Human Reproduction Vol.18, No.10 pp. 2082–2088, 2003 DOI: 10.1093/humrep/deg430

Biological variation of seminal parameters in healthy subjects

C.Álvarez[1,4], J.A.Castilla[2,5], L.Martínez[2], J.P.Ramírez[3], F.Vergara[3] and J.J.Gaforio[4]

[1]Unidad de Reproducción, Clínica Avicena, Jaén, [2]Unidad de Reproducción, Hospital Virgen de las Nieves, Granada, [3]CEIFER, Granada and [4]Departamento de Ciencias de la Salud, Universidad de Jaén, Jaén, Spain

[5]To whom correspondence should be addressed at: Unidad de Reproducción, Hospital Virgen de las Nieves, E-18014, Granada, Spain. E-mail: josea.castilla.sspa@juntadeandalucia.es

BACKGROUND: A study was undertaken to assess the components of biological variation of seminal parameters in healthy subjects. METHODS: Twenty donor candidates were included in a 10-week follow-up study. Within- and between-subject biological variation, indices of individuality and heterogeneity, coefficient of reliability, critical differences, analytical goals and the lowest value observed with a <5% probability of having a true value less than the World Health Organization (1999) reference value were estimated for the following seminal parameters: concentration, total motility (WHO grades a + b + c), progressive motility (grades a + b), rapid progressive motility (grade a), sperm morphology and vitality. All analysis was performed by a single technician according to WHO 1999 guidelines for routine semen analysis. Analytical variation was assessed on different types of quality control material (frozen straws, sperm suspension, videotape, and slides) and at different (low, medium, high) quality levels. RESULTS: The analytical variation observed depended on the quality control material used and the level of semen quality. Concentration was the semen parameter with highest within- and between-subject variation, and vitality the lowest. Indices of individuality were all <0.7, and coefficients of reliability were high (0.68–0.84). The critical difference for sequential values significant at $P < 0.05$ for vitality, progressive motility and morphology (34.4, 49.2 and 58.0% respectively) were lower than for concentration (77.8%). CONCLUSIONS: The study results showed that conventional reference values for seminal parameters have little diagnostic value because of their marked individuality, though seminal parameters can be useful for assessing differences in an individual's serial results, in particular of progressive motility, morphology and vitality.

Key words: analytical goals/analytical variation/biological variation/semen

Introduction

If a series of samples is taken from one individual for a particular laboratory test, then the results are not all exactly the same. The test results of any person vary over time, due to three factors: (i) pre-analytical influences (in the case of semen, sexual abstinence period, transport of the sample to the laboratory, etc.); (ii) analytical random (precision) and systematic error (bias); and (iii) inherent biological variation around the homeostatic setting point (this is called within-subject biological variation). If the same tests were to be performed on different individuals, the means of the results would not be found all to be exactly the same. Individual homeostatic setting points usually vary, and this difference between individuals is called between-subject biological variation. When for an analyte, within-subject biological variation is much less than between-subject biological variation, the analyte is said to have marked individuality (Fraser, 2001). Although substantial published data are available on biological variation for many analytes, the most recent and extensive listing (Ricos *et al*., 1999) does not include data on semen parameters.

Applications in the clinical laboratory of biological variation data include setting analytical quality specifications, determining the change that must occur in an individual's serial results before the change is significant, calculating the reliability coefficient to be used in epidemiology, and deciding the utility of traditional population-based reference values. In relation to the latter application, for analytes with marked individuality the dispersion of values for any individual spans only a small part of the reference interval. Reference values are then of little utility, particularly for deciding whether any change has occurred. For analytes with little individuality, the distribution of values from a single individual covers much of the entire distribution of the reference interval derived from reference subjects. Thus, conventional reference values are of significant value in many clinical settings.

Numerous studies have been reported on the biological variation for semen parameters. The controversies concerning

much of the published data are partly due to the fact that previous studies dealt with the interval between samples and the duration of the study (Hotchkiss, 1941; MacLeod and Heim, 1945; McLeod and Gold, 1951; Freund, 1962; Heuchel *et al*., 1983; Poland *et al*., 1985; Schwartz *et al*., 1986; Bostofte *et al*., 1987; Knuth *et al*., 1988; Schrader *et al*., 1988; Mallidis *et al*., 1991).

The aim of the present study was to determine the within- and between-subject biological variation of seminal parameters by using individual means (homeostatic values) obtained from week-to-week data over 10 weeks, a period containing only one spermatogenic cycle (Sharpe, 1994).

Materials and methods

Subjects

For this study of the biological variation of seminal parameters, 33 donor candidates (aged 18–24 years) with no previous history of andrological pathology were invited to participate, and each gave their informed consent for participation. Semen quality was not a limiting factor for inclusion. The subjects agreed to maintain uniform weekly habits (no medication, and no unusual meals, physical exercise or sexual habits). Only 20 subjects free of disease and taking no medication for the duration of the study provided a semen specimen each week (± 1 day) for 10 weeks. All semen samples were collected by masturbation in the laboratory after 3–4 days of sexual abstinence. In order to avoid seasonal influence on the seminal parameters, samples were collected between October 2001 and March 2002. Once collected, the samples were kept at 37°C for 45 min before assessment.

In order to avoid between-observer variability, all analyses were performed by a single experienced technician who rigorously followed the WHO (1999) guidelines for routine semen analysis. Internal quality control was performed throughout the study period, with acceptable agreement between the experienced technician and another technician (between-technician coefficient of variation <13%). In addition, the centre participated in two external quality control programmes, the European Society of Human Reproduction and Embryology (ESHRE) scheme based in Stockholm and the Spanish programme of external quality control on semen analysis under the auspices of the Association for the Study of Reproductive Biology (ASEBIR). The semen characteristics evaluated were sperm concentration, percentage of total motile spermatozoa (grades a + b + c), percentage of progressive motile spermatozoa (grades a + b), percentage of rapid motile spermatozoa (grade a), percentage of normal forms of spermatozoa and percentage of living spermatozoa.

Analytical variation

The assessment of components and total analytical variation in semen analysis was made with quality control materials and with semen samples obtained from healthy volunteers who gave their informed consent. Total analytical variation was calculated according to published recommendations (National Committee for Clinical Laboratory Standards, 1985). Two aliquots of quality control materials were analysed in each batch, with two batches per day for 10 days. In total, 40 values from each quality control material for each seminal parameter were obtained. The quality control material for the assessment of analytical variation of sperm concentration used was: frozen straws kept at –196°C with cryoprotectant added, pools of formalin (1%) semen suspension, and aliquots from a formalin semen suspension. Frozen straws and videotape were used for the study of analytical variation of sperm motility. All straws for the motility assessment were thawed for 10 min at 37°C before analysis. Analytical variation of sperm morphology was studied using frozen straws and unstained and stained smears (Diff-Quick). Analytical variation of sperm vitality was analysed on Eosin Y-Nigrosin smears and frozen straws. All quality control materials were prepared at different levels of semen quality: low (sperm parameter lower than WHO, 1999 reference value), medium (sperm parameter near WHO reference value) and high (sperm parameter higher than WHO reference value). Analytical variation was expressed by means of the coefficient of variation (CV = 100 × standard deviation/mean).

Biological variation

Before analysing the data from the candidate donors, the Dixon method was used (Dixon, 1953) to exclude outlying values from a single subject and to eliminate mean outlying values. One datum from concentration and one from rapid motility were rejected as outliers. Outliers were replaced by their respective means. The Shapiro–Wilk test was applied separately to the set of results from each individual to analyse the normality of results. All results fitted a Gaussian distribution. Although the Cochran test indicated that within-subject variances were moderately non-homogeneous for concentration, rapid motility and morphology, the bias produced by applying ANOVA to compute the total within- and between-subject variances was considered negligible (Fraser and Harris, 1989) because the indices of heterogeneity (see below) for all semen parameters studied were non-significant.

It was assumed that the within-subject variation on seminal parameters (as for most analytes and cell parameters in healthy subjects) was a random independent fluctuation around a homeostatic set-point, depending only on analytical imprecision and biological variation (Harris, 1970; Fraser, 1988). The total variability, total within-subject variance (s^2_{Bw+a}) (including both analytical variance and within-subject biological variance) and total between-subject variances (s^2_{Bb+a}) (including both analytical variance and between-subject biological variance) were separated with ANOVA according to published methods (Fraser and Harris, 1989). The computed variances for the respective semen parameters were transformed into the corresponding coefficients of variation (CV_{Bw+a} and CV_{Bb+a} respectively) using the overall means of the seminal parameters. The within-subject (CV_{Bw}) and between-subject (CV_{Bb}) values were calculated from the CV_{Bw+a} and CV_{Bb+a} values using the formulas:

$$CV^2_{Bw} = CV^2_{Bw+a} - CV^2_a$$

$$CV^2_{Bb} = CV^2_{Bb+a} - CV^2_a$$

The subtracted CV_a corresponded to the CV obtained to the similar concentration or percentage in the experiment on analytical variation, as indicated in Table I.

The index of heterogeneity (IH) (ratio of CV_{Bw+a} to the theoretical CV) was obtained by applying the following formula:

$$IH = CV_{Bw+a}/[(2/n - 1)^{1/2} \times 100]$$

where n is the number of samples per individual. In the present study, with 10 data for each subject, an IH < 1.45 indicated that the within-subject data were homogeneous because the index was $<\{1 + 2[1/(2n)^{1/2}]\}$ (hypothesis of non-heterogeneity of true within-subject variances) (Harris, 1970).

Table I. Mean values and analytical variation with different control materials at different quality levels

Analytical parameter	Low level		Medium level		High level	
	Mean	CV_a	Mean	CV_a	Mean	CV_a
Sperm concentration ($\times 10^6$/ml)						
Frozen straws	8.9	10.4	40.6	11.8	115.6	11.6
Semen pool	8.8	8.2	43.3	8.4[a]	123.7	7.9
Aliquots semen pool	9.2	9.9	53.0	10.9	102.8	10.8
Motility 'a+b+c' (% motile forms)						
Frozen straws	15.4	12.5	51.8	8.9[a]	75.2	8.7
Videotape	10.4	7.4	51.9	5.8	71.4	5.3
Motility 'a+b' (% motile forms)						
Frozen straws	10.5	13.5	43.7	9.2[a]	67.9	9.6
Videotape	10.4	9.7	40.6	7.3	4.4	6.9
Motility 'a' (% motile forms)						
Frozen straws	4.5	16.8	23.9	12.9[a]	44.4	12.8
Videotape	4.8	10.8	21.3	7.5	38.2	7.8
Morphology (% normal forms)						
Frozen straws	3.6	17.9	14.6	14.2	26.2	13.7
Unstained smears	4.7	9.2	15.1	7.3[a]	24.9	6.6
Stained smears	6.7	9.9	13.8	7.0	23.4	7.1
Vitality (% live forms)						
Frozen straws	23.2	13.4	40.3	9.6	82.1	9.7
Stained smears	15.1	10.3	52.2	6.9[a]	76.2	7.4

[a]Values used to calculate components of biological variability.

The index of individuality (II), CV_{Bw+a}/CV_{Bb}, describes the relationship between within-subject and between-subject variation and was used to evaluate the usefulness of conventional population-based reference values. If the index was low (< 0.60), an abnormal value for a person was likely to appear within the reference interval; conversely, the marker was useful for assessing differences in an individual's serial results. A high index (> 1.4) distinguished abnormal and normal values.

The critical differences (CD) at a significant level of $P < 0.05$ were calculated according to the following formula:

$$CD\ (\%) = 2.77 \times [(CV_{Bw})^2 + (CV_a)^2]^{1/2}$$

The CD, also called the least significant change and reference change value, is the minimum difference between two successive measurements in an individual that can be considered to reflect a true biological change.

The desirable imprecision (CV_D), bias (expressed as relative deviation RD_D, %) and total error (TE_D) for each WHO (1999) procedure for routine semen analysis, were calculated using the following criteria according to Fraser *et al*. (1997):

$$CV_D < 0.75\ CV_{Bw}$$

$$RD_D \leq 0.375\sqrt{CV_{Bw}^2 - CV_{Bb}^2}$$

$$TE_D \leq 1.65 * CV_D + RD_D$$

From the above variance components, the reliability coefficient (R) was computed; this was the ratio of the between-subject-person variance to the total observed population variance:

$$R = CV_{Bb}^2/(CV_{Bw}^2 + CV_{Bb}^2 + CV_a^2)$$

To compute the lowest value observed (x_{obs}) from seminal parameters with a <5% probability of having a true value less than WHO (1999) reference values (RV), the following approach was used:

$$x_{obs} = RV + 1.96\sqrt{s_{Bw+a}^2}$$

Results

The values obtained for analytical imprecision are shown in Table I. The CV_a for semen parameters depended on the quality control material used and the level of semen quality. Cryopreserved sperm always showed the highest CV_a. Among semen parameters at medium level of quality, vitality showed the lowest CV_a, whereas rapid motility showed the highest.

The means and the components of biological variation of the semen characteristics found in the study, expressed in terms of CV, are listed in Table II. Sperm vitality showed a lower within-subject variability than other semen characteristics, and sperm concentration showed the highest. Sperm concentration, rapid motility and percentage of normal forms showed high between-subject variabilities compared with other semen parameters. All semen characteristics studied showed an II < 0.70 (Table III). Total and progressive motility, morphology and vitality showed CD lower than that of concentration and rapid motility. All semen parameters showed homogeneity since all IH were < 1.45. The reliability coefficients ranged from 0.68 (index of individuality = 0.68) for total motility to 0.83 (index of individuality = 0.44) for rapid progressive motility.

The lowest values observed for semen parameters with a <5% probability of having a true value less than the WHO (1999) reference value are listed in Table III. These data were not calculated for total motility because there was no WHO (1999) reference value available. The analytical goals based on biological variability are shown in Table IV.

Discussion

The analytical variations observed for semen parameters were similar to the intra-technician variations reported by other authors in programmes of internal quality control (Neuwinger *et al*., 1990; Cooper *et al*., 1992). Although the present results

Table II. Mean and components of biological variation for semen parameters

	Mean	CV_{Bw+a}	CV_{Bw}	CV_{Bb}
Concentration ($\times 10^6$/ml)	68.1	28.1	26.8	56.4
Total motility (%)	45.8	20.4	18.4	29.8
Progressive motility (%)	36.1	17.8	15.2	32.8
Progressive rapid motility (%)	17.8	22.8	18.8	51.8
Morphology (%)	12.7	20.9	19.6	44.0
Vitality (%)	62.4	12.4	10.3	25.8

CV_{Bw+a} = total within-subject variation; CV_{Bw} = within-subject variation; CV_{Bb} = between-subject variation.

Table III. Indices of individuality and heterogeneity, critical differences, coefficient of reliability and the lowest value observed with a <5% probability of having a true value less than the WHO (1999) reference value for semen parameters.

Parameter	II	CD (%)	IH	R	x_{obs}
Concentration	0.50	77.8	0.60	0.80	57.5 ($\times 10^6$/ml)
Total motility	0.68	56.5	0.43	0.68	*
Progressive motility	0.54	49.2	0.38	0.77	62.6%
Progressive rapid motility	0.44	63.1	0.48	0.83	32.9%
Morphology	0.48	58.0	0.44	0.82	20.2%
Vitality	0.48	34.4	0.26	0.81	90.2%

*, not calculated, as no WHO (1999) reference value available.
CD = critical differences; IH = index of heterogeneity; II = index of individuality; R = reliability coefficient; x_{obs} = lowest value observed with a <5% probability of having a true value less than the WHO (1999) reference value.

Table IV. Analytical goals for semen parameters using biological variation

Parameter	CV_D	RD_D	TE_D
Concentration	<20.1	<23.4	<56.6
Total motility	<13.8	<13.1	<35.8
Progressive motility	<11.4	<13.5	<32.3
Progressive rapid motility	<14.1	<20.7	<43.9
Morphology	<14.7	<18.1	<42.3
Vitality	<7.7	<10.4	<23.2

CV_D = desirable imprecision; RD_D = desirable relative deviation; TE_D = desirable total error.

on CV_a for concentration and vitality were in agreement with those of others (Auger *et al*., 2001), the motility results were lower in the present study than in the cited paper. Reproducibility in the assessment of sperm parameters is dependent upon operator training and continuous monitoring. For this purpose, the technician herein participated in internal quality control programmes, with acceptable agreement between technicians, and also in external quality control schemes (ESHRE and ASEBIR programmes).

For total within-subject variability for percentage of normal forms, a CV_{Bw+a} value of 20.9% was found, which was higher than previously reported values of 16% (Poland *et al*., 1985) and 9% (Knuth *et al*., 1988). The criteria used to assess normal forms by these authors were different from those in the present study, and this had a direct impact on subject variability. The present percentage of normal forms was lower than that obtained by other authors but, as has been shown, it was in agreement with that of another group (Cooper *et al*., 1992); the CV_a of the percentage of normal forms is inversely related to the percentage of normal forms. A comparison of the present findings with those of other authors who studied within-subject variability for morphology is not possible as they determined the percentage of abnormal forms (Hotchkiss, 1941; MacLeod and Heim, 1945; McLeod and Gold, 1951; Freund, 1962; Schwartz *et al*., 1986; Bostofte *et al*., 1987).

Today, quality control plays a very important role in the andrology laboratory (World Health Organization, 1999), but before laboratory quality can be controlled it is essential to know exactly what level of quality is needed to ensure satisfactory clinical decision making. Different proposals have been published regarding the quality requirements of methods of analyte measurement (Petersen *et al*., 1999). One group (Cotlove *et al*., 1970) proposed that minimally acceptable CV_a for assays should be equal to or less than half of the within-subject variation (< 0.5 CV_{Bw}). However, for assays that with currently available techniques could not easily meet the analytical goals (< 0.5 CV_{Bw}), others (Fraser *et al*., 1997) suggested a new analytical goal (< 0.75 CV_{Bw}). The present authors recommend the use of the former goals because the WHO (1999) techniques are carried out manually. The goals for precision (CV_a) were met for all semen parameter assays, except for some assays in which frozen straws were used as quality control material. For this reason, the present authors agree with others (Clements *et al*., 1995) in rejecting cryopreserved samples as part of routine quality control in the laboratory. Recently, it has been shown (Johnson *et al*., 2003) that Westgard quality control rules are repeatedly violated when frozen sperm pellets are used as quality control material and in computer-assisted semen analysis (CASA) systems. Therefore, the present authors do not consider these rules acceptable as a daily-use quality control material for semen analysis.

These analytical goals can be used to study other semen analysis techniques. For example, one group (Martínez *et al*., 2000) evaluated sperm quality analyser-IIB, and in the imprecision assay observed mean CV_a values of 17.2, 11.1 and 6.3% for concentration, total motility and normal forms respectively. All assays met the analytical goals previously established. However, the goal for precision assays of sperm motility was not met by two groups (Clements *et al*., 1995; Johnson *et al*., 2003) who used frozen straws and CASA (CV_a = 18.4% and 28.4% respectively). These results suggest a lack of stability of frozen sperm pellets, most likely due to technical factors (such as sample preparation, storage conditions, thawing, etc.). Analytical goals, nevertheless, are targets to aim for, rather than inflexible criteria for acceptance or rejection (Fraser, 2001).

It appears that the less specific the semen parameters, the greater the biological variation. The greatest biological variations, both within- and between-subjects, were found for sperm concentration. The large fluctuations of sperm concentrations in an individual may reflect variations in sperm production in normal germinal epithelium, while the high between-subject variability reflects differences in homeostatic set-points among individuals, probably arising from such

factors as genetic characteristics and environment (Auger *et al*., 2001). The best homeostatic regulation within an individual and also the smallest between-subject variability were found for vitality and total motility. These findings were in accordance with previous studies on within-subject variations of semen parameters over short periods (3–6 months) (Poland *et al*., 1985; Knuth *et al*., 1988). However, over longer periods (> 9 months to years) results are contradictory (Heuchel *et al*., 1983; Bostofte *et al*., 1987; Schrader *et al*., 1988; Mallidis *et al*., 1991).

The within-subject (10.3–26.8%) and between-subject (25.8–56.4%) variability for semen parameters observed in the present study were similar to those observed by others (Valero-Politi and Fuentes-Arderiu, 1993) for sexual hormones in young men (within subjects, 9.4–24.0%; between subjects, 21.3–51.6%), suggesting that there is a similar homeostatic regulation of reproductive system components.

Semen parameters showed lower levels of variability within-subjects than between-subjects, and this was in agreement with a previous report (Heuchel *et al*., 1983). Thus, the II for semen characteristics was low (< 0.7). The II is used to assess objectively the utility of population reference values for diagnostic purposes, and can range—at least theoretically—from zero to infinity. When the II of an analyte is low (< 0.6), a typical individual's test results stay within a relatively narrow interval of the population-based reference value, and comparison of an individual's test results with population-based reference intervals to determine disease status or a change in disease status is of limited use (Harris, 1974; Fraser and Harris, 1989; Solberg and Grasbeck, 1989). Semen parameters showed strong individuality (a low II). As shown in Figure 1, this means that individual subjects can present semen parameters that are very unusual for them, yet these values are still well within conventional reference values.

The role of semen parameters as a prognostic factor of a male's fertility potential is a matter of ongoing debate (Ombelet *et al*., 1997a, b; Guzick *et al*., 2001; Menkveld *et al*., 2001). Receiver operating characteristic (ROC) curves are often used to determine the sensitivity and specificity of different cut-offs of semen parameters in order to distinguish between the fertile and subfertile population, on the basis of the largest area under the curve. However, numerous authors have observed that sensitivity of the semen parameters is poor, indicating a large overlap in the distribution of these variables in the fertile and infertile groups (Ombelet *et al*., 1997a, b; Gunalp *et al*., 2001; Menkveld *et al*., 2001).

As commented above, it cannot be concluded from the present results—unlike other authors—that either morphology (Ombelet *et al*., 1997b; Gunalp *et al*., 2001; Guzick *et al*., 2001) or progressive motility (Gunalp *et al*., 2001) are better than sperm concentration to distinguish between fertile and subfertile individuals. However, as the former criterion showed strong individuality and critical differences between consecutive values ($P < 0.05$) lower than sperm concentration, it can be concluded that these parameters are better than concentration for assessing differences in an individual's serial results.

Repeated measurements on individual subjects allow an analysis to be made of whether a critical difference has been exceeded or whether a non-random trend has occurred. The calculated critical differences, which are close to 50% for both motility and morphology, indicate that relatively large changes in these parameters (less than those for concentration, close to 80%) are required before any certainty can be felt that the change is significant. As shown in Figure 1, for certain sperm parameters (concentration, morphology and rapid motility) some subjects appear to have a small CV_{Bw}, and some a large CV_{Bw}. This apparent heterogeneity of within-subject variation can be examined by the index of heterogeneity proposed previously (Harris, 1970). Indices of heterogeneity for all semen parameters studied were non-significant. In consequence, for commonly performed semen analysis, the critical difference detailed here will be generally applicable and useful in clinical practice. It should be emphasized that the observed variation is based on measurements in the semen of healthy individuals. The variation in semen parameters in pathological men (subfertile) with very low levels cannot be predicted from the present data (Fraser and Harris, 1989). Moreover, because the analytical variance must be used for the calculation of the critical differences, each laboratory should make its own calculations for these values. Therefore, the critical difference shown in Table III is shown only for a sample.

Epidemiological studies attempt to relate between-subject differences in various measured 'risk factors' for disease occurrence. Therefore, it is the between-subject variability in a measured risk factor that is important in establishing epidemiological risk factor–disease associations (i.e., hopefully, CV_{Bb} is large in relation to CV_{Bw} and CV_a). The reliability coefficient is the ratio of the between-subject variability to the total observed population variability, and it can range from zero to 1.0 for an epidemiological study, but optimally is near unity. When the reliability coefficient is very high, a single measurement of an analyte concentration in a study participant will correctly classify that participant with respect to his or her homeostatic set point. The relatively high reliability coefficient found for semen parameters, as shown previously (Poland *et al*., 1985), suggests that basic semen measures are relatively stable within subjects. These authors propose that individuals do have characteristic patterns of spermatogenesis.

The present findings establish that a man with $>58\times10^6$ spermatozoa/ml, 63% sperm progressive motility, 33% rapid motility or 20% normal forms has a <5% probability of having a true value less than the WHO (1999) reference value. As for other studies (Overstreet and Davis, 1995; Guzick *et al*., 2001), it was estimated in the present investigation that instead of a single value for each semen parameter which supposedly distinguishes between 'normal' and 'abnormal', it is better to have two values that allow for the description of three groups. Using WHO (1999) reference values, and only from a theoretical point of view, these results allow three groups to be defined: (i) normal or fertile (men with values higher than these cut-off values); (ii) pathological or subfertile (lower than the WHO values); and (iii) intermediate fertility (between the WHO value and the above cut-off values). This classification system is clinically meaningful and is appropriate for what is, biologically, a continuous function. However, this hypothesis should be checked with clinical data.

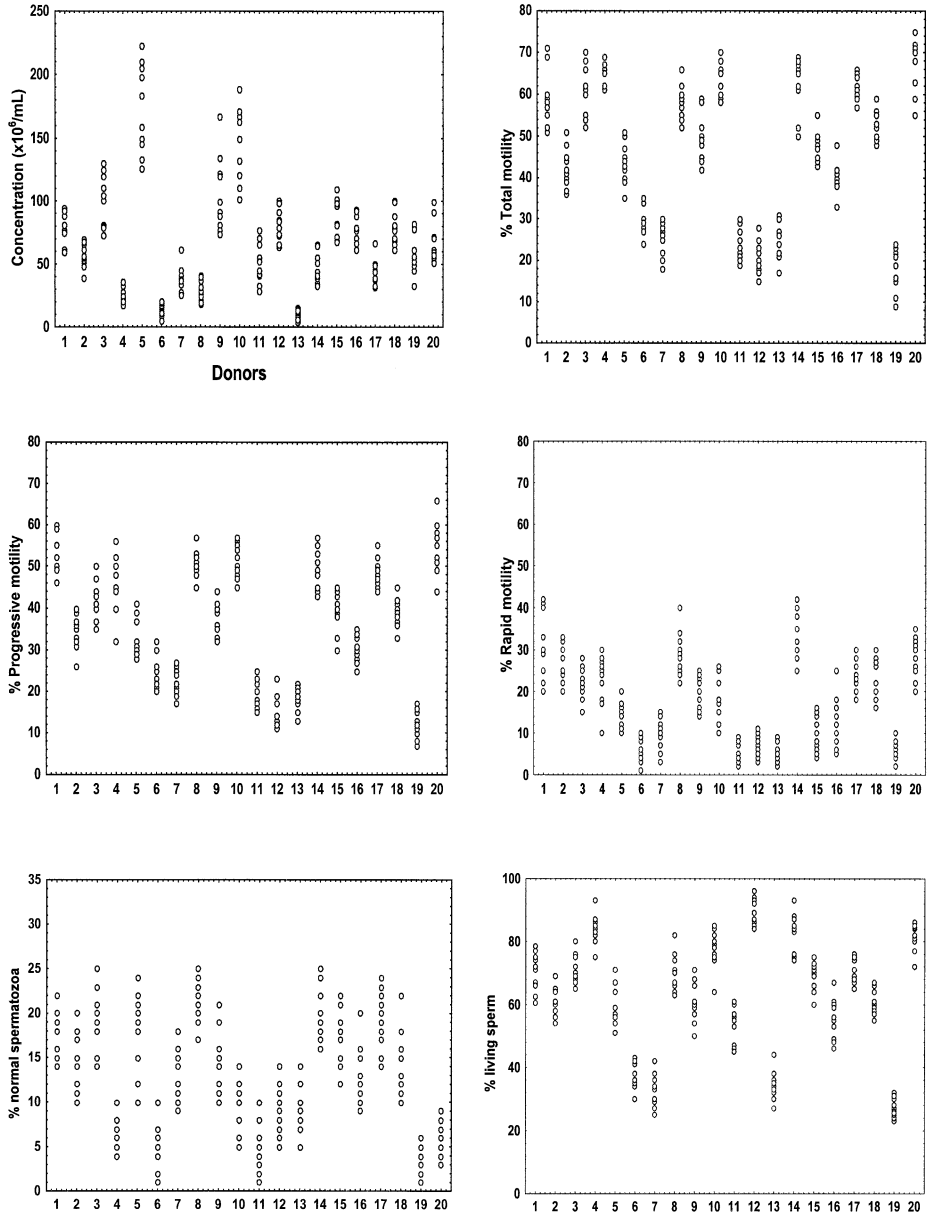

Figure 1. Scatterplots for semen parameters in 20 healthy donor candidates for a 10-week follow-up study. Samples were obtained on a week-to-week basis.

It is interesting that when the minimum observed value was calculated for semen parameters with a <5% probability of obtaining a true value less than the reference value for subfertile patients from a more recent report (Guzick *et al*., 2001) rather than previous reference values (World Health Organization, 1999), a concentration of $\geqslant 51\times10^6$/ml and $\geqslant$13% normal morphological features were obtained. These values were very similar to those reported by these authors ($\geqslant 49.0\times10^6$/ml and $\geqslant$13% respectively) to define the fertile group. Moreover, these values are in accordance with the results of three studies on the relationship between time to pregnancy and semen concentrations. In a study on partners of pregnant women who obtained their pregnancy without medical assistance (MacLeod and Gold, 1953), the median time taken to conceive was shorter among subjects with a sperm concentration $>40\times10^6$/ml than among those with a sperm concentration between 20 and 39×10^6/ml. A study of first pregnancy planners (Bonde *et al*., 1998) also indicated that sperm concentration could indeed influence fecundity up to 40×10^6/ml. Finally, a recent cross-sectional study showed that increasing sperm concentration up to 55×10^6/ml influenced time to pregnancy (Slama *et al*., 2002).

In conclusion, the results of the present study showed that conventional reference values for seminal parameters have little diagnostic value because of their marked individuality, although seminal parameters may be useful for assessing differences in an individual's serial results, in particular of progressive motility, morphology and vitality.

References

Auger, J., Eustache, F., Andersen, A.G., Irvine, D.S., Jørgensen, N., Skakkebæk, N.E., Suominen, J., Toppari, J., Vierula, M. and Jouannet, P. (2001) Sperm morphological defects related to environment, lifestyle and medical history of 1001 male partners of pregnant women from four European cities. *Hum. Reprod.*, **16**, 2710–2717.

Bonde, J.P., Ernst, E., Jensen, T.K., Hjollund, N.H., Kolstad, H., Henriksen, T.B., Scheike, T., Giwercman, A., Olsen, J. and Skakkebaek, N.E. (1998) Relation between semen quality and fertility: a population based study of 430 first-pregnancy planners. *Lancet*, **352**, 1172–1177.

Bostofte, E., Forrest, M. and Andersen, B. (1987) Regression toward the mean in fertility research an empirical study of semen quality. *Andrologia*, **19**, 76–79.

Clements, S., Cooke, I.D. and Barratt, C.L.R. (1995) Implementing comprehensive quality control in the andrology laboratory. *Hum. Reprod.*, **10**, 2096–2106.

Cooper, T.G., Neuwinger, J., Bahrs, S. and Nieschlag, E. (1992) Internal quality control of semen analysis. *Fertil. Steril.*, **58**, 172–178.

Cotlove, E., Harris, E.K. and Williams, G.Z. (1970) Biological and analytic components of variation in long term studies of serum constituents in normal subjects: III. Physiological and medical implications. *Clin. Chem.*, **16**, 1028–1032.

Dixon, D.J. (1953) Processing data for outliers. *Biometrics*, **9**, 74–89.

Fraser, C.G. (1988) The application of theoretical goals based on biological variation in clinical chemistry. *Arch. Pathol. Lab. Med.*, **112**, 404–415.

Fraser, C.G. (2001) *Biological Variation: From Principles to Practice*. 1st edn. AACC Press, Washington, USA.

Fraser, C.G. and Harris, E.K. (1989) Generation and application of data on biological variation in clinical chemistry. *Crit. Rev. Clin. Lab. Sci.*, **27**, 409–437.

Fraser, C.G., Hyltoft-Petersen, P., Libeer, J.C. and Ricós, C. (1997) Proposals for setting generally applicable quality goals solely based on biology. *Ann. Clin. Biochem.*, **34**, 8–12.

Freund, M. (1962) Interrelationships among the characteristics of human semen and factors affecting semen-specimen quality. *J. Reprod. Fertil.*, **4**, 143–159.

Gunalp, S., Onculoglu, C., Gurgan, T., Kruger, T.F. and Lombard, C.J. (2001) A study of semen parameters with emphasis on sperm morphology in a fertile population: an attempt to develop clinical thresholds. *Hum. Reprod.*, **16**, 110–114.

Guzick, D.S., Overstreet, J.W., Factor-Litvak, P., Brazil, C.K., Nakajima, S.T., Coutifaris, C., Carson, S.A., Cisneros, P., Steinkampf, M.P., Hill, J.A. *et al.* (2001) Sperm morphology, motility, and concentration in fertile and infertile men. *N. Engl. J. Med.*, **345**, 1388–1393.

Harris, E.K. (1970) Distinguishing physiologic from analytic variation. *J. Chronic Dis.*, **23**, 469–480.

Harris, E.K. (1974) Effects of intra- and interindividual variation on the appropriate use of normal ranges. *Clin. Chem.*, **20**, 1535–1542.

Heuchel, V., Schwartz, D. and Czyglik, F. (1983) Between and within subject correlations and variances for certain semen characteristics in fertile men. *Andrologia*, **15**, 171–176.

Hotchkiss, R. (1941) Factors in stability and variability of semen specimens. Observations on 640 successive samples from 23 men. *J. Urol.*, **45**, 875–888.

Johnson, J.E., Blackhurst, D.W. and Boone, W.R. (2003) Can Westgard quality control rules determine the suitability of frozen sperm pellets as a control material for computer assisted semen analyzers? *J. Assist. Reprod. Genet.*, **20**, 38–45.

Knuth, U.A., Kuhne, J., Bals-Pratsch, M. and Nieschlag, E. (1988) Intra-individual variation of sperm velocity, linearity, lateral head displacement and beat frequency in healthy volunteers. *Andrologia*, **20**, 243–248.

MacLeod, J. and Heim, L. (1945) Characteristics and variations in semen specimens in 100 normal young men. *J. Urol.*, **54**, 474–482.

MacLeod, J. and Gold, R.Z. (1951) The male factor in fertility and infertility. IV. Sperm morphology in fertile and infertile marriage. *Fertil. Steril.*, **2**, 394–414.

MacLeod, J. and Gold, R.Z. (1953) The male factor in fertility and infertility. VI. Semen quality and certain other factors in relation to ease of conception. *Fertil. Steril.*, **4**, 10–33.

Mallidis, C., Howard, E.J. and Baker, H.W. (1991) Variation of semen quality in normal men. *Int. J. Androl.*, **14**, 99–107.

Martínez, C., Mar, C., Azcárate, M., Pascual, P., Aritzeta, J.M. and López-Urrutia, A. (2000) Sperm motility index: a quick screening parameter from sperm quality analyser-IIB to rule out oligo- and asthenozoospermia in male fertility study. *Hum. Reprod.*, **15**, 1727–1733.

Menkveld, R., Wong, W.Y., Lombard, C.J., Wetzels, A.M.M., Thomas, C.M.G., Merkus, H.M.W.M. and Steegers-Theunissen, R.P.M. (2001) Semen parameters, including WHO and strict criteria morphology, in a fertile and subfertile population: an effort towards standardization of in-vivo thresholds. *Hum. Reprod.*, **16**, 1165–1171.

National Committee for Clinical Laboratory Standards (1985) *Evaluation of Precision Performance of Clinical Chemistry Devices; Approved Guideline*. NCCLS Publication EP5-A.

Neuwinger, J., Behre, H.M. and Nieschlag, E. (1990) External quality control in the andrology laboratory: an experimental multicenter trial. *Fertil. Steril.*, **54**, 308–314.

Ombelet, W., Bosmans, E., Janssen, M., Cox, A., Vlasselaer, J., Gyselaers, W., Vandeput, H., Gielen, J., Pollet, H., Maes, M. *et al.* (1997a) Semen parameters in a fertile versus subfertile population: a need for change in the interpretation of semen testing. *Hum. Reprod.*, **12**, 987–993.

Ombelet, W., Wouters, E., Boels, L., Cox, M., Janssen, M., Spiessens, C., Vereecken, A., Bosmans, E. and Steeno, O. (1997b) Sperm morphology assessment: diagnostic potential and comparative analysis of strict or WHO criteria in a fertile and subfertile population. *Int. J. Androl.*, **20**, 367–372.

Overstreet, J.W. and Davis, R.O. (1995) Methods and interpretation of semen analysis. In Key, W.R., Chang, R.J., Rebar, R.W. and Soules, M. (eds), *Infertility: Evaluation and Treatment.* W.B. Saunders, New York, pp. 580–591.

Petersen, P.H., Fraser, C.G., Kallner, A. and Kenny, D. (1999) Strategies to set global analytical quality specifications in laboratory medicine. *Scand. J. Clin. Lab. Invest.*, **59**, 1–164.

Poland, M.L., Moghissi, K.S., Giblin, P.T., Ager, J.W. and Olson, J.M. (1985) Variation of semen measures within normal men. *Fertil. Steril.*, **44**, 396–400.

Ricos, C., Álvarez, V., Cava, F., Garcia-Lario, J.V., Hernández, A., Jiménez, C.V., Minchinela, J., Perich, C. and Simon, M. (1999) Current database on biologic variation: pros, cons and progress. *Scand. J. Clin. Lab. Invest.*, **59**, 491–500.

Schrader, S.M., Turner, T.W., Breitenstein, M.J. and Simon, S.D. (1988) Longitudinal study of semen quality of unexposed workers. I. Study overview. *Reprod. Toxicol.*, **2**, 183–190.

Schwartz, D., Ducot, B., Auroux, M. and Collin, C. (1986) Within-subject variability of the percentage of morphologically abnormal spermatozoa among fertile men: biological and measurement components. *Hum. Reprod.*, **6**, 369–371.

Sharpe, R.M. (1994) Regulation of spermatogenesis. In Knobil, E. and Neill, J.D. (eds), *The Physiology of Reproduction*. Raven Press, New York, pp. 1363–1434.

Slama, S., Eustache, F., Ducot, B., Jensen, T.K., Jørgensen, N., Horte, A., Irvine, S., Suominen, J., Andersen, A.G., Auger, J. *et al.* (2002) Time to pregnancy and semen parameters: a cross-sectional study among fertile couples from four European cities. *Hum. Reprod.*, **17**, 503–515.

Solberg, H.E. and Grasbeck, R. (1989) Reference values. *Adv. Clin. Chem.*, **27**, 1–79.

Valero-Politi, J. and Fuentes-Arderiu, X. (1993) Within- and between-subject biological variations of follitropin, leutropin, testosterone, and sex-hormone-binding globulin in men. *Clin. Chem.*, **39**, 1723–1725.

World Health Organization (1999) *WHO Laboratory Manual for the Examination of Human Semen and Semen-Cervical Mucus Interaction*. 4th edn. Cambridge University Press, Cambridge.

Submitted on November 18, 2002; resubmitted on March 26, 2003; accepted on July 9, 2003

QUALITY, INNOVATION, AND SERVICE—IT'S ALL AT THE CENTER OF EVERYTHING WE DO.

From developing assisted reproductive technologies that maximize performance, like the first ART media and cultures, to expertise that streamlines productivity, FUJIFILM Irvine Scientific brings together decades of industry expertise with a powerhouse of innovation, turning opportunities into realities. Together, we're working to support healthy futures—from retrieval to realization.

ALL IN *FOR LIFE.*
irvinesci.com/*ALLIN*

FUJIFILM
Value from Innovation

IrvineScientific