Testicular Toxicity: Evaluation During Drug Development

Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to http://www.regulations.gov. Submit written comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document contact Eufrecina Deguia at 301-796-0881.

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

July 2015
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
Tel: 855-543-3784 or 301-796-3400; Fax: 301-431-6353; Email: druginfo@fda.hhs.gov
http://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)

July 2015
Clinical/Medical
Pharmacology/Toxicology

TABLE OF CONTENTS

Testicular Toxicity: Evaluation During Drug Development Guidance for Industry[1]

This draft guidance, when finalized, will represent the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not create 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 staff 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 regarding drug products that may have potential adverse effects on the testes:

- Nonclinical findings that might raise a concern for testicular injury in men and a general approach on how to weigh the relevance of the nonclinical findings
- Common nonclinical approaches used to verify the potential relevance of adverse testicular findings in animals
- Clinical monitoring that can be employed when these drug products are initially administered to human subjects
- 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 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 might be considered based on the results of the clinical trials.

[1] This guidance has been prepared by the Division of Bone, Reproductive, and Urologic Products in the Center for Drug Evaluation and Research (CDER) 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.

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.

II. DIFFICULTIES EVALUATING TESTICULAR TOXICITY IN HUMANS

A thorough evaluation of a drug product's adverse effects on testes 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 there is a latency period of several months 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.

Because it is neither practical nor feasible to conduct a trial assessing male fertility by pregnancy rate, the main outcome measures of a clinical trial 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 (ICH) 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. Whether there is a need for an evaluation of testicular toxicity in men is based upon the weight of evidence of the adverse drug-

[3] We update guidances periodically. To make sure you have the most recent version of a guidance, check the FDA Drugs guidance Web page at http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm.

related findings on the male reproductive system in animals, specifically the accumulating evidence in appropriate species. Testicular toxicity is routinely assessed using:

- Repeat-dose toxicology studies with at least 4 weeks of drug exposure in two species
- Assessment of male fertility in rodents
- Comparative evaluation of pharmacokinetics in animals and humans[4]

Additional information may come from embryo/fetal reproductive and developmental toxicity studies, and fertility assessment after prenatal, neonatal, or juvenile exposure.

B. Nonclinical Study Design Considerations

A rationale should be provided 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 to document the background incidence and severity of the finding and the potential relation of the finding to treatment. Unless studies are intended to support dosing in pediatric patients, the use of sexually immature animals in acute/subchronic toxicity studies is not recommended because histology findings in immature animals may incorrectly suggest that fertility is impaired.

Histological evaluation of the reproductive organs is considered the most 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] Histopathology assessment of the reproductive tissues in the nonclinical male fertility study/studies is recommended if adverse findings in gonadal tissues were observed in repeat-dose toxicity studies. Assessing the persistence versus the reversibility of adverse effects on the reproductive system after drug withdrawal in the repeat-dose toxicology and male fertility studies is an important consideration in the risk assessment.

[4] See the ICH guidances for industry *M3(R2) Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals*, *S5A Detection of Toxicity to Reproduction for Medicinal Products*, and *S5B Detection of Toxicity to Reproduction for Medicinal Products: Addendum on Toxicity to Male Fertility.* (In November 2005, ICH combined the S5A and S5B guidances and titled the combined document *S5(R2) Detection of Toxicity to Reproduction for Medicinal Products & Toxicity to Male Fertility.* The contents of the guidances were not revised.)

[5] Chapin RE, 1988, Morphologic Evaluation of Seminiferous Epithelium of the Testis, in: Physiology and Toxicology of Male Reproduction, JC Lamb and PMD Foster (eds), Academic Press, San Diego, California, pp. 155-178; Hess RA and Moore BJ, 1993, Histological Methods for Evaluation of the Testes, in: Methods in Toxicology, Vol. 3, Pt. A. Male Reproductive Toxicology, RE Chapin and JJ Heindel (eds), Academic Press, San Diego, California, pp. 52-85.

C. Nonclinical Findings That Raise Concern for Male Fertility

In general, reproductive toxicity findings in male animals that increase concern for impaired fertility include, but are not limited to, testicular atrophy, seminiferous tubule degeneration or necrosis, or other pathology that may suggest impaired reproductive function.

Adverse findings in the toxicology and fertility studies cause more concern under certain circumstances. The significance of the findings 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 and/or bilateral tissues
- The adverse histopathology correlates with effects on reproductive organ weight
- A finding does not resolve after a period of one spermatogenic cycle following the last drug dose
- The adverse findings occur at all of the doses evaluated
- The adverse findings are seen at pharmacokinetic exposures that do not provide a reassuring safety margin compared 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 repeat-dose toxicity studies. The level of concern increases if reproductive toxicity occurs following exposures during multiple stages of life (e.g., fetal, peri/postnatal, juvenile, and/or adult stages). Findings that are suggestive of perturbations of the endocrine system are also a concern because endocrine disruptions 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.

Findings in nonclinical studies that may increase the level of concern for impaired fertility are summarized in Table 1.

Table 1. General Nonclinical Findings to Consider in Male Fertility Risk Assessment

Nonclinical Findings That May Increase the Level of Concern for Infertility in Men
Finding is dose-dependent
Similar findings in multiple species
Finding persists or increases in severity with increasing duration of exposure
Finding persists after drug withdrawal, especially if withdrawal period is an entire spermatogenic cycle
Finding occurs in bilateral tissues
Finding is rare in healthy untreated animals
Maximum dose without adverse effect occurs at exposures that are clinically relevant
Reproductive organ weight change (increased or decrease weight) correlates with adverse histology
Decreased male fertility and impaired mating behavior
Sperm quality adversely affected (count, motility, or morphology)
Adverse effects on reproductive tissues and function at multiple stages of life (repeat-dose study in adults, adult fertility assessment, effects in adulthood after exposure during pre/postnatal period, toxicity to the reproductive tissues during development)
Anti-androgenic signs — reduced body weight, decreased weight and maturation of male sexual organs, clinical signs suggestive of reduced aggressiveness (e.g., lethargic or reduced mating behavior)
Androgenic signs — masculinization of females (decreased fertility, female sexual organ pathology, or estrus cyclicity), decreased testes size, and impaired spermatogenesis
Confounding Issues
Use of immature animals
Pharmaceuticals that cause weight loss — in some cases, findings observed only in animals with weight loss may be difficult to ascribe to the drug exposure because weight loss alone may adversely affect male fertility independent of drug exposure. Also, weight loss may be secondary to overt toxicity and may not be clinically relevant.
Pharmaceuticals that impair mating behavior or neuromuscular function
Inappropriate animal model — pharmaceutical is not active in the species or has different metabolite profile, tissue distribution, or extent of elimination

D. Confounding Factors

Numerous factors can confound apparent male reproductive toxicities. The use of immature animals or the use of pharmaceuticals that cause a reduction in body weight or impair neuromuscular/neurological function may result in signals consistent with impaired reproductive function. When azospermia or decreased spermatogenesis is detected in testicular histopathology examinations it is important to document the reproductive age of the nonclinical model. If animals were immature at the beginning of treatment but should have attained maturity by the end of the study, then it is important to determine if the drug can have temporary or permanent effects on testicular development and spermatogenesis. Drugs affecting body weight, neuromuscular function, and/or mood can also appear to affect mating and fertility. Clinical evaluation of testicular function should be considered only for direct-acting testicular toxicants, where decreased reproductive function is accompanied by adverse histopathology.

E. Follow-Up Investigations

Additional nonclinical studies to characterize an observed male reproductive toxicity should be considered on a case-by-case basis. Omission of follow-up studies to further characterize adverse findings should be justified. Follow-up studies could contain some of the following assessments:

- A demonstration of the reversibility of the adverse finding after cessation of dosing

- A reproductive hormone analysis, although hormone concentrations can vary between animals and over time

- A determination of the target cell type (e.g., germ cell, Leydig cell, or Sertoli cell)

- A sperm quality assessment including number, motility, and morphology to characterize findings further

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 premating period of the male fertility study could be increased to cover an entire spermatogenic cycle (for example, 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.

F. Conclusion

These nonclinical discussions are not intended to be comprehensive, but rather serve as a starting point for evaluating the risk of testicular injury. The potential for risk to humans should be evaluated when drug-related adverse effects in nonclinical studies are identified in male reproductive organs, semen analysis, and/or fertility assessment. This evaluation should consider the mechanism of action, route of exposure, duration of therapeutic use, exposure multiples for the expected clinical exposure, and indication of use. Finally, the nonclinical evaluation should conclude with a determination of whether clinical assessment of semen parameters is recommended.

IV. MONITORING OF THE TESTES DURING CLINICAL TRIALS

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. This plan can be discussed with the appropriate review division as part of a pre-investigational new drug application (IND) meeting or be developed by the sponsor and included with the original IND.

It is not possible to provide a single risk minimization and monitoring plan for all drugs with a potential for human testicular toxicity. Each plan should be individualized based on:

- The specific nonclinical findings
- The proposed clinical investigation
- The exposures anticipated in the clinical trial
- The target population and the indication
- The overall risk versus benefit assessment

Risk mitigation may include restricting the population to be studied, if appropriate, for the drug in question. For example, the drug could be initially investigated only in females, vasectomized men, or men with no interest in future procreation. Initial use in females and vasectomized men will not contribute any clinical data relevant to testicular toxicity, but will make initial pharmacokinetic, safety, and efficacy evaluations of the drug possible while additional nonclinical testicular safety data are obtained.

In circumstances where men who may desire future fertility will be exposed to the drug, the potential risk of testicular injury should be 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 generally should include semen analyses:

- At baseline
- At one spermatogenic cycle (13 weeks) after starting the investigational drug
- If significant adverse changes are seen at the 13-week evaluation, at one spermatogenic cycle (13 weeks) after final exposure to the investigational drug to assess for recovery of changes in semen parameters

Subjects should abstain from ejaculating for at least 48 hours 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, other biomarkers of testicular injury (such as serum concentrations of testosterone, follicle stimulating hormone (FSH), luteinizing hormone (LH), and inhibin B) should be assessed.

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. The following basic features should be considered in the trial.

A. Subject Selection

Trial subjects should be men considered to have normal potential for fertility as reflected by semen parameters. Normal range semen parameters can serve as a guide for subject selection.

We recommend that subjects have semen parameters that equal or exceed the World Health Organization reference values. As of 2010 these values are:[6]

[6] Cooper TG, Noonan E, von Eckardstein S, et al., 2010, World Health Organization Reference Values for Human Semen Characteristics, Human Reproduction Update, Vol. 15, No. 3 pp. 231-245.

- Semen volume — 1.5 milliliters (mL)
- Total sperm per ejaculate — 39 million
- Sperm concentration — 15 million per mL
- Sperm progressive motility — 32 percent
- Sperm morphology — 4 percent normal using strict 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 before each semen collection. For enrolled subjects, the average of the two specimens should be considered the baseline semen characteristics. If feasible, subjects should be representative of the population for whom the drug is intended.

B. Trial Design

A randomized, double-blind, placebo-controlled, parallel-arm trial is recommended. We recommend that the trial randomize approximately 200 men in a 1:1 ratio to receive either the investigational drug or placebo. 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 re-treatment 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.

Semen analyses should be obtained 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 re-treatment, sponsors should perform these analyses at baseline and 13 weeks after administration of the investigational drug. Subjects should abstain from ejaculating at least 48 hours 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 consistency and quality assurance.

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 re-treatment 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 all parameters in a semen analysis should be considered. Therefore, changes from baseline in sperm concentration, ejaculate volume, total sperm per ejaculate, motility, and morphology should be evaluated as 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.

Evaluation of hormones, such as serum testosterone, FSH, and LH, can be considered in cases where changes in semen parameters are suspected to be related to hormonal perturbation. In addition, these hormonal evaluations, including serum inhibin B, may help to inform 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.

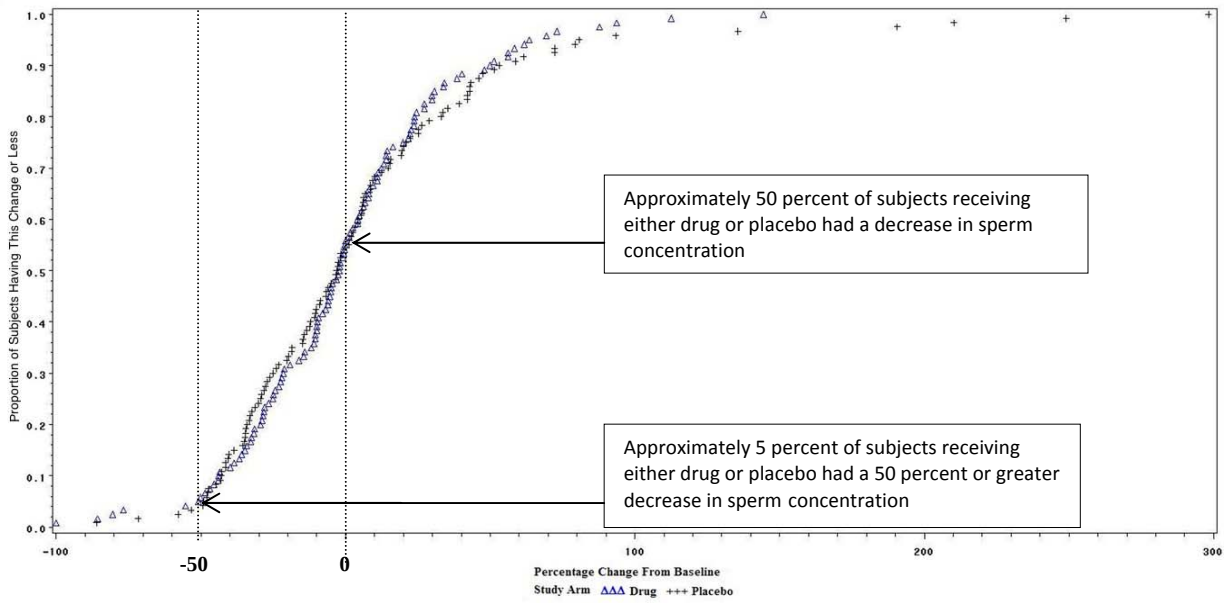
C. Presentation of Results

The primary analysis should be based on all subjects who have baseline and at least one post-baseline semen sample and should include a prespecified approach for the handling of 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, a cumulative distribution plot for the primary endpoint should be constructed for each treatment group. The x-axis should display changes from baseline in sperm concentration ranging from 100 percent decrease (i.e., −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.

A sample plot is shown in Figure 1. 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



The median change from baseline in sperm concentration and for each secondary endpoint should be calculated and shown for each treatment group. 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 including 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.

D. Evaluation of Results

The purpose of the clinical semen trial is to evaluate human testicular function based on nonclinical findings of testicular toxicity that cause concern. The trial does not directly evaluate drug effect on human male fertility.

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. Each drug, its intended use, and the results of a semen trial as outlined in this guidance should be individually

376 evaluated. Ultimately, the acceptability of the adverse effects of a drug on testicular function
377 should be based on the overall risk-benefit assessment of the particular drug and indication being
378 sought.
379