Human Reproduction Vol.16, No.6 pp. 1165–1171, 2001

Semen parameters, including WHO and strict criteria morphology, in a fertile and subfertile population: an effort towards standardization of in-vivo thresholds

Roelof Menkveld[1], Wai Yee Wong[2,3], Carl J.Lombard[4], Alex M.M.Wetzels[2], Chris M.G.Thomas[2,5], Hans M.W.M.Merkus[2] and Régine P.M.Steegers-Theunissen[2,3,6]

[1]Reproductive Biology Unit, Department of Obstetrics and Gynaecology, Tygerberg Hospital and University of Stellenbosch, Tygerberg, 7505 South Africa, Departments of [2]Obstetrics and Gynaecology and [3]Epidemiology, University Medical Centre St Radboud, 6500 HB Nijmegen, The Netherlands, [4]Biostatistics Division, Centre for Epidemiological Research in South Africa, Medical Research Council, South Africa and [5]Department of Chemical Endocrinology, University Medical Centre St Radboud, 6500 HB Nijmegen, The Netherlands

[6]To whom correspondence should be addressed at: Department of Obstetrics and Gynaecology/Epidemiology, University Medical Centre Nijmegen, PO Box 9101, 6500 HB Nijmegen, The Netherlands. E-mail: R.steegers@mie.kul.nl

In this study, the semen analysis results of a fertile population were compared with those from a subfertile population, in order to establish normal cut-off values for the standard semen parameters with the aid of receiver operating characteristic (ROC) curve analysis. The fertile group comprised healthy males (n = 107) without any history of fertility problems, the partners of whom had had a spontaneous pregnancy within one year of unprotected intercourse and were pregnant at the time of the male's inclusion into the study. A total of 103 males from couples attending the infertility clinic, and with an initial sperm count of $<20 \times 10^6$/ml were recruited to form the subfertile population. The best discriminating parameter between the two populations was sperm morphology evaluated according to WHO criteria at a cut-off point of 31% normal spermatozoa. The other cut-off values were at 8% for the acrosome index, 45% for motility, and 4% normal spermatozoa for strict criteria. Recalculating the ROC curve cut-off values based on an assumed 50% prevalence of subfertility in an assisted reproductive setting, the cut-off points were reduced to 21% and 3% normal spermatozoa for WHO and strict criteria respectively. For motility, the new cut-off value was at 20% motile spermatozoa, for motility quality at 3.5 (on a scale of 1–6), the acrosome index at 3% normal acrosomes, and the teratozoospermia index at 2.09.

Key words: fertile and subfertile population/semen analysis/sperm morphology/strict criteria/WHO criteria

Introduction

Minimum requirements for semen analysis and semen parameter standards were established in 1951 by the American Fertility Association (Abarbanel *et al*., 1951), in 1966 by Freund (Freund, 1966), and in 1971 by Eliasson (Eliasson, 1971). This was followed by publications from the World Health Organization (WHO) in 1980, 1987, 1992 and 1999. However, the so-called normal values provided by the WHO manuals for the basic semen parameters, viz. volume, qualitative and quantitative motility and morphology, were mostly obtained through studies performed on so-called fertile populations (WHO, 1987, 1992). This may result in a situation that if the semen parameters (variables) are not within the normal range as given in these publications, the male may be regarded as subfertile or not capable of fertilizing his spouse. This can lead to unnecessary treatment procedures (Comhaire *et al*., 1992; Ombelet *et al*., 1995) or to social problems and stress among couples, for example in cases where spontaneous pregnancies occur after a male has been pronounced as subfertile according to WHO guidelines for normality.

This may be one of the reasons why the clinical value of traditional semen analysis in the assessment of male fertility potential is now a subject of considerable debate. Especially with the introduction of intracytoplasmic sperm injection (ICSI), the role of the standard semen analysis is becoming an even greater point of discussion. The feeling is now that one needs only to establish the presence of sterilizing defects such as globozoospermia and the immotile spermatozoa (or Kartagener's syndrome), or the presence of azoospermia (and if so the reason for such azoospermia). In this regard, an editorial was published (McDonough,1997) in *Fertility and Sterility* suggesting that 'Traditional sperm analysis as a clinical test may become nothing more than an ancestral heirloom. It may be performed spasmodically by those who know how to do it, like a 1940-air show or laparotomy, to remind us of the good old days. We have come to the end of something. Surely

someone will want to carve a headstone for traditional sperm analysis or perhaps a mausoleum will be more fitting'.

As semen analyses are mostly used to investigate and establish the fertility potential of males with subfertility problems, the datum point for establishing standards for fertility evaluation should, therefore, not be based on what is average in a normal population. The viewpoint should rather be on what are the minimum semen parameter values needed to still give a reasonable chance for conception. These values will be much lower than the values for normality of a population with proven fertility as has been demonstrated previously (MacLeod, 1950, 1951; MacLeod and Gold, 1951a,b). These authors were the first to report differences in semen values between males of fertile and subfertile marriages. A number of reports have recently been made confirming this observation and presenting redefined values for semen parameters to provide a better diagnosis and prognosis for males seeking help for their fertility problem (Menkveld and Kruger, 1990, 1996; Comhaire *et al*., 1992; Ombelet *et al*., 1997).

It is, therefore, clear that the values for semen parameters—and especially sperm morphology as given in the WHO (1992, 1999) manuals—need to be re-evaluated (Menkveld and Kruger, 1996; Ombelet *et al*., 1997). Because of the confusion on normal values for sperm morphology, the 1999 WHO manual does not give a normal value for this parameter. This confusion with regard to sperm morphology is complicated by the fact that many different evaluation methodologies are used worldwide, as shown recently (Ombelet *et al*., 1998). These differences that are used to evaluate human sperm morphology have prevented the achievement of any consensus on the clinical diagnostic value of sperm morphology evaluation. Two of the most widely used evaluation systems are the WHO (1992) and strict criteria (Menkveld *et al*., 1990). To our knowledge, no studies have been published where these two methods have been compared in the same population with the aim of contributing to the establishment of new minimum normal values that might be used in the evaluation of a male's fertility potential for the in-vivo situation.

The aim of this study was therefore, to compare data of semen parameters obtained from fertile and subfertile populations and thereby to contribute towards the setting of new clinical thresholds for standard semen parameters, including sperm morphology as evaluated by WHO and strict criteria for the in-vivo situation.

Materials and methods

Study population

The study population consisted of 107 fertile and 103 subfertile males. The fertile group comprised healthy males without any history of fertility problems whose partners had a spontaneous pregnancy within one year of unprotected intercourse and who were pregnant at the time of the male's inclusion in the study. The fertile participants were recruited from nine midwifery practices in Nijmegen and surroundings. Subfertile males were recruited after referral to the fertility clinic of the University Medical Centre, Nijmegen and Canisius Wilhelmina Hospital, Nijmegen, The Netherlands. All the males were from couples who had failed to conceive after one year of regular unprotected intercourse and who had an initial sperm count, before admission to the trial, of $<20\times10^6$/ml, as determined by routine semen analysis at the infertility clinic. The choice of the latter criterion was due to the fact that the initial recruitment was for a prospective intervention study that was initiated in 1997 (W.Y.Wong *et al*., unpublished data). On the day that a male began participating in the trial, a semen analysis was again performed and the results from these analyses were used for this study without any further selection performed here. Subjects with known causes for their infertility problems such as chromosomal disorders related to fertility problems, cryptorchidism and vasectomies were excluded. The ethics committee of the Medical Centre, Nijmegen approved the study protocol, and all participants provided their informed consent.

Specimen collection and analytical procedures

Participants provided semen samples produced by masturbation at home after abstaining from sexual relations for a period of 3–5 days. Samples were placed into polypropylene containers and delivered within 1 h of production to the fertility laboratory. Semen analyses were performed in blinded fashion with regard to the study group. Subsequently, semen analyses were performed mainly according to the WHO guidelines (WHO, 1992) to obtain volume, pH, sperm concentration, motility and morphology. Sperm concentration was determined with the use of a Makler counting chamber. Motility was expressed as the percentage of motile spermatozoa and their mean speed, or motility quality (scale 1 to 6, where 1 = immotile and 6 = very fast progressive motile, i.e. >100 μm/s). For sperm morphology evaluation, two slides were prepared of each sample after incubation of the semen samples with trypsin (10 min at room temperature); one slide was used for routine morphology evaluation by WHO criteria and the other for strict criteria evaluation. For evaluation according to WHO criteria, smears were flame-fixed and stained with methylene blue/eosin. At least 100 cells were examined per slide, with a final magnification of ×1000. Each slide was evaluated independently by two technicians. As there was no statistical significant difference (by Parsons correlation matrix analysis) between the results of the two observers it was decided to use the results from the observer with the most complete data set (n = 207). The slides for evaluation by strict criteria were stained according to the Papanicolaou method and evaluated by one observer (R.M.) as described previously in detail, including the repeatability of the method (Menkveld *et al*., 1990). In addition to the morphology evaluation according to strict criteria, the acrosome index (AI) (Menkveld *et al*., 1996) and teratozoospermia index (TZI), as described in detail in the 1992 WHO manual, were also determined.

Teratozoospermia index

The TZI was performed as described previously (WHO, 1992; Menkveld and Kruger, 1996). The TZI is an indication of the number of abnormalities present per abnormal spermatozoon. According to the WHO manual (1992), each abnormal spermatozoon can have one to four abnormalities, viz. a head abnormality, a neck/midpiece abnormality, a tail abnormality, or the presence of a cytoplasmic residue. These abnormalities can occur as a single defect, or in a combination of two, three or all four abnormalities simultaneously. The classification of spermatozoa for the TZI is recorded simultaneously, on a five-key laboratory counter, with the recording of spermatozoa as normal or abnormal, in the specific class(es). The total number of abnormalities recorded are added together and divided by the total number of abnormal spermatozoa, i.e. 100 minus the percentage of morphologically normal spermatozoa.

Acrosome index

Sperm acrosomal morphology was evaluated by light microscopy at ×1250 oil magnification (Menkveld *et al*., 1996) based on acrosomal size and form as well as staining characteristics. Results were expressed as the AI (% normal acrosomes). For the evaluation of acrosome morphology, the same principles as for the evaluation of normal sperm morphology according to strict criteria are applicable. For an acrosome to be regarded as normal the acrosome must have a smooth normal oval shape, with the same dimensions as for a normal spermatozoon. Acrosomes must be well-defined and comprising about 40–70% of the normal-sized sperm head. The post-acrosomal part of the sperm head can be abnormal, but the rest of the spermatozoon must be normal; thus no neck/midpiece and tail abnormalities and no cytoplasmic residue may be present. If the spermatozoon is classified as normal, the acrosome must always be classified as normal. The acrosome evaluation can be performed simultaneously with the routine morphology evaluation and the TZI, with the use of two laboratory counters. As with the normal sperm morphology, at least 100 spermatozoa are evaluated. The repeatability of the AI as determined by one person has been shown to be within acceptable limits. A coefficient of variation (CV) of 9.9% was obtained for repeated evaluations of AI on the same slide. A correlation coefficient (r) of 0.8728 ($P < 0.0001$) was obtained for duplicate determinations of the AI on a set of 20 slides, with no significant differences ($P > 0.50$) between the two evaluations, as determined with the paired t-test (R.Menkveld, unpublished data).

Statistical analysis

Basic descriptive statistics, viz. means, standard deviations, medians and ranges, were calculated for the two populations separately. Semen variables from the two groups were compared for statistically significant differences (at $P < 0.05$) by means of the Wilcoxon two-sample test.

The predictive ability of the different semen variables to differentiate between the fertile or subfertile status of a male was analysed using receiver operating characteristics (ROC) curve analysis and a logistic regression model with the group's status as outcome variable. Only individual predictive patterns were evaluated. Sperm concentration was not included in the ROC curve analysis as this semen parameter was used in the initial selection of the infertile group. The areas under the curves (AUC) were estimated as well as their standard errors and confidence intervals. Pairwise comparisons were made between the six examined semen variables to test for statistically significant differences. The standard errors of the AUCs and differences were estimated through a bootstrap procedure (Efron and Tibshirani, 1993). The predictive ability was estimated assuming a 15% prevalence of subfertility in the population (Ombelet *et al*., 1997). An assumed prevalence of 50% male subfertility in an assisted reproductive setting was used in estimating the false-positive and false-negative predictive probabilities. This predictive level (of 50%) was used since it is believed that this reflects the prevalence of the male's contribution to subfertility in an assisted reproductive setting according to the literature (Wong *et al*., 2000).

Results

The mean (± SD) age of males in the fertile and subfertile groups was 33.8 ± 4.3 and 33.7 ± 3.9 years respectively (P = NS). Neither were any statistically significant inter-group differences found in the mean (± SD) days of abstinence (3.8 ± 1.9 versus 3.5 ± 1.7 days), and in semen volume (3.56 ± 1.7 versus 3.59 ± 1.7 ml) and pH (7.73 ± 0.2 versus 7.78 ± 0.2).

The descriptive statistics for the other semen parameters of the two groups are presented in Table I. All semen parameters investigated, namely sperm concentration, overall motility, motility quality, percentage of morphologically normal spermatozoa, according to WHO and strict criteria, AI index and TZI showed statistically significant differences between means of the two groups as presented in Table I.

The combined ROC curves are illustrated in Figure 1. All parameters had very similar AUC. The smallest area (78.2%) was found for sperm morphology evaluated according to SC, while the largest area (84.0%) was for sperm morphology evaluated according to WHO criteria. As sperm concentration was used as a selection criterion in the subfertile group, this semen parameter was not taken into further consideration for this study and was not displayed in Figure 1. More detail of the AUCs, 95% confidential intervals (CI) and cut-off points with their sensitivities and specificities are presented in Table II. For sperm morphology evaluated according to WHO criteria, the best cut-off point to identify the males with a possible subfertility problem based on the results of the fertile and subfertile populations investigated in this study was ⩽30% morphologically normal spermatozoa with a sensitivity and specificity of 74.5% and 76.6% respectively. For strict criteria, the cut-off point was ⩽4% morphologically normal spermatozoa with a sensitivity of 74.5% and a specificity of 77.4%. For the AI, the cut-off point was ⩽8% normal acrosomes and for the TZI ⩽1.64, with sensitivities of 73.5 and 70.8% and specificities of 70.8 and 71.7% respectively.

No statistically significant differences were found between the pairwise comparison of AUCs for the semen parameters investigated as presented in Table III.

The cut-off points, together with their false-positive and false-negative predictive probabilities as calculated for the different semen parameters based on an assumed prevalence of 50% male subfertility expected in an assisted reproductive setting are presented in Table IV. The aim was to classify as few males as possible who were probably fertile as presenting with subfertility, i.e. a high specificity where a subfertile population was investigated. The cut-off points found were lower compared with those presented in Table II. The cut-off point to identify the group of males with a possible fertility problem was ⩽21% morphologically normal spermatozoa according to WHO criteria, and ⩽3% according to strict criteria. The AI was drastically lowered to ⩽3% normal acrosomes, and the TZI value was increased to ⩾2.09. Thus, according to Table IV, the cut-off point for sperm morphology—evaluated according to WHO (1992) criteria—was at ⩽21% morphologically normal spermatozoa. This resulted in a false-positive rate of 10.6% (high specificity), and a corresponding false-negative rate of 32.2%, based on the assumed 50% prevalence rate of male subfertility in an assisted reproductive setting.

Table V was compiled to compare the current results with those in other reports, and shows normal 1999 WHO values, the ROC curve cut-off points and the 10th percentile cut-off values of the current study and another study (Ombelet *et al*., 1997), and our recalculated ROC curve cut-off points.

Table I. Descriptive statistics of the semen parameters for the fertile (n = 107) and subfertile (n = 103) groups

Parameter	Fertile				Subfertile			
	Mean ± SD[a]	Range	Median	P10	Mean ± SD[a]	Range	Median	P10
Morphology (WHO) (% normal)	40.1 ± 14.1	9–69	38	24	21.7 ± 10.9	0–47	20	8.5
Morphology (SC) (% normal)	6.5 ± 3.9	1–19	5	2	3.0 ± 2.6	0–12	2	1
TZI	1.51 ± 0.2	1.17–2.07	1.54	1.33	1.81 ± 0.3	1.26–2.64	1.81	1.74
AI (% normal)	11.2 ± 5.9	1–33	10	5	5.6 ± 3.7	1–20	5	1
Concentration ($\times 10^6$/ml)	81.07 ± 49.7	1.30–230.0	75.0	20.0	18.97 ± 26.5	0.30–130.0	8.00	1.71
Motility (% motile)	53.1 ± 15.9	20–90	57.5	30.0	31.9 ± 19.2	2–80	30	10
Motile quality (0–6)	4.4 ± 0.6	2.5–6.0	4.5	3.5	3.5 ± 0.8	2.0–5.0	3.5	2.4

[a]Fertile versus subfertile populations. $P < 0.0001$ for differences between means of all semen parameters presented in the table.
AI = acrosome index; P10 = 10th percentile; SC = strict criteria; TZI = teratozoospermia index.

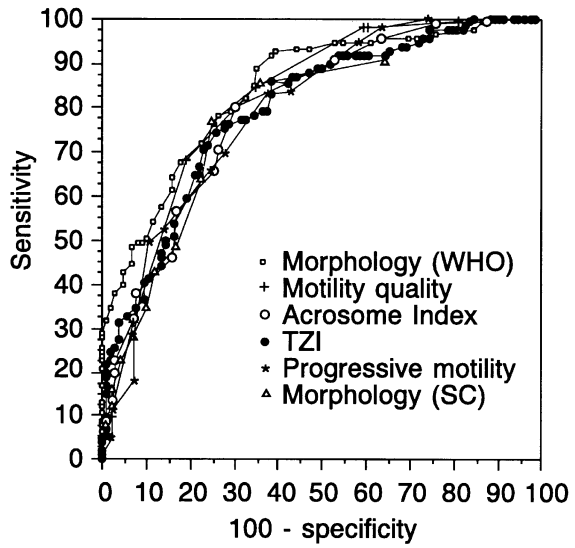

Figure 1. Receiver operating characteristics for different semen variables. SC = strict criteria; TZI = teratozoospermia index.

Discussion

The role of traditional semen analysis and semen parameters (including sperm morphology) as a prognostic factor of a male's fertility potential is a matter of on-going debate (Comhaire *et al.*, 1987; McDonough, 1997; Ombelet *et al.*, 1997). Especially for the in-vivo situation, there is a lack of information on normal and minimal values on sperm morphology, sperm concentration and motility, for the establishment of a male's fertility potential *in vivo*. Several methods have been used to establish these normal values or minimum cut-off points, including the comparison of fertile and subfertile populations without or with the aid of the ROC curve analysis method (MacLeod, 1950, 1951; Comhaire *et al.*, 1987; Ombelet *et al.*, 1997). Using the ROC curve analysis in this way, the cut-off point can be obtained to classify the male as possible fertile or subfertile. To determine the minimum values below which the occurrence of a spontaneous in-vivo pregnancy will be significantly reduced, cut-off points obtained by the lower 10th (Ombelet *et al.*, 1997) or even the lower 5th percentile (Comhaire *et al.*, 1987) values of a fertile population can be used, i.e. at 90 or 95% specificity respectively. Another option in this regard is to recalculate the ROC curve cut-off points by increasing the prevalence of the male factor, as was done in the current study.

In this study the results of the original ROC curve analysis indicated that sperm morphology evaluated according to WHO criteria was the best predictive parameter to distinguish between the fertile and subfertile populations, on the basis of the largest AUC (84.0%; Figure 1). The cut-off point was at 30% morphologically normal spermatozoa. Although sperm morphology evaluated by strict criteria showed the smallest AUC (78.2%) with a cut-off value of 4% morphologically normal spermatozoa, this AUC did not differ significantly (P = 0.0824) from morphology evaluated according to WHO criteria and all other semen parameters (Table III). The sensitivity and specificity (Table II) of the cut-off points with WHO and strict criteria were very similar, viz. 74.5 and 76.6% respectively for WHO criteria, and 74.5 and 77.4% respectively for strict criteria. The predictive score of 78.2% for strict criteria sperm morphology found in the current study is exactly the same as the value of 78% obtained by others (Ombelet *et al.*, 1997), although the cut-off value set by these authors was at 10% compared with the 4% for the current study.

The aim of the present study was not only to find cut-off points to classify males as fertile and subfertile, but also to establish possible lower cut-off points below which male fertility potential will be significantly reduced. Hence, use was made of the lower 10th percentile values of the fertile population and the ROC curve cut-off points recalculated, based on an assumed male factor prevalence of 50% in an assisted reproductive setting. From Table V it can be seen that the 10th percentile values from the current and another study (Ombelet *et al.*, 1997) are considerably lower than the original ROC curve cut-off values of this study. This can be interpreted that the lower limits for semen parameters where pregnancies can still be expected to occur are much lower compared with the cut-off point distinguishing between fertile and subfertile populations. This strengthens the previous argument (Van Zyl *et al.*, 1976; Menkveld and Kruger, 1990, 1996) that the lower limits needed still to obtain a pregnancy *in vivo* are much lower than the values for normality of a semen sample as published by the WHO.

Thus, in an assisted reproductive environment the aim should not merely be to distinguish between males who can be classified as possible fertile or subfertile. The aim must be

Table II. Estimated area under curve (AUC) and predictive cut-off point values for the individual semen parameters as obtained with the receiver operating characteristics (ROC) curve analysis between fertile and subfertile populations.

Variable	AUC	95% CI	Cut-off point	Sensitivity (%)	Specificity (%)
Morphology (WHO) (% normal)	0.840	0.752–0.898	30	74.5	76.6
Motility quality (1–6)	0.819	0.774–0.864	4.5	80.6	68.2
Acrosome index (% normal acrosomes)	0.797	0.737–0.857	8	73.5	70.8
TZI	0.794	0.731–0.857	1.64	73.8	71.7
Motility (% motile)	0.791	0.727–0.855	45	71.7	69.8
Morphology (SC) (% normal)	0.782	0.719–0.845	4	74.5	77.4

SC = strict criteria; TZI = teratozoospermia index.

Table III. Results (*P*-value) of pairwise comparison of areas under curve (AUC) to determine statistically significant differences between the semen parameters investigated

Parameter	Morphology (SC)	TZI	AI	Progressive motility	Motility quality
Morphology (WHO)	0.0824	0.2249	0.1682	0.2424	0.5518
Morphology (SC)	–	0.7653	0.4072	0.8267	0.3703
TZI	–	–	0.9346	0.9395	0.5177
AI	–	–	–	0.8729	0.5565
Motility (% motile)	–	–	–	–	0.3456

AI = acrosome index; SC = strict criteria; TZI = teratozoospermia index.

Table IV. Cut-off points for the different semen parameters to identify males who can be classified as possible subfertile based on 50% prevalence for a possible subfertile population

Semen parameter	Cut-off point	False positive	False negative
Morphology (WHO) (% normal)	⩽21	10.6	32.2
Motility quality (0–6)	⩽3.5	4.5	38.0
Acrosome index (% normal)	⩽3	11.5	40.1
TZI	⩾2.09	11.5	46.5
Motility (% motile)	⩽20	5.1	39.2
Morphology (SC) (% normal)	⩽3	18.2	29.7

SC = strict criteria; TZI = teratozoospermia index.

Table V. Comparison of normal WHO (1999) values, receiver operating characteristics (ROC) curve analysis cut-off values and 10th percentile values of fertile populations according to published reports and the current study

Semen parameter	WHO 1999 value	10th percentile values		ROC curve values		Adjusted cut-off points[b]
		Published[a]	This study	Published[a]	This study	
Concentration ($\times 10^6$/ml)	⩾20.0	14.3	20.0	34.0	Any value	–
Motility (grade a+b)	⩾50	28	30[c]	45	45[c]	⩽20[c]
(grade a)	⩾25	3	–	8	–	–
Motility quality (1–6)	–	–	3.5	–	4.5	⩽3.5
Morphology (WHO) (% normal)	⩾30[d]	–	24	–	30	⩽21
Morphology (SC) (% normal)	–	5	2	10	4	⩽3
Acrosome index (% normal acrosomes)	–	–	5	–	8	⩽3
TZI	⩽1.60	–	1.33	–	1.64	⩾2.09

[a]Data from Ombelet *et al*. (1997)
[b]From Table IV, for the classification as possible subfertile.
[c]% motile spermatozoa.
[d]WHO (1992) value, no value given in WHO (1999) manual.

to identify that proportion of males who will be contributing to the infertility problem of the couple. At the same time, as many males as possible must be excluded from being pronounced as subfertile with the subsequent unnecessary treatment procedures and possible sociological problems as mentioned previously. For this reason the ROC curve cut-off

points were recalculated, assuming a 50% prevalence of male subfertility in an assisted reproductive setting, and based on the total study population ($n = 210$). This resulted in lower cut-off values as presented in Table IV, and low false-positive rates (i.e. a high specificity). The readjusted cut-off point for sperm morphology, according to WHO criteria was at ⩽21% morphologically normal spermatozoa, with a false-positive rate of 10.6% and a false-negative rate of 32.2%. This means that if these recalculated cut-off points are used in a population of men attending an infertility clinic, one in ten men classified as subfertile will actually be fertile, while one in three men classified as fertile will be subfertile. It is more ethical to diagnose subfertile males falsely as fertile (false negative, on the basis of a semen analysis result above the recalculated cut-off values) than to diagnose fertile males as subfertile (false positive, on basis of a semen analysis result below the cut-off values). This approach will prevent over-treatment of potential fertile males, for instance referring the couple for ICSI treatment in cases where IVF might have been employed.

The adjusted ROC curve cut-off value for WHO sperm morphology found in the current study compared well with cut-off values found in other studies of fertile populations that were not based on the mean semen parameter values from the populations studied. In one such study (Bonde *et al*., 1998) a linear increase was found, independently of spermatozoa concentration, in the likelihood of pregnancies with an increasing proportion of sperm with normal morphology (WHO criteria) from 10 to 60%. Others (Barratt *et al*., 1995) concluded that the WHO (1992) cut-off point of 30% for normal sperm forms is not appropriate, as approximately half of the men in the fertile group they studied had a normal sperm morphology below this limit. It would therefore, appear that the cut-off value for sperm morphology evaluated according to WHO will be between 10 and 20% morphologically normal spermatozoa.

The adjusted cut-off value for sperm morphology evaluated according to strict criteria of ⩽3% morphologically normal spermatozoa found in this study (Table V) was near to the 10th percentile values of 5% morphologically normal spermatozoa found by others (Ombelet *et al*., 1997). It has been found (Van Zyl *et al*., 1990) that a definite cut-off point could be established at <4% morphologically normal spermatozoa with an in-vivo pregnancy rate of 11.5% and a pregnancy rate of 21.5% for the group of men with 4–9% normal spermatozoa. In a more recent study (Eggert-Kruse *et al*., 1996) it was found that, under in-vivo conditions, the pregnancy rate was significantly higher when semen samples had a better sperm morphology, the lowest thresholds being at >4% of strictly normal forms with a pregnancy rate of 21.5%. Therefore, it appears that the cut-off value for strict criteria sperm morphology may be in a range of 3–4% morphologically normal spermatozoa.

Using the 10th percentile values, it was found that the lower limit for the sperm concentration was 14.3×10^6/ml, and total motility was 28.0% motile spermatozoa (Ombelet *et al*., 1997) (see Table V). Although in the current study the choice was made not to use the sperm concentration in the ROC curve analysis (due to a possible bias in the selection of the subfertile population), this does not influence the results found among the current fertile population. The lower 10th percentile value of the fertile population (20.0×10^6/ml) can, therefore, be used with confidence. According to the recalculated ROC curve value, the cut-off point for sperm motility was at 20% motile spermatozoa. It has been reported (Van Zyl *et al*., 1990) that in 98% of in-vivo conceptions, sperm motility was ⩾30%, and the speed of forward progression (or motility quality) was ⩾2.0 (on a scale of 0–4). Others (Bonde *et al*., 1998) found that the likelihood of pregnancy was significantly decreased if the proportion of motile spermatozoa was <30%. It will therefore seem that a possible cut-off for sperm motility will be in the range of 20–30% motile spermatozoa.

The AI as an additional criterion in the diagnosis of a male's fertility potential was introduced during the mid-1990s (Menkveld *et al*., 1996), but has not yet been used for in-vivo diagnostic purposes. Based on the recalculated ROC curve cut-off point value, it would appear that the cut-off point may be at an AI of 3–5% normal acrosomes. The 1999 WHO manual sets the abnormal TZI value at >1.60, compared with the TZI value of 2.09 calculated in the current study. No other reports could be found where the TZI was used as a prognostic parameter in the in-vivo situation.

When interpreting the results of a study such as the current investigation, it must be borne in mind that the fertility status of the male—as proved by the achievement of a pregnancy—also depends on the relative fertility of the female partner (Van Zyl *et al*., 1990; Eggert-Kruse *et al*., 1996). Another important factor is the time of exposure, i.e. years of infertility. Hence, the terms 'fertile' and 'subfertile' are relative rather than absolute, and overlapping of semen parameter values between the two groups is unavoidable, as was found in the current study and elsewhere (Comhaire *et al*., 1987). Therefore, higher false-negative rates (with one in three of the men classified as potential fertile being actually subfertile) may be expected, as found in the current study. Therefore, the lower 'normal' values suggested as a result from the current study, based on the recalculated ROC curve cut-off points, are not absolute but serve rather as pointers where problems due to the male's subfertility might be expected. More studies are needed, based on standardized WHO protocols, and especially with regard to sperm morphology evaluation and motility parameters, to achieve the ultimate goal of correctly predicting male fertility potential.

In conclusion, the data from the current study and also from the literature [notably that of Ombelet *et al*. (1997)], indicated that cut-off values for normality as applicable to in-vivo fertilization are substantially lower than those proposed by the WHO manuals. The current data also indicated that two sets of cut-off points could be identified: one for the classification of normality, i.e. fertile or subfertile; and one indicating the lower limits of normality. This was achieved with the readjustment of the ROC curve cut-off values, based on a 50% prevalence of male subfertility. Using these guidelines, the cut-off points for males who may largely be responsible for the subfertility of the couple can now more readily be identified, and with low false-positive rates. If consensus on the lower normal limits for semen parameter values can be reached, the

basic semen analysis will—as it has been in the past—be an important cornerstone in the initial investigation of a male's fertility potential.

Acknowledgements

The authors wish to extend their gratitude to Mr Lulama Kepe, assistant in training to Dr C.J.Lombard, for help with the statistical analysis.

References

Abarbanel, A.R., Brown, W.E., Greulich, W.W. *et al.* (1951) Evaluation of the barren marriage. Minimal procedures. *Fertil. Steril*., **2**, 1–14.

Barratt, C.L.R., Naeeni, M., Clements, S. *et al*. (1995) Clinical value of sperm morphology for in-vivo fertility: comparison between World Health Organization criteria of 1987 and 1992. *Hum. Reprod*., **10**, 587–593.

Bonde, J.P.E., Ernst, E., Jensen, T.K. *et al*. (1998) Relation between semen quality and fertility: a population-based study of 430 first-pregnancy planners. *Lancet*, **352**, 1172–1177.

Comhaire, F.H., Vermeulen, L. and Schoonjans, F. (1987) Reassessment of the accuracy of traditional sperm characteristics and adenosine triphosphate (ATP) in estimating the fertilizing potential of human semen *in vivo*. *Int. J. Androl*., **10**, 653–662.

Comhaire, F.H., Huysse, S., Hinting, A. *et al*. (1992) Objective semen analysis: has the target been reached? *Hum. Reprod*., **7**, 237–241.

Efron, B. and Tibshirani, R.J. (1993) *Monographs on Statistics and Applied Probability 57. An Introduction to the Bootstrap*. Chapman & Hall, New York.

Eggert-Kruse, W., Schwartz, H., Rohr, G. *et al*. (1996) Sperm morphology assessment using strict criteria and male fertility under in-vivo conditions of conception. *Hum. Reprod*., **11**, 139–146.

Eliasson, R. (1971) Standards for investigation of human semen. *Andrologie*, **3**, 49–64.

Freund, M. (1966) Standards for the rating of human sperm morphology. A cooperative study. *Int. J. Fertil*., **11**, 97–118.

MacLeod, J. (1950) The male factor in fertility and sterility. An analysis of ejaculate volume in 800 fertile men and in 600 men in infertile marriages. *Fertil. Steril*., **1**, 347–361.

MacLeod, J. (1951) Semen quality of one thousand men of known fertility and in eight hundred cases of infertile marriages. *Fertil. Steril*., **2**, 115–139.

MacLeod, J. and Gold, R.S. (1951a) The male factor in fertility and infertility. II. Spermatozoon counts in 1000 men of known fertility and 1000 cases of infertile marriages. *J. Urol*., **66**, 436–449.

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

McDonough, P. (1997) Editorial comment: Has traditional sperm analysis lost its clinical relevance? *Fertil. Steril*., **67**, 585–587.

Menkveld, R. and Kruger, T.F. (1990) Basic semen analysis. In Acosta, A.A., Swanson, R.J., Ackerman, S.B. *et al*. (eds), *Human Spermatozoa in Assisted Reproduction*. Williams & Wilkins, Baltimore, USA, pp. 68–84.

Menkveld, R. and Kruger, F.T. (1996) Basic semen analysis. In Acosta, A.A. and Kruger, T.F. (eds), *Human Spermatozoa in Assisted Reproduction*. Parthenon Publishing, Carnforth, England, pp. 53–71.

Menkveld, R., Stander, F.S.H., Kotze, T.J.vW. *et al.* (1990) The evaluation of morphological characteristics of human spermatozoa according to stricter criteria. *Hum. Reprod*., **5**, 586–592.

Menkveld, R., Rhemrev, J.P.T., Franken, D.R. *et al*. (1996) Acrosomal morphology as a novel criterion for male fertility diagnosis: relation with acrosin activity, morphology (strict criteria) and fertilization *in vitro*. *Fertil. Steril*., **65**, 637–644.

Ombelet, W., Menkveld, R., Kruger, T.F. *et al*. (1995) Sperm morphology assessment: historical review in relation to fertility. *Hum. Reprod. Update*, **1**, 543–557.

Ombelet, W., Bosmans, E., Janssens, M. *et al*. (1997) 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., Bosmans, E., Janssens, M., *et al*. (1998) Multicenter study on reproducibility of sperm morphology assessments. *Arch. Androl*., **41**, 103–114.

Van Zyl, J.A., Menkveld, R., Kotze, T.J.vW. *et al.* (1976) The importance of spermiograms that meet the requirements of international standards and the most important factors that influence semen parameters. *Proceedings of the 17th International Urology Congress, Johannesburg, July 25–30, 1976*. Diffusion Doin, Paris, vol. **2**, pp. 263–271.

Van Zyl, J.A., Kotze, T.J.vW., Menkveld, R. (1990) Predictive value of spermatozoa morphology in natural fertilization. In Acosta, A.A., Swanson, R.J., Ackerman, S.B. *et al*. (eds), *Human Spermatozoa in Assisted Reproduction.* Williams & Wilkins, Baltimore, USA, pp. 319–324.

Wong, W,Y., Thomas, C.M.G., Merkus, J.M.W.M. *et al*. (2000) Male factor subfertility: possible causes and impact of nutritional factors. *Fertil. Steril*., **73**, 435–442.

World Health Organization (1980) *WHO Laboratory Manual for the Examination of Human Semen and Semen-Cervical Mucus Interaction*. Press Concern, Singapore.

World Health Organization (1987) *WHO Laboratory Manual for the Examination of Human Semen and Semen-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.

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

Received on September 8, 2000; accepted on March 1, 2001