

Utility of the sperm chromatin structure assay as a diagnostic and prognostic tool in the human fertility clinic

D.P.Evenson^{1,6}, L.K.Jost¹, D.Marshall²,
M.J.Zinaman³, E.Clegg⁴, K.Purvis⁵, P.de Angelis⁵
and O.P.Claussen⁵

^{1,6}Department of Chemistry and Biochemistry and ²Department of Animal and Range Sciences, Box 2170, South Dakota State University, Brookings, SD 57007, USA, ³Department of Obstetrics and Gynecology, Loyola Medical Center, Maywood, IL 60153, USA, ⁴Reproductive and Developmental Toxicology, Environmental Protection Agency, Washington DC, USA and ⁵Institute for Pathology, National Hospital, Oslo, Norway

⁶To whom correspondence should be addressed

The sperm chromatin structure assay (SCSA) was used to measure over 500 human semen samples from two independent studies: Study I, 402 samples from 165 presumably fertile couples wishing to achieve pregnancy over 12 menstrual cycles; Study II, samples from 115 patients seeking fertility counselling. The SCSA measures susceptibility to DNA denaturation *in situ* in spermatozoa exposed to acid for 30 s, followed by acridine orange staining. SCSA data from the male partners of 73 couples (group 1) achieving pregnancy during months 1–3 of Study I were used as the standard of 'sperm chromatin compatible with high fertility' and were significantly different from those of 40 couples (group 3) achieving pregnancy in months 4–12 ($P < 0.01$) and those of male partners of 31 couples (group 4) not achieving pregnancy ($P < 0.001$). Group 2 contained couples who had a miscarriage. SCSA values for Study II were almost twice that of the Study I fertility standards. Within-couple repeatability tended to be less for group 3 than for groups 1, 2 or 4. Based on logistic regression, spermatozoa with denatured DNA (cells outside the main population, $COMP\alpha_t$) were the best predictor for whether a couple would not achieve pregnancy. Some 84% of males in group 1 had $COMP\alpha_t < 15\%$, while no couples achieved pregnancy in group 1 with $\geq 30\%$ $COMP\alpha_t$, a threshold level considered not compatible with good fertility. Using selected cut-off values for chromatin integrity, the SCSA data predicted seven of 18 miscarriages (39%).

Key words: chromatin structure/human fertility/pregnancy loss/SCSA

Introduction

Numerous studies have sought a key factor(s) that would be predictive for male fertility potential, including sperm count, motility, acrosome status, cell membrane integrity, morphology and morphometry of whole spermatozoa and/or sperm heads

and integrity of nuclear chromatin. However, to date, no single laboratory test can assess a man's total fertility (Amann, 1989). Initially, many of these tests used light microscope observations, but these suffered from labour intensiveness, intra-observer variations and low numbers of spermatozoa analysed, leading in turn to poor statistical power. These difficulties helped spawn the development of computer-driven instrumentation, most notably the computer-assisted sperm analysis (CASA) systems for motion analysis, elementary morphometry and sperm concentration measures.

Computer-interfaced flow cytometry (FCM) provides a powerful advantage over light microscopy techniques in terms of speed, multiple parameters/cell measured, objectivity, lack of bias in sample selection, and thousands of cells measured per sample, and thus provides a very high degree of statistical power (Shapiro, 1995). Additionally, FCM can potentially sort out unique cells that comprise $< 1\%$ of the total population for further study and/or use in assisted reproduction treatment, e.g. sexing sperm (Johnson *et al.*, 1987, 1989).

This study utilized the sperm chromatin structure assay (SCSA), first described by Evenson *et al.* (1980), and since refined for both animal and human spermatozoa (Evenson and Jost, 1994; Spano *et al.*, 1998). The SCSA utilizes the metachromatic properties of acridine orange (AO) to distinguish between low pH- or heat-denatured (red fluorescence = single-stranded) and native (green fluorescence = double-stranded) DNA in sperm chromatin. As first discussed by Evenson *et al.* (1980), human and bull sperm nuclear DNA was more resistant in fertile than subfertile subjects. Data from subsequent animal heterospermic insemination experiments showed strong correlations between SCSA data and fertility ranking in bulls (Ballachey *et al.*, 1988; $r = -0.94$, $P < 0.01$) and boars (Evenson *et al.*, 1994; $r \leq 0.93$, $P < 0.01$), thus providing strong evidence that mammalian sperm chromatin structure was highly correlated with pregnancy outcome.

Other investigators (Tejada *et al.*, 1984) have used a modified SCSA procedure by treating smears of human spermatozoa on glass microscope slides with acid, staining with AO, and observing sperm nuclei under fluorescence light microscopy. The sperm nuclei were generally scored as fluorescing green or red; in some cases, cells have also been scored as yellow, being intermediate between the red and green [see cover of *Science* (Evenson *et al.*, 1980)]. Although this method provides a general picture of the status of the sperm DNA denaturation susceptibility, it is limited to two or three classifications rather than the 1024 discrete channel levels of red and green fluorescence/cell by FCM. Also, since the metachromatic staining of AO is strictly dependent on exact equilibrium conditions (Darzynkiewicz *et al.*, 1975), artefacts exist in the

light microscopy method due to the unevenness of the surface on glass microscope slides providing differential micro-environment staining conditions. Fluorescence fading presents an additional problem, as well as varied time intervals between staining and scoring (D.P.Evenson, unpublished observation). Nevertheless, Tejada *et al.* (1984) reported a strong correlation between human fertility and the percentage of spermatozoa which fluoresced green. In addition, others (Sterik *et al.*, 1989) have reported solid correlations between the percentage of green-fluorescing sperm and fertility potential, thus providing support to our FCM studies.

In a study by Ibrahim *et al.* (1988), AO staining was the most discriminatory test ($P = 0.0001$) in a study of three groups (unexplained infertility, habitual abortions and normal fertile donors) when compared with the zona-free hamster egg penetration test and conventional semen analyses. The frequency of sperm chromatin heterogeneity as detected by AO red fluorescence was highest in habitual abortion (39.4%), followed by unexplained infertility (16.4%) and fertile donors (9.4%). The percentage of penetration was highest in habitual abortion (50.7%), followed by fertile donors (43.1%) and unexplained infertility (33.9%). Conventional semen tests (concentration, motility, morphology and vitality) were the least reliable in discriminating between the three groups.

Human SCSA data are more constant over time than the classical measures of spermatozoa, as shown in a study of 45 men who were unexposed to environmental insult and who provided monthly semen samples over a 9-month period (Schrader *et al.*, 1988; Evenson *et al.*, 1991). The coefficient of variation of SCSA variables within a man were much lower than for the classical measures (Schrader *et al.*, 1988; Evenson *et al.*, 1991; Spano *et al.*, 1998). Thus, for cross-sectional, or single sample analysis of a patient, SCSA data can be considered to characterize one component of semen quality with a higher degree of certainty than the classical measures. Human spermatozoa obtained from cancer patients (Evenson *et al.*, 1984) or those with high fever (D.P.Evenson, unpublished data) show increased SCSA values, followed by either a return to normal (D.P.Evenson, unpublished data) or continued abnormality (Evenson *et al.*, 1984; Fossa *et al.*, 1997). Numerous animal toxicology studies have consistently shown the same very high level of repeatability of SCSA measures and convincing biological dose-response interpretations (Evenson *et al.*, 1989, 1993a,b). All human studies to date have indicated that SCSA data are poorly correlated with the classical semen measures; thus, SCSA data are considered independent variables that are of diagnostic and prognostic value in the human andrology clinic, as shown in these studies (Evenson *et al.*, 1991; Spano *et al.*, 1998).

Materials and methods

Semen donors

Study I: this study design has been described previously (Zinaman and Katz, 1996) and abstracts have been presented on the SCSA data (Evenson *et al.*, 1996; Evenson and Jost, 1998). This was a prospective study of 200 presumably fertile couples wishing to achieve pregnancy and who were discontinuing contraception. Couples were counselled

to have intercourse centred on the predicted day of ovulation. Phase 1 included the first three cycles in which women collected daily morning urine samples, underwent mid-cycle post-coital tests and, if late for their menses, presented for serum human chorionic gonadotrophin (HCG) testing. Phase 2 encompassed the next nine cycles in which the remaining women continued to attempt pregnancy and underwent serum HCG testing if menstruation was delayed. Urine samples from phase 1 were used for sensitive HCG testing to detect occult pregnancies. Clinically recognized pregnancies were followed until delivery in order to ascertain outcome. During phase 1, semen samples were collected from the 200 male partners in the clinic early in each cycle until pregnancy was achieved. A minimum of two semen samples was obtained from each man. Some 82% of the 200 couples followed for the entire study period conceived, the maximal fertility rate being ~ 30% per cycle in the first two cycles. This rate rapidly tapered with subsequent cycles as the more fertile couples were removed. Pregnancy loss occurred in 31% of the first pregnancies detected; 41% (15/36) of these losses were seen only by urine HCG testing and were categorized as occult. Eleven of these same couples later achieved clinically recognized conceptions during the study. Semen samples from the male partners of only 165 couples were measured by the SCSA due to a freezer failure which resulted in the loss of some samples.

Study II: 115 consecutive men appearing at the Andrology Clinic, National Hospital, Oslo, Norway, for fertility counselling provided a semen sample by masturbation following an instructed abstinence period of 3 days. One sample per patient was collected in the clinic, and analysed by routine criteria and the SCSA.

Sample handling

In both studies, 200 μ l aliquots of semen were admixed with 300 μ l TNE buffer (0.15 M NaCl, 0.01 M Tris-HCl, 0.001 M EDTA, disodium pH 7.4 and 10% glycerol) and frozen directly at -80°C . (N.B. current protocol freezes raw semen directly without cryoprotectants with the same data resulting.) Study I samples were shipped on dry ice to South Dakota State University for SCSA analysis, while Study II samples were measured in the flow cytometry laboratory in the Institute for Pathology, the National Hospital, Oslo, Norway.

Sperm quality measures

Study I: the semen quality measures of volume, sperm concentration, motility by CASA and morphology by WHO and Kruger strict criteria were performed for all samples following liquefaction. These data will be the subject of a separate manuscript (E.Clegg *et al.*, in preparation).

Study II: following liquefaction at room temperature for 30 min, an aliquot of spermatozoa was diluted and immobilized in 5% chloramine-T (1:10) and counted using a Makler chamber (Sefi Medical Instruments Ltd, Haifa, Israel). Additional aliquots (50 μ l) of the samples were subjected to vital staining using eosin Y and nigrosin, as well as haematoxylin, for the assessment of morphology. Microcephalic, macrocephalic, bicephalous and bicaudal spermatozoa, or cells possessing a coiled tail or a deformed or abnormally small acrosome were all classified as abnormal (WHO, 1987). The total numbers of spermatozoa with head and tail abnormalities in the ejaculate were computed. Sperm motility was assessed using a Hamilton Thorn sperm motility analyser (Danvers, MA, USA) which, in addition to other motility parameters, enabled the percentage of spermatozoa exhibiting any form of movement ($>10 \mu\text{m/s}$) and the percentage of progressive spermatozoa with movement (velocity $>25 \mu\text{m/s}$) to be computed.

SCSA

Frozen aliquots of semen were placed in a 37°C water bath until just thawed, after which samples were diluted with TNE buffer to $1\text{--}2 \times 10^6$ sperm cells per ml. 0.20-ml aliquots of diluted samples were mixed with 0.40 ml of acid-detergent solution (0.08 M HCl, 0.15 M NaCl, 0.1% Triton X-100, pH 1.2). After 30 s, the cells were stained by adding 1.2 ml acridine orange (AO) stain solution containing 6 µg AO (chromatographically purified; Cat. # 04539, Polysciences Inc., Warrington, PA, USA) per ml buffer [0.037 M citric acid, 0.126 M Na_2HPO_4 , 0.0011 M EDTA (di-sodium), 0.15 M NaCl, pH 6.0 (Darzynkiewicz *et al.*, 1976; Evenson *et al.*, 1985)]. At 3 min after the staining procedure started, fluorescence measurements were collected on 5000 cells per sample. Both flow laboratories used Cytofluorograf II flow cytometers (Ortho Diagnostics Inc., Westwood, MA, USA) equipped with ultrasense optics. The Oslo instrument used a Coherent 95, 5 W laser operated at 200 mW, and the South Dakota instrument used a Lexel 100 mW argon ion laser operated at 35 mW; both used an excitation wavelength of 488 nm. Histogram files were transferred via MULTLINK software (Phoenix Flow Systems, San Diego, CA, USA) to a PC for analysis with MULTI2D software (Phoenix Flow Systems) which included the calculation of alpha t (α_t) parameters. The South Dakota State University (SDSU) Ortho Cytofluorograf is currently interfaced to a Cicero System (Cytomation, Inc., Fort Collins, CO, USA).

The extent of DNA denaturation was quantified by the calculated parameter α_t [$\alpha_t = \text{red}/(\text{red} + \text{green})$ fluorescence; Darzynkiewicz *et al.*, 1975]. Normal, native chromatin remains structurally sound and produces a narrow α_t distribution. DNA in spermatozoa with abnormal chromatin structure has increased red fluorescence (Evenson *et al.*, 1980, 1985) which yields an α_t distribution which is usually broader, having a higher mean channel ($X\alpha_t$) and a larger percentage of cells outside the main population of cells ($\text{COMP}\alpha_t$). Standard deviation of α_t ($\text{SD}\alpha_t$) describes the extent of chromatin structure abnormality within a population. Mean green fluorescence reflects DNA content and/or degree of sperm chromatin condensation, the latter because it excludes DNA stainability.

Statistical analysis

Statistical tests used are described in the corresponding results section.

Results

Representative examples of cytograms of AO-stained spermatozoa from couples in different pregnancy outcome categories are shown in Figure 1. Individual cells were represented by dots for a total of 5000 per sample. The position of each dot was relative to the amount of green and red fluorescence. Cellular debris was demarcated by the computer-gated area in the lower left hand corner, and was excluded from data analysis. Cytogram E1 provided an example of gating between the main population without denatured DNA, and the cells outside the main population ($\text{COMP}\alpha_t$) which had decreased green fluorescence and increased red fluorescence, i.e. cells with denatured DNA. Cytogram G1 is an example of the high green fluorescence (HGRN) threshold, dictated by the upper edge of the main population of the reference sample. These cells had an increased DNA stainability, probably due to a lack of chromatin condensation. Cytogram I3 showed the presence of bacteria (arrow) in the sample. The extended line was due to the random number of clumps of individual, AO-stained bacteria in the semen.

The couple represented in cytograms A1 and A2 conceived in month 2 of the study, and had a normal pregnancy. This was an excellent quality sample with minimal cellular debris and very low $\text{COMP}\alpha_t$ values of 8.7% (A1) and 7.2% (A2). The A2 sample was obtained during month 2, within about 3 days of conception.

The couple represented in cytograms B1 and B2 conceived in month 1 of the study, and also had a normal pregnancy. This example was an exception to the high month-to-month repeatability of SCSA values (Evenson *et al.*, 1991). Sample B1, with a $\text{COMP}\alpha_t$ value of 13.3%, was obtained within 1–3 days of conception. The $\text{COMP}\alpha_t$ value for the following month (B2) was 39.5%, which is considered not compatible with fertility. This type of dramatic deterioration has previously been seen after exposure to high fever (D.Evenson, unpublished observations) or prescription drugs (Evenson *et al.*, 1991).

The couple represented in cytograms C1 and C2 experienced a miscarriage after conceiving in month 1. The $\text{COMP}\alpha_t$ value of 34.0% (C1) was just over our current threshold of 30.0%, and considered not compatible with successful pregnancy. Otherwise, this sample had a low amount of debris and a high repeatability between the two months.

Cytograms D1, D2 and D3 were obtained from a couple who conceived in month 3. There was a dramatic change in sample quality, as shown by $\text{COMP}\alpha_t$ values from months 1 to 3, as very poor (64.4%), improved (34.4%) and excellent (8.0%). The month 3 sample would have been a full spermatogenic cycle away from a possible biostress event that may have caused the very low quality in month 1. Note that the main population had essentially the same level of quality, the only difference between the samples being the percentage of cells with denatured DNA. Note also that the average $\text{COMP}\alpha_t$ was 35.5%, a mean value not considered compatible with good fertility, though SCSA data obtained near the conception date were highly compatible with our fertility threshold.

The couple represented in cytograms E did not conceive throughout the 12 months of the study. Note the high month-to-month repeatability in cluster arrangements, with moderately high $\text{COMP}\alpha_t$ values of 19.7, 22.7 and 20.0% for months 1, 2 and 3 respectively.

Cytograms F1 to F3 represent a couple who became pregnant in month 5. Samples from months 1–3 would be characterized as of good quality, with respective $\text{COMP}\alpha_t$ values of 13.1, 18.8 and 15.0%. There was a high degree of repeatability, i.e. the $\text{COMP}\alpha_t$ region had two populations of cells with approximately the same number of cells in each and whose positions were essentially identical. This implied that these abnormal cells were produced at a steady rate, and were the likely progeny of the same stem cells.

Couple G did not conceive. Monthly $\text{COMP}\alpha_t$ values between months 1 and 3 were 25.4, 22.5 and 19.0% respectively, with an overall mean of 22.3%, considered of moderate quality on that characteristic alone. However, a second abnormality was present, namely, a population with high DNA stainability (HGRN). These two characteristics are often seen to be mutually exclusive; however, this is an example where both were present and repeatable over time.

Couple H experienced an occult pregnancy which was

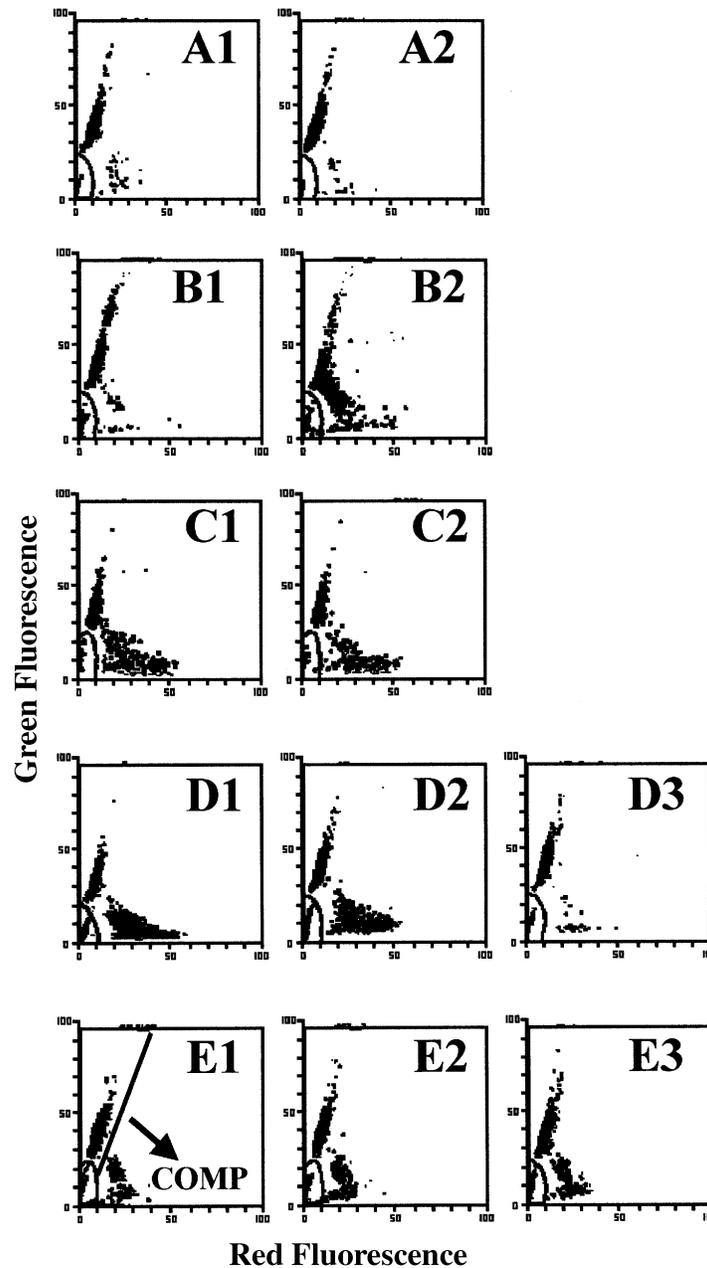


Figure 1. Legend on facing page.

detected 13 days before the semen sampling in month 3, the month with the highest $COMP\alpha_t$ value near the end of month 2. No later pregnancies occurred during months 3 to 12. There was a high month-to-month repeatability, and $COMP\alpha_t$ values were 13.0, 14.5 and 19.1% in months 1, 2 and 3 respectively.

Couple I did not conceive. This series of cytograms showed an increasing relative proportion of bacteria (arrow) to spermatozoa from month to month.

Study I

SCSA data were collected from duplicate measurements on semen samples collected monthly. Of the 165 couples with initial semen samples, 163 contributed a second sample one month later, and 84 provided a sample in month 3; only 402 of these 412 samples were available for SCSA measurement.

Each of 402 semen samples from 165 men were thawed independently and sequentially measured twice (total of 804 measurements). The number of monthly samples contributed by each man depended on pregnancy outcome. Throughout the course of these sample measurements, 136 separately frozen and thawed aliquots of a single reference sample were measured in duplicate ($n = 272$ independent measurements) for quality control. This reference sample was not characterized as fertile or infertile, but was used only for quality control. As shown previously (Evenson *et al.*, 1991), sonication of random samples produced the same data as on whole cells (data not shown).

The descriptive statistics for SCSA measurements for overall means are shown in Table I, which also includes reference sample data. Table II shows data for within-couple

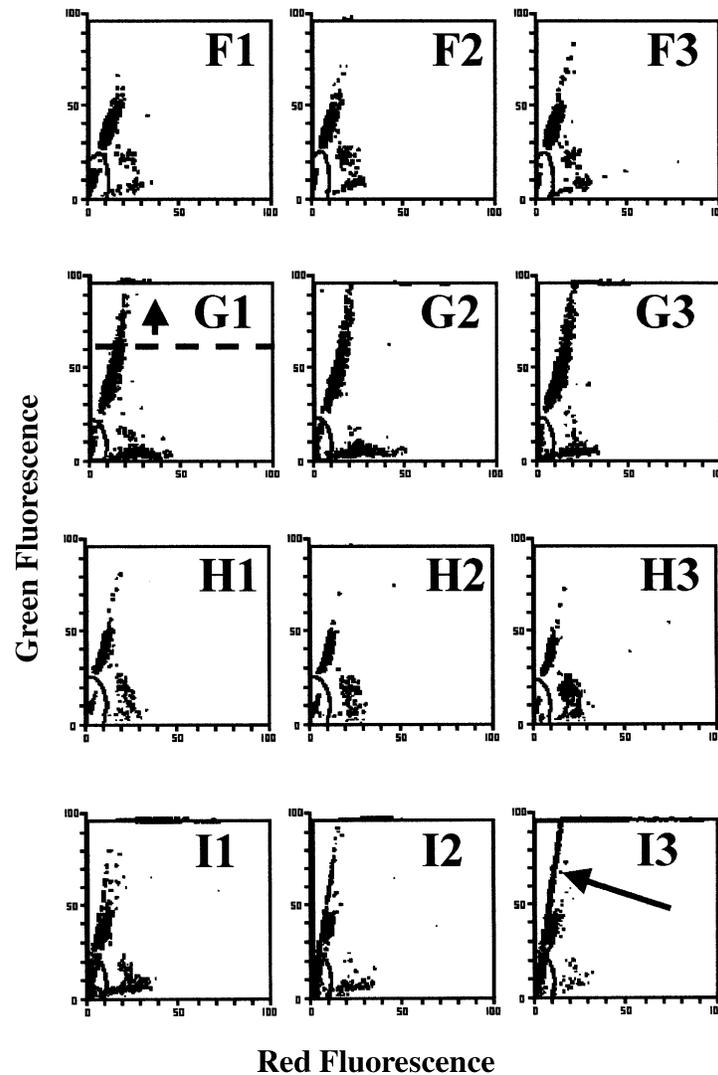


Figure 1 (continued). Sperm chromatin structure assay (SCSA)-derived cytograms, showing green (normal double-stranded DNA) versus red (denatured single-stranded DNA) fluorescence of 5000 acridine orange-stained spermatozoa from semen obtained from nine representative couples from Study I. Each couple shows a specific fertility/infertility characteristic or trait as measured by the SCSA. A = pregnancy, excellent quality, highly repeatable; B = pregnancy month 1, uncharacteristically unrepeatable; C = miscarriage, highly repeatable; D = pregnancy month 3, improved quality; E = not pregnant, highly repeatable; F = pregnancy month 5, highly repeatable; G = not pregnant, moderate quality, increased HGRN, highly repeatable; H = occult pregnancy, repeatable; I = not pregnant, repeatable. COMP = cells outside the main population. Arrows: E1, indicates COMP; G1, indicates HGRN threshold; I3, indicates presence of bacteria.

Table I. Descriptive statistics of sperm chromatin structure assay (SCSA) variables for reference sample ($n = 136$) and Study I overall within-couple means ($n = 165$ couples)

Descriptive statistic	$X\alpha_t$	$SD\alpha_t$	$COMP\alpha_t$	Red (mean)	Green (mean)	Total (mean)	HGRN
<i>Reference</i>							
Mean	252.1	142.6	12.7	127.28	390.2	257.7	8.4
SD	6.3	7.8	1.0	3.7	6.3	3.7	1.8
<i>Study I</i>							
Mean	247.5	150.8	13.7	118.2	382.3	248.6	9.9
SD	39.0	40.5	7.2	17.8	25.7	15.5	7.2
Coeff. of variation (%)	15.7	26.9	52.3	15.1	6.7	6.2	71.9

$SD\alpha_t$ = variation of the α_t distribution; $X\alpha_t$ = mean of the α_t distribution; $COMP\alpha_t$ = percentage of cells outside the main population of the α_t distribution; HGRN = percentage of cells with high green fluorescence.

means. Each observation was assigned to one of four pregnancy outcome groups: group 1 = normal pregnancy in months 1–3; group 2 = miscarriage; group 3 = normal

pregnancy in months 4–12; and group 4 = no pregnancy in months 1–12. SCSA values obtained by measuring the 148 semen samples from the 73 couples achieving pregnancy

Table II. Least-squares means and SE by pregnancy outcome group for selected SCSA measurements from Studies I and II ($n = 165$)

Group	Pregnancy outcome	n^a	n^b	$X\alpha_t$	$SD\alpha_t$	$COMP\alpha_t$	HGRN
Study I ^c							
1	PG 3 months	165	402	234.6	137.9	11.2	8.95
	SE	73	148	4.6	4.6	0.85	0.80
2	Miscarriage	18	39	246.9	151.4	13.4	9.20
	SE			9.2	9.1	1.69	1.61
3	PG 4–12 months	40	118	255.1**	157.9**	15.5**	8.78
	SE			5.9	6.0	1.09	1.07
4	Increase (%)			8.7	14.7	38.3	-1.9
	No pregnancy	31	89	270.3***	173.7***	17.2***	15.03***
	SE			6.7	6.8	1.24	1.22
Study II	Increase (%)	115	115	15.2	30.0	53.6	68.0
	SE			308.6***	194.2***	24.0***	NA
	Increase (%)			31.5	40.8	114.2	-

^aCouples; ^bSamples; ^cOccult group not shown ($n^a = 3$; $n^b = 8$); ** $P < 0.01$; *** $P < 0.001$. PG = pregnancy; NA = not applicable. For abbreviations, see Table I.

Table III. Least-squares means and SE by pregnancy outcome group for log of selected SCSA measurements from Studies I and II ($n = 165$)

Fertility group	n^a	n^b	$X\alpha_t$	$SD\alpha_t$	$COMP\alpha_t$	% HGRN
Study I ^c						
1 PG in 3 months	73	148	5.45	4.90	2.36	2.06
SE			0.016	0.029	0.052	0.064
2 Miscarriage	18	39	5.50	4.97	2.49	2.08
SE			0.032	0.059	0.105	0.129
3 PG 4–12 months	40	118	5.52**	5.03**	2.60**	1.99
SE			0.021	0.039	0.068	0.086
4 No pregnancy	31	89	5.58**	5.12**	2.69**	2.46**
SE			0.024	0.044	0.077	0.098

^aCouples; ^bsamples; ^coccult group not shown ($n^a = 3$; $n^b = 8$); ** $P < 0.01$. PG = pregnancy. For abbreviations, see Table I.

within months 1–3 (group 1) are shown in row 1; these values serve as the ‘fertile’ reference standard for the other data in both Studies I and II.

In the first three months of study, 10 couples experienced an occult pregnancy. Seven of these subsequently achieved a successful pregnancy while on study and were analysed in that latter outcome group rather than as occult. With only three remaining couples in the occult group, statistical comparisons were not feasible and are not shown.

Least-squares means by pregnancy outcome group were calculated for selected flow cytometry parameters by mixed linear model methods using the MIXED procedure of the SAS® system software (Littell *et al.*, 1996). This procedure takes into account the covariance structure among repeated measurements from a given subject. Observations for this analysis were the within-couple means (averaged across duplicates and months). The least-squares mean for pregnancy outcome group 1 was compared with the means of each of the other groups, using the MIXED procedure. The MIXED model least-squares analysis was repeated using the log of flow cytometer values, as the data had non-normal distribution (Table III).

Repeatability (intraclass correlation) of flow cytometer measurements was calculated between duplicate measurements of an individual semen sample and between monthly samples

of a given couple (Table IV). Components of variance (between couples, between months within couple, and between duplicates within month) used in the calculation of repeatability were computed by restricted maximum likelihood (REML) methods using the VARCOMP procedure of the SAS® system software (SAS Institute Inc., 1985). Duplicate measurements from an individual semen sample were highly repeatable (97.3% to 99.1% across all pregnancy outcome groups), indicating a high degree of consistency for flow cytometer measurements. Repeatability of measurements taken in different months from the same man varied somewhat across the different flow cytometer parameters, as seen in Figure 1. For several parameters, month-to-month measurements tended to be somewhat less repeatable among couples becoming pregnant in months 4–12 (group 3) than among couples in group 1 (pregnancy in months 1–3) or group 4 (not pregnant). This was consistent with the observation that several couples in group 3 had relatively poor flow cytometer values in the first sample month, followed by improved values in the following two months (Figure 1, cytograms D1–3). Compared with group 3, a greater proportion of couples in all other groups tended to have consistently good or poor readings across months (Table IV).

Due to previous (Evenson and Jost, 1994) SCSA analysis of a variety of semen samples from couples having experienced spontaneous abortions, we postulated that the SCSA data from

Table IV. Repeatability (%) of SCSA measurements by pregnancy outcome group^a

Pregnancy outcome group ^b		X α_t	SD α_t	COMP α_t	Red (mean)	Green (mean)	Total (mean)	HGRN
All groups ^c	REPm	74.8	84.3	73.7	80.5	77.6	90.4	94.2
	REPd	98.4	99.1	98.8	98.2	97.3	98.4	98.5
1 PG month 1–3	REPm	76.7	80.4	71.8	86.5	82.4	88.8	89.9
	REPd	95.7	98.3	97.4	96.4	95.4	97.4	98.1
2 Miscarriage	REPm	87.2	90.8	84.2	86.1	88.6	89.2	88.1
	REPd	98.0	99.2	98.2	98.1	97.3	98.5	98.5
3 PG month 4–12	REPm	64.4	80.9	65.4	68.6	74.3	92.4	95.7
	REPd	98.8	99.1	99.2	98.7	98.4	98.7	99.2
4 No pregnancy	REPm	85.8	86.8	84.0	90.8	79.6	89.8	95.2
	REPd	98.8	99.4	99.0	98.3	97.4	98.7	98.2

^aREPm = between-month repeatability (within couple); REPd = between-duplicate repeatability (within couple and month); ^boccult group not shown; ^c*n* = 165 couples.

PG = pregnancy. For abbreviations, see Table I.

Table V. Frequency (%) of observations with low, moderate, or high COMP α_t values for each pregnancy outcome group (*n* = 165 couples)

Pregnancy outcome group ^a	COMP α_t value		
	Low (≤ 15)	Moderate (15–30)	High (≥ 30)
1 = PG month 1–3	83.6	16.4	0.0
2 = Miscarriage	61.1	38.9	0.0
3 = PG month 4–12	57.7	32.5	10.0
4 = No pregnancy	51.6	29.0	19.4

^aOccult group not shown. PG = pregnancy. For abbreviations, see Table I.

the miscarriages would have overall statistically significant values. Although this was not the case for mean values (Table II), using cut-off values for both levels of DNA denaturation and increased DNA stainability, SCSA data predicted seven of 18 miscarriages (Tables VI and VII), consistent with a theoretical but unproven 1:1 ratio for male:female caused miscarriages.

Of the variables evaluated, COMP α_t appeared to be the best indicator to predict whether a couple would not become pregnant based on logistic regression results. All couples with abnormally high values of COMP α_t (≥ 30) experienced either delayed pregnancy or no pregnancy (Table V). About 84% of couples in group 1 had relatively low values of COMP α_t (< 15). Of importance, none of the men in group 1 had a COMP α_t value above 30% (with one explainable exception), which is the cut-off value currently adopted as being categorized as not compatible with good fertility. A low COMP α_t value for natural fertility obviously does not guarantee conception because of the many other factors involved in the control of fertilization. As seen in Table V, 52% of the couples in group 4 (no pregnancy) also had low COMP α_t values; thus, the chromatin was normal at the time of measurement, but other reasons prevailed for lack of fertility. In our other animal and human fertility experiments, a COMP α_t > 30 % and good fertility have not been seen together. Similarly, with regard to the HGRN staining samples, of which we have little understanding at this point, group 4 included a much greater proportion (32.3%) of abnormally high HGRN values than did

Table VI. Frequency (%) of observations with normal or high % high green (HGRN) fluorescence values for each pregnancy outcome group (*n* = 165 couples)

Pregnancy outcome group ^a	HGRN value	
	Normal (< 15)	High (≥ 15)
1 PG month 1–3	89.0	11.0
2 Miscarriage	94.4	5.6
3 PG month 4–12	92.5	7.5
4 No pregnancy	67.7	32.3

^aOccult group not shown. PG = pregnancy. For abbreviations, see Table I.

group 1 (Table VI). The frequency percentage of observations of COMP α_t and HGRN for each group are shown in Table VII.

One or more diagnostic tests could be developed to identify couples likely to experience reproductive problems due to poor sperm chromatin. A positive test result could be defined as a prediction of possible reproductive failure (pregnancy outcome groups 2 and 4), whereas a negative test would indicate likely reproductive success (groups 1 or 3). Using only SCSA data presented here, about 46% [24 of 52 (including three couples failing to become pregnant after experiencing occult pregnancy)] of couples who actually experienced reproductive problems in this study would have been predicted to have a problem (i.e. tested positive), assuming cut-off values of ≤ 15 % (negative) versus > 15 % (positive) for mean COMP α_t (Table VIII). If it were to be assumed, for example, that a test result for possible reproductive success was positive if the value of either HGRN (≥ 15 %) or COMP α_t (> 15 %) was above a cut-off point, then about 52% (27 of 52) of couples who actually experienced reproductive problems in this study would have been predicted to have a problem (Table IX). These predicted values may have been higher if the values of HGRN and COMP α_t near the time of conception had been used rather than the mean value over several months (Figure 1, cytograms D1–3). The sensitivity of the test may have been affected by changing cut-off values, or by adding additional criteria.

Study II

Semen samples obtained from 115 consecutive men arriving at the Andrology Laboratory, the National Hospital, Oslo,

Table VII. Frequency (%) of observations by combination of COMP α_t and HGRN levels for each pregnancy outcome group ($n = 165$ couples)

Pregnancy outcome group ^a	Normal HGRN (<15)			High HGRN (≥ 15)		
	Low COMP α_t	Moderate COMP α_t	High COMP α_t	Low COMP α_t	Moderate COMP α_t	High COMP α_t
1 PG month 1–3	76.7	12.3	0.0	6.9	4.1	0.0
2 Miscarriage	61.1	33.3	0.0	0.0	5.6	0.0
3 PG month 4–12	55.0	30.0	7.5	2.5	2.5	2.5
4 No pregnancy	41.9	19.4	6.5	9.7	9.7	12.9

^aOccult group not shown. PG = pregnancy. For abbreviations, see Table I.

Table VIII. Numbers of couples experiencing reproductive failure or success by COMP α_t prediction groups

Prediction	Actual reproductive outcome	
	Failure (groups 2 or 4)	Success (groups 1 or 3)
Failure (COMP $\alpha_t > 15\%$)	24	29
Success (COMP $\alpha_t \leq 15\%$)	28	84

For abbreviations, see Table I.

Table IX. Numbers of couples experiencing reproductive failure or success by COMP α_t /HGRN cut-off groups

Prediction	Actual reproductive outcome	
	Failure (groups 2 or 4)	Success (groups 1 or 3)
Failure (COMP $\alpha_t > 15$ or HGRN ≥ 15)	27	35
Success (COMP $\alpha_t \leq 15$ or HGRN <15)	25	78

For abbreviations, see Table I.

Table X. Norway infertility clinic data

Fertility variable	n	Mean	Minimum	Maximum
X α_t	115	302.1	213.3	492.0
SD α_t	115	181.8	77.9	283.5
COMP α_t (%)	115	24.4	5.3	68.9
Volume (ml)	115	3.8	0.5	11.5
Concentration ($\times 10^6$)	115	33.4	0.0	270.0
Motility (%)	109	55.3	4.0	98.0
Progressive motility (%)	110	25.1	1.0	60.0
Abnormal (%)	113	60.7	32.0	95.0

For abbreviations, see Table I.

Norway were measured for the classical parameters of semen volume, sperm concentration, and percentage motility, progressive motility and normal morphology by WHO criteria. An aliquot of all samples was frozen, thawed and analysed by the SCSA.

The range of sperm quality criteria was much broader than observed in Study I, where the couples were selected for no known risk of infertility. As shown in Table X, these samples ranged from near-azoospermic to very low motility to nearly

all with abnormal morphology. Note in Table XI that there were no biologically meaningful correlations between classical criteria of sperm quality and SCSA parameters; for example, a correlation of $r = -0.23$, $P < 0.05$ between percentage of cells showing DNA denaturation (COMP α_t) and progressive motility practically stated that dead cells were not synonymous with cells having abnormal chromatin structure; these data agree with other studies with human spermatozoa (Schradler *et al.*, 1988; Evenson *et al.*, 1991).

Comparison of SCSA values within Study I, and between Studies I and II

The SCSA data variables in Study I had increased values (lower quality) related to increased time to pregnancy and no pregnancy. These increases were shown as significant percentage increases relative to the fertile standard values (pregnancy in 3 months) for two of the other three groups in Study I, as well as the mean of the values for Study II (see Table II). The SCSA values from 118 semen samples from 40 couples achieving pregnancy in months 4–12 were significantly ($P < 0.01$) higher, representing 9, 15 and 38% increased values for X α_t , SD α_t and COMP α_t respectively. The 89 semen samples obtained from 31 couples not achieving pregnancy by month 12 had highly significant ($P < 0.001$) increases in SCSA values, representing 15, 30, 54 and 68% increased values for X α_t , SD α_t , COMP α_t and % immature (HGRN) respectively. The 115 samples from Study II had mean values for X α_t , SD α_t and COMP α_t that were 32, 41 and 114% greater respectively, than the fertile reference group in Study I.

Discussion

Sperm chromatin integrity, measured by the SCSA, and defined as the susceptibility to DNA denaturation *in situ*, is predictive of time to conception or, primarily, failure to conceive. The SCSA provides an easily and rapidly obtained diagnostic and prognostic evaluation of one measure of semen quality as shown in three animal studies (Ballachey *et al.*, 1987, 1988; Evenson *et al.*, 1994), a preliminary human study (Evenson *et al.*, 1980), and the present study.

In Study I, using the SCSA values from the couples conceiving during the first 3 months as a 'standard' for 'highly fertile' individuals, the increased SCSA values obtained from individuals conceiving during the next 9 months were significantly different ($P < 0.01$). More dramatically, those indi-

Table XI. Norway infertility clinic semen variable correlations

	SD α_t	COMP α_t	Volume (ml)	Concentration ($\times 10^6$)	Motility (%)	Progressive motility (%)	Abnormal (%)
X α_t	0.84***	0.86***	-0.11	-0.17	-0.31***	-0.28**	0.21*
SD α_t		0.64***	-0.10	-0.05	-0.40***	-0.45***	0.31***
COMP α_t			-0.02	-0.27**	-0.30**	-0.23*	0.21*

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. Note: ≤ 15 = not correlated; 0.15–0.39 poorly correlated; 0.4–0.45 moderately correlated.

For abbreviations, see Table I.

viduals not conceiving over the 12-month study yielded SCSA values that had greater significant differences ($P < 0.001$) from the 'standard' group, providing strong evidence that sperm chromatin structure was reflective of fertility potential. This viewpoint is strengthened by SCSA values from Study II, in which the quality of semen samples from men of couples suspected of possible infertility were significantly ($P < 0.001$) different from those of the standard fertile group. Even so, in Study II and previous studies (Schrader *et al.*, 1988; Evenson *et al.*, 1991) the SCSA data were not well correlated with classical semen parameters, emphasizing that SCSA data are independent to a significant degree. Of greatest importance to the andrology clinic are those cases where the classical criteria are within normal ranges, but the SCSA values are poor, and probably not compatible with good fertility (COMP $\alpha_t \geq 30\%$). For many reasons, before embarking on extensive assisted reproduction treatment, it would be of great value to the patient and the clinic to know the quality of the sperm chromatin, not measurable by any other efficient method.

In Study I, we had postulated that a significant correlation would be found between sperm chromatin quality and occult pregnancy. There were not sufficient numbers ($n = 3$) to detect differences, and only a trend was noted between occult pregnancy and reduced total DNA stainability, probably due to chromatin hypercondensation factors and/or hyperoxidation of cysteine -SH groups.

High-quality SCSA data are not directly predictive of good fertility potential, since natural fertility requires many other positive traits such as motility, morphology, acrosome integrity, etc. The inverse, i.e. poor SCSA data, are predictive of subfertility/infertility. For the purposes of intracytoplasmic sperm injection (ICSI), there is ample justification for an SCSA measurement since there is generally poor correlation between SCSA measures and motility, morphology and viability. Thus, in those cases reported in the literature where there is a concern whether to utilize an ejaculate with little or no motility or to obtain a testicular sample in hope of better quality spermatozoa, if the SCSA data are good then it would be instructive to consider using the ejaculate for ICSI. On the other hand, if the SCSA data are of poor quality, even in the face of 'normal' motility and morphology, then serious consideration should be given not to utilize that sample for fertilization. In cases where the testis may have been exposed to a perturbing agent, e.g. high fever (D.P.Evenson, unpublished observations), and the resulting data are of poor quality, it is suggested that another sample be evaluated by the SCSA after one spermatogenic cycle.

For some time we have wrestled with the question of what percentage of spermatozoa with denatured DNA is the cut-off level for considering the sample as compatible or possibly non-compatible with fertility. From data in this study and other unpublished data, that threshold is currently set at 30%. However, it is important to discuss that this does not mean that the other 70% of spermatozoa have fully normal chromatin. This value means only that, given the physical conditions imposed on the spermatozoa to induce DNA denaturation, that 30% of the spermatozoa crossed that threshold. As has been discussed previously (Evenson and Jost, 1994), stronger inducing conditions for DNA denaturation will cause a higher percentage of sperm nuclei to cross that threshold. It is important to note, however, that the shift in different samples is proportional, i.e. the dose–response curves are parallel, and thus no new valuable information is obtained by varying physical conditions. Thus, an exact set of conditions as detailed in Materials and methods for treatment of spermatozoa and measurement by flow cytometry (Evenson and Jost, 1994) strictly defines the SCSA, and all sample data can be related to other samples.

Furthermore, whereas the variable COMP α_t describes only the percentage of cells demonstrating DNA denaturation, SD α_t describes the variation from the mean. In other studies, e.g. effects of smoking on sperm head morphometry (Rubes *et al.*, 1998), the variation around the actual mean values rather than the mean values was indicative of smoke-induced changes. The SD α_t has been the parameter most highly correlated with rodent reproductive toxicants, dose and time parameters. A high SD α_t is further suggestive that damaged chromatin exists in the 'main population', and even though any single cell may not have crossed the 'threshold' to be counted as a DNA denatured (COMP α_t) cell, the extent of damage may be sufficient to have negative fertility/pregnancy outcome factors.

What is the physical damage in sperm chromatin that is directly or indirectly being detected by the SCSA? A recent study (Aravindan *et al.*, 1997) selected 23 human sperm samples with COMP α_t values ranging from 5% to 95%. Aliquots of these samples were analysed for DNA strand breaks by two independent methods: (i) flow cytometric TUNEL assay for 3'-OH broken ends of DNA (Gorzcyca *et al.*, 1993); and (ii) DNA fragments by single cell gel electrophoresis analysis (modified COMET assay; Singh *et al.*, 1988). Correlations between the percentage of cells demonstrating denatured DNA by SCSA (COMP α_t) were very strong with TUNEL data ($r = 0.859$, $P < 0.001$) and COMET data ($r = 0.973$, $P < 0.001$). The latter assay, being the more

sensitive to DNA strand breaks, showed a near 1:1 relationship between the percentage of cells with denatured DNA and percentage of cells with 'comets' consisting of DNA fragments. These data provide strong evidence that the SCSA data are showing DNA strand breaks which, by definition, would be considered negative for fertility and pregnancy outcome. The origin of these strand breaks is not clear; however, recent studies in our laboratory show that oxidative stress agents, e.g. exposure to H₂O₂, produce the same SCSA pattern in a dose-response relationship (Larson *et al.*, 1998), and we currently favour the idea that most damage is due to oxidative stress to DNA. Also, consideration has been given to the idea that they may represent an altered form of apoptosis where the DNA fragments can exist in cells that are otherwise 'healthy' from the standpoint of mitochondrial function (motility) and membrane viability (viable) (Gorzcyca *et al.*, 1993; Aravindan *et al.*, 1997). Other studies are in progress to evaluate this possibility.

In summary, these two studies, as well as previous studies, provide a solid rationale for andrology clinics to measure SCSA in all or selected semen samples before counselling patients on fertility issues, especially expensive assisted reproduction procedures.

How useful is the SCSA?

As discussed by Matson (1997), for a test to be useful, it must first be reproducible such that similar results are obtained each time a man is tested, assuming that all other factors are unchanged. Two aspects of this are: (i) the robustness of the method in performing to the same level each time; and (ii) the stability of the parameter being measured. The SCSA is among the best of semen assays with regard to both of these aspects. For the first, as shown in this study—and in measurements of thousands of other mammalian sperm samples—the repeatability is invariably in the 0.98 to 0.99 range. Correlations between measurements of frozen-thawed aliquots of the same samples on six different flow cytometers at six different institutions and times were also in the 0.97 to 0.98 range; this even included using both orthogonal laser-driven FCM versus mercury arc epi-illumination FCM. Furthermore, the use of different computer software for calculations of SCSA variables showed correlation coefficients of the same variable from different programs to all be 0.98 or higher ($P < 0.001$; Evenson *et al.*, 1995). These results far exceed the repeatability of many tests, even computer-driven CASA systems, done on different instruments and in different laboratories. With regard to the second aspect, i.e. stability of the parameter being measured and also biological variability, SCSA variables were the most stable over 8 months in an unexposed population (Schrader *et al.*, 1988; Evenson *et al.*, 1991). It is noted, however, that biological and physical stress to mammalian testes does produce dramatic shifts in SCSA variables which have been shown to be highly reproducible and dose/time responsive.

The technical variability of the SCSA is practically negligible; repeatability of the same samples in this study was in the range of 0.98–0.99. More importantly, the repeatability of the reference sample frozen in many aliquots and used

repeatedly throughout the weeks of measurements was 0.98 ($P < 0.001$), thus showing excellent internal quality control. External quality assurance has been addressed above, with aliquots of the same sample being measured in different laboratories with different instruments, and by different technicians.

The whole purpose of performing a diagnostic test in the evaluation of the male partner of a suspected infertile couple is to learn whether his fertility is impaired (Matson, 1997). The test must then have a threshold above and below which it will provide discrimination and predictive capabilities, and with little overlap between fertile and infertile men. As an example of this problem, a 'normal' sperm concentration is regarded as being $>20 \times 10^6/\text{ml}$ (WHO, 1992); however, this was derived from a study (Macleod and Gold, 1951) where the median sperm concentrations in the fertile and infertile groups were $90 \times 10^6/\text{ml}$ and $74 \times 10^6/\text{ml}$ respectively. Obviously, this criterion has little power to identify prospectively infertile men. In sharp contrast, in this study, men with 30% or more spermatozoa showing DNA denaturation were subfertile/infertile, thus showing reasonable sensitivity. With regard to specificity, i.e. the proportion of normal individuals correctly predicted to have a fertility problem and who actually experienced a fertility problem was 52% (27/52). If however, SCSA data had been based on a single measurement taken within several days of attempted conception, rather than on a mean of several monthly samples, the predictive power would have significantly increased.

Using a combination of selected cut-off values for percentage spermatozoa with denatured DNA and/or increased DNA stainability (HGRN), the SCSA predicted 39% of miscarriages (seven of 18; six from COMP α_4 and one from HGRN) (Tables V and VI). Since over 50% of miscarriage problems are likely to be attributed to the female, this result was strongly related to the prediction. Again, these results were based on within-couple mean values for SCSA parameters. Inspection of individual (monthly) samples can be more informative. For example, some had poor values in month 1, followed by improved values in later months. Even though the overall mean COMP α_4 was $>15\%$, the value at the time of pregnancy was $<15\%$.

Acknowledgements

This work was supported in part by EPA Grant Numbers R820968-01 and R827019, Norwegian Council for Science and the Humanities (NAVF), March of Dimes Grant Number FY91-0528 and National Science Foundation Grant Number EHR-9108773 and the South Dakota Futures Fund. This is South Dakota Agricultural Experiment Station Publication Number 3095 of the journal series.

References

- Amann, R.P. (1989) Can the fertility potential of a seminal sample be predicted accurately? *J. Androl.*, **10**, 89–98.
- Aravindan, G.R., Bjordahl, J., Evenson, D.P. *et al.* (1997) Susceptibility of human sperm to *in situ* DNA denaturation is strongly correlated with DNA strand breaks identified by single-cell electrophoresis. *Exp. Cell Res.*, **236**, 231–237.

- Ballachey, B.E., Hohenboken, W.D. and Evenson, D.P. (1987) Heterogeneity of sperm nuclear chromatin structure and its relationship to fertility of bulls. *Biol. Reprod.*, **36**, 915–925.
- Ballachey, B.E., Saacke, R.G. and Evenson, D.P. (1988) The sperm chromatin structure assay: relationship with alternate tests of sperm quality and heterospermic performance of bulls. *J. Androl.*, **9**, 109–115.
- Darzynkiewicz, Z., Traganos, F., Sharpless, T. *et al.* (1975) Thermal denaturation of DNA *in situ* as studied by acridine orange staining and automated cytofluorometry. *Exp. Cell Res.*, **90**, 411–428.
- Darzynkiewicz, Z., Traganos, F., Sharpless, R. *et al.* (1976) Lymphocyte stimulation: a rapid multiparameter analysis. *Proc. Natl Acad. Sci. USA*, **73**, 2881–2884.
- Evenson, D.P. and Jost, L.K. (1994) Sperm chromatin structure assay: DNA denaturability. In: Darzynkiewicz, Z., Robinson, J.P. and Crissman, H.A. (eds), *Methods in Cell Biology, Vol. 42, Flow Cytometry*, Second Edition. Academic Press, Inc., Orlando, pp. 159–176.
- Evenson, D. and Jost, L. (1998) Utility of the sperm chromatin structure assay in the infertility clinic. XX Congress of the International Society for Analytical Cytology, March 1–6, Colorado Springs, CO.
- Evenson, D.P. and Melamed, M.R. (1983) Rapid analysis of normal and abnormal cell types in human semen and testis biopsies by flow cytometry. *J. Histochem. Cytochem.*, **31**, 248–253.
- Evenson, D.P., Darzynkiewicz, Z. and Melamed, M.R. (1980) Relation of mammalian sperm chromatin heterogeneity to fertility. *Science*, **240**, 1131–1133.
- Evenson, D.P., Klein, F.A., Whitmore, W.F. *et al.* (1984) Flow cytometric evaluation of sperm from patients with testicular carcinoma. *J. Urol.*, **132**, 1220–1225.
- Evenson, D.P., Higgins, P.H., Grueneberg, D. *et al.* (1985) Flow cytometric analysis of mouse spermatogenic function following exposure to ethylnitrosourea. *Cytometry*, **6**, 238–253.
- Evenson, D.P., Baer, R.K. and Jost, L.K. (1989) Long-term effects of triethylenemelamine exposure on mouse testis cells and sperm chromatin structure assayed by flow cytometry. *Environ. Mol. Mutagen.*, **14**, 79–89.
- Evenson, D.P., Jost, L.K., Baer, R.K. *et al.* (1991) Individuality of DNA denaturation patterns in human sperm as measured by the sperm chromatin structure assay. *Reprod. Toxicol.*, **5**, 115–125.
- Evenson, D.P., Jost, L.K. and Baer, R.K. (1993a) Effects of methyl methanesulfonate on mouse sperm chromatin structure and testicular cell kinetics. *Environ. Mol. Mutagen.*, **21**, 144–153.
- Evenson, D.P., Jost, L.K. and Gandy, J.G. (1993b) Glutathione depletion potentiates ethyl methanesulfonate-induced susceptibility of rat sperm DNA denaturation *in situ*. *Reprod. Toxicol.*, **7**, 297–304.
- Evenson, D.P., Thompson, L. and Jost, L. (1994) Flow cytometric evaluation of boar semen by the sperm chromatin structure assay as related to cryopreservation and fertility. *Theriogenology*, **41**, 637–651.
- Evenson, D., Jost, L., Gandour, D. *et al.* (1995) Comparative sperm chromatin structure assay measurement on epillumination and orthogonal axes flow cytometers. *Cytometry*, **19**, 295–303.
- Fossa, S.D., De Angelis, P., Kraggerud, S.M. *et al.* (1997) Prediction of post-treatment spermatogenesis in patients with testicular cancer by flow cytometric sperm chromatin structure assay. *Commun. Clin. Cytom.*, **30**, 192–196.
- Gorzycza, W., Gong, J. and Darzynkiewicz, Z. (1993) Detection of DNA strand breaks in individual apoptotic cells by the *in situ* terminal deoxynucleotidyl transferase and nick translation assays. *Cancer Res.*, **53**, 945–951.
- Ibrahim, M.E., Moussa, M.A.A. and Pedersen, H. (1988) Sperm chromatin heterogeneity as an infertility factor. *Arch. Androl.*, **21**, 129–133.
- Johnson, L.A., Flook, J., Look, M. and Pinkel, D. (1987) Flow sorting of X and Y chromosome-bearing spermatozoa into two populations. *Gamete Res.*, **16**, 1–9.
- Johnson, L.A., Flook, J. and Hawk, H. (1989) Sex preselection in rabbits: live births from X and Y sperm separated by DNA and cell sorting. *Biol. Reprod.*, **41**, 199–203.
- Larson, K., Jost, L. and Evenson, D. (1998) Detection of oxidative stress in mammalian sperm DNA. XX Congress of the International Society for Analytical Cytology, March 1–6, Colorado Springs, CO.
- Littell, R.C., Milliken, G.A., Stroup, W.W. *et al.* (1996) *SAS® System for Mixed Models*. SAS Institute Inc., Cary, NC.
- Macleod, J. and Gold, R. (1951) Male factors in fertility and infertility: spermatozoon counts in 1000 men of known fertility and in 1000 cases of infertile marriage. *J. Urol.*, **66**, 436–449.
- Matson, P.L. (1997) Clinical value of tests for assessing male infertility. *Bailliere's Clin. Obstet. Gynaecol.*, **11**, 641–654.
- Rubes, J., Lowe, X., Moore, D. *et al.* (1998) Cigarette-smoking lifestyle is associated with increased sperm disomy in teenage men. *Fertil. Steril.*, **70**, 715–723.
- SAS Institute Inc. (1985) *SAS® Users Guide: Statistics*, Version 5 Edition. SAS Institute Inc., Cary, NC.
- Schrader, S.M., Turner, T.W., Breitenstein, M.J. *et al.* (1988) Longitudinal study of semen quality of unexposed workers. I. Study overview. *Reprod. Toxicol.*, **2**, 183–190.
- Shapiro, H. (ed.) (1995) *Practical Flow Cytometry*. 3rd edn. Wiley-Liss, New York.
- Singh, N.P., McCoy, M.T., Tice, R.R. *et al.* (1988) A simple technique for quantitation of low levels of DNA damage in individual cells. *Exp. Cell Res.*, **175**, 184–191.
- Spano, M., Kolstad, A.H., Larsen, S.B. *et al.* (1998) The application of the flow cytometric sperm chromatin structure assay in epidemiological studies. *Hum. Reprod.*, **13**, 2495–2505.
- Sterik, K., Rosenbusch, B., Sasse, V. *et al.* (1989) The acridine orange test. A new parameter in assessing the fertilizing capacity of spermatozoa. *Zentralbl. Gynakol.*, **111**, 1361–1367.
- Tejada, R.I., Mitchell, J.C., Norman, A. *et al.* (1984) A test for the practical evaluation of male fertility by acridine orange (AO) fluorescence. *Fertil. Steril.*, **42**, 87–91.
- World Health Organization (1987) *WHO Laboratory Manual for the Examination of Human Semen and Sperm–Cervical Mucus Interaction*. 2nd edn. Cambridge University Press, Cambridge.
- World Health Organization (1992) *WHO Laboratory Manual for the Examination of Human Semen and Sperm–Cervical Mucus Interaction*. 3rd edn. Cambridge University Press, Cambridge.
- Zinaman, M.J., Clegg, E.D., Brown, C.C. *et al.* (1996) Estimates of human fertility and pregnancy loss. *Fertil. Steril.*, **65**, 503–509.

Received on September 2, 1998; accepted on December 23, 1998