Variability of Semen Parameters with Time in Placebo Treated Men

Jonathan P. Jarow,* Xin Fang and Tarek A. Hammad

From the Offices of New Drugs (JPJ); Biometrics (XF) and Surveillance and Epidemiology (TAH), Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, Maryland

Purpose: We describe the variability of semen parameters with time in normal men receiving placebo. We also report the impact of season and geographic region, among other variables, on these parameters.

Materials and Methods: Data from the placebo arms of 5 randomized, controlled trials were pooled. All trials set minimum standards for semen parameters as an eligibility criterion for entry. Semen parameters examined include volume, density, motility, total count, total motile count and morphology. Mixed model repeated measure analysis was used for statistical analysis. Coefficients of variation for each semen parameter and the percent change from baseline were calculated.

Results: The mean within-subject coefficient of variation for each semen parameter ranged from a low of 10% to a high of almost 50%. The contribution of season and region to variability was negligible. The reduction in variability with an increasing number of samples per time point had decreasing returns beyond 2 samples.

Conclusions: There was considerable variation in semen parameters with time in subjects who received placebo. Variation could not be attributed to season or region. We observed a general negative trend in semen parameters in this population selected for normal baseline semen parameters, which was likely due to the placebo response or to regression toward the mean.

Key Words: testis; semen; reference values; infertility, male; placebo effect

Abbreviations and Acronyms

CV = coefficient of variation

LS = least squares

MMRM = mixed model repeated measure

Accepted for publication November 8, 2012.

* Correspondence: WO22 Room 5387, 10903 New Hampshire Ave., Silver Spring, Maryland 20993 (telephone: 301-796-0149; FAX: 301-796-9897; e-mail: jonathan.jarow@fda.hhs.gov).

A finding of testicular toxicity in animal models is not an uncommon event during the early phase of drug development.[1] A human study is often then required to determine the presence, extent and reversibility of the testicular toxicity observed in nonclinical studies if the product is intended for use in men. This human study typically relies on measurement of semen parameters and serum hormone concentrations to detect potential toxicity to the male reproductive system as a surrogate for fertility since measuring pregnancy rates is neither feasible nor practical.

Semen analysis measures sperm production and quality. Parameters that are measured directly include semen volume, sperm concentration (density), percent motility and percent normal morphology. Parameters that are calculated include total sperm count and total motile sperm count per ejaculate. Sperm concentration is commonly used as a generic measure of testicular sperm output, although a more accurate measure of testicular output may be total sperm per ejaculate per abstinence time.[2]

The typical clinical semen study uses a noninferiority design based on

0022-5347/13/1895-1825/0
THE JOURNAL OF UROLOGY®

http://dx.doi.org/10.1016/j.juro.2012.11.077
Vol. 189, 1825-1829, May 2013
Printed in U.S.A.

a responder analysis against placebo.[3] The semen parameter and cutoff point chosen to define a responder, ie toxicity, as well as the inferiority margin from comparators, in this case placebo, are arbitrarily defined. Few previous studies have determined the extent of variability of all semen parameters,[2] which leaves a significant knowledge gap. We sought to provide the evidence needed to inform these decisions, including the underlying variability of semen parameters with time and the distribution of the percent change from baseline of semen parameters in a group of volunteers on placebo during clinical trials.

METHODS

Data sets from 1 unpublished and 4 published trials in men receiving placebo were submitted to the Food and Drug Administration for analysis.[3–6] All trials set minimum standards for semen parameters as an entry eligibility criterion. The most significant variation among the 5 trials was the minimum standard for sperm density, which was greater than 40 million per ml in 3 trials and greater than 20 million per ml in the other 2. The specific methodology used for semen analysis varied among the 5 trials.[7–9] All trials used central reading for morphology but the other parameters were assessed on site. The semen parameters evaluated included semen volume, sperm concentration (density), percent motility, total sperm count per ejaculate, total motile sperm count per ejaculate and percent normal morphology. The abstinence duration was fixed in all trials at 2 to 3 days.

In general, all MMRM analyses in this study used a compound symmetry variance-covariance structure for repeat measurements with time. For each MMRM analysis we calculated the average of measurements for each semen parameter in a defined visit period as an observation for that visit unless specified otherwise.

We combined all data and used MMRM analysis to assess the effect of season on semen quality. The MMRM model for season included fixed effects of the visit (counted by month), season and trial. The baseline value was added to the model as a covariate. Season was defined using dates for the Northern Hemisphere.

The CV of each semen parameter was calculated using the SD divided by the mean for each subject, which was then averaged for the population. For this CV analysis each measurement was used rather than creating an average for each interval.

The effect of region in the United States on semen parameters was determined based on the state in which each investigator site was located. We did not have data on the location of individuals. We combined the states into regions east and west of the Mississippi River due to the few subjects per state. Mean baseline values of the 2 regions were comparable based on an ANCOVA model including fixed effects of the region and trial. For each parameter regional effects were compared using MMRM analysis with fixed effects of the region, visit (defined by the majority of measurements) and trial. The baseline value was added to the model as a covariate.

We examined the mean percent change from baseline in all semen parameters. Baseline was defined as the average of all pretrial randomization measurements. Each individual measurement was used to determine the overall distribution of the percent change from baseline. LS means in the percent change from baseline for 3, 6, 9 and 12 months were based on the same model used for the regional effect. We plotted the LS means of the percent change from baseline for 52 weeks of followup.

The effect of the number of samples per visit interval on the reduction of variance in the percent change from baseline was evaluated. For this purpose we only used sampling intervals with at least 3 measurements available. Within-subject variance was estimated by residual variance based on the same model used for the regional effect.

RESULTS

Table 1 lists the design characteristics and entry criteria of each trial. The pooled database included 333 subjects who received placebo. A total of 4,173 semen samples were obtained (average 12.5 per subject). Average followup was 57 weeks.

Table 2 shows the overall within-subject CV for each semen parameter measured. The semen parameter with the lowest variability was morphology and the next lowest parameters were motility and semen volume. Sperm density (concentration) had the lowest variability among the parameters measuring testicular sperm output. The median CV for sperm density measured with time in 333 men on placebo was 36%.

Table 3 shows semen parameter seasonal averages. The general pattern observed in this population was a nadir in the winter and a peak in the spring for almost all parameters. However, this observed pattern may have been due to random variation.

Table 1. *Design characteristics of each trial used for analysis*

	Unpublished Data	Overstreet et al[4]	Amory et al[5]	Hellstrom et al[3]	Jarvi et al[6]
Trial dates	1993–1995	1995–1997	1998–2000	2004–2005	2005–2006
No. subjects	68	39	32	128	66
Entry criteria:					
Sperm density (million/ml)	Greater than 40	Normal	Greater than 40	Greater than 20	Greater than 20
Age range	18–50	18–40	18–55	Greater than 45	25–64
Duration (wks)	108	108	52	66	24
No. samples/interval	3	3	3	2	2
No. intervals	4	4	3	8	3