Research Article
Predictivity of Nonclinical Male Reproductive Findings for Human Effects

Anthony R. Scialli *[1], Richard V. Clark[2], and Robert E. Chapin[3]

Background: Testing of pharmaceutical products for reproductive toxicity in male laboratory animals is required for registration. **Methods:** We evaluated whether the results of studies showing male reproductive toxicity in experimental animals was predictive of reproductive effects in men participating in clinical trials. We surveyed companies for information on pharmaceutical candidates that had shown male reproductive toxicity in nonclinical studies for which there was information on male reproductive effects in clinical trials. **Results:** Among 12 pharmaceutical candidates submitted by five companies, only one compound that had shown male reproductive toxicity in experimental animals also demonstrated reproductive toxicity in men. **Conclusion:** In this sample of compounds, nonclinical studies appeared to over-predict reproductive toxicity in men. We identified possible reasons for the apparent lack of predictivity of the experimental animal studies.

Birth Defects Research 00:000–000, 2017.

Key words: male reproduction; reproductive toxicity; animal-human concordance; testicular toxicity; nonclinical studies

Introduction

Male reproductive toxicity testing in experimental animals is required for the approval of pharmaceutical products in the United States, Europe, and elsewhere. Implicit in experimental animal testing is the expectation that human risk can be predicted by data from these nonclinical studies. Specific requirements for reproductive toxicity testing of pharmaceuticals is contained in the guidelines of the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH S5(R2), 2005), which are currently under revision (ICH S5(R3), 2015).

In conjunction with the development of the ICH guidelines in the 1990s, a literature review was prepared on the detection of effects of chemicals on male reproduction (Ulbrich and Palmer, 1995). The paper summarized the results of studies on 117 compounds or groups of compounds in which male reproductive findings were reported in experimental animals. The authors considered any adverse finding at any exposure level to be potentially predictive, and they concluded that histopathology and weight of male reproductive organs were the best nonclinical endpoints for the detection of reproductive toxicity. In a comparison of experimental animal results for 46 compounds with available human data, adverse outcomes were shown in at least one experimental species for most of the compounds considered to have adverse male reproductive effects in men.

A survey of manufacturers indicated that most companies encounter evidence of testicular toxicity in nonclinical studies at least occasionally during drug development (Sasaki et al., 2011). The responses to these findings varied, with many companies performing additional testing on a case-by-case basis. In some instances, the additional testing included monitoring of male reproductive endpoints in clinical studies. Although the Ulbrich and Palmer (1995) review suggests that some problematic exposures in men can be predicted by exposure levels in experimental animal studies, it is not known how often male reproductive findings in non-clinical drug development programs correctly predict human effects when male reproductive endpoints are included in clinical programs. The survey reported here attempted to estimate how often experimental animal findings might predict reproductive effects in men.

Materials and Methods

This study was performed by the Developmental and Reproductive Toxicology Technical Committee of the International Life Sciences Institute Health and Environmental Sciences Institute. Questionnaires were sent to member companies requesting information on pharmaceutical products for which there were experimental findings of male reproductive toxicity that led to the inclusion of any type of male reproductive safety monitoring in the clinical program.

The survey document requested the compound class but no other identifying information, and respondents were encouraged to provide what information they could, recognizing that there might be proprietary concerns in the provision of complete responses. Information was requested on

[1]Scialli Consulting LLC, Washington, DC
[2]Endocrinology Consulting, Durham, North Carolina
[3]Retired

*Correspondence to: Anthony Scialli, Scialli Consulting LLC, 2737 Devonshire Place NW #120, Washington, DC 20008. E-mail: ascialli@scialliconsulting.com

Published online 0 Month 2017 in Wiley Online Library (wileyonlinelibrary.com).
Doi: 10.1002/bdr2.1102

the nature of the non-clinical findings and the dose level at which they were observed. The clinical findings and associated dose level also were requested.

The analytical plan included a calculation of the predictive value of male reproductive toxicity findings in an experimental animal study at any dose level for any abnormal male reproductive findings in the clinical program.

Results

Responses were received regarding 12 pharmaceutical candidates from five different companies (Table 1). There was one case (Compound 11) in which the experimental study predicted the human response. In this case, there were abnormalities of testicular histology and decreases in serum testosterone in dogs, and men had decreases in serum testosterone. The remaining 11 compounds produced adverse effects in one or more experimental species, but no adverse effects were identified in men. Considering all 12 compounds, the predictive value of a finding in an experimental animal study was 1/12 or 8.5%.

Among the 11 compounds that gave potentially non-predictive responses in the non-clinical studies, 5 compounds were evaluated in men using serum hormone concentrations without semen analysis or testicular histology (Compounds 3, 5, 6, 7, and 9). Removing these cases as possibly including inadequate assessment of human reproduction potential leaves a predictive value for the remaining set of 1/7 or 14%.

One of the compounds (Compound 8) demonstrated adverse testicular histology in dogs but not in rats, monkeys, or men. The dog findings occurred at 0.1 times the human plasma concentration (area under the curve, or AUC) basis. The rat and monkey studies were conducted at up to 18 and 11 times, respectively, the human AUC exposure basis. In this instance, a risk assessment based on the most sensitive species would have been misleading.

An examination of the exposure multiples at which the experimental animal findings occurred shows that no-observed-adverse-effect levels (NOAELs) were usually less than 10 times the human therapeutic concentration on a plasma AUC basis. Lowest-observed-adverse-effect levels (LOAELs) were 0.1, 7, 12, 40, or 200 times the human dose for the five compounds for which this information was provided, including Compound 8 mentioned previously.

Discussion

This survey was intended to address the question of how often adverse male reproductive findings in experimental animal studies are predictive of adverse male reproductive effects in clinical trials. We accepted any adverse effect at any exposure level in an experimental animal study as potentially predictive of human reproductive risk, similar to the method of Ulbrich and Palmer (1995) in their evaluation of the published literature on male reproductive toxicity in experimental animals. The predictive value of experimental animal findings for adverse effects in men in our sample was 8 to 14%, depending on whether the studies in men used semen analysis, testicular histopathology, or only reproductive hormone concentrations in peripheral blood. Although serum testosterone and luteinizing hormone (LH) would be expected to reflect Leydig cell function, and follicle stimulating hormone (FSH) and inhibin-B would be expected to reflect Sertoli cell function and the general health of the seminiferous tubule, these hormone measurements are probably insufficiently sensitive surrogates for semen analysis endpoints in men (Green et al., 2013). It is not known how many compounds are withdrawn from development because of testicular toxicity observed in non-clinical studies and a reluctance to carry out further clinical studies that might not provide definitive results. Recommendations in the United States Food and Drug Administration draft guidance on the evaluation of testicular toxicity during drug development (U.S. FDA, 2015) may also influence company development decisions when there have been non-clinical signals of testicular toxicity. Clinical studies to evaluate possible effects on spermatogenesis require adequate duration of exposure of at least two to three spermatogenic cycles and adequate sampling to compensate for individual variation in semen endpoints, for example, three samples over 2 weeks at each study assessment time point (Amory et al., 2007).

The sample of studies described in the present manuscript was limited to voluntary submission by members of the Developmental and Reproductive Toxicology Technical Committee of the International Life Sciences Institute Health and Environmental Sciences Institute and may not be representative of the universe of pharmaceutical candidates that have undergone experimental animal testing with male reproductive end-points. Failure of studies in men to identify adverse male reproductive effects could have been due to some or all of the following four factors, or perhaps additional factors:

1. *High dose levels or exposure concentrations in the experimental animal studies*. Compound 7, for example, produced findings in experimental animals at 200 times the human AUC concentration, and the NOAEL was 12 times the human dose on a plasma concentration basis. It may be unrealistic to expect men treated with this compound at $1/12^{th}$ the animal NOAEL to demonstrate adverse reproductive effects;
2. *Insufficient sensitivity (or relevance) of the endpoints used in men*. Semen analysis or reproductive hormones may not be capable of identifying modest histopathological testicular changes such as delayed spermiation that can be readily appreciated in experimental animal studies. Indeed, the variation between a pair of same-

TABLE 1. *Survey Responses*

	Class	Species	Non-clinical findings	Margin based on AUC unless otherwise noted		Clinical design	Clinical findings
				LOAEL	NOAEL		
1	Antiviral	Rat	Seminiferous tubule degeneration, 8 of 15 at high dose, 23% reduction in motility, no change in count. Recovered after exposure stopped.	12	7	Semen sub-study. 20 men/arm, 3 arms (standard care [ribavirin/interferon] + placebo or low dose or high dose). 2 semen samples pre-start and then 2 samples each at week 12, 24, 36, and 60. Enrolled n=20/arm; ~10/arm finished	No change from baseline in any sperm/semen measure or reproductive hormones
2	Pain/CNS (published as Sikka et al., 2015)	Rat	Reduced sperm count, motility, fertility, and litter size. Increased abnormal sperm and pre-implantation embryo loss. Histological lesions in testis and epididymis		3	Non-inferiority design; 903 men screened, randomized to placebo ($n = 109$) and treatment ($n = 111$). $N = 70$ and 75 completed to 10 weeks treatment, respectively. Two semen samples collected at beginning and end and also 14 weeks after end of dosing.	Treatment was non-inferior for number of men with >50% reduction in count or motility or testosterone concentrations. No change in FSH, abnormal forms, or semen volume
3	CNS active	Dog	3 months: testis (seminiferous tubule degeneration characterized by multi nucleate giant cell formation, degeneration/death of spermatids, tubular dilatation, and occasional Sertoli cell vacuolation) and epididymis (reduced intraluminal sperm and intraluminal desquamated seminiferous epithelial	12 (AUC); 22 (C_{max})		30 treated and 10 controls in single ascending dose (SAD) study (Japanese) 66 treated and 22 controls in SAD study	No clinically significant abnormalities in testosterone, prolactin, LH, FSH, or TSH at doses up to 6.25x human therapeutic dose No clinically relevant treatment-related changes in testosterone, prolactin, LH, FSH, or TSH at doses up to 7.75× human therapeutic dose

TABLE 1. *Continued*

	Class	Species	Non-clinical findings	Margin based on AUC unless otherwise noted		Clinical design	Clinical findings
				LOAEL	NOAEL		
			cells). Decreased sperm motility & counts and increased abnormal mor phology. Testicular toxicity at mid and high dose; generalized toxicity at high dose only.			34 treated men in 14-day multi ple ascending dose study	No clinically relevant dose-related changes or trends in testosterone, prolactin, LH, FSH, or TSH at doses up to
						42 treated male patients in 28 day Phase 2a study	1.5× human therapeutic dose No evidence of seminiferous tubule dysfunction was observed using serum inhibin B and FSH at human therapeutic dose
		Rat	1 month: reduced weight of testis, epi didymis and prostate with no histo pathological changes. Testicular toxic ity at high dose; generalized toxicity at mid and high dose.	2.5			
		Rat	3 months: testis (tubular atrophy, Sertoli cell vacuolation, spermatogenic giant cells), epididymis (cellular debris). Testicular and generalized toxicity at the high dose.	3.2			
4	Squalene synthase inhibitor	Rat	Degeneration of germ cells, vacuoliza- tion of Sertoli cells, atrophy of semi- niferous tubules	7	3	50 healthy adult/double-blind, placebo-controlled; men treated for 12 weeks in the arm with semen evaluation.	No significant effect on major semen variables (sperm morphology, motil- ity, viability, count, and concentra- tion, and ejaculate volume)
5	TrkA inhibitor	Rat	13 weeks: Treated at 30% of lethal dose. No change in body weight or evidence of general toxicity. Slight to moderate atrophy of the seminiferous tubules. Low sperm count, abnormal sperm, reduced fertility. Reversible.	40	13	Multiple ascending dose study ($n = 45$); 2-week duration.	No changes in FSH, inhibin B, LH, testosterone

TABLE 1. *Continued*

	Class	Species	Non-clinical findings	Margin based on AUC unless otherwise noted		Clinical design	Clinical findings
				LOAEL	NOAEL		
6	FAAH inhibitor	Rat	13 weeks. Spermatid retention in the seminiferous tubules at stages IX to XI. Low sperm count, abnormal sperm, and reduced fertility. Reduced testosterone and prostate weight and generalized toxicity at higher dose levels. Reversible within 2 weeks.	0.002		Vasectomized men, $N = 9$ in single ascending dose study, $N = 9$ in multiple ascending dose study; 2-week duration. Proof of concept study, 4–7 weeks, 82 subjects at start of study	No change in FSH, inhibin B, testosterone, or LH at 4× the proof of concept study dose Inclusion of fertile men based on information in informed consent.
7	CB2 antagonist	Dog	Degeneration and atrophy of seminiferous tubules, absence of sperm in the epididymis, immaturity in the prostate after 4 weeks, associated with generalized toxicity	200	12	6 healthy male in single asending dose study	No effect on testosterone, LH, FSH, inhibin B, and SHBG
8	Endothelin antagonist	Dog	Atrophy of seminiferous tubules that became more severe with dose and duration. After 52 weeks of dosing and 52 weeks of recovery, slight changes remained in 1/2 animals; NOAEL at 0.1% of generally toxic dose level, LOAEL at 1% of generally toxic dose level.	0.1	0.012	Dedicated study in men scheduled for orchiectomy for prostate cancer with 4 weeks of dosing; treatment up to 3 months in some subjects	Histopathology of human testis showed no drug-related changes
		Rat	No treatment-related findings; NOAEL at 70% of generally toxic dose level.		18		
		Monkey	No treatment treatment -related find ings; NOAEL at 40% of generally toxic dose level.		11		

TABLE 1. *Continued*

	Class	Species	Non-clinical findings	Margin based on AUC unless otherwise noted		Clinical design	Clinical findings
				LOAEL	NOAEL		
9	5-HT1B/5-HT2A receptor antagonist	Rat	Retention of seminal fluid at 200 mg/kg/day in 1-month study	Not determined		40–46 male or female patients given 5, 10, or 20 mg daily for 24 weeks	No adverse effects on LH, FSH, or total testosterone at 1 year post-treatment
		Rat	Small, flaccid testes; testicular tubular atrophy, Leydig cell hypertrophy, hypertrophy of tunica albuginea, empty epididymal lumen, cellular debris and eosinophilic globules in epididymis at 150 mg/kg/day in 6-month study. Changes only partially reversible after 22 week recovery period (partial repopulation of testicu lar tubular epithelium); generalized toxicity at 100 mg/kg/day for 6 months.	Not determined			
		Rat	Decreased fertility index, sperm motility, sperm count, and number of sperma tids. Altered sperm morphology, marked testicular tubular atrophy, epi didymal and prostatic changes at 150 mg/kg/day in male fertility study (males treated for 4 weeks prior to cohabitation)	Not determined			
10	Glucagon-like peptide-1 agonist	Dog	Focal dilation of seminiferous tubular lumen with tubular vacuolation and variable germ cell layer loss (hypo spermatogenesis) and minimal seg mental sperm stasis in testes at 300/100 μg/kg twice daily and 1000/400/250 μg/kg twice daily for 13 weeks.		372	138 patients given 10 μg titration dose for 3 days, then 15 μg titration dose for 4 days, then 20 μg maintenance dose on Days 8–182	No effects on percent subjects with ≥50% decrease in sperm concentration, ejaculate sperm count, sperm morphology, sperm motility, or concentrations of FSH, LH, free and total testosterone, neutral α-glucosidase, or inhibin B.

TABLE 1. *Continued*

	Class	Species	Non-clinical findings	Margin based on AUC unless otherwise noted		Clinical design	Clinical findings
				LOAEL	NOAEL		
			Reversible after 4-week recovery period; generalized toxicity at ≥200 μg/kg/day for up to 12 months.				
		Dog	Hypospermatogenesis with focal to multifocal vacuolation and atrophy of seminiferous tubules and focal sperm stasis in the testis at ≥200 μg/kg BID for 12 months; generalized toxicity at ≥200 μg/kg/day for up to 12 months. Epididymal oligospermia or aspermia with tubular dilation and epithelial degeneration observed at ≥200 μg/kg BID.		17		
		Dog	Minimal to moderate testicular tubular dilation, increased vacuolation, increased sperm stasis at 20 μg/kg twice daily, 200 μg/kg daily, or 200 μg/kg twice daily given for 8 months to juvenile (age 16–17 weeks) dogs; generalized toxicity at ≥200 μg/kg/day for up to 12 months. Epididymal tubular dilation, epithelial degeneration, oligospermia at 200 μg/kg daily or twice daily. Reversible after 2-month recovery period		6.8		
11	Tachykinin NK3/NK2 receptor antagonist	Dog	Decreased testosterone concentration at ≥0.5 mg/kg/day and LH levels at 0.5 mg/kg/day in 1-month toxicity study accompanied by testicular tubular		0.04	6 men/group received single doses at 2, 5, 12, 25, 50 mg	Decreased testosterone concentrations at 50 mg at 8 hours and 12 hours post-dose. Cohort 1: 5/6 subjects met pre-defined stopping criteria:

TABLE 1. *Continued*

	Class	Species	Non-clinical findings	Margin based on AUC unless otherwise noted		Clinical design	Clinical findings
				LOAEL	NOAEL		
			degeneration/atrophy and Sertoli cell/ Leydig cell vacuolation, decreased epididymal sperm content, and/or prostatic atrophy at ≥0.5 mg/kg/day.				testosterone <6.9 mM and/or decrease >50% from mean time-matched Day-1 values. Cohort 2: 4/ 6 subjects had decreased testosterone.
		Dog	Decreased to absent testosterone and decreased LH levels from 10 mg/kg/ day in 3-month tox study; reversible after 6-month recovery period. Accompanied by decreased testes, epididymides, prostate weights and impaired spermiogenesis and/or absence of elongated spermatids in testis, decreased epididymal sperm content, and hypoplasia/atrophy of prostate at ≥10 mg/kg/day. All sexual organ changes were reversible after 6-month recovery period (100 mg/kg/ day). Generalized toxicity at ≥10 mg/ kg/day for 3 months.		0.02		
12	Pyrimidine synthesis inhibitor	Mouse	Testicular atrophy/degeneration; similar effects in prostate, seminal vesicles at 60 or 100 mg/kg/day for 2 weeks; generalized toxicity at ≥30 mg/kg/day for up to 3 months	Not determined		14 patients/group – 100 mg daily for 3 days, then 20 mg daily maintenance dose for 12 weeks, then 12-week treatment-free phase	Possible trend towards decreased sperm count but relationship to treatment could not be excluded due to small sample size ($n = 14$) and high variability of semen data. Decreased sperm density was transient and not apparent 3 months post-treatment.

TABLE 1. *Continued*

Class	Species	Non-clinical findings	Margin based on AUC unless otherwise noted		Clinical design	Clinical findings
			LOAEL	NOAEL		
	Rat	Germinal epithelium atrophy in testes at 20 mg/kg/day for 1 month	Not determined			
	Rat	Decreased epididymal sperm count at 4 mg/kg/day; no effect on sperm motility or fertility endpoints at up to and including 4 mg/kg/day for 10 weeks prior to cohabitation, through cohabitation, until euthanasia (~ 13–14 weeks total); generalized toxicity at ≥4 mg/kg/day for up to 3 months	Not determined			
	Dog	Decreased prostate and testes weights at 16 mg/kg/day for 3 months gener alized toxicity at ≥16 mg/kg/day for up to 3 months.	Not determined			

CNS, central nervous system; TSH, thyroid stimulating hormone.

donor human semen samples is quite large (Schrader et al., 1991), and this variation makes human semen studies quite insensitive unless large numbers of men are sampled. There is also the question of whether hormones are appropriate or sufficiently sensitive to detect a treatment-induced effect on spermatogenesis. The hypothesis that FSH and inhibin B are good markers of spermatogenesis (Meachem et al., 2001; Kumanov et al., 2006) is not universally supported in the literature, and many acute treatment-induced changes appear the most problematic for a hormone response (Kolb et al., 2000; Salenave et al., 2012; Rendtorff et al., 2012);

3. *Small sample sizes and high variability*. The human studies sometimes included fewer than 20 men per dose group, which was likely insufficient to identify statistically significant changes in measures with high variability such as semen analysis endpoints. Reproductive hormones are produced episodically and vary across a wide range of normal, perhaps obscuring subtle but important alterations in hormone production;
4. *Insufficient duration of exposure*. Some of the clinical studies in this sample involved single-dose treatment or treatment on the order of a month in duration. These exposures may not have resulted in alterations that would have occurred during more prolonged exposure.

Although limited by a relatively small number of studies, we have examined the available data and found only a poor correlation between the male reproductive toxicity produced in non-clinical animal studies and results from human clinical trials. There may be several reasons for this discrepancy, despite the belief that testicular and spermatogenic function are generally conserved across these various mammalian species. It is possible that newer techniques such as transcriptional or metabolic profiling of semen will improved predictivity. We hope continued collection and analysis of relevant data will lead to conclusions with greater confidence.

References

Amory JK, Wang C, Swerdloff RS, et al. 2007. The effect of 5α-reductase inhibition with dutasteride and finasteride on semen parameters and serum hormones in healthy men. J Clin Endocrinol Metab 92:1659–1665.

Green DM, Zhu L, Zhang N, et al. 2013. Lack of specificity of plasma concentrations of inhibin B and follicle-stimulating hormone for identification of azoospermic survivors of childhood cancer: a report from the St Jude lifetime cohort study. J Clin Oncol 2013;31:1324–1328.

ICH. 2005. ICH harmonised tripartite guideline. Detection of toxicity to reproduction for medicinal products & toxicity to male fertility S5(R2). http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S5/Step4/S5_R2__Guideline.pdf. Accessed July 23, 2017.

ICH. 2015. Final concept paper. S5(R3): detection of toxicity to reproduction for medicine products & toxicity to male fertility. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S5/S5_R3__Final_Concept_Paper_27Mar2015.pdf. Accessed July 23, 2017.

Kolb BS, Stanczyk FZ, Sokol RZ. 2000. Serum inhibin B levels in males with gonadal dysfunction. Fertil Steril 74:234–238.

Kumanov P, Nandipati K, Tomova A, Agarwal A. 2006. Inhibin B is a better marker of spermatogenesis than other hormones in the evaluation of male factor infertility. Fertil Steril 86:332–338.

Meachem SJ, Nieschlag E, Simoni M. 2001. Inhibin B in male reproduction: pathophysiology and clinical relevance. Eur J Endocrinol 145:561–571.

Rendtorff R, Beyer M, Muller A, et al. 2012. Low inhibin B levels alone are not a reliable marker of dysfunctional spermatogenesis in childhood cancer survivors. Int J Androl 44:219–225.

Salenave S, Trabado S, Maione L, Brailly-Tabard S, Young J. 2012. Male acquired hypogonadotropic hypogonadism: diagnosis and treatment. Ann Endocrinol (Paris). 73(2):141–146.

Sasaki JC, Chapin RE, Hall DG, et al. 2011. Incidence and nature of testicular toxicity findings in pharmaceutical development. Birth Defects Res B Dev Reprod Toxicol 92:511–525.

Schrader SM, Turner TW, Simon SD. 1991. Longitudinal study of semen quality in unexposed workers: sperm motility characteristics. J Androl 12:126–131.

Sikka SC, Chen C, Almas M, et al. 2015. Pregabalin does not affect sperm production in healthy volunteers: a randomized, double-blind, placebo-controlled, noninferiority study. Pain Pract 15:150–158.

Ulbrich B, Palmer AK. 1995. Detection of effects on male reproduction: a literature survey. J Am Coll Toxicol 14:293–327.

U.S. Food and Drug Administration. 2015. Testicular toxicity: evaluation during drug development. Draft Guidance. July, 2015. https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm455102.pdf. Accessed July 23, 2017.