-related* events, often related to events that form the clinical trial endpoint, are of little value and, if anything, diminish the likelihood of recognizing unexpected drug toxicity. Such events should be monitored by the sponsor or a DMC, but individual reports should not be submitted to the FDA. Therefore, for some trials,

¹⁰⁵

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM227351.pdf>

particularly in large trials with mortality or major morbidity endpoints and in which a DMC is monitoring for safety concerns, the sponsor should be encouraged to identify in the protocol (or other document submitted to the IND) adverse consequences of the disease that will be monitored in the trial, and compared for treatment groups by the sponsor or external group (DMC or other), but not reported as individual safety reports. The FDA can accept such modification of usual procedures under 21 CFR 312.32. A sponsor who repeatedly files IND safety reports that aren't required can be reminded of the possibility of prospectively identifying AEs that would not be reported individually.

If the reviewer becomes aware that IND safety reports have not been filed as required (e.g., by seeing in the sponsor's annual reports, marketing applications, or FDA inspectional reports events for which IND safety reports should have been filed but were not), he or she should consider the need for educational or compliance actions.

For more information on safety reports, see the guidance for clinical investigators, sponsors, and IRBs *Adverse Event Reporting to IRBs — Improving Human Subject Protection*.¹⁰⁶

The definitions of terms used to determine the need for an IND safety report are found in 21 CFR 312.32 and are summarized in the following subsections.

10.3.1 Definition: Adverse Event

Adverse event means any untoward medical occurrence associated with the use of a drug in humans, whether or not considered drug-related.

An AE (also referred to as an adverse experience) can be any unfavorable and unintended sign (e.g., an abnormal laboratory finding), symptom, or disease temporally associated with the use of a drug, without any judgment about causality. An AE can arise from any use of the drug (e.g., off-label use, use in combination with another drug) and from any route of administration, formulation, or dose, including an overdose.

10.3.2 Definition: Adverse Reaction

An adverse reaction means any AE caused by a drug. Adverse reactions are a subset of all suspected adverse reactions for which there is reason to conclude that the drug caused the event.

For the purposes of prescription drug labeling, the term *adverse reaction* is defined to mean “an undesirable effect, reasonably associated with use of a drug, that may occur as part of the pharmacological action of the drug or may be unpredictable in its occurrence. This definition does not include all adverse events observed during use of a drug, only those adverse events for which there is some basis to believe there is a causal relationship

106

<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm079753.pdf>

between the drug and the occurrence of the adverse event.” (See 21 CFR 201.57(c)(7) and 201.80(g).)

10.3.3 Definition: Suspected Adverse Reaction

Suspected adverse reaction means any AE for which there is a reasonable possibility that the drug caused the AE. For the purposes of IND safety reporting, *reasonable possibility* means there is evidence to suggest a causal relationship between the drug and the AE. A suspected adverse reaction implies a lesser degree of certainty about causality than adverse reaction, which means any AE caused by a drug.

Some examples of a reasonable possibility include:

- A single occurrence of an event that is uncommon and known to be strongly associated with drug exposure (e.g., angioedema, hepatic injury, Stevens-Johnson Syndrome)
- One or more occurrences of an event that is not commonly associated with drug exposure, but is otherwise uncommon in the population exposed to the drug (e.g., tendon rupture)
- An aggregate analysis of specific events observed in a clinical trial (such as known consequences of the underlying disease or condition under investigation or other events that commonly occur in the trial population independent of drug therapy) that indicates those events occur more frequently in the drug treatment group than in a concurrent or historical control group

10.3.4 Definition: Unexpected

An AE or suspected AE is considered unexpected if it is not listed in the investigator’s brochure or is not listed at the specificity or severity that has been observed; or, if an investigator’s brochure is not required or available, is not consistent with the risk information described in the general investigational plan or elsewhere in the current application. For example, under this definition, hepatic necrosis would be unexpected (by virtue of greater severity) if the investigator’s brochure referred only to elevated hepatic enzymes or hepatitis. Similarly, cerebral thromboembolism and cerebral vasculitis would be unexpected (by virtue of greater specificity) if the investigator’s brochure listed only cerebral vascular accidents.

Unexpected, as used in this definition, also refers to AEs or suspected AEs that are mentioned in the investigator’s brochure as occurring with a class of drugs or as anticipated from the pharmacological properties of the drug, but are not specifically mentioned as occurring with the particular drug under investigation.

10.3.5 Definition: Serious

An AE or suspected AE is considered serious if, in the view of either the investigator or sponsor, it results in any of the following outcomes:

- Death
- A life-threatening AE
- Inpatient hospitalization or prolongation of existing hospitalization
- A persistent or significant incapacity or substantial disruption of the ability to conduct normal life functions
- A congenital anomaly or birth defect

Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered serious when, based upon appropriate medical judgment, they may jeopardize the patient or subject and may require medical or surgical intervention to prevent one of the outcomes listed in this definition. Examples of such medical events include allergic bronchospasm requiring intensive treatment in an emergency room or at home, blood dyscrasias or convulsions that do not result in inpatient hospitalization, or the development of drug dependency or drug abuse.

10.3.6 Definition: Life-Threatening

An AE or suspected adverse reaction is considered life-threatening if, in the view of either the investigator or sponsor, its occurrence places the patient or subject at immediate risk of death. It does not include an AE or suspected adverse reaction that, had it occurred in a more severe form, might have caused death.

10.4 Assessment of Adverse Drug Reports

Reports of serious unexpected events are submitted to the FDA and investigators so that any steps needed to protect subjects can be taken. The sponsor and the investigator should have taken any needed action, but FDA reviewers bring special knowledge (e.g., general experience, information on related drugs) as well as a distinct viewpoint and contribute importantly to the evaluation.

The sponsor must report in an IND safety report any suspected adverse reaction that is both serious and unexpected. Before submitting this report, the sponsor needs to ensure that the event meets all three of the definitions contained in the requirement:

- Suspected adverse reaction
- Serious
- Unexpected

If the AE does not meet all three of the definitions, it should not be submitted as an expedited IND safety report.

To assist sponsors with determining whether an AE meets the definition of suspected adverse reaction, the requirement under 21 CFR 312.32(c)(1)(i) specifies that sponsors are to report to the FDA only if there is evidence to suggest a causal relationship between the drug and the AE and provides examples of such evidence, described below.

- **Individual occurrences (21 CFR 312.32(c)(1)(i)(A)).** Certain serious AEs are informative as single cases because they are uncommon and are known to be strongly associated with drug exposure. Some examples include angioedema, blood dyscrasias, rhabdomyolysis, hepatic injury, anaphylaxis, and Stevens-Johnson syndrome. The occurrence of even one case of such AE meets the definition of suspected adverse reaction (i.e., that there is a reasonable possibility that the drug caused the event).
- **One or more occurrences (21 CFR 312.32(c)(1)(i)(B)).** A single occurrence, or a small number of occurrences, of a serious AE that is uncommon in the trial population but not commonly associated with drug exposure may also be informative. If the event occurs in association with other factors strongly suggesting causation (e.g., strong temporal association, event recurs on rechallenge), a single case may be sufficiently persuasive to report in an IND safety report. Often, more than one occurrence from one or multiple trials would be needed before the sponsor could determine that there is a reasonable possibility that the drug caused the event. Examples include tendon rupture or heart valve lesions in young adults, or intussusception in healthy infants.
- **Aggregate analysis of specific events (21 CFR 312.32(c)(1)(i)(C)).** Certain serious AEs can be anticipated to occur in the trial population independent of drug exposure. Such events include known consequences of the underlying disease or condition under investigation (e.g., symptoms, disease progression) and events unlikely to be related to the underlying disease or condition under investigation, but common in the trial population independent of drug therapy (e.g., CV events in an elderly population). An example of the former would be a non-acute death observed in a trial in cancer patients. An example of the latter would be an acute MI observed in a long-duration trial in an elderly population with cancer. In some investigations, serious AEs that are consequences of the underlying disease may be clinical trial endpoints (e.g., mortality or major morbidity).

As noted, individually persuasive events or subsequent analyses of grouped events can affect the conduct of the trial, including stopping the trial, modifying trial monitoring, changing the consent form, and modifying the investigator's brochure.

10.5 Actions After Reviewing Adverse Drug Reports

IND safety reports of serious and unexpected AEs that show a potential causal relationship to the drug deserve close attention. After initially collecting all available relevant information, relevant team leaders and supervisors should be informed and consideration given to: what, if any, additional information should be obtained; whether to inform other reviewers and supervisors within the FDA dealing with the same or

similar drugs; and what actions to take regarding the IND and related INDs. The general guidance in [section 2.6](#), IND Safety Reports (21 CFR 312.32(c)), contains a list of actions that should be considered.

If the safety of continuing the trial becomes uncertain, or if there is insufficient information with which to accurately assess a serious AE, the reviewer should consider an immediate decision to recommend to the division director that he or she place the trial on clinical hold. Above all, it is imperative to work closely with the sponsor to obtain all relevant information to appropriately address any safety issues and protect subjects. In addition, consulting the Office of Surveillance and Epidemiology staff may be helpful, particularly if the investigational drug is related to approved drugs in the same class.

10.6 Annual Reports

Annual reports submitted by sponsors should be reviewed. Under 21 CFR 312.33, sponsors must submit in the annual report a brief report on the status of ongoing or completed trials conducted during the previous year, including the title of the trial; the patient population; the total number of subjects enrolled to date tabulated by sex, race, and age group; the number of subjects who completed the trial as planned; the number of subjects who dropped out of the trial for any reason; and a brief description of any available results. In addition, sponsors must include the following in the annual report (21 CFR 312.33):

- A narrative or tabular summary, by body system, of the most frequent adverse reaction and most serious adverse reaction
- A list of all subject deaths with cause of death
- A list of all subject dropouts in association with AEs, whether or not they are thought to be related to the investigational drug
- A summary of all IND safety reports submitted during the previous year
- A brief description of what was learned about the drug's actions (e.g., dose-response or bioavailability)
- A list of nonclinical studies (including animal studies) completed or in progress during the past year, and a summary of the major nonclinical findings
- A summary of any significant manufacturing or microbiological changes made during the past year
- A description of the general investigational plan for the coming year
- Any revised investigator's brochure

- Any significant phase 1 protocol modifications made in the previous year and not reported in an amendment
- A brief summary of significant foreign marketing developments (e.g., approval or withdrawal of marketing within any country)
- A log of any outstanding business with respect to the IND for which the sponsor requests or expects a reply, comment, or meeting

Reviewers should consider whether any of the actions described under [section 10.5](#), Actions After Reviewing Adverse Drug Reports, are appropriate based on the new safety data.

Review of the annual report is a particularly important and valuable tool for a reviewer who is newly assigned responsibility for an ongoing IND file.

10.7 Other Safety Data

Findings that suggest a significant risk generally arise from ongoing or completed clinical trials, pooled data from multiple trials, epidemiological studies, and published and unpublished scientific papers. Findings from clinical trials that are subject to this requirement are those that have not already been reported under 21 CFR 312.32(c)(1)(i). For example, any clinically important finding from a drug interaction trial or from a trial evaluating QT interval would be reported under this provision.

Findings from animal studies (such as carcinogenicity, mutagenicity, teratogenicity, or reports of significant organ toxicity at or near the expected human exposure), are examples of the types of findings that could suggest a significant risk. Before reporting a finding to the FDA, the sponsor should use judgment to determine whether the finding suggests a significant risk in humans or is too preliminary to interpret without replication or further investigation.

11. EXPEDITED DRUG DEVELOPMENT PROGRAMS

The regulations in 21 CFR part 312, subpart E, describe approaches to expediting the availability of new therapies to patients with serious conditions with unmet medical needs, while maintaining standards for safety and effectiveness. The Food and Drug Administration Modernization Act of 1997 created the fast track designation program under newly added section 506 of the FD&C Act. Title VIII of FDASIA amended section 506 of the FD&C Act to provide fast track designation to any drug that has been designated as a qualified infectious disease product under section 505E(d).¹⁰⁷ FDASIA

¹⁰⁷ Title VIII of FDASIA, GAIN, provides incentives for the development of antibacterial and antifungal drugs intended to treat serious and life-threatening infections. Under GAIN, a sponsor may be granted a QIDP designation for a drug that meets the criteria outlined in the statute. A drug that receives a QIDP designation is eligible for fast track designation and, upon submission of an NDA or supplement for that designated use, will receive a priority review. Upon approval of an application for a QIDP, a 5-year extension will be added to any exclusivity granted with that approval.

also created section 506(a) of the FC&C Act to provide for the designation of a drug as a breakthrough therapy if intended to treat a serious or life-threatening condition, and preliminary clinical evidence indicates the potential for a substantial improvement over existing therapy on clinically significant endpoint(s).

The FDA's expedited drug development programs are designed to facilitate the development and expedite the review of new drugs and biologics that are intended to treat serious or life-threatening conditions and that demonstrate the potential to address unmet medical needs.

These expedited program designations apply to a combination of the drug and the specific indication for which it is being studied. The indication includes both the condition for which the drug is intended (e.g., amyotrophic lateral sclerosis) and the anticipated or established benefits of use (e.g., improved time to ventilation or improved survival). Such a program is referred to as a fast track drug development program or a breakthrough drug development program.¹⁰⁸ Therefore, it is the development program for a specific drug for a specific indication that will receive the expedited drug development designation, not the drug itself.

For more information about the FDA's expedited drug development programs, including qualifying criteria and features, see relevant provisions of the FD&C Act and regulations, and see the draft guidance for industry *Expedited Programs for Serious Conditions — Drugs and Biologics*.¹⁰⁹ The FD&C Act, the regulations, and the guidance, when finalized, will all take priority over this document.

11.1 Serious or Life-Threatening Condition

The judgment of whether or not a condition is serious generally is based on such factors as survival, day-to-day functioning, or the likelihood that the disease, if left untreated, will progress from a less severe condition to a more serious one. For a condition to be serious, it should be associated with morbidity that has substantial effect on day-to-day functioning. Short-lived and self-limiting morbidity usually will not be sufficient, but the morbidity need not be irreversible, providing it is persistent or recurrent.¹¹⁰ Thus, in making a recommendation for a fast track or breakthrough determination, the medical

¹⁰⁸ See the guidance for industry *Fast Track Drug Development Programs — Designation, Development, and Application Review*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm079736.pdf>)

¹⁰⁹ Some language in this document also appears in the draft guidance. The FDA is currently reviewing public comments on this draft guidance. When finalized, this guidance will represent the FDA's current thinking. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM358301.pdf>)

¹¹⁰ The preamble to the rule "Restricted Marketing Under Accelerated Approval" discusses this further. (<http://www.fda.gov/ohrms/dockets/ac/00/backgrd/3627b2bm.pdf>)

reviewer must assess whether the development program is designed to demonstrate an effect on a serious aspect of the serious or life-threatening condition.

Specific examples of when the potential for a new drug (to avoid the serious sequelae of existing drugs) would qualify for expedited drug development designation have been provided in guidance.¹¹¹ Many conditions not generally considered to be serious have rare or distant serious sequelae (e.g., urinary tract infections or duodenal ulcers). Drug development programs for such conditions could be considered for expedited drug development programs if sponsors were to specifically design the development program to demonstrate an effect on those serious sequelae.

Conversely, some conditions that are serious have nonserious manifestations requiring symptomatic therapy (e.g., insomnia associated with schizophrenia, skin discoloration from Addison's disease, alopecia with lupus, subcutaneous nodules from rheumatoid arthritis). Review staff should not generally recommend designation as a fast track or breakthrough development program for a drug whose effect is to be measured in terms of nonserious manifestations unless the drug's effect on those manifestations is reasonably likely to predict benefit on a serious manifestation.

11.2 Available Therapy

Both fast track and breakthrough drugs must provide benefit that is potentially better than available therapy. *Available therapy* generally should be interpreted as a therapy that:

- Is approved or licensed in the United States for the same indication being considered for the new drug; and
- Is relevant to current U.S. standard of care (SOC) for the indication.

Approval or licensure. Only in rare cases will a treatment that is not approved for the indicated use or not FDA-regulated (e.g., surgery) be considered available therapy. In those cases, the reviewer should consider whether the therapy is supported by compelling evidence, including evidence in the published literature (e.g., certain established oncologic treatments), when making the available therapy determination.

U.S. standard of care. For a given condition, there may be a substantial number of approved therapies with varying relevance to how the disease is currently treated in the United States, including therapies that are no longer used or are used rarely. A reviewer generally should focus only on treatment options that reflect the current SOC for the specific indication (including the disease stage) for which the drug is being developed. When determining the current SOC, reviewers should consider recommendations by authoritative scientific bodies (e.g., National Comprehensive Cancer Network, American Academy of Neurology) based on clinical evidence and other reliable information that

¹¹¹ See the draft guidance for industry *Expedited Programs for Serious Conditions — Drugs and Biologics*. (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM358301.pdf>)

reflects current clinical practice. In the absence of a well-established and documented SOC, reviewers should consult with their supervisors to decide if consultation with special government employees or other experts may be warranted.

Over the course of new drug development, it is foreseeable that the SOC for a given condition may evolve (e.g., because of approval of a new therapy or new information about available therapies). The reviewer should consider what constitutes available therapy at the time of the relevant regulatory decision for each expedited program the sponsor intends to use (e.g., generally early in development for fast track and breakthrough therapy designations, at time of BLA or NDA submission for priority review designation, during BLA or NDA review for accelerated approval).

A drug approved under the accelerated approval regulations (21 CFR part 314, subpart H, or 21 CFR part 601, subpart E), based on a surrogate or clinical endpoint, is not considered available therapy. Drugs can also be approved under those regulations with restricted distribution and drugs can be approved with a risk evaluation and mitigation strategy (REMS) that includes elements to assure safe use (ETASU) under section 505-1 of the FD&C Act. Those approved drugs should be considered an available therapy only if the trial population for the new drug would be eligible to receive the approved drug under a restricted distribution program or an ETASU REMS.

11.3 Demonstrating the Potential to Address Unmet Medical Need

An unmet medical need is a medical condition whose treatment or diagnosis is not addressed adequately by existing therapy. An unmet medical need includes an immediate need for a defined population (i.e., to treat a serious condition with no or limited treatment) or a longer term need for society (e.g., to address the broad problem of development of resistance to antibacterial drugs).

If no therapy exists for a serious condition, there is clearly an unmet medical need and a new treatment intended to be effective in that condition would meet this aspect of the criteria for use of expedited programs.

When available therapy exists for a condition, a new treatment generally would be intended to address an unmet medical need in the following situations:

- Effect on serious outcomes of the condition that are not known to be influenced by available therapy (e.g., progressive disability in multiple sclerosis when the available therapy has shown an effect on symptoms but has not shown an effect on progressive disability)
- Improved effect on serious outcome(s) of the condition compared to available therapy (e.g., superiority of the new drug used alone or in combination with available therapy in an active- or historically controlled trial assessing an endpoint reflecting mortality or serious morbidity)

- Ability to benefit patients who are unable to tolerate available therapy or whose disease has failed to respond to available therapy, or an ability to be used effectively with other critical agents that cannot be combined with available therapy
- Ability to provide benefits similar to those of available therapy, while: (1) avoiding serious toxicity that occurs with available therapy; (2) avoiding less serious toxicity that is common and causes discontinuation of treatment of a serious condition; or (3) reducing the potential for harmful drug interactions
- Ability to provide benefits similar to those of available therapy but with documented improvement in some factor, such as compliance, that is expected to lead to an improvement in serious outcomes
- Ability to address an emerging or anticipated public health need, such as a drug shortage

In some disease settings, a drug that is not shown to provide the kinds of direct efficacy or safety advantage over available therapy described previously may nonetheless provide an advantage that would be of sufficient public health benefit to qualify for fast track development or accelerated approval. In diseases with limited numbers of therapies with modest efficacy and significant heterogeneity in patient response to treatment, a drug that appears to have a comparable risk-benefit profile but with a novel mechanism of action could be determined to have the potential to provide an advantage over available therapy. In such cases, the novel mechanism of action should have a well-understood relationship to the disease pathophysiology.

In addition, there should be a reasonable basis for concluding that a significant number of patients may respond differently to the new drug compared to available therapy. For example, mechanistic diversity, even without a documented efficacy or safety advantage, could be advantageous in disease settings in which drugs become less effective or ineffective over time, as occurs in infectious diseases and oncology. In these settings, infectious disease drugs or targeted cancer therapies with novel mechanisms of action, although appearing to have comparable efficacy across the disease population, could benefit patients who no longer respond to available therapy.

As discussed in [section 11.2](#), Available Therapy, reviewers may consider drugs to potentially address unmet medical need notwithstanding the availability of therapies approved under the accelerated approval regulations.

11.4 Qualifying Criteria for Expedited Drug Development Designations

11.4.1 Fast Track

The type of information needed to demonstrate the *potential* of a drug to address unmet medical need depends on the drug development stage. Before beginning human trials,

sponsors should be encouraged to design a development plan to assess the potential for a drug to address unmet medical need based on pharmacological and animal model data. At this stage, there may be little evidence of effectiveness of the drug in humans and the potential will be largely theoretical. For later fast track designation, but still before the completion of the principal controlled trials, available clinical data should begin to confirm or be consistent with the potential to address unmet medical need. Still later in the drug's development, we will normally consider whether the clinical data from controlled and uncontrolled trials, as summarized by the sponsor, support the potential of the drug to address unmet medical need.

11.4.2 Breakthrough Therapy

Unlike the information that could support fast track designation, which could include theoretical rationale, mechanistic rationale (based on nonclinical data), or evidence of nonclinical activity, breakthrough therapy designation requires preliminary clinical evidence of a treatment effect that would represent substantial improvement over available therapies for the treatment of a serious condition. Assessment of the treatment effect for the purposes of breakthrough therapy designation will be based on preliminary clinical evidence, which could include early clinical evidence of both clinical benefit and an effect on a mechanistic biomarker (generally derived from phase 1 and phase 2 trials). Nonclinical information could support the clinical evidence of drug activity. In all cases, preliminary clinical evidence demonstrating that the drug may represent a substantial improvement over available therapy should involve a sufficient number of patients to be considered credible. However, the FDA recognizes that the data cannot be expected to be definitive at the time of designation.

Ideally, preliminary clinical evidence would be derived from a study that compares the investigational drug to an available therapy (or placebo, if there is no available therapy) in clinical testing and shows superiority, or from a study that compares the new treatment plus SOC to the SOC alone. The FDA should encourage sponsors to obtain some preliminary comparative data of this kind early in development. Other types of clinical data that also could be persuasive include studies comparing the new treatment with historical experience (generally, the FDA expects such data would be persuasive only if there is a large difference between the new treatment and historical experience).¹¹²

Approaches to demonstrating preliminary clinical evidence of substantial improvement include:

- Direct comparison of a new drug to available therapy (or to no treatment if none exists) showing a much greater or more important response (e.g., complete response where the control treatment results in partial response). Such a trial could be conducted in treatment-naïve patients or in those whose disease failed to respond to available therapies either as a comparison with the failed therapy (if ethically acceptable) or as a no-treatment controlled study.

¹¹² Sponsors contemplating the use of historical controls should consult the ICH guidance for industry *E10 Choice of Control Group and Related Issues in Clinical Trials* for more detailed discussions. (<http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm129460.pdf>)

- The new drug added to available therapy results in a much greater or more important response compared to available therapy in a controlled study or to a historical control. This trial also could be conducted in treatment-naïve patients or in those whose disease failed to respond to available therapies.
- The new drug treats the underlying cause of the disease, in contrast to available therapies that treat only symptoms of the disease, and preliminary clinical evidence shows significant efficacy. In this case, the treatment effect is entirely new (i.e., has not been observed with available therapies). For example, a drug that targets a defective protein that is the underlying cause of a disease (whereas current therapies only treat the symptoms of the disease).
- The new drug reverses disease progression, in contrast to available therapies that only provide symptomatic improvement.
- The new drug has an important safety advantage that relates to serious adverse events compared to available therapies and has similar efficacy.

Clinically Significant Endpoint. For purposes of breakthrough therapy designation, the FDA considers *clinically significant endpoint* generally to refer to an endpoint that measures an effect on irreversible morbidity or mortality (IMM) or on symptoms that represent serious consequences of the disease. It can also refer to findings that suggest an effect on IMM or serious symptoms, including:

- An effect on an established surrogate endpoint
- An effect on a surrogate endpoint or intermediate clinical endpoint considered reasonably likely to predict a clinical benefit (i.e., the accelerated approval standard)
- An effect on a PD biomarker(s) that does not meet criteria for an acceptable surrogate endpoint, but strongly suggests the potential for a clinically meaningful effect on the underlying disease
- A significantly improved safety profile compared to available therapy (e.g., less dose-limiting toxicity for an oncology agent), with evidence of similar efficacy

In a breakthrough therapy designation request, the FDA should encourage sponsors to provide justification for why the endpoint, biomarker, or other findings should be considered clinically significant.

11.5 Features of Expedited Drug Development Designations

11.5.1 Features of Fast Track

There is an opportunity for frequent interactions with the review team for a fast track drug. These include FDA-sponsor meetings, including pre-IND, end-of-phase 1, and EOP2 meetings to discuss study design, extent of safety data required to support approval, dose-response concerns, use of biomarkers, and other meetings as appropriate (i.e., to discuss accelerated approval, the structure and content of an NDA, and other critical issues).

In addition, such a drug could be eligible for priority review if supported by clinical data at the time of BLA, NDA, or efficacy supplement submission.

If the FDA determines, after preliminary evaluation of clinical data submitted by the sponsor, that a fast track drug may be effective, the FDA shall evaluate for filing, and may consider reviewing portions of a marketing application before the sponsor submits the complete application.

11.5.2 Features of Breakthrough

A sponsor of a drug designated as a breakthrough drug is entitled to the same features of fast track designation. In addition, the following features apply.

- **Intensive guidance on an efficient drug development program, beginning as early as phase 1**

Breakthrough therapy designation usually means the effect of the drug is large compared to available therapies. In such cases, the development program for the breakthrough therapy could be considerably shorter than for other drugs intended to treat the disease being studied. However, the FDA notes that a compressed drug development program still must generate adequate data to demonstrate that the drug is safe and effective to meet the statutory standard for approval.¹¹³ Omitting components of the drug development program that are necessary for such a determination can significantly delay, or even preclude, marketing approval.

CDER staff should be prepared to provide special attention so that the development program for a breakthrough drug is as efficient as possible. Reviewers may suggest, or a sponsor can propose, alternative clinical trial designs (e.g., adaptive designs, an enrichment strategy, use of historical controls) that may result in smaller trials or more efficient trials that require less time to complete. Such trial designs could also help minimize the number of patients exposed to a potentially less efficacious treatment (i.e., the control group treated with available therapy). Timely review of sponsor submissions is critical, with appropriate feedback.

¹¹³ Section 505(d) of the FD&C Act; section 351(a) of the Public Health Service Act.

The review team and the sponsor should meet throughout drug development to address important issues at different phases of development.

- **Organizational commitment involving senior managers**

Senior managers and experienced review staff are expected to conduct a proactive collaborative, cross-disciplinary review. A cross-disciplinary project lead for the review team generally will be assigned to facilitate an aggressive timeline and efficient review of the development program. The cross-disciplinary project lead will serve as a scientific liaison among the members of the review team (i.e., clinical; pharmacology/toxicology; chemistry, manufacturing, and controls; compliance; biostatistics) for coordinated internal interactions and coordinated communications with the sponsor through the review division's regulatory health project manager.

If a sponsor has not requested breakthrough therapy designation, a reviewer, after consultation with the appropriate supervisors, may suggest that the sponsor consider submitting a request if: (1) after reviewing submitted data and information (including preliminary clinical evidence), the FDA thinks the drug development program may meet the criteria for breakthrough therapy designation; and (2) the remaining drug development program can benefit from the designation.

12. REFERENCES¹¹⁴

Best Pharmaceuticals for Children Act (reauthorized under Title V, section 502, of the Food and Drug Administration Amendments Act of 2007, Public Law 107-109) (<http://www.fda.gov/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticAct/FDCAct/SignificantAmendmentstotheFDCAct/FoodandDrugAdministrationAmendmentsActof2007/FullTextofFDAAALaw/default.htm>). The BPCA was made permanent in 2012 with passage of FDASIA (<http://www.gpo.gov/fdsys/pkg/BILLS-112s3187enr/pdf/BILLS-112s3187enr.pdf>).

CDER Pediatric Drug Development Web page.
<http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DevelopmentResources/ucm049867.htm>

The Consensus Trial Study Group, 1987, Effects of Enalapril on Mortality in Severe Congestive Heart Failure, *New England Journal of Medicine*, 316;1429–1435.

Draft guidance for industry *Adaptive Design Clinical Trials for Drugs and Biologics*.

Draft guidance for industry *Animal Models — Essential Elements to Address Efficacy Under the Animal Rule*.

¹¹⁴ The draft guidances listed here, when final, will represent the FDA's current thinking on the respective topics.

Draft guidance for industry *Clinical Pharmacogenomics: Premarketing Evaluation in Early Phase Clinical Studies*.

Draft guidance for industry *Drug Interaction Studies — Study Design, Data Analysis, Implications for Dosing, and Labeling Recommendations*.

Draft guidance for industry *Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products*.

Draft guidance for industry *Expedited Programs for Serious Conditions — Drugs and Biologics*.

Draft guidance for industry *General Considerations for Pediatric Pharmacokinetic Studies for Drugs and Biological Products*.

Draft guidance for industry *Integrated Summary of Effectiveness*.

Draft guidance for industry *Non-Inferiority Clinical Trials*.

Draft guidance for industry *Pediatric Oncology Studies in Response to a Written Request*.

Draft guidance for industry *Pediatric Study Plans: Content of and Process for Submitting Initial Pediatric Study Plans and Amended Pediatric Study Plans*.

Draft guidance for industry *Pharmacokinetics in Pregnancy — Study Design, Data Analysis, and Impact on Dosing and Labeling*.

Draft guidance for industry *Qualification Process for Drug Development Tools*.

Draft guidance for industry and Food and Drug Administration staff *In Vitro Companion Diagnostic Devices*.

Draft guidance for industry and investigators *Safety Reporting Requirements for INDs and BA/BE Studies*.

Draft guidance for industry and review staff *Target Product Profile — A Strategic Development Process Tool*.

Drug Development and Drug Interactions Web site.

<http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DevelopmentResources/DrugInteractionsLabeling/ucm080499.htm>

Echt, DS et al., 1991, Mortality and Morbidity in Patients Receiving Encainide, Flecainide or Placebo: The Cardiac Arrhythmia Suppression Trial, *New England Journal of Medicine*, 324:781-788.

The Food and Drug Administration Safety and Innovation Act of 2012.
<http://www.gpo.gov/fdsys/pkg/BILLS-112s3187enr/pdf/BILLS-112s3187enr.pdf>

Gelfand, JA, RJ Sherins, DW Alling, and MM Frank, 1976, Treatment of Hereditary Angioedema With Danazol – Reversal of Clinical and Biochemical Abnormalities, *New England Journal of Medicine*, 295:1444–1448.

Guidance for clinical investigators, sponsors, and IRBs *Adverse Event Reporting to IRBs — Improving Human Subject Protection*.

Guidance for clinical trial sponsors *Establishment and Operation of Clinical Trial Data Monitoring Committees*.

Guidance for industry *Acceptance of Foreign Clinical Studies*.

Guidance for industry *Antibacterial Drug Products: Use of Noninferiority Trials to Support Approval*.

Guidance for industry *Bioavailability and Bioequivalence Studies for Orally Administered Drug Products — General Considerations*.

Guidance for industry *Clinical Studies Section of Labeling for Human Prescription Drug and Biological Products — Content and Format*.

Guidance for industry *Collection of Race and Ethnicity Data in Clinical Trials*.

Guidance for industry *Content and Format of Investigational New Drug Applications (INDs) for Phase 1 Studies of Drugs, Including Well-Characterized, Therapeutic, Biotechnology-derived Products*.

Guidance for industry *Developing Medical Imaging Drug and Biological Products, Part 3: Design, Analysis, and Interpretation of Clinical Studies*.

Guidance for industry *Drug-Induced Liver Injury: Premarketing Clinical Evaluation*.

Guidance for industry *End-of-Phase 2A Meetings*.

Guidance for industry *Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers*.

Guidance for industry *Exposure-Response Relationships — Study Design, Data Analysis, and Regulatory Applications*.

Guidance for industry *Fast Track Drug Development Programs — Designation, Development, and Application Review*.

Guidance for industry *Food-Effect Bioavailability and Fed Bioequivalence Studies*.

Guidance for industry *Guideline for the Study and Evaluation of Gender Differences in the Clinical Evaluation of Drugs*.

Guidance for industry *Guideline for the Study of Drugs Likely to be Used in the Elderly*.

Guidance for industry *IND Meetings for Human Drugs and Biologics; Chemistry, Manufacturing, and Controls Information*.

Guidance for industry *INDs for Phase 2 and Phase 3 Studies; Chemistry, Manufacturing, and Controls Information*.

Guidance for industry *Nonclinical Safety Evaluation of Pediatric Drug Products*.

Guidance for industry *Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims*.

Guidance for industry *Premarketing Risk Assessment*.

Guidance for industry *Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products*.

Guidance for industry *Special Protocol Assessment*.

Guidance for industry *Using a Centralized IRB Review Process in Multicenter Clinical Trials*.

Guidance for industry, investigators, and reviewers *Exploratory IND Studies*.

Guidance for institutional review boards, clinical investigators, and sponsors *Exception from Informed Consent Requirements for Emergency Research*.

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*.

ICH guidance for industry *E2A Clinical Safety Data Management: Definitions and Standards for Expedited Reporting*.

ICH guidance for industry *E3 Structure and Content of Clinical Study Reports*.

ICH guidance for industry *E4 Dose-Response Information to Support Drug Registration*.

ICH guidance for industry *E5 Ethnic Factors in the Acceptability of Foreign Clinical Data*.

ICH guidance for industry *E6 Good Clinical Practice: Consolidated Guidance*.

ICH guidance for industry *E7 Studies in Support of Special Populations: Geriatrics*.

ICH guidance for industry *E9 Statistical Principles for Clinical Trials*.

ICH guidance for industry *E10 Choice of Control Group and Related Issues in Clinical Trials*.

ICH guidance for industry *E11 Clinical Investigation of Medicinal Products in the Pediatric Population*.

ICH guidance for industry *E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs*.

ICH guidance for industry *S9 Nonclinical Evaluation for Anticancer Pharmaceuticals*.

Influence of Adherence to Treatment and Response of Cholesterol on Mortality in the Coronary Drug Project, 1980, *New England Journal of Medicine*, 303:1038–1041.

Katz, R, 2004, Biomarkers and Surrogate Markers: An FDA Perspective, *NeuroRx*, 1 (2):189–195.

MAPP 6010.3 Rev. 1, *Good Review Practice: Clinical Review Template, Attachment B: Clinical Safety Review of an NDA or BLA*.

<http://www.fda.gov/downloads/aboutfda/centersoffices/officeofmedicalproductsandtobacco/cder/manualofpoliciesprocedures/ucm080121.pdf>

MAPP 6030.2 *INDs: Review of Informed Consent Documents*.

MAPP 6030.6 *INDs: Processing Treatment INDs and Treatment Protocols*.

MAPP 6030.8 *INDs: Exception From Informed Consent Requirements for Emergency Research*.

NIH Women's Health Initiative, 2004, *JAMA*, 291:1701-1712.

<http://gateway.ut.ovid.com/gw1/ovidweb.cgi>

Packer, M et al., 1991, Effect of Oral Milrinone on Mortality in Severe Chronic Heart Failure. The PROMISE Study Research Group, *New England Journal of Medicine*, 325 (21):1468–75.

Packer, M et al., 1993, Effect of Flosequinan on Survival in Chronic Heart Failure: Preliminary Results of the PROFILE Study (abstract), *Circulation*, 88(suppl 1).

Pediatric Research Equity Act of 2007 (reauthorized under Title IV, section 402, of the Food and Drug Administration Amendments Act of 2007)

(<http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DevelopmentResources/ucm049867.htm>). PREA was made permanent in 2012 with passage of FDASIA

(<http://www.gpo.gov/fdsys/pkg/BILLS-112s3187enr/pdf/BILLS-112s3187enr.pdf>).

Ridker, P et al., 2008, Rosuvastatin to Prevent Vascular Events in Men and Women With Elevated C Reactive Protein, *New England Journal of Medicine*, 359:2195–2207.

Steering Committee of the Physician's Health Study Research Group, 1989, Final Report on the Aspirin Component of the Ongoing Physician's Health Study, *New England Journal of Medicine*, 321(3):129–35.

The Veterans Administration Cooperative Study Group on Antihypertensive Agents, 1967, Effects of Treatment on Morbidity in Hypertension: Results in Patients with Diastolic Blood Pressure Averaging 115-129 mmHg, *Journal of the American Medical Association*, 202:1028–1034.

21 CFR part 314, subpart I, Approval of New Drugs When Human Efficacy Studies Are Not Ethical or Feasible, or 21 CFR part 601, subpart H, Approval of Biological Products When Human Efficacy Studies Are Not Ethical or Feasible.

13. GLOSSARY OF ACRONYMS

ACE	angiotensin converting enzyme
AE	adverse event (whether drug related or not)
ADR	adverse drug report
AMI	acute myocardial infarction
AUC	area under the curve
BA	bioavailability
BE	bioequivalence
BLA	biologics license application
BP	blood pressure
BPCA	Best Pharmaceuticals for Children Act
CD ₄	cluster of differentiation 4
CDER	Center for Drug Evaluation and Research
CHF	congestive heart failure
CRF	case report form
CV	cardiovascular
CYP450	cytochrome p450
DMC	data monitoring committee ¹¹⁵
ECG	electrocardiogram
EOP2	end-of-phase 2
ETASU	elements to assure safe use
FD&C	Federal Food, Drug, and Cosmetic Act
FDASIA	Food and Drug Administration Safety and Innovation Act of 2012
GAIN	Generating Antibiotic Incentives Now
GCP	good clinical practice
GHB	gamma-hydroxybutyric acid
GRP	good review practice
HER-2	human epidermal growth factor receptor 2
HNSTD	highest nonseverely toxic dose
ICD	informed consent document ¹¹⁶
ICH	International Conference on Harmonisation
IEC	independent ethics committee
IMM	irreversible morbidity or mortality
IND	investigational new drug application
IRB	institutional review board
ISE	integrated summary of effectiveness
ISS	integrated summary of safety
ITT	intent to treat
LIFE	Losartan Intervention for Endpoint reduction in hypertension trial
LOCF	last observation carried forward
MI	myocardial infarction

¹¹⁵ DMCs also can be called data and safety monitoring boards or data and safety monitoring committees.

¹¹⁶ The term *consent form* is also used to refer to the ICD.

NDA	new drug application
NOAEL	no observed adverse effect level
OATOP	organic anion transporting polypeptide
OSI	Office of Scientific Investigations
PD	pharmacodynamic
PeRC	Pediatric Review Committee
PK	pharmacokinetic
PREA	Pediatric Research Equity Act
pre-IND	pre-investigational new drug application
PRO	patient-reported outcome
PSP	pediatric study plan
QIDP	qualified infectious disease product
REMS	risk evaluation and mitigation strategy
SAP	statistical analysis plan
STD	severely toxic dose
SOC	standard of care
SOP	standard operating procedure
SPA	special protocol assessment
UGT	uridine diphosphate glucuronosyltransferase