
Testicular Toxicity: Evaluation During Drug Development Guidance for Industry

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

**October 2018
Clinical/Medical
Pharmacology/Toxicology**

Testicular Toxicity: Evaluation During Drug Development Guidance for Industry

Additional copies are available from:

*Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10001 New Hampshire Ave., Hillandale Bldg., 4th Floor
Silver Spring, MD 20993-0002*

*Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353; Email: druginfo@fda.hhs.gov
<https://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm>*

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

**October 2018
Clinical/Medical
Pharmacology/Toxicology**

TABLE OF CONTENTS

I. INTRODUCTION..... 1

II. DIFFICULTIES EVALUATING TESTICULAR TOXICITY IN HUMANS 2

III. NONCLINICAL EVALUATION..... 2

A. Introduction..... 2

B. Nonclinical Study Design Considerations 3

C. Nonclinical Findings That Raise Concern for Male Fertility 4

D. Confounding Factors 5

E. Follow-Up Investigations..... 5

IV. MONITORING OF THE TESTES DURING CLINICAL TRIALS..... 6

**V. DESIGN OF A CLINICAL TRIAL TO EVALUATE THE EFFECT OF A DRUG
ON THE TESTES 7**

A. Subject Selection 7

B. Trial Design 8

C. Presentation of Results 9

D. Conclusions..... 11

Testicular Toxicity: Evaluation During Drug Development Guidance for Industry¹

This guidance represents the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA office responsible for this guidance as listed on the title page.

I. INTRODUCTION

The purpose of this guidance is to assist sponsors who are developing drug products that may have potential adverse effects on the testes, which we refer to as *testicular toxicity*, based on findings in nonclinical studies.² This guidance discusses the following topics:

- Nonclinical findings that suggest risk of clinical testicular toxicity, and further nonclinical assessments that may be necessary to evaluate the extent of this risk
- Clinical monitoring that can be employed when these drug products are initially administered to men
- The design of a clinical trial that has as its primary purpose the evaluation of drug-related testicular toxicity

The guidance provides general considerations for when clinical trials to assess the risk of testicular toxicity may be needed but does not cover all possible scenarios that would prompt such a trial. The guidance also does not discuss the regulatory actions that FDA might consider based on the results of the clinical trials.

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

¹ This guidance has been prepared by the Division of Bone, Reproductive, and Urologic Products in the Center for Drug Evaluation and Research at the Food and Drug Administration.

² For the purposes of this guidance, all references to *drugs* and *drug products* include both human drugs and therapeutic biological products unless otherwise specified.

II. DIFFICULTIES EVALUATING TESTICULAR TOXICITY IN HUMANS

A thorough evaluation of a drug product's testicular toxicity in humans is challenging for the following reasons:

- Only a few clinical markers can reliably monitor potential changes in human testicular function that might accompany drug exposure. Examples of measurements of testicular function include semen analysis, serum testosterone concentrations, and serum gonadotropin concentrations.
- Monitoring for adverse testicular effects in humans in real time presents a challenge because a latency period of several months exists between the time of an injury to seminiferous tubules and the time when that injury can be detected using the most commonly used test: semen analysis.
- The ability to interpret changes from baseline in the previously mentioned measurements of testicular function and to correlate those changes with effects on male fertility is limited, short of extreme findings.

Conducting a trial assessing male fertility that uses pregnancy rate as an outcome is neither practical nor feasible. Thus, the main outcome measures of a clinical trial assessing testicular toxicity in men are semen parameters. This guidance provides information on the design and conduct of such a trial.

Sponsors of anticancer drugs that fall under the scope of the International Conference on Harmonisation guidance for industry *S9 Nonclinical Evaluation for Anticancer Pharmaceuticals* should consult with the Office of Hematology and Oncology Products before initiating follow-up studies evaluating testicular toxicity.³

III. NONCLINICAL EVALUATION

A. Introduction

Nonclinical evaluation of the male reproductive system is a standard component of the nonclinical safety assessment during drug development. Evidence of adverse drug-related findings on the male reproductive system in animals, specifically the accumulating evidence in appropriate species, informs whether there is a need for an evaluation of testicular toxicity in men. Testicular toxicity is routinely assessed using:

- Repeat-dose toxicology studies with 2 to 4 weeks of drug exposure in two species, unless only one species is studied based on pharmacological relevance

³ We update guidances periodically. To make sure you have the most recent version of a guidance, check the FDA guidance page at <https://www.fda.gov/RegulatoryInformation/Guidances/default.htm>.

Contains Nonbinding Recommendations

- Assessment of male fertility in rodents (when applicable)

Additional information may come from embryo/fetal reproductive and developmental toxicity studies and fertility assessment after prenatal, neonatal, or juvenile exposure. Sponsors should also consider known class effects and/or potential target-related effects.

B. Nonclinical Study Design Considerations

The sponsor should provide a rationale for the choice of doses, duration of exposure, and species used to investigate male reproductive toxicity in nonclinical studies. All studies should include a control group of animals, and early safety evaluation programs should employ both a rodent and nonrodent species, based on pharmacological relevance. The objective of early preclinical safety studies is to define pharmacological and toxicological effects using conventional approaches where immature animals are routinely used at study initiation. Unless studies are intended to support dosing in pediatric patients, sponsors should consider the limitations of using sexually immature animals in subchronic toxicity studies because histology findings in immature animals may incorrectly suggest that fertility is impaired.⁴

For chronic studies where nonhuman primates are the only relevant species based on pharmacology, the sponsor should ensure when assessing reproductive toxicity that male animals have attained sexual maturity by the end of the study. However, this is not always feasible. If impaired spermatogenesis is observed in maturing males, the sponsor may need to further investigate potential effects in fully mature males, on a case-by-case basis. FDA encourages sponsors to discuss with the review division appropriately designed reproductive toxicity assessments in nonhuman primates.

FDA considers histopathological evaluation of the reproductive organs of sexually mature animals to be an appropriately sensitive endpoint for evaluating testicular injury in animals. Toxicology studies should include an examination of the histopathology of the testes, seminal vesicle, epididymis, and prostate with appropriate fixation and staining of the testes.⁵ If adverse findings in gonadal tissues were observed in the repeat-dose toxicity studies, histopathology assessment of the reproductive tissues in the nonclinical male fertility study/studies may provide additional evidence for the human risk assessment. The persistence versus the reversibility of adverse effects in a group sacrificed after a specified period of drug withdrawal (sufficient to

⁴ We support the principles of the 3Rs (reduce/refine/replace) for animal use in testing when feasible. FDA encourages sponsors to consult with review divisions when considering a nonanimal testing method believed to be suitable, adequate, validated, and feasible. FDA will consider if the alternative method could be assessed for equivalency to an animal test method.

⁵ Creasy DM, Chapin RE, 2013, Male Reproductive System. In: Haschek WM, Rousseaux CG, and Wallig MA, editors, Haschek and Rousseaux's Handbook of Toxicologic Pathology, 3rd ed., London: Academic Press, 2541–2542; Creasy D et.al., 2012, Proliferative and Nonproliferative Lesions of the Rat and Mouse Male Reproductive System, Toxicol Pathol, 40:40S–121S; Chapin RE, 1988, Morphologic Evaluation of Seminiferous Epithelium of the Testis. In: Lamb JC and Foster PMD, editors, Physiology and Toxicology of Male Reproduction, San Diego, California: Academic Press, 155–178; Hess RA and Moore BJ, 1993, Histopathological Methods for Evaluation of the Testes. In: Chapin RE and Heindel JJ, editors, Methods in Toxicology, Vol. 3, Pt. A. Male Reproductive Toxicology, San Diego: Academic Press, 52–85.

Contains Nonbinding Recommendations

allow reconstitution/recovery of spermatogenesis) on the reproductive system is an important consideration in the risk assessment.

C. Nonclinical Findings That Raise Concern for Male Fertility

In general, reproductive toxicity findings in male animals that raise concern for impaired fertility include, but are not limited to, atrophy, degeneration, necrosis, or hypocellularity of testes; increased seminiferous tubule degeneration or necrosis; germ cell depletion; or other pathology that may suggest impaired reproductive function. In addition, findings in other associated male reproductive organs (e.g., prostate, seminal vesicles, epididymis) may be suggestive of testicular toxicity. The sponsor should consider clinical evaluation of testicular function for direct-acting testicular toxicants that are associated with decreased reproductive function and/or adverse histopathology.

The significance of adverse findings in the toxicology and fertility studies increases if:

- The incidence and/or severity of the findings increase with dose and/or duration of treatment
- The reproductive findings occur in multiple species
- The reproductive findings occur in tissues bilaterally
- The adverse histopathology correlates with effects on reproductive organ weight
- A finding does not resolve after a period of one or two spermatogenic cycles or after five half-lives following the last drug dose
- The adverse findings occur at clinically relevant exposures
- The adverse findings are seen at pharmacokinetic exposures that result in a safety margin comparable to clinical exposure

Although histology is the most sensitive way to detect testicular and sperm quality toxicities, findings of reduced fertility, impaired mating behavior, and reduced capacity to mate in male fertility studies are concerns in and of themselves. These findings are especially concerning if they are corroborated by histopathological evidence of adverse effects on reproductive tissues in repeat dose toxicity studies. Findings that are suggestive of perturbations of the endocrine system are also a concern because changes in hormone homeostasis may adversely affect male (and female) reproductive physiology and performance. For example, drug-induced alterations in endocrine function can affect testicular weight, gamete maturation and release, sperm count, and/or fertility.

Table 1 summarizes findings in nonclinical studies that may increase the level of concern for impaired fertility.

Contains Nonbinding Recommendations

Table 1. Nonclinical Findings That May Increase the Level of Concern for Impaired Fertility in Men

General Nonclinical Findings to Consider in Male Fertility Risk Assessment
Finding occurs at clinically relevant exposures or small multiples of the clinical exposure
Finding occurs in multiple species
Finding increases in incidence and severity with increasing duration of exposure
Finding does not resolve, or at least show partial recovery, after one or two spermatogenic cycles or after five half-lives following the last drug dose
Finding occurs bilaterally in paired organs
Finding is rare in healthy untreated animals
Reproductive organ weight change (increased or decreased weight) correlates with adverse histology
Decreased male fertility and impaired mating behavior
Sperm quality adversely affected (count, motility, or morphology)
Signs of hormonal perturbation: <ul style="list-style-type: none">• Anti-androgenic signs — decreased weight and maturation of male sexual organs, including seminal vesicles and ventral prostate when weighed with their secretions, clinical signs suggestive of reduced aggressiveness (e.g., lethargic or reduced mating behavior, feminization of males)• Androgenic signs — masculinization of females (decreased fertility, female sexual organ pathology, or estrus cyclicity), decreased testes size, and impaired spermatogenesis

D. Confounding Factors

Numerous factors can confound apparent male reproductive toxicities. The use of drugs that cause a reduction in body weight or impair neuromuscular/neurological function may result in signals consistent with impaired reproductive function. When decreased spermatogenesis is detected in testicular histopathology examinations, it is important to document the reproductive age of the nonclinical model and to determine if the drug can have temporary or permanent effects on testicular development and spermatogenesis. Drugs that cause adverse effects on sperm quality in rats, without an effect on mating outcome, may still represent a risk to human males because these findings may indicate undesirable effects on testicular function, independent of mating outcome. For testicular toxicants where the mechanism of action is based on changes in hormone levels, the sponsor should clinically monitor hormones.

E. Follow-Up Investigations

Based upon an assessment of the findings from the nonclinical toxicology studies and any additional findings, the sponsor should consider additional nonclinical studies to characterize an observed male reproductive toxicity on a case-by-case basis. Follow-up studies could contain some of the following assessments:

- A demonstration of the potential for recovery from the adverse finding after cessation of dosing, if not available from the initial toxicology studies
- A reproductive hormone analysis, recognizing that hormone concentrations can vary significantly between animals and over the course of a day and over the course of the study
- A determination of the target cell type (e.g., germ cell, Leydig cell, Sertoli cell)

Contains Nonbinding Recommendations

In selected cases, adding fertility and/or sperm quality analysis to repeat-dose toxicity or fertility studies may be appropriate. The length of dosing in the pre-mating period of the male fertility study could be increased to cover an entire spermatogenic cycle and epididymal transit (for example, approximately 63 days in rats) to determine the extent of expected or observed toxicities in previous studies. A confirmatory study in a second species may be useful in cases where the finding is suspected to be species dependent (e.g., when effects are caused by a species-specific metabolite).

FDA does not intend for these nonclinical discussions to be comprehensive but rather to serve as a starting point for evaluating the risk of testicular injury in humans.

IV. MONITORING OF THE TESTES DURING CLINICAL TRIALS

Whether an evaluation of testicular toxicity in men is needed depends on various factors, including the mechanism of action of the testicular toxicity, if known, duration of therapeutic use, exposure multiples for the expected clinical exposure, intended target population, and indication of use.

A plan to minimize and monitor for the risk of human testicular injury should be in place early in clinical development for drugs that have a potential to cause human testicular toxicity based on nonclinical findings at anticipated clinically relevant exposures, taking into account the other aforementioned factors. The sponsor can discuss this plan with the appropriate review division as part of a pre-investigational new drug application (pre-IND) meeting or develop the plan and provide it with the original IND or at other time points in the IND phase, as appropriate. The sponsor can also discuss these issues at milestone or other meetings.

It is not possible to provide a single risk minimization and monitoring plan for all drugs with a potential for human testicular toxicity. The sponsor should individualize each plan considering factors that may include 1) the type, severity, consistency, and reversibility of the findings in the toxicology studies; 2) the duration and dose of clinical exposure relative to the duration margin (i.e., if seen with short-term or only exposure in chronic toxicology studies) and exposure margin (minimum observed adverse effect level margins close to (i.e., less than tenfold) clinical exposures or with substantial margins to clinical exposure); and 3) consideration of benefit-risk based upon the intended indication (i.e., treatment of a serious/life-threatening disease with unmet medical need or a common disease without unmet medical need).

Risk mitigation may include restricting the population to be studied, at least in the early phases of development, if appropriate, for the drug in question. For example, the drug could be initially investigated only in females or vasectomized men who have completed family planning. Although initial use in females and these vasectomized men may not contribute clinical data relevant to semen quality, assessment of reproductive hormones could be informative. This approach may allow initial pharmacokinetic, pharmacokinetic/pharmacodynamic, safety, and efficacy evaluations of the drug while the sponsor obtains additional nonclinical testicular safety

Contains Nonbinding Recommendations

data, and better understands the likely clinical exposure range (to assess the clinical exposure margin relative to the no observed adverse effect level).

In circumstances where men will be exposed to the drug, the potential risk of testicular injury should be clearly conveyed in the informed consent. During the clinical trial in these subjects, information should be gathered on the effect of the drug on the testes. This information should be based on the specific circumstances of the subject's exposure and could include semen analyses, as follows:

- At baseline
- At one spermatogenic cycle (13 weeks) after starting the investigational drug
- At one spermatogenic cycle (13 weeks) after drug discontinuation or following complete drug elimination, whichever is later, with an assessment for recovery from changes in semen parameters, if significant adverse changes are seen at the 13-week evaluation

Subjects should abstain from ejaculating for at least 48 hours and a maximum of 7 days before each semen collection. For each assessment time point, semen analysis should be based on the average of two semen specimens collected several days apart. In addition, the sponsor could assess other biomarkers of testicular injury, such as serum concentrations of testosterone, follicle stimulating hormone (FSH), and luteinizing hormone (LH). The sponsor should factor in the diurnal and episodic nature of hormone secretion in the sampling scheme when performing hormonal testing.

V. DESIGN OF A CLINICAL TRIAL TO EVALUATE THE EFFECT OF A DRUG ON THE TESTES

Based on the nonclinical findings, the results from initial human testing, and the intended use of the drug being considered, it may be appropriate to conduct a dedicated clinical safety trial having as its primary purpose an evaluation of the effect of the drug on testicular function. Sponsors should discuss with the review division the appropriateness of conducting such a safety study before or in parallel with the phase 3 trials. A separate safety trial may not be necessary under certain circumstances, such as if the drug belongs to a class with known effects on the testes (e.g., radiomimetics, androgens, anti-androgens). The following section describes basic features that the sponsor should consider when designing such a trial.

A. Subject Selection

Trial subjects should be men considered to have normal potential for fertility as reflected by semen parameters.

Contains Nonbinding Recommendations

We recommend that subjects have semen parameters that equal or exceed the generally accepted 5th percentile values of the World Health Organization (WHO) reference values. As of 2010, these 5th percentile values are:⁶

- Semen volume — 1.5 milliliters (mL)
- Total sperm per ejaculate — 39 million spermatozoa per ejaculate
- Sperm concentration — 15 million spermatozoa per mL
- Sperm progressive motility — 32 percent
- Sperm morphology — 4 percent normal forms using strict “Tygerberg” method

These values should be equaled or exceeded in at least two semen specimens that are collected at least several days apart at baseline. Subjects should abstain from ejaculating for at least 48 hours and a maximum of 7 days before each semen collection. The sponsor should consider the average of the two specimens to be the baseline semen characteristics for each enrolled subject. To the extent feasible, subjects should be representative of the population for whom the drug is intended, with careful considerations for factors, such as disease severity and concomitant medications, that may confound the semen parameters.

B. Trial Design

We recommend a randomized, double-blind, placebo-controlled, parallel-arm trial. We recommend that the trial randomize approximately 200 men in a 1:1 ratio to receive either the investigational drug or placebo. In general, this sample size has been found to be adequate for the purposes of estimating cumulative distribution curves and producing a 95 percent confidence interval width that is reasonably narrow for the primary endpoint.

The investigational drug should be administered at a dose and frequency that is representative of its intended clinical use. In general, for drugs intended for chronic use, the drug should be administered for at least two human spermatogenic cycles, which is 26 weeks. Drugs indicated for short-term use or intermittent retreatment should be administered according to the maximum duration of intended use; sponsors may need to discuss the actual duration of investigational drug exposure with the review division.

Sponsors should obtain semen analyses at baseline, at the end of the first 13 weeks, and again at the end of the 26-week dosing interval for chronically administered drugs. For drugs intended for short-term use or intermittent retreatment, sponsors should perform these analyses at baseline and 13 weeks after administration of the investigational drug. For drugs with long half-lives (weeks to months), we recommend the sponsor discuss with the review division the appropriate timing of semen analyses to adequately evaluate the drug effect. Subjects should abstain from ejaculating at least 48 hours and a maximum of 7 days before each semen collection. For each assessment time point, two semen samples should be collected several days apart. The methods of collecting and handling of semen samples should be standardized for all sites in a trial. A single central laboratory should process and analyze all semen samples for the purposes of

⁶ Cooper TG, Noonan E, von Eckardstein S, et al., 2010, World Health Organization Reference Values for Human Semen Characteristics, *Hum Reprod Update*, 15(3):231–245.

Contains Nonbinding Recommendations

consistency and quality assurance. If there are significant challenges with adhering to these standard approaches, the sponsor should discuss alternative plans with the review division.

The primary endpoint of the trial should be the percentage of subjects in each group who experience a 50 percent or greater decline in sperm concentration, compared to baseline, 13 weeks after starting the investigational drug (short-term use or intermittent retreatment drugs) or after 26 weeks of drug exposure (chronically administered drugs). Currently, sperm concentration is considered the most reliably quantifiable semen parameter that has potential utility in providing information about male fertility. It should be noted, however, that no single semen parameter can predict fertility potential and that the sponsor should consider all parameters in a semen analysis. Therefore, the sponsor should evaluate as secondary endpoints changes from baseline in sperm concentration, ejaculate volume, total sperm per ejaculate, motility, and morphology. The sponsor could also consider comparisons to WHO reference values and population-specific values as additional secondary endpoints. The semen parameters at baseline and during treatment should represent the mean of two semen samples collected a few days apart at each time point.

The sponsor should consider evaluating hormones, such as serum testosterone, FSH, and LH, in cases where changes in semen parameters are suspected to be related to hormonal perturbation. In addition, these hormonal evaluations may help to inform on the drug's effect on testicular function.

Individual subjects who experience a 50 percent or greater decline in sperm concentration should be re-evaluated after at least a 13-week drug-free interval to assess the recovery following drug exposure. An evaluation of recovery after a longer drug-free interval may be necessary for drugs with particularly long half-lives. In these affected subjects, the mean of at least two semen analyses collected a few days apart at the end of the drug-free interval should be used to determine the change from baseline and change from the last on-treatment values of the semen parameters.

C. Presentation of Results

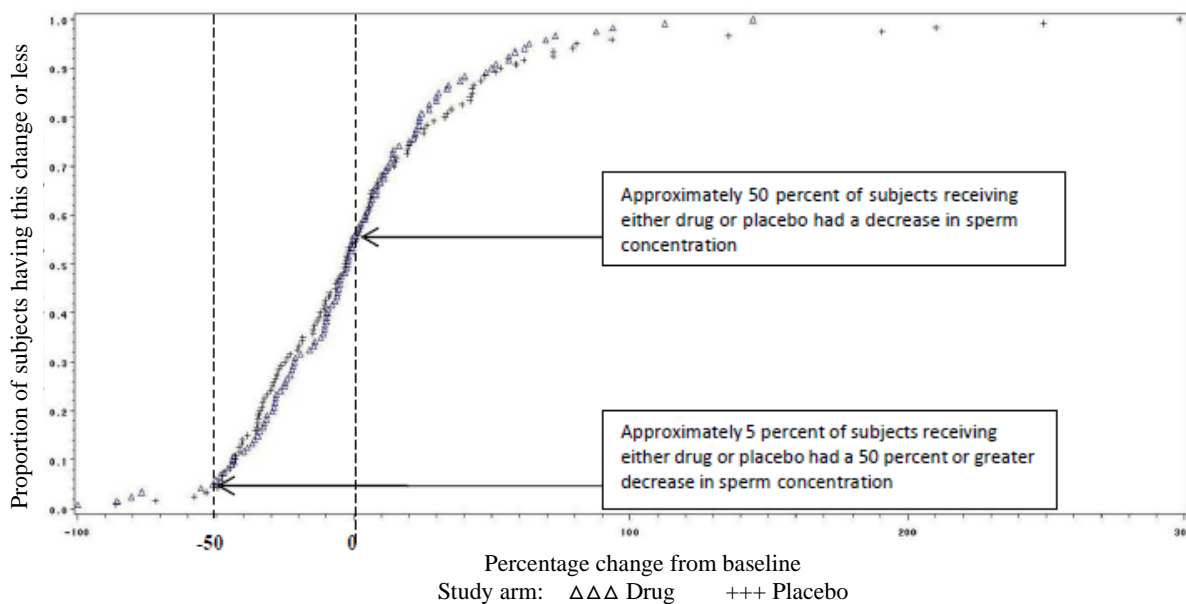
The sponsor should base the primary analysis on all subjects who have baseline and at least one post baseline semen sample and should include a prespecified approach for handling missing data. The proportion of subjects experiencing at least a 50 percent decrease in sperm concentration from baseline should be calculated together with the associated 95 percent confidence interval for the difference between the drug and placebo groups.

In addition, the sponsor should construct a cumulative distribution plot for the primary endpoint for each treatment group. The x-axis should display changes from baseline in sperm concentration ranging from 100 percent decrease (i.e., minus 100 percent or azoospermia) to the maximal observed increase. The y-axis should display the proportion of subjects who experienced a percentage change in sperm concentration, at the primary time point, equal to or less than the corresponding x-axis value.

Contains Nonbinding Recommendations

Figure 1 shows a sample plot. This plot shows that approximately 50 percent of subjects treated with either the investigational drug or placebo had a decrease in sperm concentration from baseline during treatment. It also shows that a decrease in sperm concentration of greater than 50 percent occurred in approximately 5 percent of the subjects who were treated with either the investigational drug or placebo.

Figure 1. Example of a Cumulative Distribution Plot



For each treatment group, the sponsor should calculate the median change from baseline in sperm concentration and in each secondary endpoint. The associated 95 percent confidence interval for the difference between the drug and placebo groups should be shown for all endpoints.

The percentage of subjects having individual secondary semen parameters within the normal reference range at the end of the treatment period should be presented for each treatment group. The percentage of subjects having all secondary semen endpoints within the normal reference range at the end of the treatment period should also be presented for each treatment group.

We recommend that sponsors include tables showing shift analyses from baseline to week 13 (or 26 for chronically administered drugs) for each of the primary and secondary endpoints for each treatment group. Each table would include shift analyses from within the reference range at baseline to above the reference range at week 13 (or 26 for chronically administered drugs) and from within the reference range at baseline to below the reference range at these time points.

The report should also include a discussion of reversibility of the findings during the drug-free follow-up period, if applicable.

D. Conclusions

The sponsor should consider the potential for risk to humans when nonclinical studies demonstrate drug-related adverse effects on male reproductive organs, semen analysis, and/or fertility. If assessment of testicular toxicity in men is indicated, plans to address and monitor for the risk of testicular injury should be in place early in clinical testing in male subjects or at the appropriate phase of drug development. The primary purpose of the clinical semen safety trial is to evaluate human testicular function based on nonclinical findings of testicular toxicity that cause concern. For FDA, the semen safety trial does not directly evaluate drug effect on human male fertility, but the trial can provide useful information about drug effect on commonly used measures of testicular function.

In general, it is not possible to stipulate firm guidelines for interpretation of these trial results and, a priori, specify results that would resolve the concern of testicular toxicity. Sponsors should evaluate individually each drug, its intended use, and the results of a semen trial as outlined in this guidance. Ultimately, the acceptability of the adverse effects of a drug on testicular function should be based on the overall benefit-risk assessment of the particular drug and indication being sought.