*Reproductive Toxicology*, Vol. 2, pp. 183–190, 1988
Printed in the U.S.A.

0890-6238/88 $3.00 + .00


# LONGITUDINAL STUDY OF SEMEN QUALITY OF UNEXPOSED WORKERS*

## I. Study Overview

STEVEN M. SCHRADER, TERRY W. TURNER, MICHAEL J. BREITENSTEIN, and STEPHEN D. SIMON

Department of Health and Human Services, Public Health Service, National Institute for Occupational Safety and Health, Division of Biomedical and Behavioral Science, Experimental Toxicology Branch, Cincinnati, Ohio

**Abstract — A longitudinal study of 45 men was conducted evaluating the semen quality of monthly samples collected over 9 months. The statistical variation of sperm count, semen volume, percentage of motile sperm, sperm velocity, sperm morphology, and sperm viability, assessed by both the vital stain and the hypoosmotic swelling (HOS) assay, were each evaluated using intraclass correlations and coefficients of variation. Sperm count and semen volume had large intraclass correlations (62% and 60%, respectively), indicating that if a subject has a high count or volume he will tend to continue to have high counts or volumes. On the other hand, sperm velocity had an intraclass correlation of only 16% indicating that fluctuations within a subject were nearly as large as fluctuations from subject to subject. The remaining parameters had intraclass correlations ranging from 42% to 47%. Sperm count, percent motile sperm, and semen volume each had large coefficients of variation (both between and within subjects). These variables, especially count, had relatively poor precision. Sperm velocity, percent motile sperm, percent normal morphology, the HOS assay, and the vital stain assay had lower coefficients of variation, indicating greater precision.**

*Key Words:* Male reproductive toxicity, Semen quality, Sperm count, Morphology, Motility, Testicular volume.

## INTRODUCTION

Since the discovery of the adverse effects of DBCP on the health of exposed workers (1), occupational field studies to assess reproductive hazards to male workers have utilized semen analysis to detect changes in the reproductive potential of these worker populations (2–10). In recent years, many new, more objective measures of semen analysis have become available (11), but little is known about the population stability and variability of these semen parameters. This study utilized a longitudinal design of 45 men across 9 months to evaluate semen quality parameters. The purpose of this report is to give an overview of the study population and to provide information on the variability of both between-subject and within-subject sampling of key semen parameters. Subsequent reports will provide detailed information on each parameter.

Address correspondence to: Steven M. Schrader, Department of Health and Human Services, Public Health Service, National Institute for Occupational Safety and Health, Division of Biomedical and Behavioral Sciences, Experimental Toxicology Branch, 4676 Columbia Parkway, Cincinnati, OH 45226.

*Mention of a product or company name does not constitute endorsement by NIOSH.

## MATERIALS AND METHODS

### *Population*

Analysis of the variations of semen parameters observed in previous occupational groups indicated that at least 35 subjects were needed to provide adequate precision for variance estimates. A recruitment time of two weeks was initiated. Men, 25 to 35 years old (at the onset of the study), were recruited via a newspaper ad and word of mouth. The only exclusion criteria were that the men could not be current employees of the U.S. Government nor could they have had a vasectomy. When inquiries were made, a packet of information explaining the study (including a consent form, a medical history form, and instructions to call for an appointment) was mailed to the potential participant. Forty-seven men proceeded to make appointments. At the first appointment, a consent form was signed and a medical history was taken. An occupational health physician then drew blood and conducted a brief physical exam. Testicular volume was measured using plastic templates (12). A video tape was also shown explaining how the sample was to be collected. This video tape stressed three points: 1) 2 days of sexual abstinence was requested prior to sample collection, 2) the samples had to be collected by mastur-

bation, and 3) the samples had to be delivered to our laboratory within one hour of ejaculation. Each man was assigned a code number and all personal identifiers were kept in a confidential file. The only access to this confidential file was to notify men of missed monthly appointments and to mail the final results of the study to all of the participants. All medical histories, semen samples, and blood samples contained only the code number. Payments for semen samples were graduated in increasing amounts to encourage continued participation. The payments for the nine months were: $25, $25, $30, $35, $40, $50, $60, $70, and $90. A brief medical history update was taken with the monthly delivery of the semen sample to determine if any changes in health status or occupational exposure had occurred from the previous month. Additional blood samples were taken in the fifth and the ninth month.

*Semen analysis*

The study was conducted from June 1986 to February 1987. The semen analyses were conducted during the week containing the fifteenth day of each month. Participant appointments were distributed over a five-day work week. At the time of collecting the semen sample, each subject recorded his abstinence, collection time, and spillage on the jar label.

Semen analyses were conducted in two phases. The initial evaluation of the sample was conducted when the sample arrived at the laboratory and consisted of recording the temperature, turbidity, color, liquefaction time, volume, osmolality, and pH of the semen. Video recordings, viability assessments, sperm counts, preparation of slides, and preservation of seminal plasma were also conducted at this time. Morphologic and morphometric analyses of slides and computer analyses of video tapes for motility and velocity were conducted at a later time. Sperm viability was determined by two methods, eosin y stain exclusion (13) and hypoosmotic swelling (HOS assay) (14). These techniques test for the structural and functional integrity of the membrane, respectively (15). Sperm concentration and motility characteristics were measured in a Makler chamber (16). Sperm concentration was calculated by averaging the count of two groups of ten Makler squares.

The measurements of sperm motility and velocity were conducted by first recording eight fields around the center grid of the chamber. The microscope stage was warmed to 37°C, and each field was recorded for about 15 seconds. The percent motile sperm, sperm velocity, linearity, and lateral head displacement were assayed using a computer-assisted semen analysis (Cellsoft, Cryo Resources, New York). The computer system parameters were set to analyze 15 frames at a rate of 30 frames/second. The velocity settings were a minimum of 4 frames with a maximum velocity of 120 $\mu$m/second. A minimum of 2 frames were needed to detect motility. The individual cell data were further evaluated, and any cells having a straight line velocity less than or equal to 5 $\mu$m/second or a linearity less than or equal to 1.00 were classified as immotile; these cells were not included in velocity or motility measurements.

Sperm morphology was measured after the preparation of air dried and stained smears, and sperm were classified into nine categories according to Zaneveld and Polakoski (17). This classification system is essentially identical to that recommended by the World Health Organization (18) and consists of normal forms (oval heads and normal tails), microcephalic heads, macrocephalic heads, tapered heads, absent heads, amorphous heads, double heads, abnormal tails, and immature forms. Two hundred sperm were read on each slide. One slide was read for each semen sample.

*Statistical analysis*

The data were analyzed using Statistical Analysis System (SAS) (19) on an IBM-3090 mainframe. The NESTED procedure in SAS computed between-subjects and within-subjects standard deviations as well as the intraclass correlations. The MEANS procedure produced all other descriptive statistics except for the total coefficient of variation which was hand calculated.

For all of the analyses performed for this report, we assumed that the intrasubject error terms for one month were independent of other months. There are two possible violations of this assumption. First, autocorrelations of one month with the following month may violate the independence assumption. Second, seasonal variability, if it exists, might violate the independence assumption. With only nine months of data, determining the effects of autocorrelation and seasonality is difficult. Inspection of the data, however, showed that effects, if present, would be very small.

For this study five key statistics were computed: the between-subject standard deviations ($SD_B$), the within-subject standard deviations ($SD_W$), the intraclass correlations ($R_I$), the total coefficient of variation ($CV_T$), and the average individual coefficient of variation (within-subject coefficient of variation or $CV_W$).

Table 1. Population description

| Population Variable | *N* | Mean | *SD* | Range |
|---|---|---|---|---|
| Age (years) | 46 | 28.9 | 3.0 | 25–35 |
| Height (inches) | 46 | 71.38 | 2.71 | 64.25–78.50 |
| Weight (pounds) | 46 | 183.17 | 33.67 | 131.00–295.00 |
| Testicular volume[a] | | | | |
| Right (mL) | 45 | 23.35 | 2.60 | 12.00–28.00 |
| Left (mL) | 46 | 23.38 | 2.62 | 12.00–28.00 |
| Caffeine (drinkers only) (drinks/day) | 43 | 3.72 | 2.47 | 1.00–45.00 |
| Alcohol (drinkers only) (drinks/week) | 37 | 8.92 | 10.05 | 1.00–13.00 |
| Smoking (smokers only) (smokes/day) | 24 | 13.50 | 11.08 | 1.00–50.00 |
| Number of pregnancies (fathers only) | 21 | 2.19 | 0.98 | 1.00–4.00 |

[a]Based on 2 testicles for 45 men and 1 testicle for one man (cryptorchid).

The between and within standard deviations are the square roots of the respective components of variance in a nested ANOVA model. The between-subject standard deviation measures variation from one individual to another. The within-subject standard deviation measures variation of repeated samples within an individual. Standard deviations are reported instead of variances because they are in the same units as the mean.

The intraclass correlation (20) measures the size of the between component of variance relative to the total (between plus within) component of variance. A large intraclass correlation indicates that the between-subject variance is relatively large. If we suspect such a pattern, future comparative studies of this variable may require an emphasis on obtaining more subjects or on matching subjects. Either would reduce the impact of the large between-subject variance. A small intraclass correlation indicates that the within-subject variance is relatively large. If we suspect this pattern, future studies may require an emphasis on obtaining repeat samples so as to reduce the impact of the large within-subject variance. An alternate interpretation is that the intraclass correlation is a measure of repeatability. The intraclass correlation estimates how strongly two observations from the same individual are correlated.

The two coefficients of variation, $CV_T$ and $CV_W$, provide assessments of precision relative to the mean. $CV_T$ uses the total standard deviation in the numerator and the overall mean in the denominator. This number represents the relative precision of a parameter estimate (computed once for a single subject) for representing the mean of a population over time. $CV_W$ is an average of the 45 intrasubject coefficients of variation for a given parameter. It represents the relative precision of a single measurement of a parameter for representing the mean over time of the same individual who was measured.

For a given variable, either the standard deviations or the coefficients of variation (but not both) would be relevant for power calculations in future studies. For variables where the standard deviation tends to be proportional to the mean, the coefficients of variation are preferred. For variables where the standard deviation tends to be constant, the standard deviations would be preferred. In this study, only sperm count has standard deviations proportional to the mean.

## RESULTS

*Population*

None of the men reported a significant exposure to a known reproductive toxicant, nor were any male factor fertility problems reported. One of the 47 men dropped out of the study after the physical exam and did not contribute any semen samples. This man was not included in any of the statistics. Of the remaining 46 men, one dropped out after one semen sample, one man withdrew from the study after month five, and two other men missed a semen sample for a single month. These men were included in all statistics for which there were data.

Table 1 contains information on the population studied. The average age of the participants was 28.9 ± 3.0 (mean ± SD) years at the onset of the study. The mean height was 71 ± 2.71 inches, and mean weight was 183 ± 33.67 pounds. The average left testicular volume was 23.35 ± 2.60 mL. One

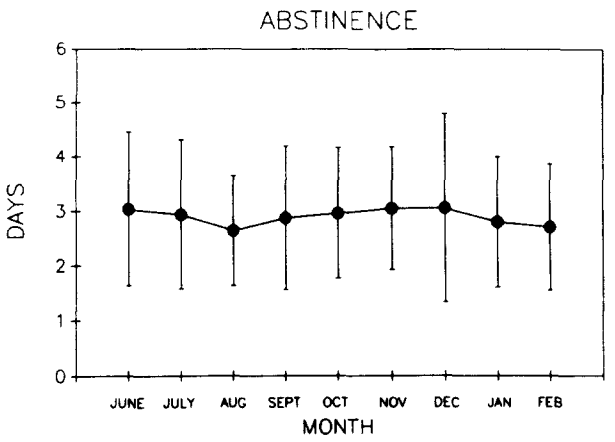


Fig. 1. Days of abstinence. The solid circles are mean number of days of abstinence. The error bar is one standard deviation.

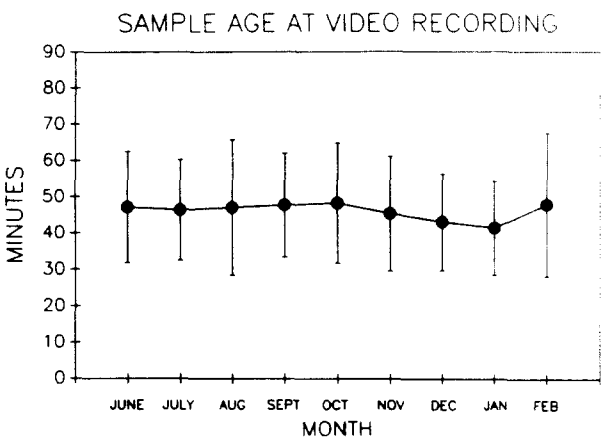


Fig. 2. Sample age at video recording. The solid circles are the mean age of the semen samples in minutes at the time they were video recorded. The error bars represent one standard deviation.

man was a unilateral cryptorchid lacking the right testicle. The right testicular volume, determined from 45 measurements, was 23.38 ± 2.62 mL. Ninety-four percent of the men consumed caffeinated drinks. Of those men drinking caffeinated drinks, an average of 3.72 ± 2.47 drinks (coffee, soft drinks, or tea) were consumed daily. Eighty percent of the participants consumed alcoholic beverages. The men drinking alcohol drank an average of 8.92 ± 10.05 alcoholic drinks per week. Fifty-two percent of the men smoked (pipe, cigar, or cigarette). The average number of smokes (pipes, cigars, and/or cigarettes) was 13.5 ± 11.08 per day. Forty-six percent of the men had fathered pregnancies. These men had averaged 2.19 ± 0.98 fathered pregnancies. These men were asked to have 2 days sexual abstinence when collecting the semen samples for analysis. Because there was no method of verifying abstinence, there was no penalty for shorter or longer times, but instead the need for the truth was emphasized. Figure 1 represents the mean number of days abstinence for each month. The monthly means ranged from 2.64 ± 1.01 days (August) to 3.07 ± 1.73 days (December). The range of abstinence for all samples was 0.5 days to 11 days.

The men were instructed to deliver the semen sample to the laboratory within 1 hour of ejaculation. Again, this could not be verified and the need for recording the actual time of ejaculation was emphasized. The mean age of the semen sample in minutes from the time of ejaculation to video recording the motility is illustrated in Figure 2. The mean ages ranged from 41.38 ± 12.84 minutes (January) to 48.20 ± 16.52 minutes (October). The overall age for all samples ranged from 14 to 118 minutes.

*Sperm count*

The descriptive statistics for sperm count are found in Table 2. The mean sperm count for all samples was 47.43 million sperm per mL with a standard deviation of 29.47 between men and 22.90 within men. The distribution of sperm counts by month and the median count for each month are shown in Figure 3. (It should be noted that the *y* axis is a log scale to better illustrate the distribution.) While the individual sperm counts were variable across time, the median sperm count for this population was quite consistent. The median sperm counts ranged from 35 million/mL (September) to 44.5 million/mL (December) for the population. The sperm count for all semen samples ranged from 0.2 million/mL to 265 million/mL.

*Semen volume*

The descriptive statistics for semen volume can be found in Table 2. The mean volume was 2.78 mL with a standard deviation of 0.91 between men and 0.74 within men. The distribution and median semen volume are plotted in Figure 4. The median semen volume ranged from 2.25 mL in June to 3.00 mL in October. The overall range of semen volumes ranged from 0.3 mL to 5.8 mL.

*Sperm motility*

The percent motile sperm and sperm velocity descriptive statistics are included in Table 2. The

Table 2. Descriptive statistics for semen parameters

| Semen Parameter | Mean | $SD_B$ | $SD_W$ | $R_I$ | $CV_T$ | $CV_W$ |
|---|---|---|---|---|---|---|
| Sperm count (millions/mL) | 47.43 | 29.47 | 22.90 | 62 | 79 | 44 |
| Semen volume (mL) | 2.78 | 0.91 | 0.74 | 60 | 42 | 27 |
| Sperm velocity (μm/sec) | 44.57 | 3.33 | 7.61 | 16 | 19 | 16 |
| Motile sperm (% motile) | 59.76 | 10.87 | 12.71 | 42 | 45 | 26 |
| Morphology (% normal) | 72.38 | 9.52 | 10.22 | 46 | 19 | 14 |
| Vital stain (% unstained) | 71.41 | 7.38 | 7.92 | 46 | 15 | 11 |
| Hypoosmotic swelling (% swollen) | 64.08 | 9.04 | 9.61 | 47 | 21 | 15 |

$SD_B$ represents between-subject standard deviation.
$SD_W$ represents within-subject standard deviation.
$R_I$ represents the intraclass correlation coefficient multiplied by 100.
$CV_T$ represents the total coefficient of variation expressed as a percentage of the overall mean.
$CV_W$ represents the average within-subject coefficient of variation expressed as a percentage of any given individual's mean.

mean velocity for all sperm samples was 44.57 μm/second with a between standard deviation of 3.33 and a within standard deviation of 7.61. The mean percent of motile sperm for all semen samples was 59.76 with a standard deviation of 10.87 between subjects and 12.71 within subjects. The distribution of sperm velocities and the percent motilities per month are found in Figures 5 and 6, respectively. Both values appear to have declined in the month of September and slowly increased over the next several months. The median sperm velocities ranged from 36.22 μm/seconds in September to 49.98 μm/second in August. The overall velocities for all samples in all months ranged from 11.69 μm/second to 76.61 μm/second. The median percent motilities ranged from 47.22% in September to 68.97% in February. The percent motile sperm in individual samples ranged from 0 to 95.16%.

*Sperm morphology*

The descriptive statistics for the percentage of normal shape are found in Table 2. The mean for all samples was 72.38% normal sperm with a standard deviation of 9.52 between and 10.22 within subjects.

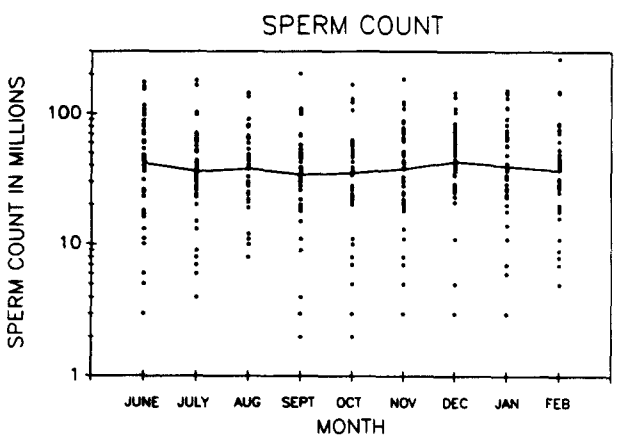


Fig. 3. Sperm count. The circles represent the sample distribution for each month. Only one circle is shown for each value, even if more than one sample had that value. The solid line represents the median value for each month. The *y* axis is a log scale.

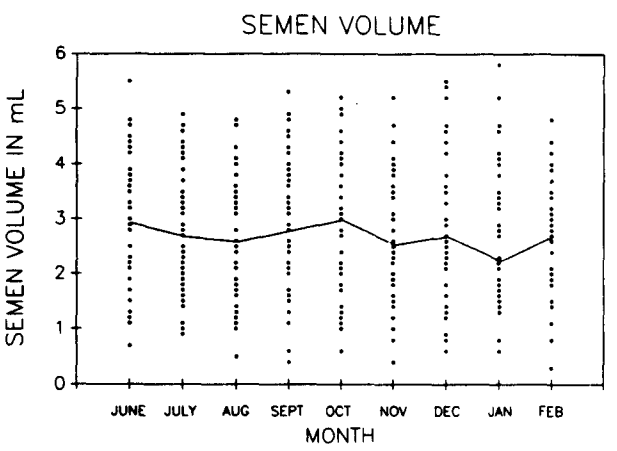


Fig. 4. Semen volume. The circles represent the sample distribution for each month. Only one circle is shown for each value even if more than one sample had that value. The solid line represents the median value for each month.

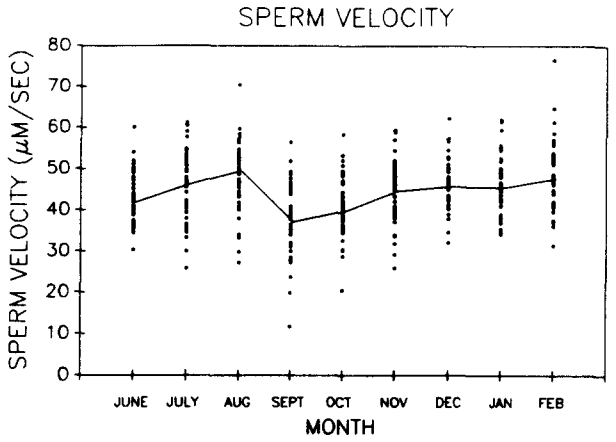


Fig. 5. Sperm velocity. The circles represent the sample distribution for each month. Only one circle is shown for each value, even if more than one sample had that value. The solid line represents the median value for each month.

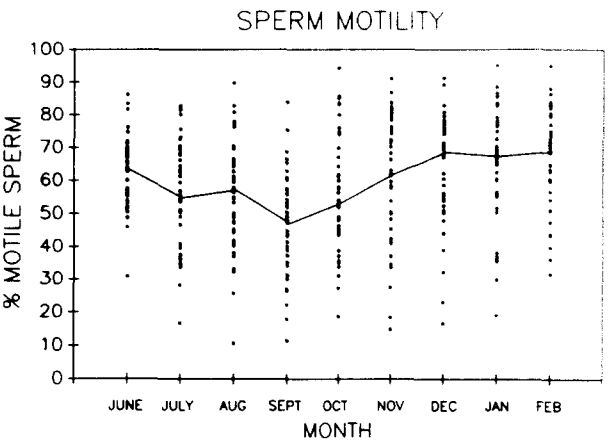


Fig. 6. Sperm motility. The circles represent the sample distribution for each month. Only one circle is shown for each value, even if more than one sample had that value. The solid line represents the median value for each month.

Figure 7 illustrates the distribution of the proportion of sperm with normal morphology for each month and the median values for each month. The median percent of normal cells ranged from 69.5% (October) to 82.5% (February). Percent normal sperm for all semen samples ranged for 4.5% to 91.5%.

*Sperm viability*

The HOS and the vital stain assay descriptive statistics are found in Table 2. The HOS assay indicated that 64.08% of the sperm from all samples were swollen (possessed a functional membrane). The standard deviation for both the between and within are quite similar (9.04 and 9.61, respectively). Figure 8 demonstrates the stability of the values across time with the median fluctuating only 7% over the nine-month period (62.85% to 69.71%).

The vital stain assay indicated that 71.41% of the sperm from all samples possessed a barrier to the stain. The standard deviations, both between subjects and within subjects, are essentially the same (7.38 between and 7.92 within). The median value for the vital stain assay was also very stable, ranging from 69.96% to 74.73%.

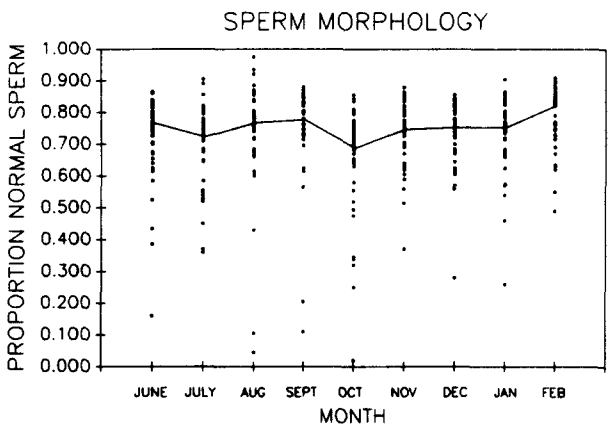


Fig. 7. Sperm morphology. The circles represent the sample distribution for each month. Only one circle is shown for each value, even if more than one sample had that value. The solid line represents the median value for each month.

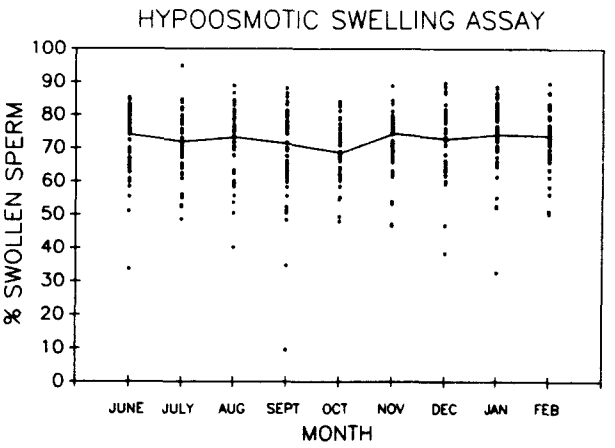


Fig. 8. The hypoosmotic swelling assay. The circles represent the sample distribution for each month. Only one circle is shown for each value, even if more than one sample had that value. The solid line represents the median value for each month.

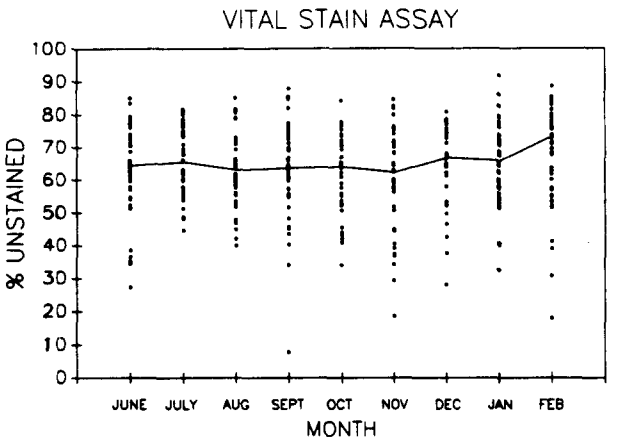


Fig. 9. The vital stain assay. The circles represent the sample distribution for each month. Only one circle is shown for each value, even if more than one sample had that value. The solid line represents the median value for each month.

## DISCUSSION

Sperm count is the most often measured variable in occupational field studies (2) assessing male reproductive effects. The standard deviations, both between and within (29.47 and 22.90, respectively), are very large in relation to the mean of 47.73 million sperm per mL. The relatively large between standard deviation indicates that careful matching of subjects and/or obtaining large numbers of subjects may be needed to obtain a reasonably powerful hypothesis test in future experiments that compare exposure groups. The intraclass correlation was surprisingly high at 62%, indicating that a man with a relatively low sperm count will continue to have a low count and a man with a high count will continue to have a high count. However, the average coefficient of variation for within subject sperm count was 44%, indicating that variability within a subject is large relative to the mean, which in turn makes a precise measurement of sperm count difficult. This variable may be useful in providing information on trends, but the variable lacks precision; thus, detecting population differences using sperm count data will be very difficult.

Semen volume also has a high intraclass correlation and moderate coefficients of variation. These data suggest that the semen volume from men is somewhat repeatable, the between and within coefficients of variation both being smaller than those for sperm count.

Sperm velocity measurements have very low between and within coefficients of variation, indicating that a single measurement is representative of both the population and the individual; however, there is greater variation within samples from individuals than between individuals. This causes the intraclass correlation to be very low. This variable will be useful for determining population shifts, but a single measurement may provide little information about an individual within a population.

Morphology, hypoosmotic swelling, and vital stain all have intraclass correlations around 50% and coefficients of variation near or below 20%. These parameters have good precision and should be useful in detecting trends and differences in populations.

*Acknowledgements* — The authors would like to thank Mr. Al Stine and Ms. Jill Lawson for programming support. Gratitude is also extended to Dr. Danny Brown, who served as the occupational health physician, to Ms. Janice Lubbers, who was the logistics coordinator, and to Mr. Dennis Lynch, who provided extensive input and critical review of this manuscript.

## REFERENCES

1. Whorton D, Krauss RM, Marshall S, Milby TH. Infertility in male pesticide workers. *Lancet* 1977; **2**:1259–1260.
2. Wyrobek AJ, Gordon LA, Burhart JG, Francis MW, Kapp RW, Letzl G, Malling HV, Topham JC, Whorton MD. An evaluation of human sperm as indicators of chemically induced alteration of spermatogenic function. *Mutat Res* 1983; **115**:73–148.
3. Ratcliffe JM, Meinhart TJ, Schrader SM, Turner TW, Steenland K, Leffingwell SE. Reproductive effects of ethylene dibromide in pine beetle fumigators. Cincinnati, OH: NIOSH — Health Hazard Evaluation Interim Report 1984: No. 1 TA 83244.
4. Ward LB, Hokanson JA, Smith ER, Chang LW, Pereira MA, Whorton EB, Legator MS. Sperm count, morphology, and fluorescent body frequency in autopsy service workers exposed to formaldehyde. *Mutat Res* 1984; **130**:417–424.
5. Heussner JC, Ward JB, Legator MS. Genetic monitoring of aluminum workers exposed to coal tar pitch volatiles. *Mutat Res* **155**:143–155.
6. Rosenberg MJ, Wyrobek AJ, Ratcliffe J, Gordon LA, Watchmaker G, Fox SH, Moore DH, Hornung RW. Sperm as an indicator of reproductive risk among petroleum refinery workers. *Br J Ind Med* 1985; **42**:123–127.
7. Ratcliffe JM, Clapp DE, Schrader SM, Turner TW, Tanaka S, Oser J, Halperin WE. Semen study of workers exposed to 2-ethoxyethanol. Washington, DC: NIOSH-Health Hazard Evaluation Report 1986: HETA 84-416-1688.
8. Ratcliffe JM, Schrader SM, Steenland K, Clapp DE, Turner TW, Hornung RW. Semen quality in papaya workers with long term exposure to ethylene dibromide. *Br J Ind Med* 1987; **44**:317–326.
9. Schenker MB, Samuels SJ, Perkins C, Lewis EL, Katz DF, Overstreet JW. Prospective surveillance of semen quality in the workplace. *J Occup Med* 1988; **30**:336–344.
10. Welch LS, Schrader SM, Turner TW, Cullen MR. Effects of exposure to ethylene glycol ethers on shipyard painters: I. Male reproduction. *Am J Ind Med* 1988 (in press).
11. Schrader SM, Ratcliffe JM, Turner TW, Hornung RW. The use of new field methods of semen analysis in the study of occupational hazards to reproduction: the example of ethylene dibromide. *J Occup Med* 1987; **29**:963–966.

12. Takihara H, Sakatoku J, Fujii M, Nasu T, Cosentino MJ, Cockett ATK. Significance of testicular size measurements in andrology. I. a new orchimeter and its clinical application. *Fertil Steril* 1983; **39**:836–840.
13. Eliasson R, Treichl L. Supravital staining of human spermatozoa. *Fertil Steril* 1971; **22**:134–137.
14. Jeyendran RS, Van den Ven HH, Perez-Palaez M, Crabo BG, Zaneveld LJD. Development of an assay to assess the functional integrity of the human sperm membrane and its relationship to other semen characteristics. *J Reprod Fertil* 1984; **70**:219–228.
15. Schrader SM, Platek SF, Zaneveld LJD, Perez-Pelaez M, Jeyendran RS. Sperm viability: a comparison of analytical methods. *Andrologia* 1986; **18**:530–538.
16. Makler A. The improved ten-microliter chamber for rapid sperm count and motility evaluation. *Fertil Steril* 1980; **33**:337–338.
17. Zaneveld LJD, Polakoski KL. Collection and physical examination of the ejaculate. In: Hafez ESE, ed. *Techniques of human andrology*. New York: Elsevier; 1977:147–172.
18. Belsey MA, Moghissi KS, Eliasson A, Paulsen CA, Gallegos AJ, Prasal MRN. Laboratory manual for the examination of human semen and semen cervical mucus interaction. Singapore: Press Concern; 1980.
19. *Statistic Analysis Systems: User's Guide: Statistics*. 5th ed. Cary, NC: SAS Institute; 1985.
20. Sokal RR, Rohlf FJ. *Biometry*. New York: WN Freeman and Company; 1981:215.