WHAT CONSTITUTES MORPHOLOGICAL NORMAL AND ABNORMAL HUMAN SPERM HEADS

BY

Tholoana 'M' atahleho Leubane - Janse van Rensburg

Submitted In Partial Fulfillment of the Requirements

For the Ph.D. Degree in

Physiology

Department of Physiological Sciences

University of the Western Cape

Promotor: **Prof. G. van der Horst**
Co-promotor: **Prof. M.S. Bornman**

May 1998

DECLARATION

I, the undersigned hereby declare that the work contained in this dissertation is my own original work and has not previously in its entirety or in part been submitted at any university for a degree.

Signature J.M. Lewba - van Rensburg Date 05.05.1998

DEDICATION

This dissertation is dedicated to my mother , my husband and my daughter.

" 'M'e le Kedibone

Ke ka kholiso , likhothaletso le lerato la lona ke tsebileng ho fihla boemong bona, 'me kea le leboha kese ke kopa hore Morena Molima a le tsireletse, ' me a lefe khotso le bophelo bo botelele. Kea lebeho, **Maphuthing.** "Ngoana oa khotso ha a lebale tsa khotsong "

To my husband for his love, encouragement, support and tedious typing of this manuscript . Thanks a million for your patience and giving me a basic course in computer usage. God bless you.

" I will not forsake you, but I will be always with you " says the Lord of Host.

ABSTRACT

The percentage morphologically normal sperm appears to be of predictive value in the *in vitro* fertilization laboratory. However, the methodology used in this context is technically inaccurate and imprecise (subjective) and needs to be improved. Nevertheless, most laboratories continue to use such techniques. It is therefore not surprising to find that three different sperm morphology classification systems are in use. Some of the published methods include, the World Health Organization system (WHO), the Tygerberg strict criteria (TSC) and the Dusseldorf criteria (DC) to define morphologically normal sperm.

Each of these methods use different criteria and different cut-off values to define sperm morphology normality in patients. For WHO it is >30%, for TSC it is >14% and for DC it is >30%. Consequently, there is no objective morphological criteria for defining normal spermatozoa in human semen at present. Such criteria can be established only on the basis of extensive studies that assess the morphometric characteristics of spermatozoa. Therefore, there is a great need to standardize methodology in this context.

The second problem lies with the training of technicians. Unfortunately implementation of visual semen analysis often differs between laboratories. Moreover, few laboratories systematically train their technicians by one standard method, and monitor within and between technician variability. Accurate and precise visual semen analysis will only be achieved by implementing a program of international standardization and technician training and proficiency testing.

The third problem is that specimen handling and preparation for evaluation of sperm morphology are not standardized and this needs to be done to the highest degree possible.

In this investigation four microscopic techniques were used to study sperm morphology. The purpose of this part of the investigation was to test whether Papanicolaou stained (PAP) sperm smears studied by means of bright field microscopy represents a reliable method to study sperm morphology when compared to more sophisticated microscopic techniques. Consequently Bright field microscopy (PAP staining), Normaski differential interference microscopy (NDIM), Scanning electron microscopy (SEM) and Confocal microscopy of normal and abnormal sperm types were compared. However, a new technique had to be developed for preparing sperm for confocal microscopy. For this purpose both unfixed and fixed sperm were embedded in agarose to avoid motion artifacts during confocal imaging. Both groups of unfixed and fixed sperm were placed in PBS buffer containing 0.874mM dihexaoxacarbocyanine iodide (DIOC6(3)) at room temperature. Sperm were pre-loaded for 30 minutes with Tetramethyl rhodamine methylester (TMRM). DIOC6(3) is a lipophilic fluorescent dye, and it has been found that the fluorescent intensity of sperm increased over time.

For Laser Scanning Confocal Microscopy 3D sperm morphology was reconstructed from a series of 20 - 30 optical sections taken at 0.2:m intervals through each sperm. In order to examine 3D morphology a series of projected (rotated) views were reconstructed to allow visualization of a 3D animation set.

All forms of abnormal sperm could be clearly identified by all four microscopic techniques. Despite the fact that human sperm structure can be visualized and studied more comprehensively by means of both NDIM and confocal microscopy than with bright field microscopy, the latter technique of PAP stained smears is adequate to identify abnormal sperm morphology on a routine basis in the clinical laboratory. However, confocal microscopy of human sperm reveals that sperm may be scored normal/abnormal on the basis of its orientation. Only slight rotation of a 3D constructed image of a spermatozoon by means of confocal microscopy can change the classification of a sperm from normal to abnormal.

In the second part of the investigation bright field microscopy of PAP stained smears were used to test the reliability/repeatability of scoring among three technicians from three different laboratories using the Tygerberg Strict Criteria. All three technicians scored the same 77 patients. While two technicians scored within a relatively close range, the third technician varied more when TSC was employed. One technician classified 43% of the samples in a different TSC category when compared to a second technician. A fourth technician scored the same sperm smears but used the WHO criteria (1992). The scores for TSC and WHO were consequently compared. It was found that a 3-5% TSC range of abnormal sperm corresponded to a 6-30% WHO range.

In the third part of the investigation various dimensions of normal and abnormal sperm were quantified by means of quantitative image analysis. The Flexible Image Processing System (FIPS) was used in this context. The purpose of this part of the investigation was to develop a set of quantitative criteria which could be employed in automated computer analysis that will allow discrimination between normal and abnormal sperm forms. FIPS has a software program that draws on a complementary library of image processing routines that allows the image to be manipulated and characterised based on the grey scale analysis. FIPS can save images and data either as printed (hard copy) or as computer files. To suit specific requirements, the FIPS library of image processing routine can be expanded without changing the main program. For routine image analysis, FIPS has a macro facility that may be used to execute enhancement and measurement routines using standard parameters. The dimensions of specifically the sperm head and sperm acrosome were measured using the FIPS systems for parameters, surface area, perimeter, the minimum ferret (width), maximum ferret (length), average ferret, aspect ratio and the shape factor. The term ferret briefly entails 360 measurements (each time at a different angle) from the centre of the sperm head or acrosome.

ROC-curve analysis was used to establish cut-off points between normal sperm

and each abnormal sperm type for the parameters listed above of both sperm heads and sperm acrosomes . The best parameters for discriminanting between normal and abnormal sperm heads were sperm head length (Maxfer) and the average for sperm head length and width (Avefer). The poorest discriminators was aspect ratio and shape factor as determined by ROC-curves. By using the quantitative data of this investigation it will be possible in future to use a combination of the cut-off values for several parameters that will discriminate quantitatively and objectively among normal and abnormal sperm with a high level of specificity and sensitivity.

LIST OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 General aspects

Much attention has been paid to the question of infertility in women and men. Up to 15% of those trying to conceive experience infertility (Appleton *et al.* 1994). Very little is known about the extent of infertility in men. Hence, the development of suitable methods for selecting sperm for in vitro fertilization is important and sperm morphology appears to be important in this context.

Particularly since 1930 the debate on normal sperm morphology started. Moench and Holt (1930) concluded that when the number of abnormal forms is above 25% there is always sterility. Lane-Roberts *et al.* (1939) stated that no man with a satisfactory reproductive record produced more than 18% of sperm with abnormal head forms

Hotchkiss *et al.* (1938) found that 2.5% of the men in his study had more than 25% abnormal sperm. Macleod and Gold (1951) found that 39% of men in their study had more than 20% abnormal forms, 17% had more than 30%, and 3% had more than 50% abnormal forms. Falk Kaufman (1950) found that in 100 fertile men 11.5% had abnormal forms. Human sperm differs from that of rabbit (Curtis & Granance, 1995) and ram (Hendricks 1996) in having a relatively high natural variation in shape of sperm which makes it a problem to establish which forms are associated with infertility and which are normal variants. Freund (1968) and Hargreave and Nilson (1983), have echoed that it is rather difficult to guess which forms are associated with infertility.

Mcleod and Gold (1951) were among the first to show that sperm morphology was significantly different in fertile men than in infertile men. Since then different criteria have been used to evaluate sperm morphology and several classification systems have been developed.

Currently there is a general lack of uniformity among laboratories in classification system (Freund, 1966; David *et.al.* 1975; Fredricsson, 1979; Berenyi and Corradi, 1982). Many authors have indicated (Comhaire *et* al.,1995 and Macleod and Gold, 1951; Eliasson, 1971; Van Duijn *et al.*,1972; Helinga, 1976) criteria for defining sperm normality, while cells with border morpholgy must be considered abnormal. To the contrary, Page and Houlding (1951) have defined the criteria of abnormal spermatozoa and considered all other cells normal, but do not claim any advantage of this approach except for a reduction in the errors of judgement between the observers.

1.2 Classification systems for abnormal sperm and relationship to fertility

On the basis of the above information it is not surprising to find that three different systems are currently in use. Most follow published methods, such as the World Health Organization (WHO, 1987), Tygerberg Strict criteria (1986) and Dusseldorf criteria (1985) to define morphologically normal sperm.

Menkveld *et.al.* (1990) have claimed that the application of strict criteria for normality would give better results in terms of reproducibility, clinical accuracy and predictive power than the more liberal criteria described by the World Health Organization (WHO, 1987. 2nd).

Jeyendran *et al.* (1986) stated that the amorphous sperm head shape was the best predictor of fertility, in both the modified William system(1946) and the WHO system, predicting the proper fertility status in 65% of the cases. They further stated that using only normal sperm in the statistical model, normal sperm became a correct predictor in 66% of the cases in the modified William system. They further acknowledged that morphometric measurements were not useful in predicting the fertility status of the samples and that morphology can not be used as a sole predictor of fertilizing capacity of human spermatozoa.

Kruger *et al.* (1986-1990) have introduced the cut off values of >14% normal forms and described the fertilization rate of 7.6% during *IVF* cycle if the normal morphology was < 4% and 64 % fertilization rate when the normal morphology was > 4 %. Also, Oehninger *et al.* (1988) evaluated the pregnancy outcome in accordance with strict criteria. Their data confirmed the original Kruger study in that the normal (>14 %) group had 94 % fertilization compared to only 44.5 % in the poor TSC group with <14%

Frank *et.al.* (1994) stated that there has been a continuous debate on which criteria should be applied to define normal spermatozoa and which classification of abnormal forms is most appropriate. Many authors advocated the use of Tygerberg Strict Criteria for sperm normality, whereas cells with borderline morphology must be considered abnormal. On the contrary Page and Houlding (1951) defined the criteria of abnormal spermatozoa and considered all other cells normal.

The percentage morphologically normal sperm appears to be of predictive value in the in vitro fertilization laboratory (Gravance *et al.* 1995) However, the methodology used in this context is technically inaccurate and imprecise (subjective) and needs to be improved. Nevertheless, most laboratories continue to use such techniques.

Each of these methods use different criteria and has different cut- off values. Until recently at least 50% normal- shaped sperm were required according to WHO standards, but this considered sperm heads almost exclusively. In the latest edition of the WHO manual (WHO 1992) this value has been changed to 30%, with reference to flagellar disturbances, however, not supported by pertinent studies.

The Tygerberg cut- off value of normal sperm is >14%. This percentage was determined after correlation with *in vitro* fertilization, and is not comparable with percentage of sperm from *in vivo* (natural) conception and resulting pregnancy.

In contrast to the WHO classification and Tygerberg Strict Criteria group, the Dusseldorf classification is not confined to the number of normal shaped sperm(Haidi, 993), but rather allows conclusions about the kind and severity of malformation found (Hofmann and Haider 1985).

Perez-Sanchez *et al.* (1994) have reported that the recent strict criteria for the assessment of sperm morphology have improved the predictability of *IVF* outcome.

Hendricks (1996) has outlined that for human sperm especially, evaluation has been difficult due to the high percentage and variation of abnormal sperm forms prevalent in the semen sample. There are many forms of head abnormalities in a semen sample. There are ten sperm classes the oval "normal" , as well as nine classes shape and size head abnormalities including double, small, large / megalo or big, round, pear, narrow, taper, tear drop or peanut and flame or pencil heads. Some of these abnormal heads are due to environmental factors such as exposure to pesticides DBCP (Whorton, 1977) and kepones (Cannon, 1978) have been reported to result in oligozoospermia and some cases of azoospermia, Evans *et al.* (1981).

1.3 Sperm morphology in relation to external factors and relation with other semen parameters

Algren *et al.* (1974) reported that between 79 to 98% of human spermatozoa recovered from the ampulla were morphologically normal, while Asch (1976) also echoed that no abnormal forms were encountered. However, Mortimer (1982) has argued that this idea of morphologically normal sperm at the site of fertilization was an over simplification and that few abnormal sperm may reach the site of fertilization. They concluded that the selection of

morphologically normal sperm is not a direct function of the female tract but that sperm effect their own selection because of their motility.

Ragni *et al.* (1985) supported Mortimer's work that the mucous acts as a " passive filter " with selection depending on spermatozoa themselves in relation to motility.

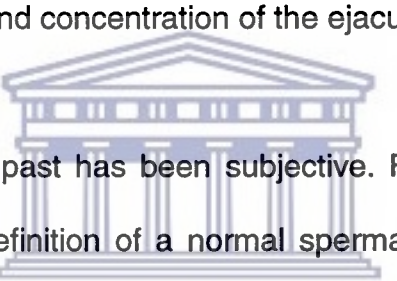
However, Hall, *et al.* (1995) echoed Mortimer's work that human cervical mucus has been reported to differentially select viable spermatozoa and act as a barrier to non-viable spermatozoa (Sujan *et al.*, 1963; Perry *et al.* 1977), but, with the development of assisted reproduction technology as treatment for both male and female factor infertility, this natural selection procedure is by-passed.

Damage to spermatogenesis has been reported in workers working with lead (Dancranjan, 1975), patients treated with a synthetic antifertility agent cyproterone acetate in prisoned volunteers (Mclead, 1974), cancer patients treated with chemotherapeutic agents (Qureshi, 1972), and patients who have had X-irradiation of the gonads. Also Viczian, (1969) reported an increase in the number of morphologically abnormal sperm in cigarette smokers compared with non- smokers. Chia *et al.*, (1994) added that the possible adverse effects of environmental factors on sperm quality has been a concern.

Vine (1996) also concluded that cigarette smoke contains known mutagens and carcinogens, and that there has been concern that smoking may have adverse effects on the male reproduction . He added that cigarette smoking is associated with modest reductions in semen quality including sperm concentration, motility and morphology.

Dibromochloropropane (DBCP) (Olsen, 1990), pesticides (Whorton, 1981), lead (Lancranjan, 1975), carbon disulfide and radiation are some of the better known reproductive toxicants.

In a study on the determination of fragmentation of DNA in human sperm and the correlation of detected DNA damage with semen analysis parameters and fertilization rates in the *in vitro* fertlization (*IVF*) Sun *et al.* (1997) found that the fragmented DNA was less than 4% in the majority of samples but ranged from 5% to 40% in approximately 27% of the samples. So they determined that there was a negative correlation between the percentage of DNA fragmentation and motility, morphology and concentration of the ejaculated sperm.



Normal morphology assessment in the past has been subjective. Fredricsson (1979) has accurately confirmed that even if the definition of a normal spermatozoa may be beyond dispute (that the normal sperm according to WHO classification is considered normal if it complies with the following criteria: the head has a smooth oval configuration with a well defined acrosome involving about 40% to 70% of the sperm head, as well as an absence of the neck, midpiece or tail defects; no cytoplasmic droplets of more than the size of the sperm head should be present; the length of a normal sperm head is 3 to 5µm, the width of 2 to 3 µm, midpiece should be between 5 to 7 µm length, 1.0 µm in width and the tail should be 45 mm long), the application of the criteria is subjective. He acknowledged that the situation has changed in that the criteria of the normal sperm has been more accurately defined by Eliasson (1971).

Schmassmann (1982) agreed that sperm morphology, however, is still being estimated in a subjective, unreliable and unstandardized way. Freund (1966) stated that apart from their

publication no method allowing an objective determination of sperm morphology has been done..

Katz *et al.* (1989) stated that morphologically abnormal sperm, as a group have inferior motility compared with normal sperm in the same ejaculate but that this differential swiming ability is not large enough to account for the exclusion of such a large population of abnormal sperm from mucous penetration.

While the above methodologies are subjective or semi-quantitative Katz *et al.* (1986) showed that some measurements of sperm may assess normality more accurately. Katz *et al.* (1990) have used videomicrography to simultaneously analyze the motion and morphology of individual human spermatozoa. In their qualitatitive study, sperm head length and width were analyzed for the penetration of normal and abnormal sperm into fresh human cervical mucus. They concluded that abnormal sperm swam slower in mucus than the normal sperm. When human spermatozoa in semen come into contact with cervical mucus only, a small fraction succeed in penetration. Thus, it appears that these abnormal sperm are less able to penetrate into and through the mucus than their morphologically normal counterparts, which means that cervical mucus acted as a biological "filter" that restricts the migration of abnormal sperm.

Davis and Gravance (1993) have echoed that no study has evaluated the relationship between objective measurements of sperm morphology and human fertility. However, the threshold for the percentage of normal sperm required to produce a clinical outcome *in IVF* is not necessarily an accurate or complete model for sperm fertility in the population at large.

So until the studies can be done that evaluate the relationship between objective sperm measurements, sperm classification based on these measurements, and in vivo reproductive outcome, the definition of normal sperm morphology will remain arbitrary and subjective.

From the above discussion it is apparent that there is still much uncertainty as to what constitutes a normal sperm and the potential relationship with fertility / infertility.

1.4 Preparation techniques and subjective interpretation of normal sperm morphology

It appears that many variations exist in preparing sperm smears for morphology assessment. These variations may further complicate standardization in sperm morphology assessment. The various preparation techniques and their effect on the percentage normal morphology will be subsequently described.

Considerable technical variation exists in preparing specimens for morphological analysis. These limitations have been long recognized in the laboratory evaluation of human semen (Davis and Gravance, 1995). In an attempt to reduce some of this technical variation, the World Health Organization developed a laboratory manual to standardize the human semen analysis (WHO, 1992). However, these procedures of semen analysis are predominantly subjective or semi-quantitative. Today, most laboratories have adopted some version of the WHO procedures.

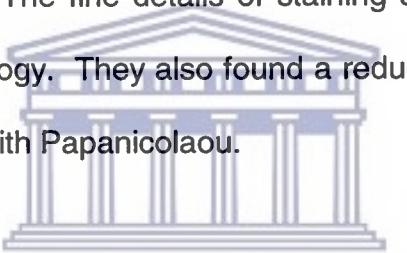
However, recent studies have shown that implementation of these methods remains difficult, because considerable variation within and between technicians and laboratories has been found (Zaini *et al.*, 1985; Baker & Clarke, 1987; Dunphy *et al.*, 1989; Neuwinger *et al.*, 1990)

This variability was demonstrated by studies of Chong *et al.* (1983); Mortimer *et al.*(1986); Jequior and Ukombe (1983) and Dunphy *et al.*(1982). In addition, Knuth *et al.* (1989) demonstrated that long term variability in semen analysis may pass unnoticed without appropriate quality control. A reason for the high variability are the subjectiveness of the techniques and a lack of standardized laboratory procedures (Neuwinger *et al.*, 1990). Meschede *et al.* (1993) have analyzed how the results of sperm morphology assessment were influenced by different staining techniques. They concluded that only one standard method should be recommended for the preparation slides for morphology in order to ensure inter-laboratory comparability.

The morphologic evaluation of spermatozoa is a matter of personal opinion according to Fredricsson (1978). Even if the definition of normal spermatozoa may be beyond dispute, the application of the criteria is subjective. This means that the tradition of analysis may develop in different directions in different laboratories as shown by Freud (1966). Fredricsson (1978) has agreed that different initiatives have been taken to improve the situation and the criteria of normal spermatozoa were more accurately defined (Eliasson, 1971). In spite of this there is often reason to suspect that laboratory data, particularly with regard to sperm morphology, cannot be accurately translated from one laboratory to another. Yang *et al.* (1995) confirmed that the judgement of sperm morphology is thought to be subjective and comparisons are difficult to make among the different laboratories and sometimes even within the same laboratory.

In many cases the same methods are used to fix and stain specimens for the morphology assay from different species and no data exists that verifies that a particular specimen preparation method is optimal for any species. In order to avoid this subjectivity over many years numerous studies that incorporate image analysis techniques in the sperm morphometric analysis have appeared (Schmassmann *et al.*, 1982; Katz *et al.*, 1986; Jagoe *et al* , 1990; Wang *et al*, 1991 a, b; Davis *et al.*, 1992 a, b; Kruger *et al.*, 1993).

However, Harasymouycz *et al.*, (1976) suggested that methods that include dehydration protocols will results in cell shrinkage. The fine details of staining and fixation techniques can influence the fine details of morphology. They also found a reduction in the dimensions of human sperm head when assessed with Papanicolaou.



Kruger *et al.* (1986, 1988) and Menkveld *et al.* (1990) came out strongly that the methodology required careful training of technicians and subjectivity cannot be avoided. Hence in a recent paper they have extented variations in staining technique comparing Diff Quick and Papanicolaou (Menkveld *et al.*, 1997). They concluded, however, that Diff-Quik has the advantage in having a clear background of the stained smears. Furthermore, Diff-Quik staining is time-saving procedure compared to other staining techniques.

Most of the above studies show that visual semen analysis is technically inaccurate and imprecise. Nevertheless, most laboratories continue to use such techniques. Unfortunately, the implementation of some technique often differ between laboratories, because such protocols can be vague, contradictory or confusing (Davis & Gravance 1993). Moreover,

few laboratories systematically train their technicians by any standard method and monitor within and between technicians variability.

Based on this data, it is clear that accurate results of visual semen analysis will only be achieved by implementing an international guideline or protocol (Katz, 1993) and training technicians, given the certificates and proficiency testing. Unfortunately the cost maybe high but it would reduce the technicians variation.

Davis and Gravance (1993) have also confirmed that even a well executed protocol for the quality of sperm morphometric analysis would be self limiting, and would not lead this biological field into the 21st century.

1.5 Specimen handling

It seems particularly desirable that specimen handling and preparation for evaluation of morphology be standardized to the highest degree possible (Meschede *et al.,* 1993). It has not been clarified, however, to what degree the results of morphology analysis are influenced by different preparations and whether the methods recommended by WHO (1987) yielded comparative data. Katz *et al.* (1993) concluded that when sperm are prepared for morphologic examination, most laboratories use smears in which the cells are dried and stained. It is likely that methods incorporating dehydration will cause cell shrinkage.

Not only the standardization of sperm morphology Frank *et al.,* (1964) have asserted that the microscopic study of spermatozoa, the morphology of sperm head constituted the greatest single source of information as to the fitness of these cells for reproduction. So, it appears

necessary to try to define predictive criteria of sperm performance and preparative standard and staining procedures in order to have a single uniform standard of sperm morphology.

Improved specimen preparation procedures and objective computer sperm morphometric analysis methods have been developed recently to reduce these technicians variations in sperm analysis (Perez-Sanchez *et al.*).

The aim of this present study was to critically compare the existing methodologies and employ computer assisted techniques to describe what constitutes the morphological normal and abnormal human sperm head. Five techniques were used and included bright field microscopy, Nomarski differential interference microscopy, Flexible Image Processing System (FIPS), Scanning Electron Microscopy (SEM) and Confocal Microscopy.

The main aims of this investigation using the above techniques were:

1.6 AIMS

1. to compare sperm head morphology using different microscopic techniques and assess their advantages and disadvantages

2. to monitor technicians methodology for evaluating sperm morphology. Since considerable technical variation exists in evaluating sperm morphology the same samples were analyzed by different technicians using both strict criteria as well as

WHO. The aim was to assess inter-laboratory concordance as well as the relationship of TSC with WHO.

3. to identify a method of preparation and specimen staining that optimized, the accuracy and precision of the sperm morphometry analysis and the selection of a standard uniform method for confocal microscopy.

4. to develop quantitative and semi-quantitative criteria for assessing sperm head morphology and determine the criteria for identifying what constitutes a normal sperm.

5. Once an objective assessment of sperm morphology is available it can be used with other quantitative methods for a wide range of applications:

   a) Reproductive toxicology.
   b) In the *in vitro* fertilization laboratory.
   c) Outcome of male contraceptives.
   d) Assessment of fertility *in vivo.*

CHAPTER 2

2. MATERIALS AND METHODS

2.1 Sperm collection and Preparation

Semen samples were obtained from 3 groups of men.

The first group consisting of 500 individuals aged between 25-40 years were men attending the Groote Schuur hospital infertility clinic with their wives aged between 25-38 years. At this clinic it was estimated that 40% of the cause of infertility was female tubal factors, 20% was endometriosis and ovarian dysfunction (in the females) and 40% of the causes were related to male problems.

The second group, was the control group ("normal group"), consisting of 50 clinically healthy male volunteers aged between 25-45 years who regularly supplied ejaculates for cross-examination and testing in comparison with semen from fertility patients. All ejaculates within this control group were completely normal according to the routine clinical standard, with high sperm densities, high progressive motility and good longevity of the sperm.

The third group was another control group consisting of 77 semen samples which were donated by proven fathers who claimed to have had one or more children in the past 6 to 18 months and who did not suffer from any urogenital or systemic diseases.

Semen specimens were produced after three days of sexual abstinence by masturbation, and then analyzed within 1 hour of production. Because semen samples present a possible biohazard as they may contain harmful viruses e.g. hepatitis viruses, human Auto Immune Deficiency viruses (AIDS) and herpes viruses, extreme care was taken in handling the semen samples, e.g. wearing gloves, masks, spectacles and working in fume cupboards. Semen samples were liquefied by placing in an incubator at 33°C for 20 minutes before routine semen analysis (SA) by means of light microscopy. Routine SA included measurement of sperm concentration, motility, forward progression, mixed antibody reaction (MAR) and morphology. Papanicolaou stain was used with bright field oil immersion optics. Tygerberg Strict Criteria were used to assess sperm morphology of all samples. WHO sperm analysis assessment was prepared on the 77 proven fathers only.

2.2 Subjective evaluation by three technicians

From each semen sample three slides were prepared. One was an unstained "wet" preparation, the second was stained according to Papanicolaou (1942) and the third one was left unstained (to be used later if necessary). Air dried stained slides according to Papanicolaou (1942) were viewed by three independent technologists, unaware of the results obtained by each other. At least 100 sperm were scored per patient.

One was from the Centre for Fertility Studies at Pretoria Central Hospital, the other was from Andrology at Groote Schuur Hospital and the third was the author from the University of the Western Cape. This was done to establish the extent of conformity with different technologists and laboratories. All three technologists had similar training in Tygerberg Strict Criteria. One techonologist scoring the same slides had training from a certified WHO group.

2.3 PREPARATION OF SEMEN FOR VARIOUS MICROSCOPIC METHODS.

Four techniques for analysis were used: Bright field and interference light microscopy. For all four techniques the methods for preparation and isolation are the same as below. Flexible Image Processing System (FIPS) (which is a recent tool for morphometric analysis using image analysis), scanning electron microscopy (SEM) and confocal microscopy.

2.3.1 LIGHT MICROSCOPY

Evaluation of microscopic slides.

Most slides were evaluated under bright field using a research microscope (Zeiss) equipped with a 100x oil (N.A. 1.3) immersion lens. The slides were also examined by phase contrast optics using a 40x objective.

2.3.2 INTERFERENCE LIGHT MICROSCOPY

For the "wet" preparation a drop of undiluted semen (5 µl) was spread out gently with the unused glass slides onto a 76 x 25 slide. The slide was mounted and immediately examined. The samples were viewed with the (Nomarski Differential Interference) using a Zeiss Research Microscope (D-7082 Oberkochen). Spermatozoa were photographed at various magnifications using an MC 63 automatic photomicrographic camera (for 35mm film) which was mounted on the microscope.

2.3.3 PROCESSING FOR SCANNING ELECTRON MICROSCOPY (SEM)

Part of the sample was used primarily for the investigation and remainder of the ejaculate was handled as follows in order to carry out the other assays. Sperm cells were separated from seminal plasma by centrifugation at 200 xG for 10 minutes (Beckman TJ-6 centrifuge)

at room temperature. The sperm pellet was re-suspended in phosphate buffer saline (PBS) at pH 7.4 and then fixed in 2.5% gluteraldehyde in 0.1 M Phosphate buffered saline (PBS) until required for microscopy.

Sperm were separated from seminal plasma by centrifugation at 1000 x G max. for 10 minutes at room temperature. The sperm pellet was re-suspended in phosphate buffer saline (PBS) at pH 7.4 and then fixed in 2.5% gluteraldehyde in 0.1 M PBS until required for microscopy. The sperm suspension in the syringe was slowly (0,5ml over 30 seconds) injected into the filter which allowed fluid and small particles through but retained the sperm which are larger than the pore sizes of 3µm employed. Subsequent processing involved slowly injecting the processing fluid with a tuberculin syringe on the membrane filter. Sorenson's phosphate buffered 1% Osmiumtetroxide was then injected into the filter holder. Osmiumtetroxide fixative was washed with Sorenson's phosphate buffer for the routine dehydration with 70%, 90%, 95%, absolute alcohol (15 minutes each) using the injection method described above. The specimens were always submerged in the desired fixative (Van der Horst *et al.*,1989). Each membrane filter with trapped sperm was rapidly transferred into a metal basket and placed in the precooled chamber of a critical point drier (Hitachi X650). The membrane filters had to be kept soaked in absolute alcohol during these transfers. However this necessitated removal of ethyl alcohol by three to four washings with liquid carbon dioxide in a chamber of the critical point drier.

After critical point dying the membranes were attached to SEM specimen stubs with press on adhesive tabs. The material was sputter-coated with gold (10-18nm) using an Edwards 5150B sputter coater.

The sputter-coated specimens were viewed with a Hitachi X650 scanning electron microscope. Photograph of the SEM images were taken with a camera which was incorporated into the SEM.

2.3 4. LASER SCANNING CONFOCAL MICROSCOPY

For confocal microscopy sperm were either viewed live or after fixation.

Both unfixed and fixed sperm were embedded in agarose to avoid motion artefacts during confocal imaging.

Both groups of fixed and unfixed sperm were placed in PBS buffer containing 0.874 mM dihexaoxacarbocyanine iodide ($DIOC_{6(3)}$) at room temperature. $DIOC_{6(3)}$ is a lipophilic fluorescent dye, and it has been found that the fluorescent intensity of the sperm increased over time up to 3 days. Therefore for all studies of 3D morphology, 3 days staining gave optimal fluorescence for sperm.

Living sperm were examined after 1 hour of staining and fixed sperm were stained for 1-3 days before examination. Stained sperm were embedded in solution of 1% agarose in PBS. To prepare the agarose/PBS, 1% agarose was heated to 60°C to dissolve in PBS and then cooled to room temperature at which time the sperm were added and re-suspended in the solution. A small volume of agarose suspension was placed onto a coverslip in the microscope chamber and cooled on ice for 5 minutes to cause the suspension to gel and thus immobilise the sperm.

Sperm were preloaded for 30 minutes with 1ml Tetramethyl rhodamine methylester (TMRM). 200µl TMRM was also added to the agarose solution in order to compensate for dye loss and photobleaching. TMRM staining of mitochondria in cells indicates their energised status, and thus a good indicator of viability.

The sperm, stained and embedded in agarose in a stainless steel chamber were viewed, on a Zeiss LSM 410 Inverted microscope. A 63x oil immersion lens (1.4 NA) was used for imaging and the image zoomed further by a factor of 6 (using the Zeiss LSM software), to give a suitable magnification. For $DIOC_{6(3)}$, the filter setting for excitation and emission was a 488nm band pass and 515 long pass filter respectively; and for TMRM, 568nm and 590 long pass. For morphological studies, 3D sperm morphology was reconstructed (using the Zeiss LSM 3D software) from a series of 20 - 30 laser optical sections taken at 0.2µm intervals through each sperm. Several examples of all the normal and abnormal sperm categories were studied

Optical sectioning was accomplished by changing the stage (Z) height using the LSM (see fig. 3. 4 A). In order to examine the 3D morphology a series of projected (rotated) images were reconstructed to allow visualization of a 3D animation set. Since the animation cannot be shown here some examples of 3D projection are provided (See fig. 3.4 A).

2.3.5 FLEXIBLE IMAGE PROCESSING SYSTEM (FIPS)

Stained preparations according to Papanicolaou (1942) of semen samples of 500 patients from Groote Schuur andrology clinic were analyzed and examined under oil immersion using a 100x plan apochromatic bright field objective to determine sperm morphology. The images were relayed from Grundig video camera mounted on a Zeiss D - 7082 (Oberkochen)

transmitted light research microscope to a computer containing the FIPS software. Measurement of the above were calibrated with respect to a videotaped micrometer slide. An image of sperm was captured and analyzed with a flexible image processing system (FIPS) (CSIR, South Africa) in the automated mode. At least 100 per sperm morphology type was analyzed by means of FIPS.

Cells were displayed on the video monitor and each sperm head image was processed by separating the head from the tail using a specific function of FIPS, for image enhancement and thresholding. Analysis of the sperm midpiece and tail was excluded.

FIPS is a powerful image processing system that allows video image to be captured, characterised and stored using widely available desk top personal computing equipment. The system comprises a basic image handling program that allows images to be displayed, zoomed and positioned. This program draws on a complementary library of image processing routines that allow the image to be manipulated and characterised based on the grey scale analysis. FIPS can save images and data either as printed (hard copy) or as computer files. To suit specific requirements, the FIPS library of image processing routine can be expanded without changing the main program. For routine image analysis, FIPS has a macro facility that may be used to execute enhancement and measurement routines using 'standard' parameters. The dimensions of specifically the sperm head and sperm-acrosome, were measured using the FIPS systems for parameters, surface area, perimeter, the minimum ferret (width), maximum ferret (length), average ferret, aspect ratio and the shape factor. The term ferret briefly entails 360 measurements (each time at a different angle) from the centre of the sperm head. This central area was determined by the computer after thresholding to entire sperm head, or sperm acrosome.

A key feature of FIPS is its ability to analyse an image automatically, pixel by pixel according to the pixel colour within a scale ranging from black to white. All measurements made on the screen image are related to actual physical dimesions. This is done by calibrating the image screen using an object (displayed) of a known actual size or by using a calibration slide. The image can be stored as a *.TIF image, and displayed as a stored image. The ENHANCEMENT option allows the displayed image to be enhanced using several routines, before actual measurements can be performed. A valid calibration file must be loaded before any measurements can be done. If a valid calibration file is loaded, a prompt for a 'results' file name is displayed. The prompt allows the results to be appended to, or overwritten on, an existing file. When the result filename is selected, the following window will be displayed in the dialogue box: area, perimeter, minimum ferret, maximum ferret, average ferret, aspect ratio and shape factor. Measurement is done by moving the mouse cursor over the object and pressing the left mouse button. The colour of the selected object will change to blue and relevant measurement values will be displayed in the dialogue box. A prompt to SAVE the measurement will be given after each operation. The results file will bear a .ARE extension. A prompt to print the results file is also displayed.

2.3.6 Statistics analysis

Descriptive statistics and ANOVA analysis for comparisons among sperm morphology types were performed using the MedCalc statistical program (Schoonjans, 1996) and Statgraphics Plus for Windows. Box and Whisker plots as well as ROC-curve analysis were performed using the MedCalc statistical package (Schoonjans, 1996).

CHAPTER 3

RESULTS

3.1 Description of normal and abnormal sperm morphology using different microscopic techniques

3.1.1 General

The aim of this part of the investigation was to study human sperm morphology using different microscopic techniques in order to assess whether routine bright field microscopy of Papanicolaou stained smears represent a reliable method to distinguish between normal / abnormal spermatozoa . Four different microscopic techniques were accordingly compared; bright field microscopy of Papanicolaou stained sperm smears made according to the method of WHO (1987); Nomarski differential interference microscopy of fresh or glutaraldehyde fixed spermatozoa: Scanning electron microscopy of glutaraldehyde / osmium tetroxide fixed spermatozoa and Laser Scanning Confocal Microscopy of flourochrome treated sperm.

3.1.2 Bright field microscopy (Papanicolaou staining)

Figs. 3.1.1 (A to H) are photomicrographs representative of Papanicolaou stained sperm smears as viewed by bright field microscopy. These figures show normal and most types of abnormal sperm heads. Fig. 3.1.1 (A, C, D and H) show normal spermatozoa which typically have a smooth head, an oval configuration with a well defined acrosome involving about 40% to 70% of the sperm head, as well as an absence of the neck, midpiece or tail defects. No cytoplasmic droplets of more than

size of the sperm head were present (Figs. 3.1.1 A, C, D and H). The length of a normal sperm head is 3 to 5μm, the width of 2 to 3 μm, midpiece should be between 5 to 7 μm length, 1.0 μm in width and the tail should be 45 μm long). Fig. 3.1.1 (A) shows an example of a sperm with a round head and shows that the length and the width of the head are the same dimensionally. Fig. 3.1.1 (B) shows an example of an amorphous head. These are spermatozoa that are difficult to define in terms of the length or the width. They are therefore irregular in shape. Fig. 3.1.1 (C and F) shows an example of a tapered head, long-headed spermatozoa have hitherto been characterized by a distinct increase in head length and the head is accompanied by slight thinning. Fig. 3.1.1 (D) shows an example of double, small and flame heads. Double headed spermatozoa are characterized by two small sperm-cells that are joined and small headed spermatozoa show a decrease in the length and the width of the head while flame headed spermatozoa also show a reduction of the acrosome and post-acrosome and the general decrease of the length and width. Fig. 3.1.1 (E) shows an example of pin, big, flame heads. These big headed spermatozoa have been characterized by clear distinct increase in the lengths and widths of the head while the pin headed spermatozoa show the total reduction of the length and width of head and the scattering of the mitochondria and occasionally of remnants of the cytoplasmic droplets. Fig. 3.1.1 (G) shows an example of double and big heads and Fig. 3.1.1 (H) shows an example of amorphous head. However, in some cases it is difficult to classify all sperm on a smear clearly. The major difficulty appears to distinguish between what may be considered as normal and borderline amorphous (Fig. 3.1.1 H). By and large, however, normal versus abnormal can be distinguished by the trained technician

with relative ease using the above technique. However, care should be taken to evaluate sperm head morphology when the sperm smear is too thick or sperm concentration is too high (Fig. 3.1.1 B) or the smear is either over or understained (Fig. 3.1.1 A).

Description of different sperm forms by using different microscopic techniques to established which complimenting techniques indicate similarities in the definition of various sperm forms

3.1.2 Bright field microscopy of Papanicolaou stained smears.

Figs. 3.1.2: A - H show the main types of normal and abnormal sperm as encountered in routine smears

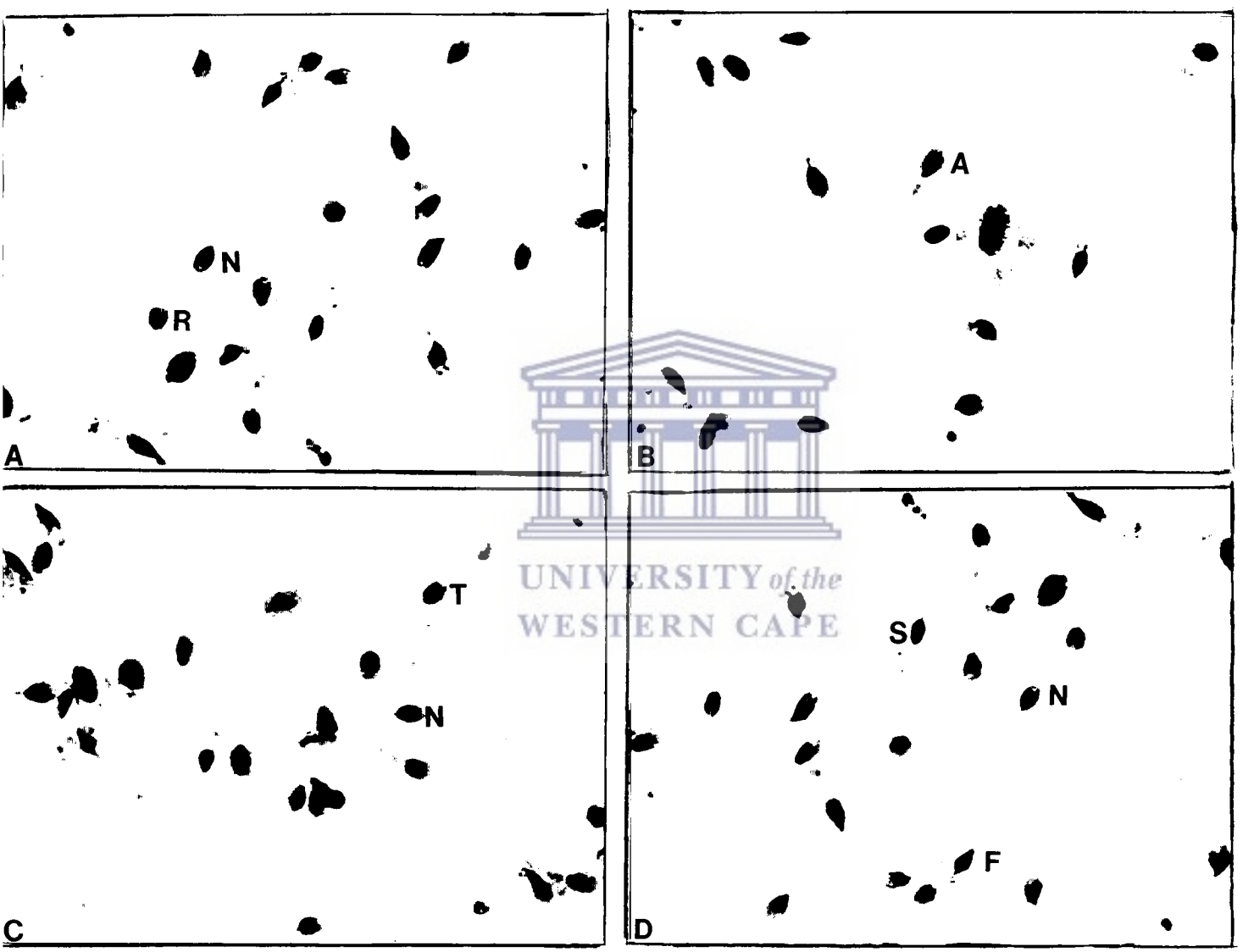


Fig. 3.1.2: A -H: Bright field micrographs of human sperm from fertile and infertile patients. Showing all forms of normal and abnormal human sperm. Sperm were stained according to Papanicolaou (Mag. X100 using bright field oil-immersion optics) .

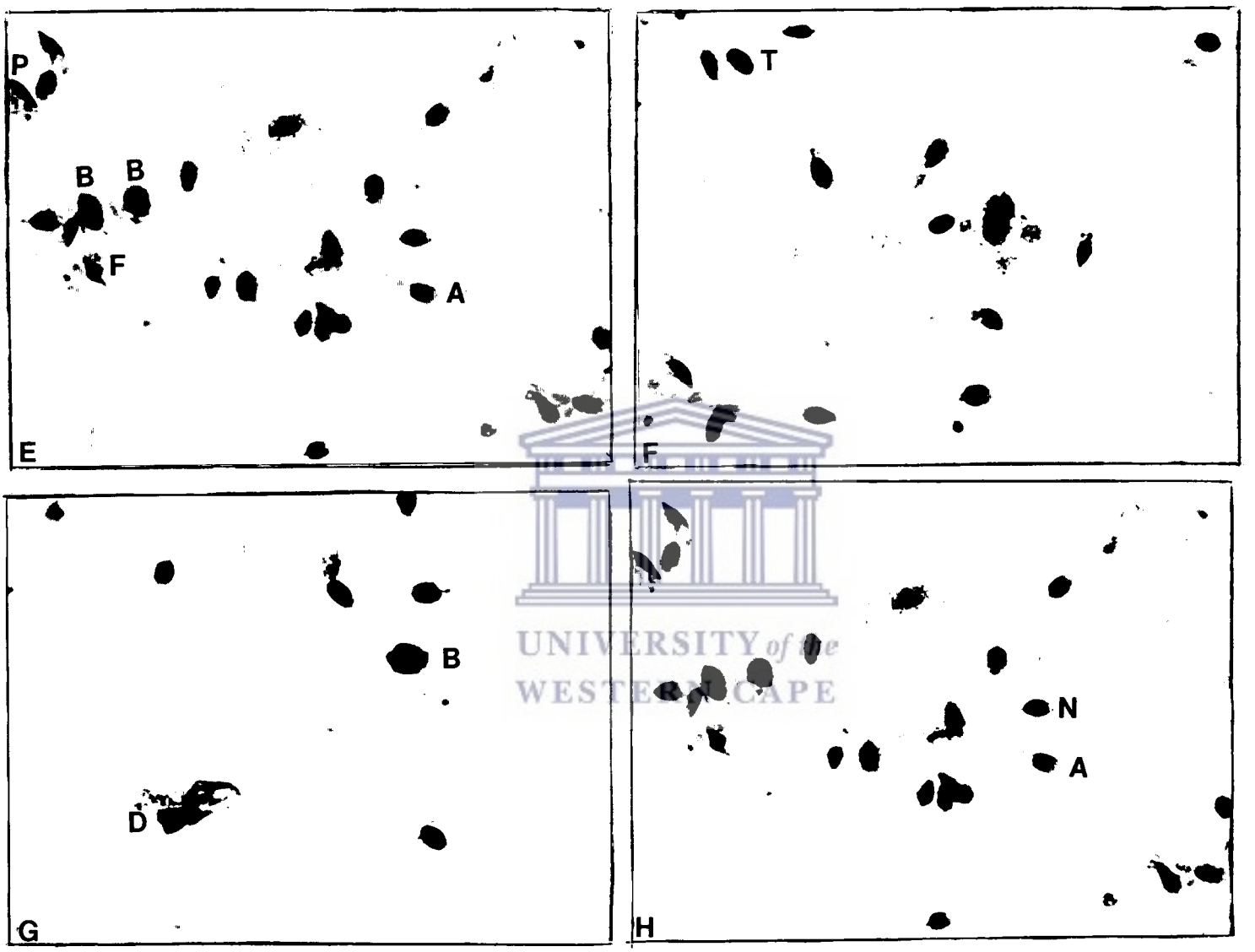
(A) Round / Normal head

(B) Amorphous head

(C) Normal / Tapered head

(D) Small head / Flame / Normal head

3.1.2: Bright field microscopy of Papanicolaou stained smears



(E) Pin / Big / Flame / Amorphous head

(F) Tapered head

(G) Double / Big head

(H) Normal / Amorphous head

3.1.3 Nomarski Differential interference

Figs. 3.1.2 (A-H) represent photomicrographs of Nomarski differential interference microscopy. In all cases unstained spermatozoa are viewed and their outlines can be seen with much greater clarity and the three dimensional appearance furthermore assists to distinguish among the different sperm forms. It is also noticeable that in most cases the acrosome can be distinguished despite the fact that the spermatozoa are not stained. All types of sperm head abnormalities were also evident using this type of microscopy Fig. 3.1.2 (A-H). A feature observed in most human sperm using this technique were small indentations in different parts of the head and these resembled cytoplasmic vacuoles. However, these may represent artefacts of fixation or preparation. The fact that sperm head outline can be seen with such great clarity may make this technique a good candidate for future routine application in the clinical andrology laboratory. The major advantage of this technique is that live or glutaraldehyde fixed but unstained sperm are visualized instantaneously. The disadvantage is that the acrosome can in some instances not be clearly seen due to protein coatings.

3.1.3 Nomarski Differential Interference

Figs. 3.1.3: A - H: show the main types of normal and abnormal sperm as observed in routine smears

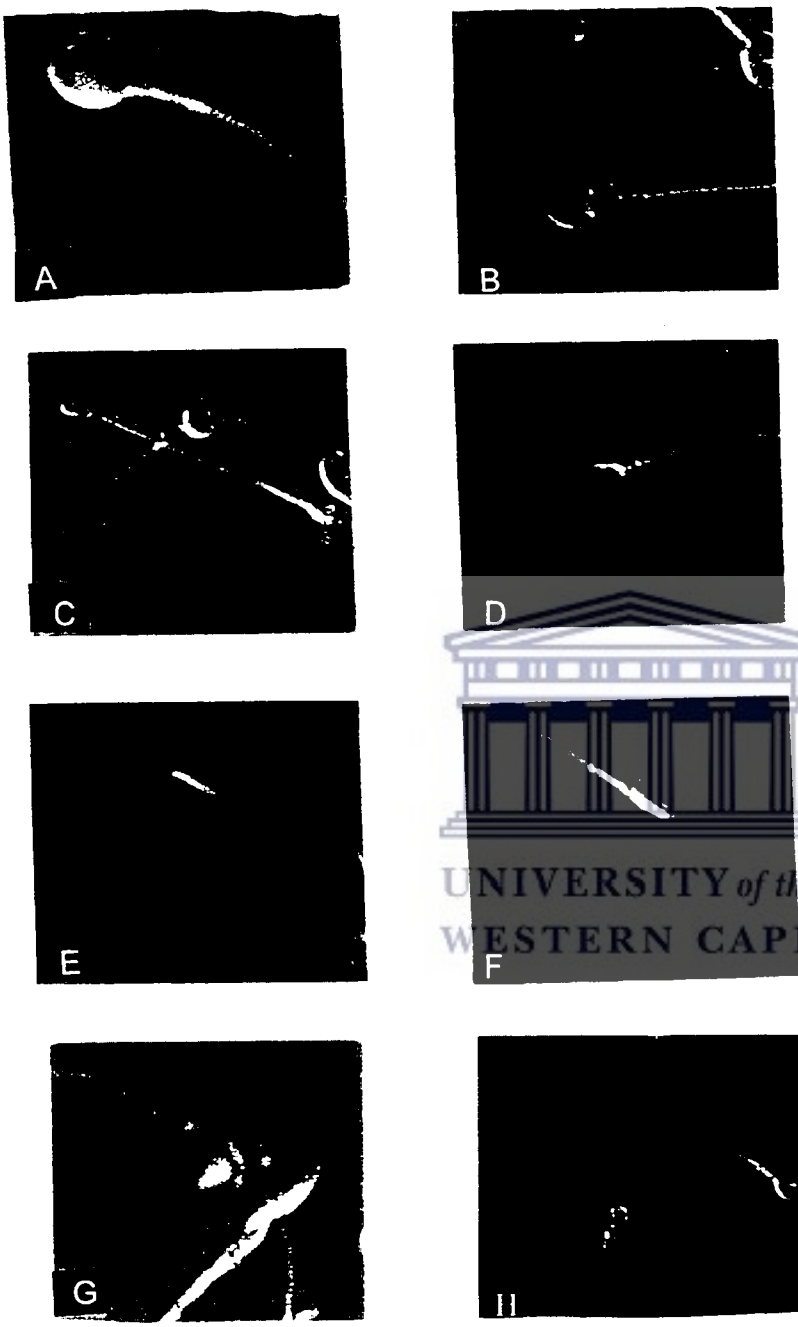


Fig 3.1.3: A - H Differenetial Interference micrographs of human sperm from infertile patients (Mag. X 100 using bright field oil - immersion optics) showing human spermatozoa fresh and unstained.

(A) Normal head
(B) Big head
(C) Round and Small heads
(D) Tapered / Elongated heads
(E) Pin head
(F) Flame head
(G) Normal head with cytoplasmic droplet still attached
(H) Small head with dag-defect

3.1.4 Scanning electron microscopy

Figs. 3.1.3 (A to P) are photomicrographs representative of scanning electron microscopy of glutaraldehyde / osmium tetroxide fixed spermatozoa. These figures show normal and most types of abnormal spermatozoa heads. Fig. 3.1.3 (A and B) shows examples of normal spermatozoa. Fig. 3.1.3 (C) is an example of spermatozoa that have attached cytoplasmic droplets, but the head is amorphous with the acrosome occupying between 40% - 70% of the sperm head while the small sperm head can be seen with clarity. Fig. 3.1.3 (D, O and P) is an example of a round headed sperm. Here the length and width of the head are the same dimensionally. Fig. 3.1.3 (E) shows pin headed spermatozoa with the midpiece defect and a scattering of the mitochondria and occasionally of remnants of the cytoplasmic droplets. Fig. 3.1.3 (F) shows an example of tear drop-shaped sperm head and it can be seen that there is a reduction of the length and width together with the scattering of the mitochondria and remnants of cytoplasmic droplets. Fig. 3.1.3 (G) shows pear-shaped head spermatozoa . Fig. 3.1.3 (H) represents of double headed spermatozoon. This sperm appears as if two sperm cells have been joined. Fig. 3.1.3 (I and L) represent examples of tapered spermatozoa showing an increase in length and width of the head. Fig. 3.1.3 (J) is an example of a big headed spermatozoon which clearly shows a distinct increase of the length and width of the head. Fig. 3.1.3 (L) represents an amorphous sperm head showing an irregular sperm head. Fig. 3.1.3 (M and N) respectively represent the examples of spermatozoa with normal head with dag-defect while the other photomicrograph shows flame shaped-spermatozoa also with reduction in length and width of sperm head.

While most normal /abnormal sperm forms can be seen with a great deal of clarity, the technique is not only cumbersome and tedious but often sperm are covered in protein coatings. These coatings obscure surface details and makes classification of normal/ abnormal difficult So, for the above reasons the technique can not be used as a method for routine semen analysis in the clinical andrology laboratory.

3.1.4: Scanning Electron Microscopy

Figs. 3.1.4: A - P show the main types of normal and abnormal sperm as encountered in routine smears

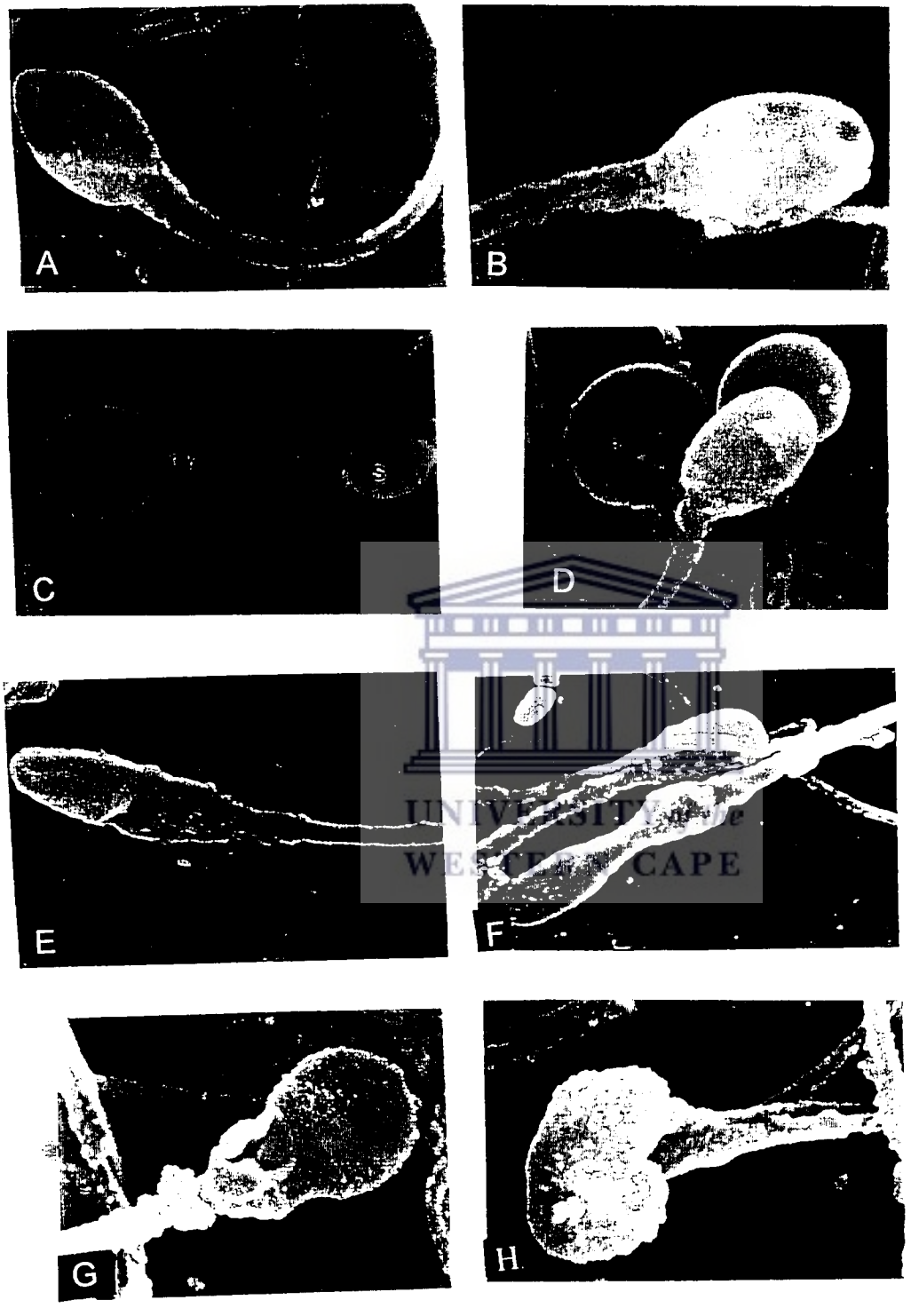
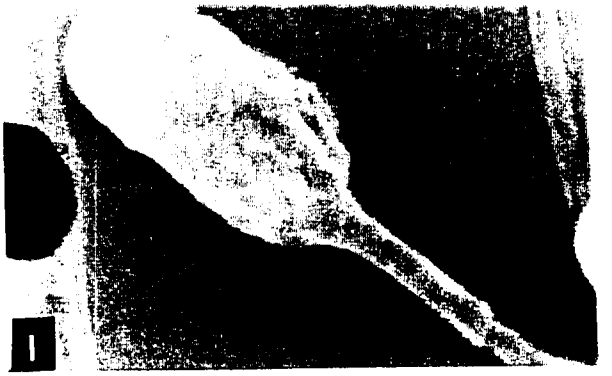
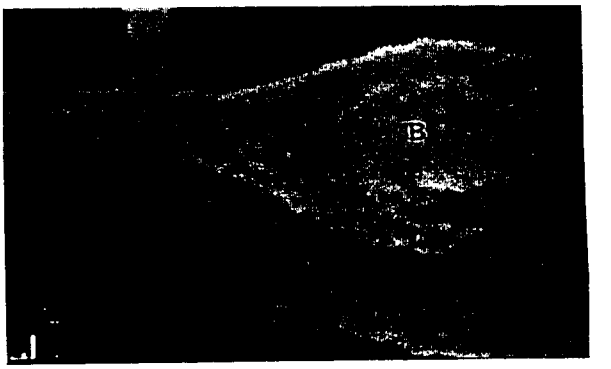


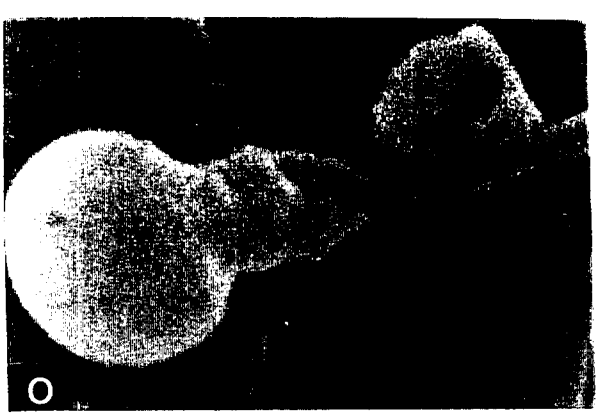
Fig. 3.1.4: A- P: Scanning Electron Microphotographs of human sperm heads

(A) Normal head
(B) Normal head
(C) Normal head with cytoplasmic droplets (CY) / Small (S) head
(D) Round head
(E) Pin head
(F) Tear drop- shaped head
(G) Pear- shaped head
(H) Double head

3.1.4: Scanning Electron Microscopy







(I) Elongated / Tapered head
(J) Big head (B)
(K) Round head
(L) Amorphous (A) head / Tapered (T) head
(M) Normal head with dag - defect
(N) Flame head
(0) Round head
(P) Round head

3.1.5 Confocal Laser Scanning Microscopy

Figs. 3.1.4 (A-J) are photomicrographs representative of laser scanning confocal microscopy of flourochrome treated sperm. These figures show normal and most types of head abnormalities. Each one of these photomicrographs show 3D reconstructed images on the basis of, optical sections and series of rotated images through 360 °C (see plates A) 1. Represents the optical sections, 2. Represents the rotated images though 360°C and 3. Represents the 3D sperm morphology reconstruction). For morphological studies 3D sperm morphology was reconstructed from a series 20-30 laser optical sections, and in order to examine the 3D morphology a series of rotated images were reconstructed (see material and methods). Figs. 3.1.4 (A) shows the normal headed spermatozoa. The series are rotated to 360°C, but looking at the sperm at 30°C to 40°C angles respectively, the sperm can be scored as abnormal. Rotation of the spermatozoon demonstrates that interpretation of morphology of the human sperm is more complex than previously thought. The sperm may adhere to the surface of a glass-slide in various positions. Therefore, it is not surprising to find different results on the morphology of the human sperm in the light of the above discovery. Figs. 3.1.4 (B-J) show the examples of many abnormal forms of spermatozoa. Only the amorphous in Fig. 3.1.4 (B) resembles the normal spermatozoa but all others are different. Fig. 3.4 (I) shows a dag-defect abnormality, while Fig. 3.1.4 (C) shows the double headed spermatozoa and clearly two cells can be observed that are joined and this shows that they did not separate during the meiosis process. Fig. 3.1.4 (D) shows the examples of a big-headed spermatozoa which demonstrate the distinct increase in length and width of the spermatozoa. Fig. 3.1.4 (E) indicates examples of small headed spermatozoa with a reduction in the length and width of the spermatozoa. Fig. 3.1.4 (F)

are examples of tapered / elongated spermatozoa or long-headed spermatozoa. They are characterized by a clear increase in length of the sperm head and obvious thinning. Fig. 3.1.4 (G) shows an example of a flame headed spermatozoon. Fig. 3.1.4 (H) shows normal spermatozoa with a cytoplasmic-droplet. Fig. 3.1.4 (J) shows pin headed spermatozoa where the acrosome and post-acrosome cannot be differentiated.

Laser scanning confocal microscopy is a powerful tool, insofar that it enables the researcher to view successfully deeper layers in a specimen with great clarity, without first having to undertake the laborious task of cutting the specimen into thin sections. It also allows reconstructed images to be rotated around an axis and viewed from different angles. Despite the fact that the useful to technique is useful to researchers for solving problems of normality, it is extremely costly and time consuming and hence is not suitable for routine semen analysis investigations.

3.1.5 Laser Scanning Confocal Microscopy

Figs. 3.1.5: A - J show the main types of normal and abnormal sperm as encountered in routine semen analysis.

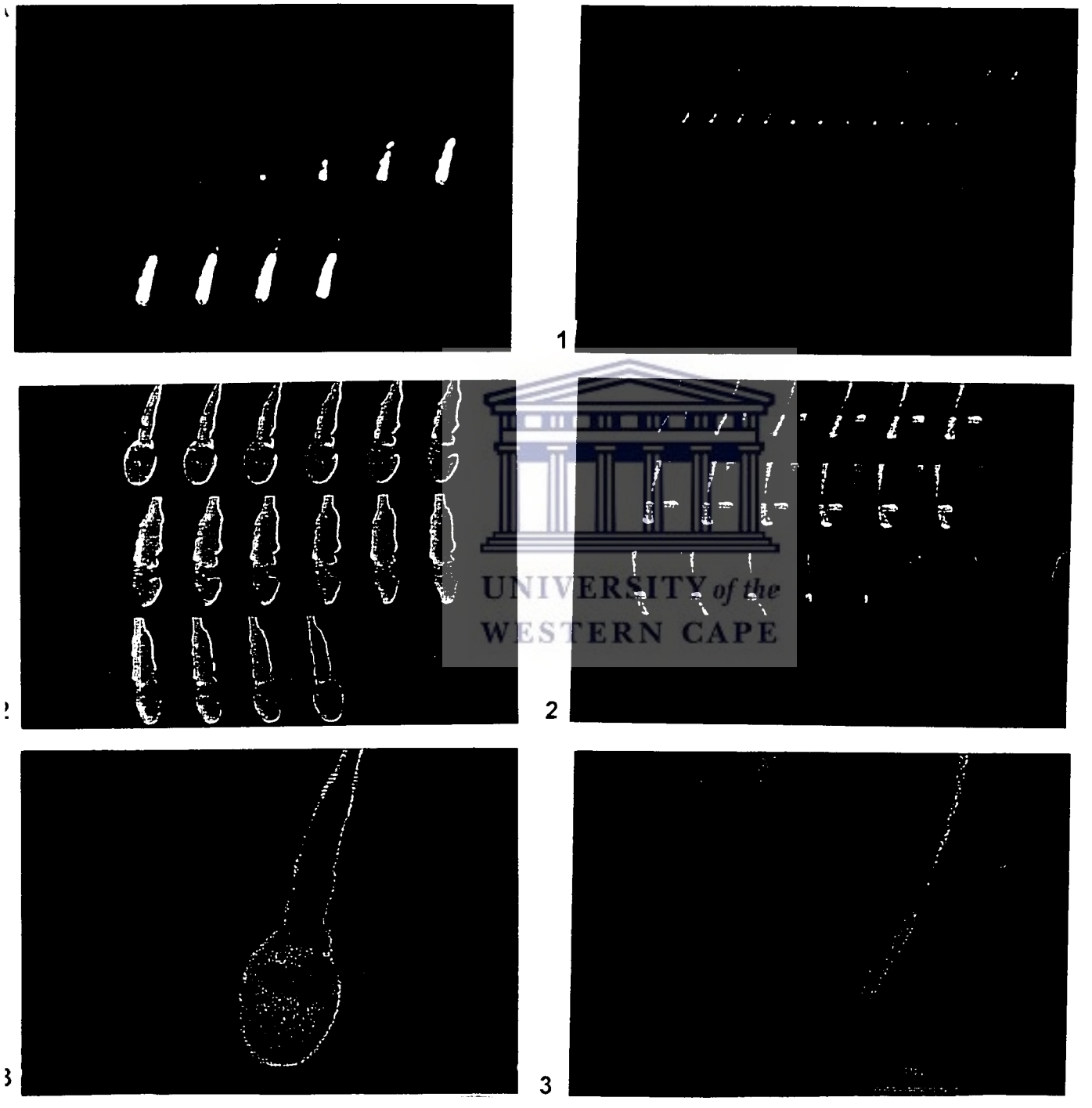


Fig. 3.1.5 A: Laser Scanning Confocal Microphotographs of morphologically normal spermatozoa.

(1). Optical sectioning
(2). A series of 3D rotated sperm images
(3). 3D morphology

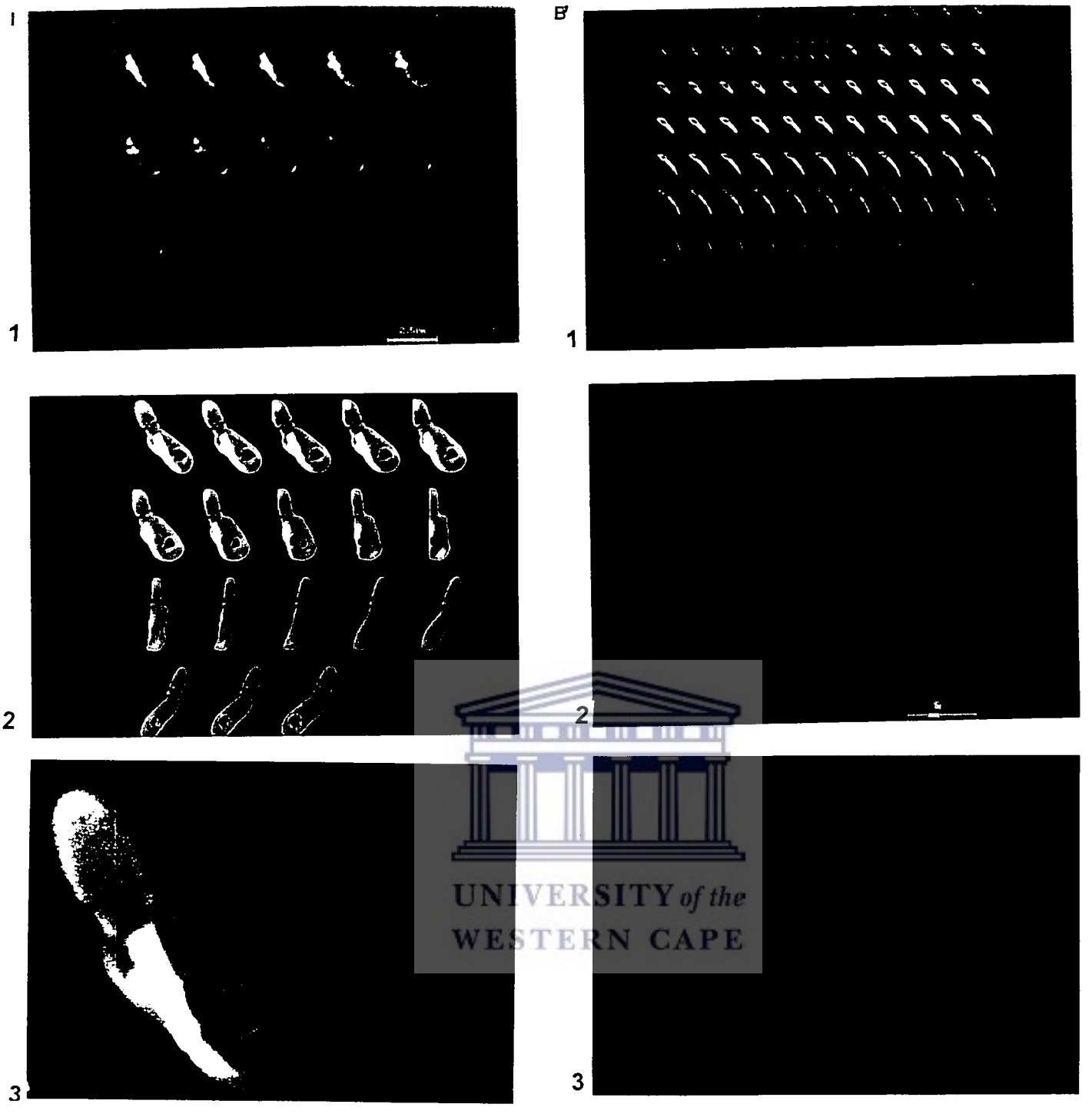


Fig. 3.1.5 I: Laser Scanning Confocal Microphotographs of morphologically abnormal spermatozoa with dag defect

(1). Optical sectioning
(2). A series of 3D rotated sperm images
(3). 3D morphology

Fig. 3.1.5 B: Laser Scanning Confocal Microphotographs of morphologically abnormal (amorphous) spermatozoa

(1). Optical sectioning
(2). A series of 3D rotated sperm images
(3). 3D morphology

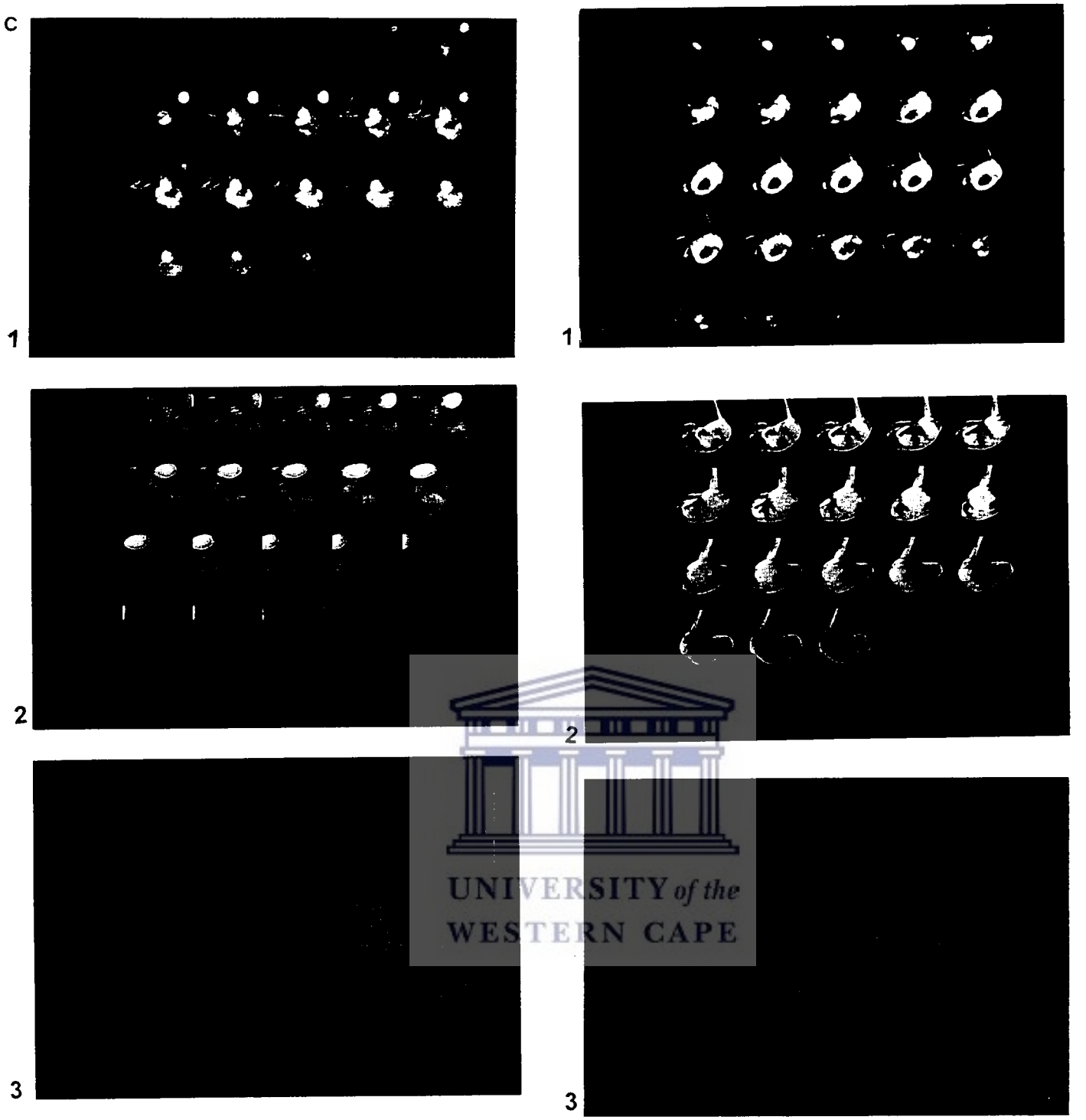


Fig. 3.1.5 C: Laser Scanning Confocal Microphotographs of morphologically abnormal (double head) spermatozoa

(1). Optical sectioning
(2). A series of 3D rotated sperm images
(3). 3D morphology

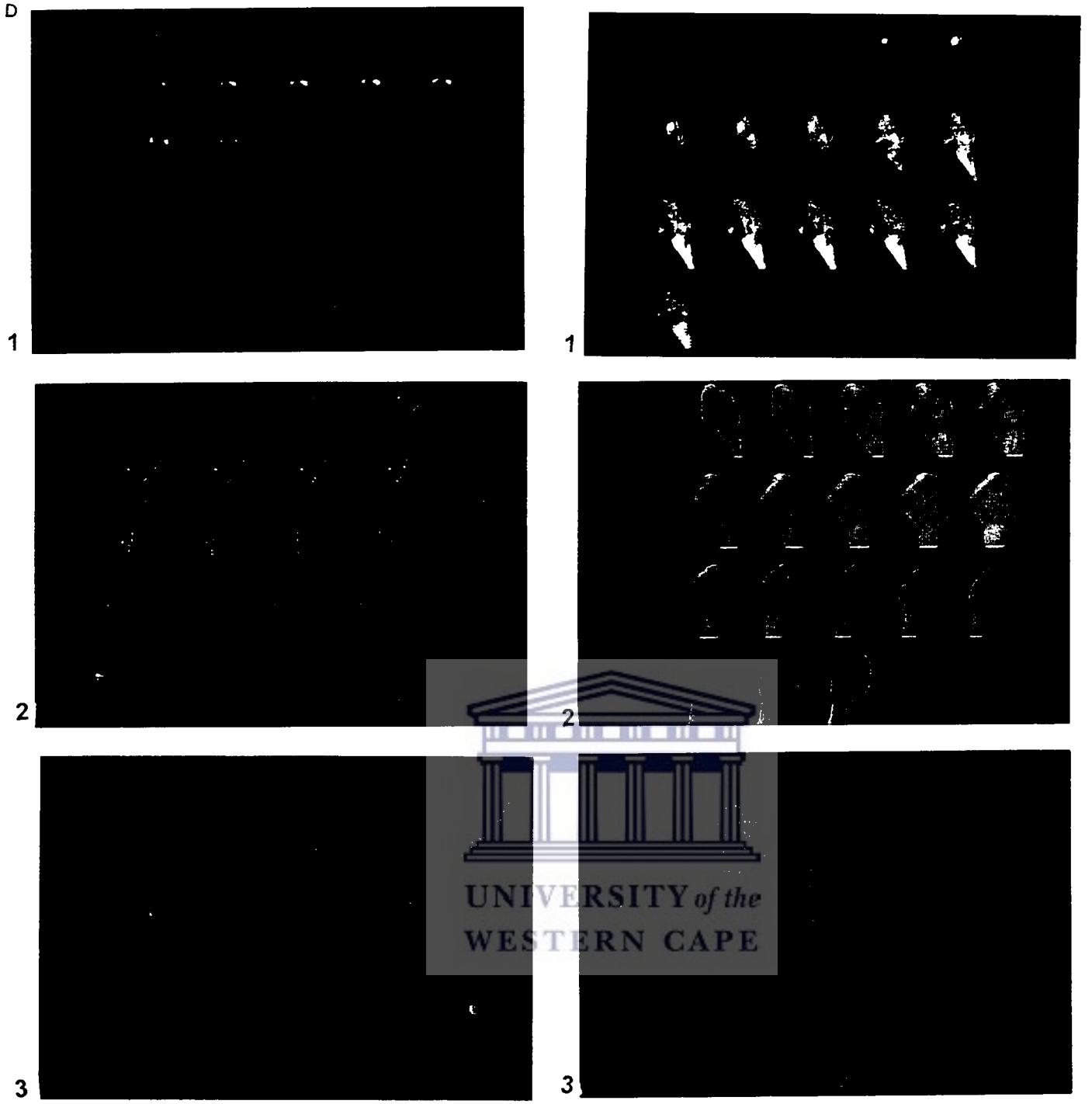


Fig. 3.1.5 D: Laser Scanning Confocal Microphotographs of morphologically abnormal (big head) spermatozoa

(1). Optical sectioning
(2). A series of 3D rotated sperm images
(3). 3D morphology

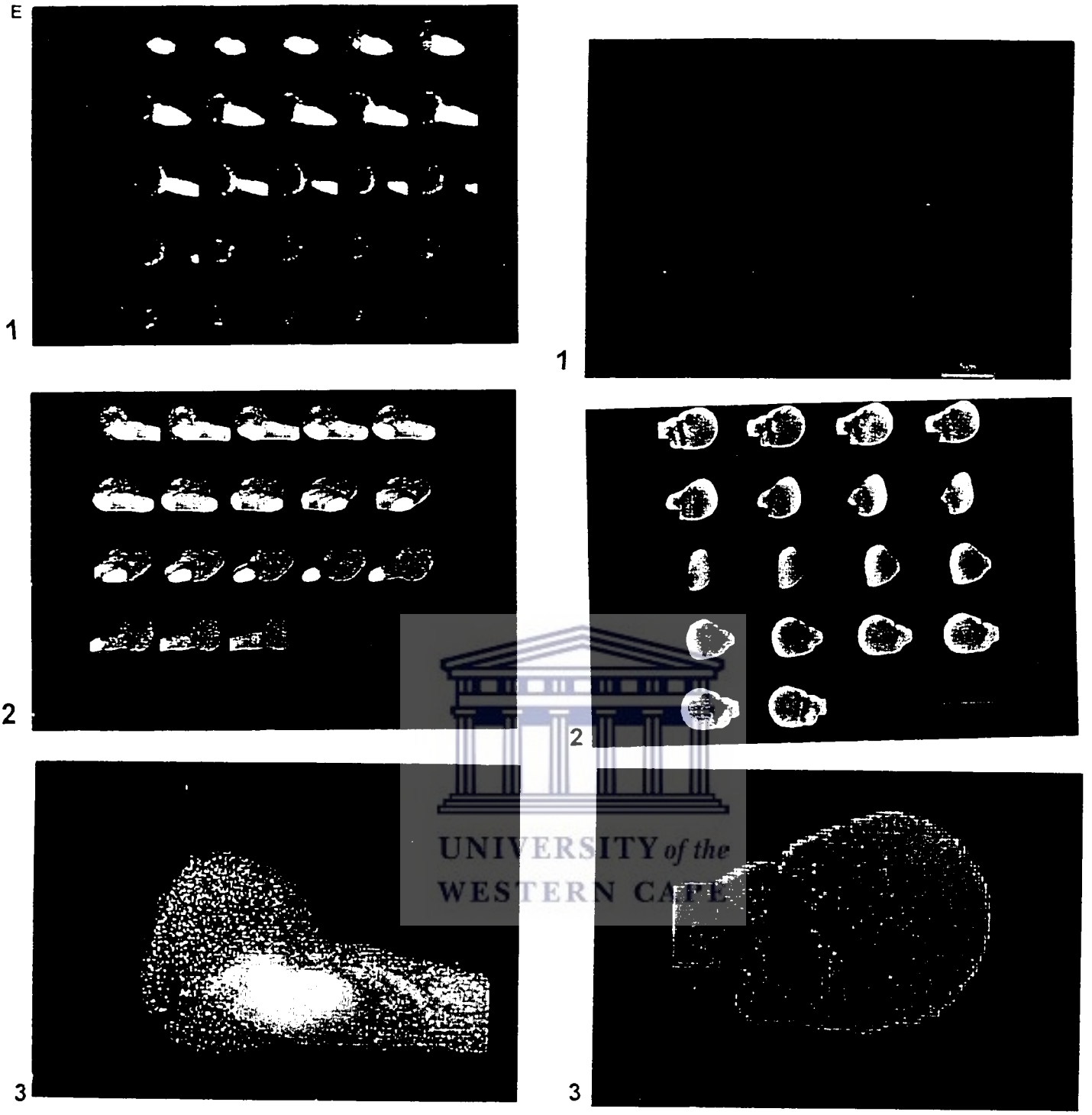


Fig. 3.1.5 E: Laser Scanning Confocal Microphotographs of morphologically abnormal (small head) spermatozoa

(1). Optical sectioning
(2). A series of 3D rotated sperm images
(3). 3D morphology

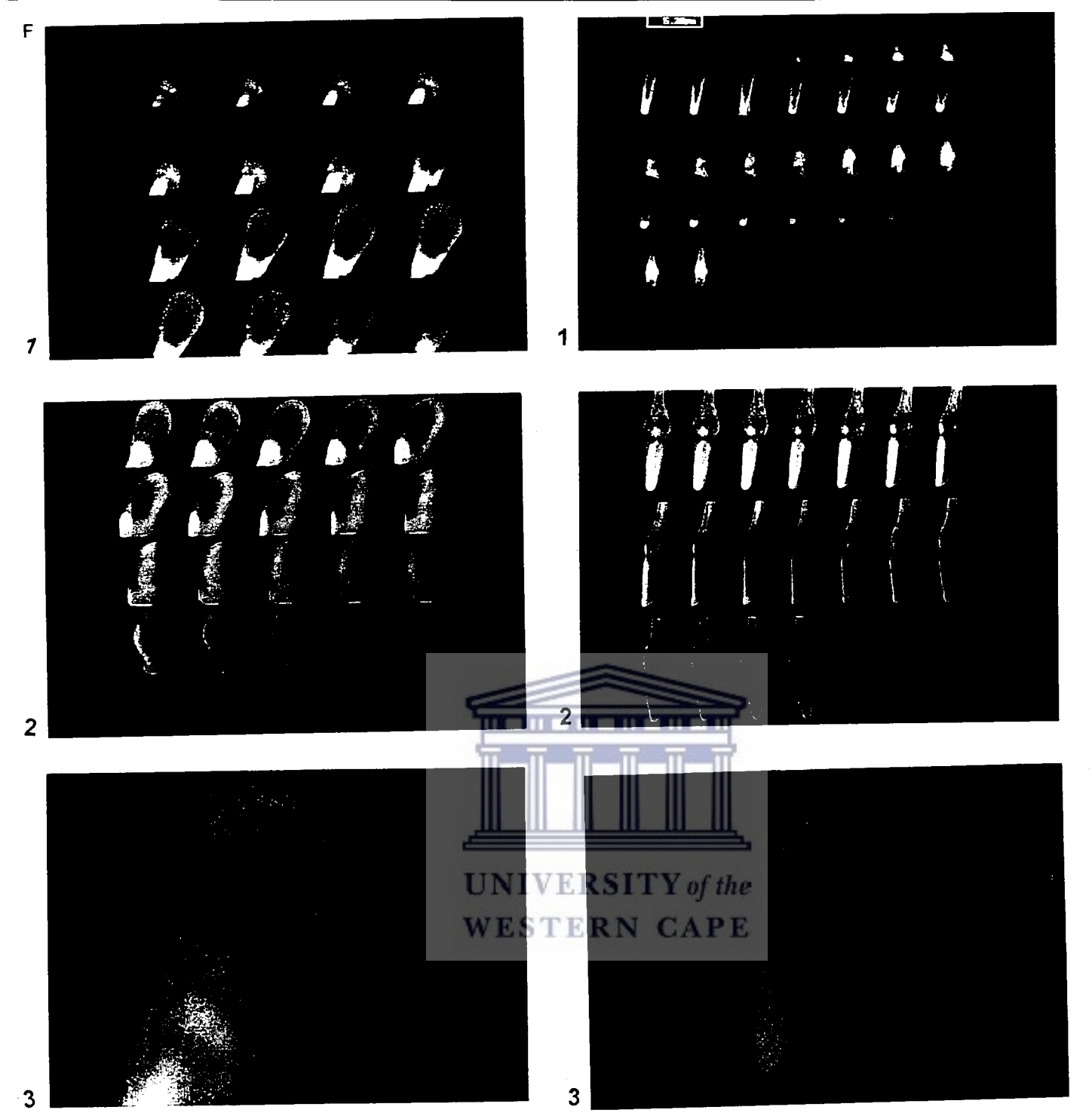


Fig. 3.1.5 F: Laser Scanning Confocal Microphotographs of morphologically abnormal (tapered head) spermatozoa

(1). Optical sectioning
(2). A series of 3D rotated sperm images
(3). 3D morphology

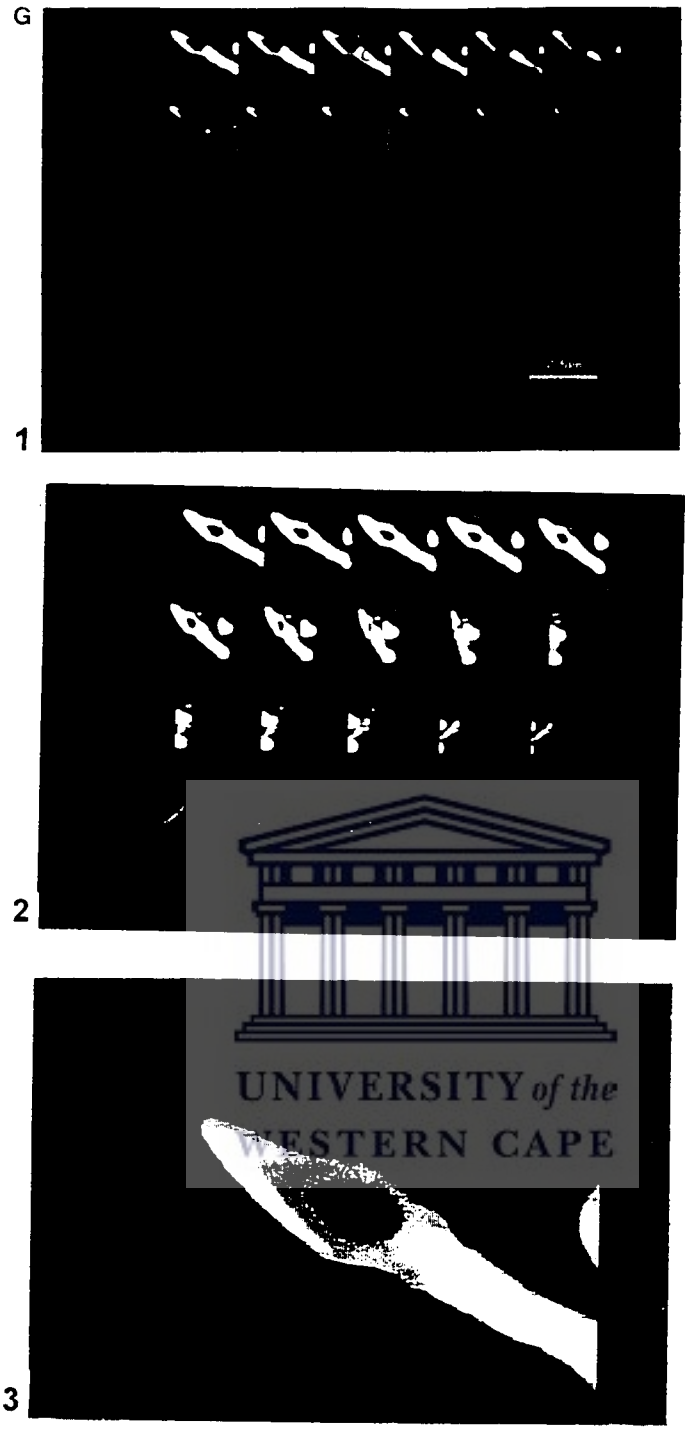


Fig. 3.1.5 G: Laser Scanning Confocal Microphotographs of morphologically abnormal (flame head) spermatozoa

(1). Optical sectioning
(2). A series of 3D rotated sperm images
(3). 3D morphology

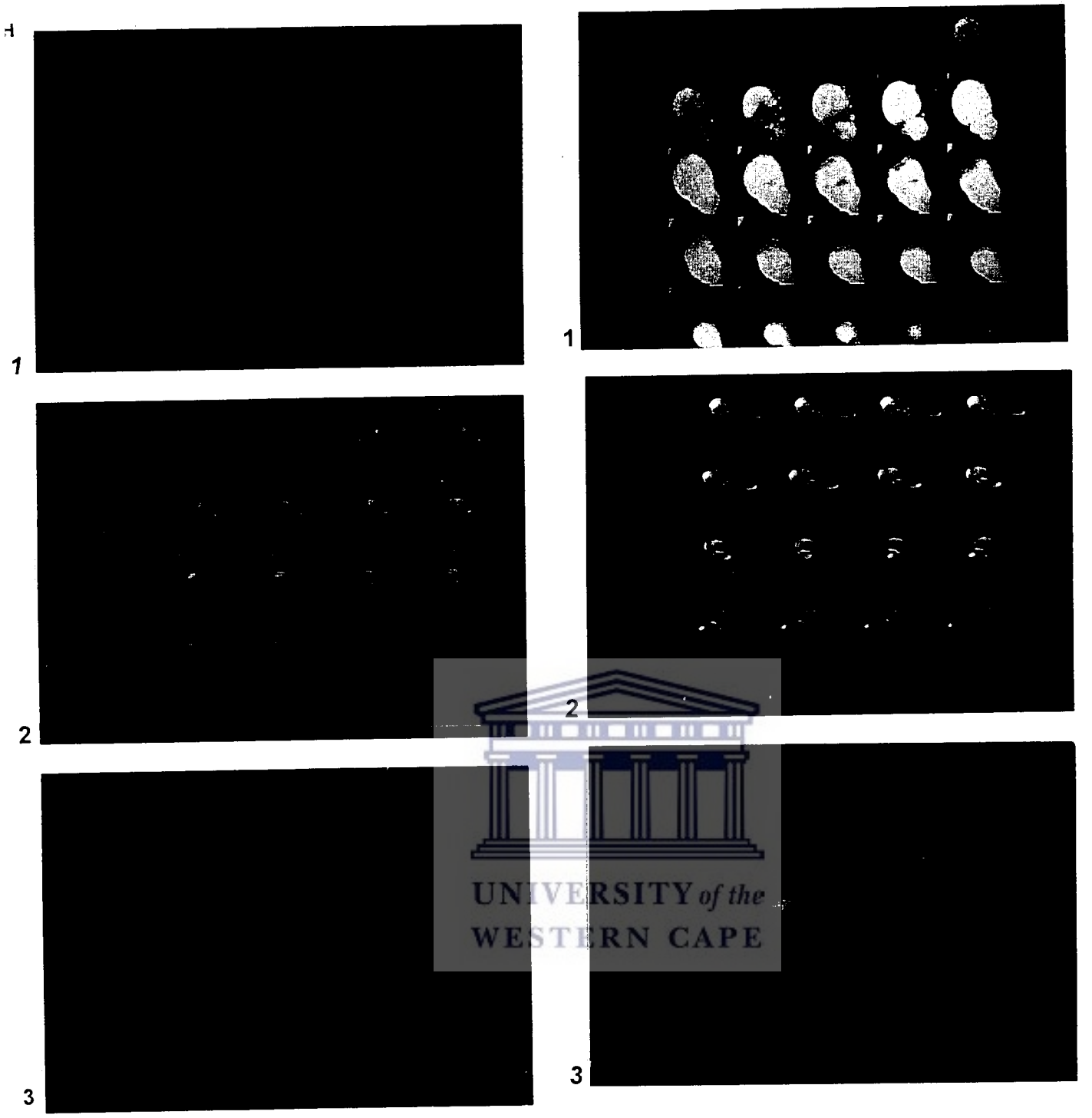


Fig. 3.1.5 H: Laser Scanning Confocal Microphotographs
of morphologically abnormal spermatozoa with cytoplasmic droplets

(1). Optical sectioning
(2). A series of 3D rotated sperm images
(3). 3D morphology

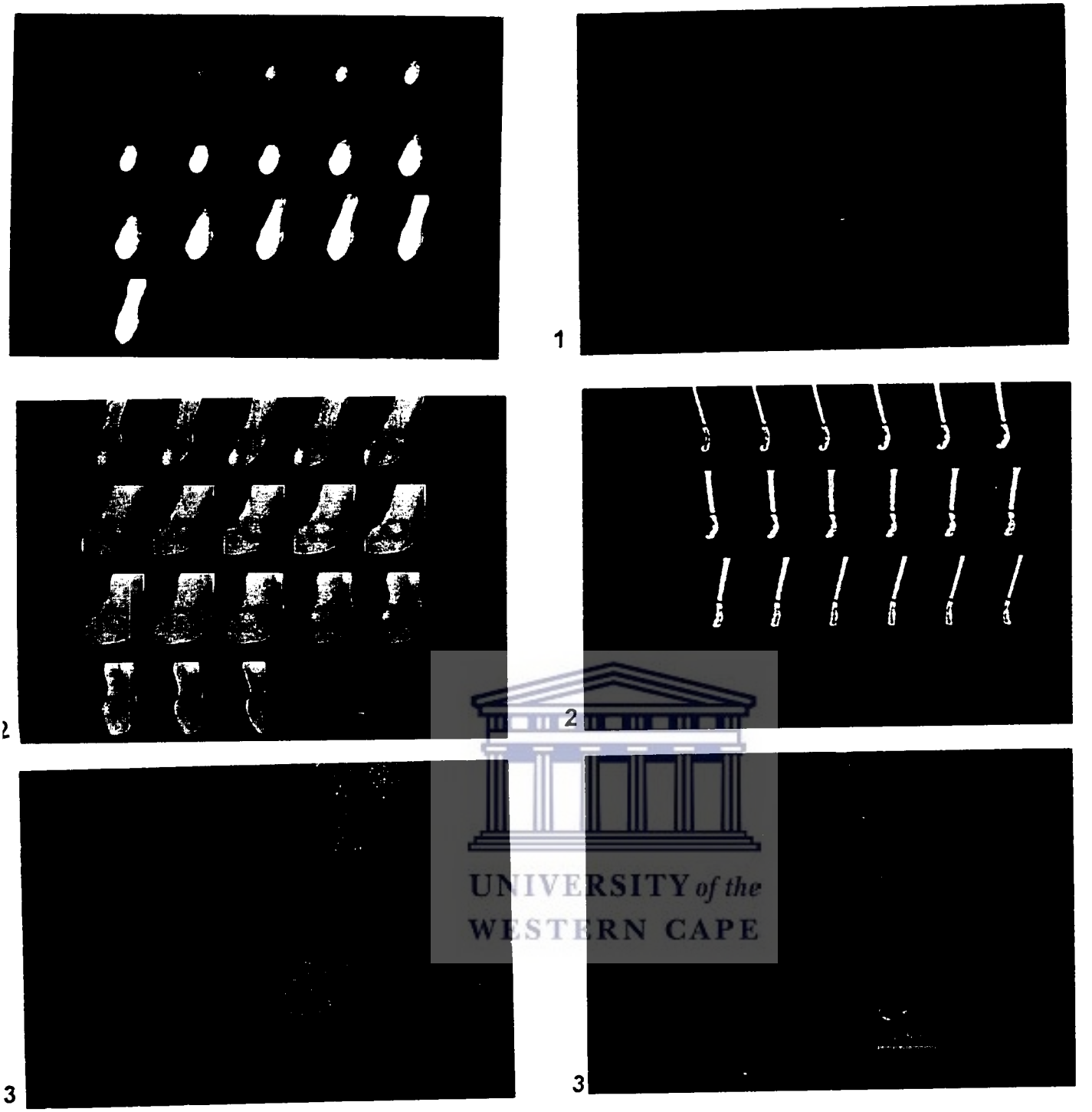


Fig. 3.1.5 J: Laser Scanning Confocal Microphotographs of morphologically abnormal (pin head) spermatozoa

(1). Optical sectioning
(2). A series of 3D rotated sperm images
(3). 3D morphology

3.1.6 Conclusion

All types of sperm can be clearly discerned with the four techniques. A comparative study of Confocal and bright field is particularly important in relation to amorphous versus normal sperm. The present investigation has clearly shown that these two types of sperm may in many cases be defined purely on the basis of the position of the sperm in relation to the sperm surface attachment. However, the complementary techniques above show that Papanicolaou / Bright field microscopy remains a good tool to identify the various human sperm forms. The Normarski Differential interference technique is within the realm of the routine andrology laboratory (e.g. it is used in most *IVF* laboratories to do *ICSI*) and can be also tested against bright field in the future.

Quantitatively it appears that some techniques show certain features better, but laser scanning confocal microscopy may be useful to distinguish borderline amorphous. The complementary techniques by and large show that all the major abnormal sperm head forms can be identified. Bright field microscopy of Papanicolaou stained smears, however appear to be acceptable in identifying normal / abnormal sperm when compared to the other three techniques in routine semen analysis.

In the next section (3.2) the Papanicolaou (1947) was accordingly applied to compare in a blinded trial how three technologists score the same samples for normality.

3.2 Comparison of technicians assessment of sperm morphology on the basis of Tygerberg Strict Criteria and WHO criteria

3.2.1 General aspects

While it is relatively simple for most trained staff in the routine andrology laboratory to distinguish among the major abnormal sperm forms (see section 3.1), the assessment of what technologists regard as normal in terms of subtle differences (e.g. amorphous versus normal sperm as pointed out under 3.1.4 - confocal microscopy) is open to criticism, because of the subjective nature. It was furthermore emphasized in the introduction that sperm morphology assessment is subjective and open to different interpretations and can not globally be accepted as reliable (Davis & Gravance,1993). The main emphasis in this section was to establish how technologists in three different andrology laboratories assess sperm morphology of 77 patients using Tygerberg Strict Criteria (TSC). The 77 sperm smears were all from proven fathers (see Materials and Methods). In addition one technician with no prior knowledge of TSC, assessed the same sperm smears using WHO criteria (1992).

Four different approaches were used to analyze the results. Firstly, results were compared among the three laboratories using Bland and Altman Plots. This is one of the recognized methods to test repeatability among technologists for a given technique or for testing the reliability of different methodologies (Schoonjans, 1996). Secondly, Dot and line diagrams assisted to show the spread of coinciding and non coinciding points among three laboratories and possible relationships between the TSC and WHO criteria. Thirdly, regression analysis was performed as well as correlation coefficients determined among

the different laboratories to test the goodness of fit / predictability and to indicate possible relationships between TSC and WHO sperm morphology criteria.

Notched Box and Whisker plots were constructed to establish basic statistical summary data for proven fathers.

3.2.2 Bland and Altman Plots, Dot and Line Diagrams

Figs 3.2.1 to 3.2.3 show Bland and Altman Plots when each laboratory's results were compared to the others. These results indicate that despite the fact that both UWC and GSH differed more from PCH than UWC from GSH, most data points were evenly distributed and were within the 1.96 S.D. interval. This suggests a high level of concordance among the three laboratories. However, when Bland and Altman Plots were constructed on the basis of percentage differences, large discrepancies were noted despite the fact that most data points were within the 1.96 S.D. In this context it can be seen that many points in all comparisons represent a deviation from the average from 0 to as much as 145%.

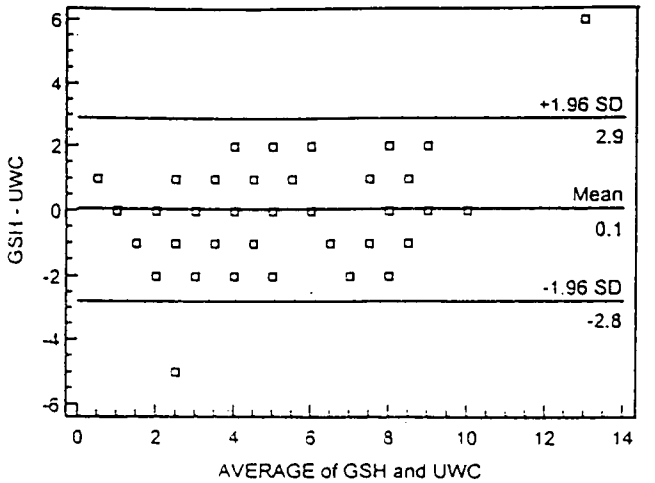
Table 3.2.1 shows the distribution/frequency of similarities and differences in readings among the different laboratories. It can be seen that in only 5% of proven fathers all three technicians scored exactly the same; in 38% of proven fathers two technicians scored exactly the same, and in 57% of cases all three technicians scored differently. It was surprising that among the three technicians range of scoring for 43% samples coincided with different TSC classes (e.g. >5% and < 5%).

In order to further demonstrate differences and similarities among technicians in scoring the same patients, Dot and Line diagrams were constructed (Figs. 3.2.7 - 3.2.9). The Dot

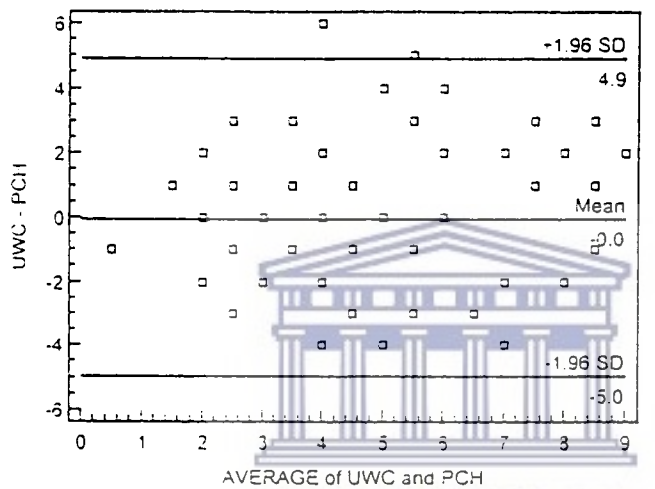
and Lines were first constructed for joining lines for the same samples as evaluated by two different technicians (laboratories). It is evident from these diagrams that one technician may score a smear representing 1% normality, a second technician may score it 7% and a third technician may score it 9%. The Dot and Line diagram in Fig. 3.2.10 connects the scores for the the same samples as evaluated by three technicians in different laboratories by lines and indicate the large variations that exist among different laboratories.

The results above (Table 3.2.1 and Figs 3.2.7-3.2.10) further support the information on the Bland and Altman Plots as it relates to percentage deviation from the average. This approach suggests that the concordance as viewed from the frequency of obtaining a similar or a close reading or by connecting points (Dot and Line Diagrams) or expressed as a percentage deviation from the average (Bland Altman Plots, Figs. 3.2.4- 3.2.6) show poor concordance among laboratories.

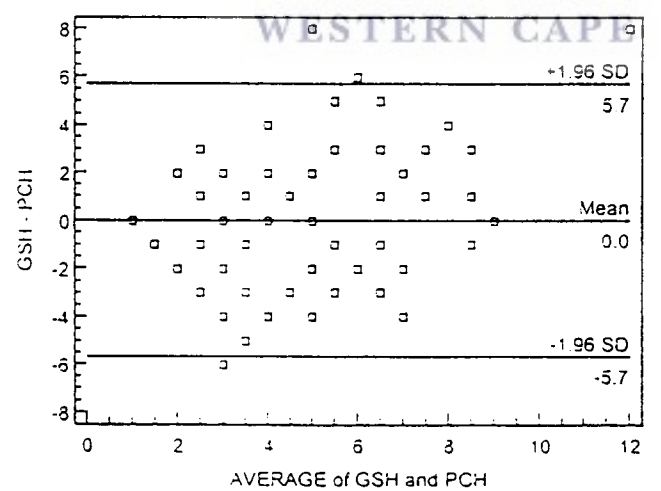
3.2.1



3.2.2

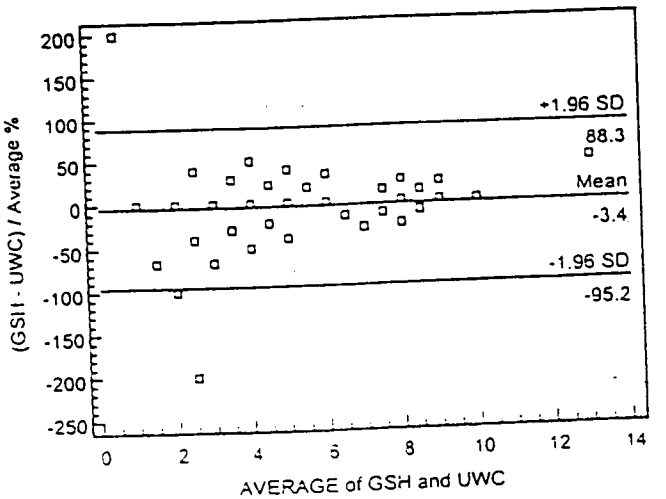


3.2.3

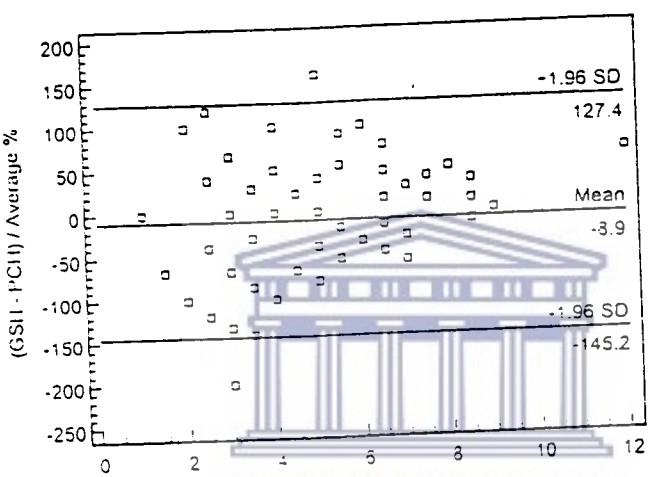


Figs 3.2.1 - 3.2.3: Bland and Altman Plots to assess in each case the repeatability between two technicians scoring of the same smear. In these figures the difference between the two results is plotted against the average of the two results.

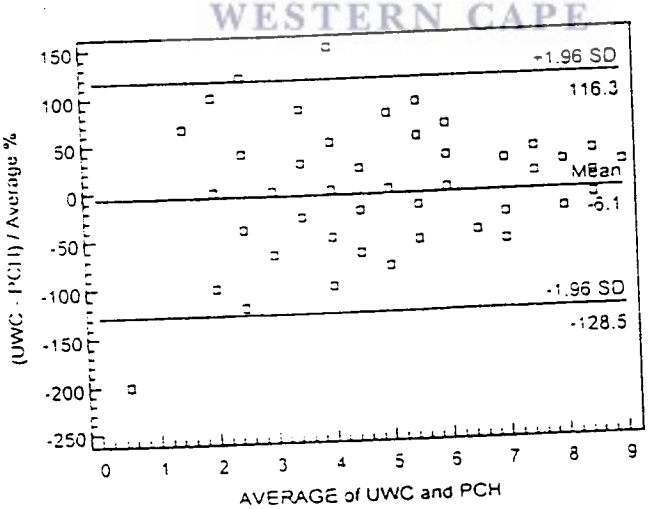
3.2.4



3.2.5



3.2.6



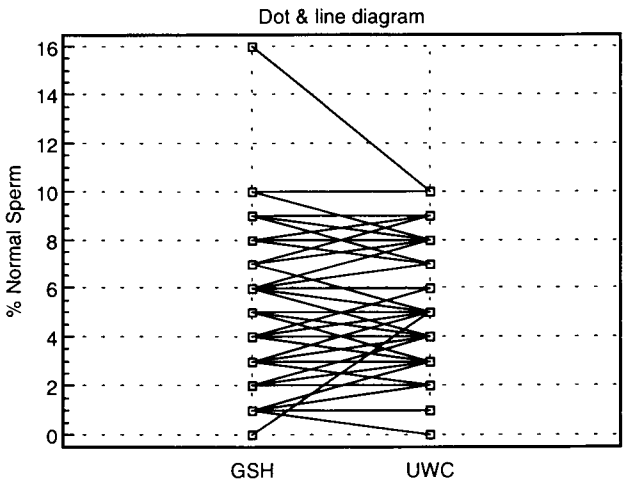
Figs 3.2.4 - 3.2.6: Bland and Altman Plots to assess in each case the repeatability between two technicians scoring of the same smear. In these figures the difference between the two results is plotted as a percentage of the average.

Table 3.2.1: The level of agreement when three technicians from three different laboratories score the percentage normal sperm on the basis of Tygerberg Strict Criteria (TSC) of sperm smears from 77 proven fathers.

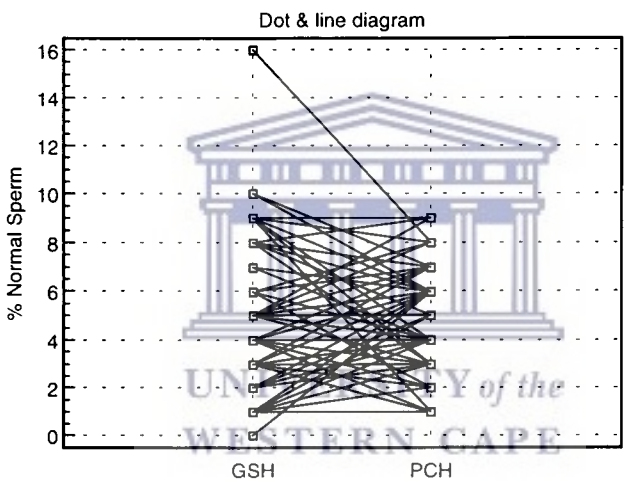
LEVEL OF AGREEMENT BETWEEN TECHNICIANS		
100% Agreement among three technicians	100% Agreement among two technicians	No agreement among three technicians
5%	**38%***	**57%***

***In both these groups, at least one technician classified 43% of samples in a different TSC category.**

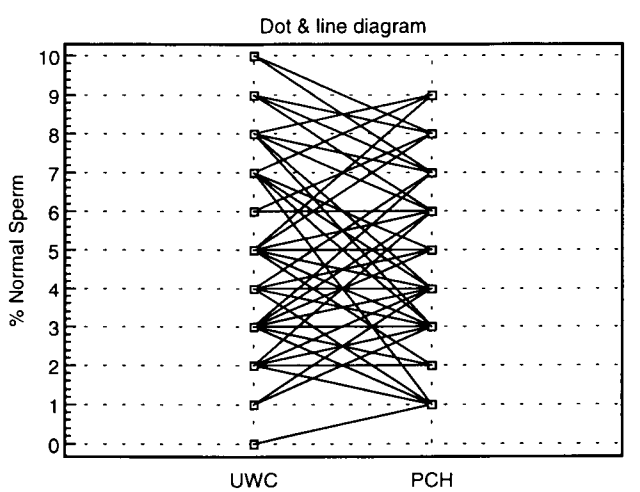
3.2.7



3.2.8



3.2.9



Figs. 3.2.7 - 3.2.9: Dot and line diagrams connecting readings for the same sample by lines (UWC = Technician from UWC; PCH = Technician from Pretoria Central Hospital, GSH = Technician from Groote Schuur Hospital).

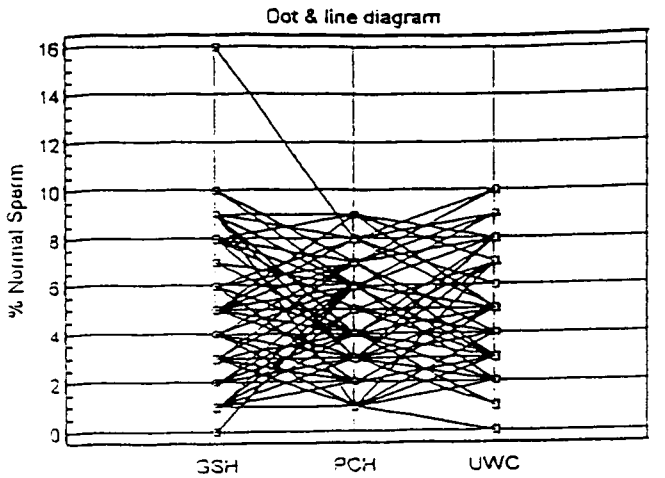


Fig. 3.2.10: Dot and line diagram connecting technician scoring of the same samples among three technicians from three laboratories.

3.2.3 Correlations and Regression analysis for Strict Criteria

Table 3.2.2 shows the correlation coefficients for the TSC- scoring of the three technicians for the same sperm smears of 77 proven fathers. A high correlation was found between GSH and UWC which was highly significant but the correlations for GSH/PCH and UWC/PCH was lower than for UWC/GSH. In the latter case the differences between the correlation coefficients of GSH/UWC and GSH/PCH ($P<0.001$) as well as between GSH/UWC and UWC/PCH ($P<0.001$) were significantly different. However, the correlations for both GSH/PCH and UWC/PCH were highly significant (Table 3.2.2).

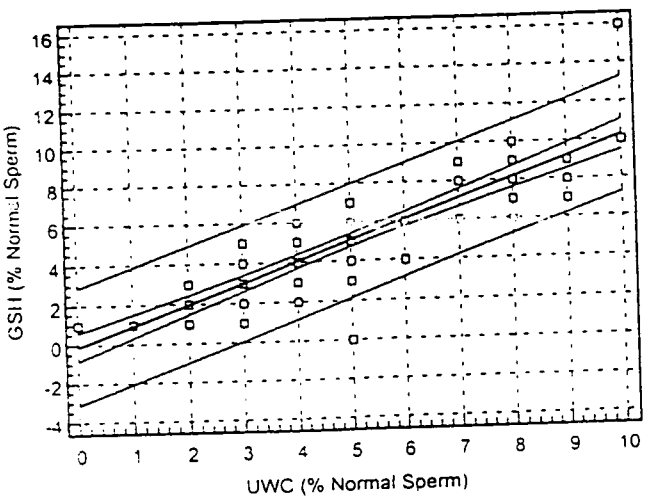
Table 3.2.2: Correlation coefficients among three technicians from different laboratories for scoring the same 77 sperm smears for proven fathers using the TSC.

Technician comparisons	**Correlation coefficient**	**Significance**
GSH/UWC	0.87	*P*<0.001
GSH/PCH	0.385	*P*<0.0005
PCH/UWC	0.41	*P*<0.0004

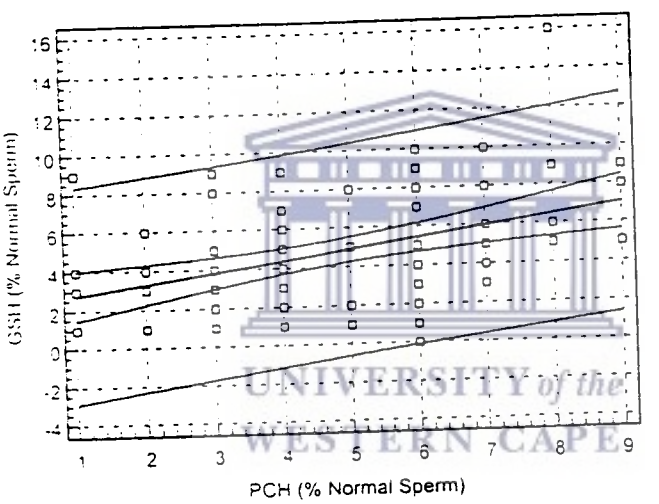
The correlation analysis above showed that the three technicians scores for the percentage normal sperm are associated but that the association of two technicians from two laboratories were different (*P*<.001) from a third technician on the basis of differences in terms of the correlation coefficient. These results indicate difficulties in accepting the accuracy of scoring sperm smears according to the TSC on an interlaboratory basis or as an international standard for comparing the percentage normal sperm. This is particularly relevant in the TSC system where patients may be placed in different fertility classes.

Regression analysis for technicians scoring 77 sperm smears for normal morphology according to TSC was performed to establish how well results can be predicted among different laboratories. Figures 3.2.11 - 3.2.13 are graphic representations of the regression analysis and in each regression both the 95% confidence interval as well as the 95% prediction are indicated. It is evident that a high level of prediction is evident among GSH and UWC on the basis of the regression line as well as the fact that the intercept is at the zero point. The prediction for UWC and PCH as well as for GSH and PCH is less favourable. On the basis of the regression line a 2% score for PCH coincides with a 3% score for GSH and a 6% PCH score coincides with a 8% score for GSH. Furthermore, the regression line originates at 2 or 3 in these latter two regressions. However, for all regressions most of the points on the scatter diagram fall within the 95% prediction interval and indicate a statistically acceptable level of prediction among the three laboratories.

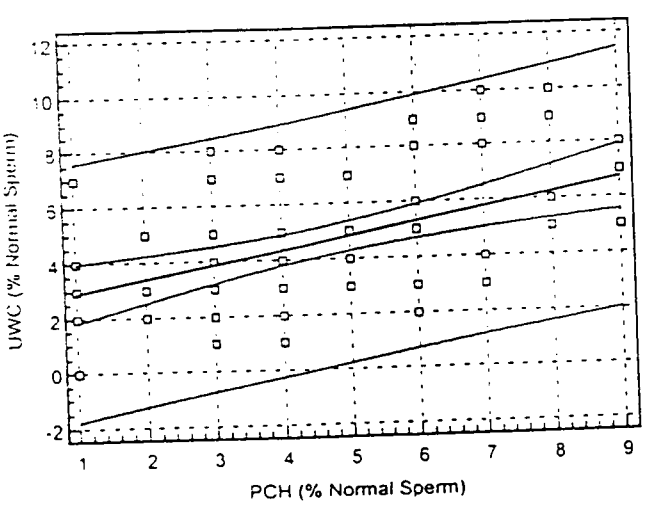
3.2.11



3.2.12



3.2.13



Figs 3.2.11 - 3.2.13: Scatter diagrams with regression lines and 95% confidence intervals, and 95% probability intervals for three technicians in three different laboratories.

3.2.4 Correlations and Regression analysis for Strict Criteria and WHO

Table 3.2.3 shows the correlations between each of the three technicians for TSC criteria and one technician for WHO criteria. The correlations between each technician scoring according to TSC and WHO were better than the correlations between UWC/PCH and GSH/PCH. All correlations for technician scores according to TSC were highly significant when correlated with WHO (Table 3.2.3).

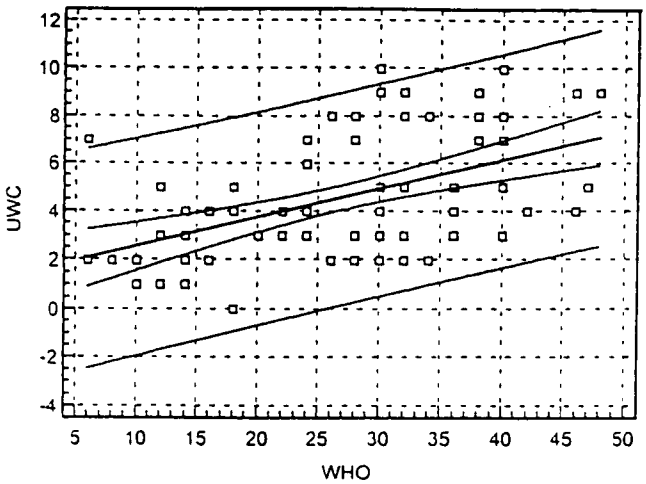
Table 3.2.3: Correlation coefficients for TSC and WHO.

Technician comparisons	**Correlation coefficient**	**Significance**
GSH/WHO	0.4655	*P*<0.0001
UWC/WHO	0.5010	*P*<0.0001
PCH/WHO	0.4505	*P*<0.0001

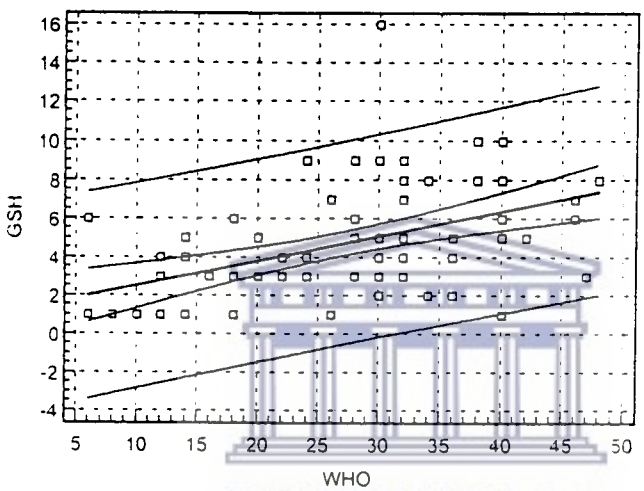
Figs 3.2.14 - 3.2.16 represent scatter diagrams with regression lines and 95% prediction intervals when UWC, GSH, and PCH are compared to sperm morphology scores for WHO. For all regressions the intercepts are not favourable, however, it appears that a good level of prediction is evident. It appears that a 2% TSC score coincides with a 6% WHO score, a 4% TSC coincides with a 22% WHO score and a 5% TSC score coincides with a 30% WHO score.

Fig. 3.2.17 shows Dot and Line when 3% to 5% TSC were connected to the corresponding WHO scores. This supports the data for the regression lines in Figs. 3.2.11 - 3.2.13. The data furthermore shows that the TSC range of 3 to 5% corresponds to a 6 to 30% range for WHO. It has been shown in earlier analysis that one technician may score a smear as 3%, a second technician may score it as 5% and a third technician may score it as 7%. It therefore appears that the very narrow TSC range may be representative of a wide range of possible differential fertilities on the basis of the WHO scale.

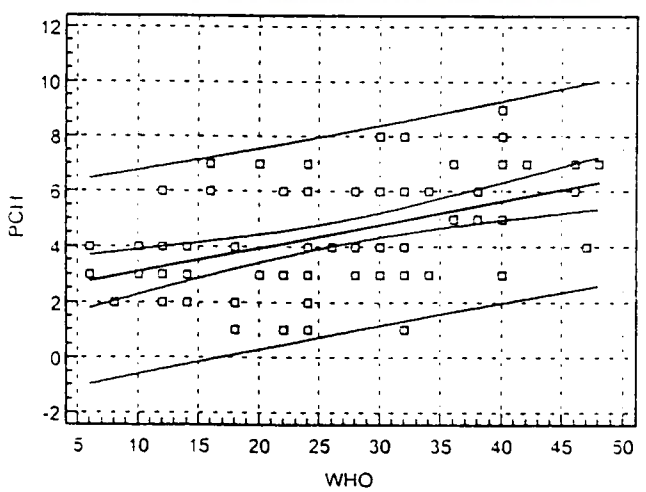
3.2.14



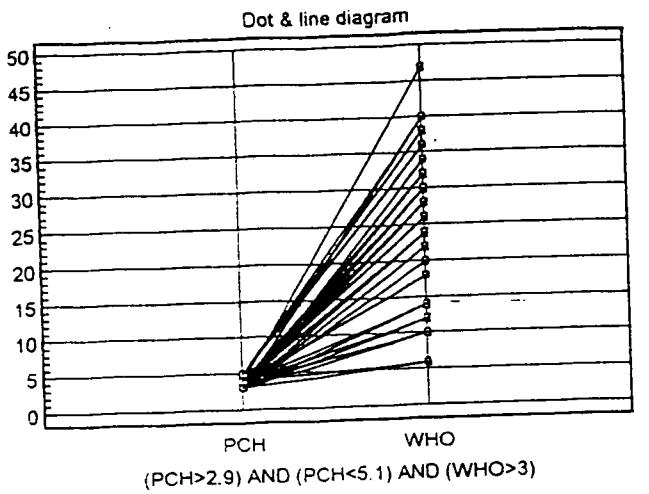
3.2.15



3.2.16



Figs 3.2.14 - 3.2.16: Represent scatter diagrams with regression lines and 95% prediction intervals when UWC, GSH and PCH are compared to sperm morphology scores for WHO.



Figs 3.2.17: Shows Dot and line when 3% to 5% TSC were connected to the corresponding WHO scores.

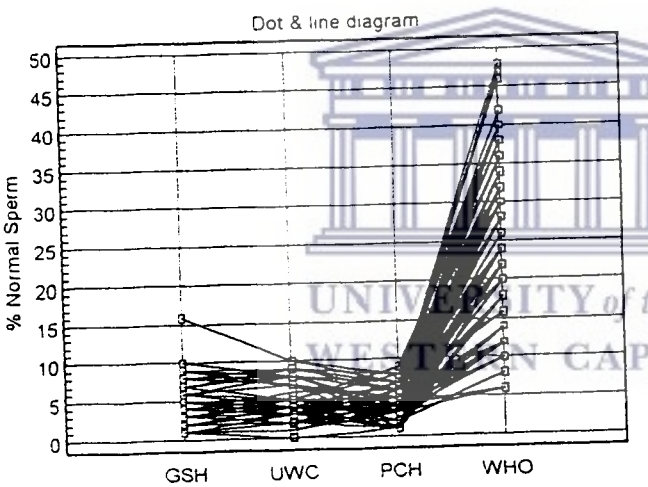
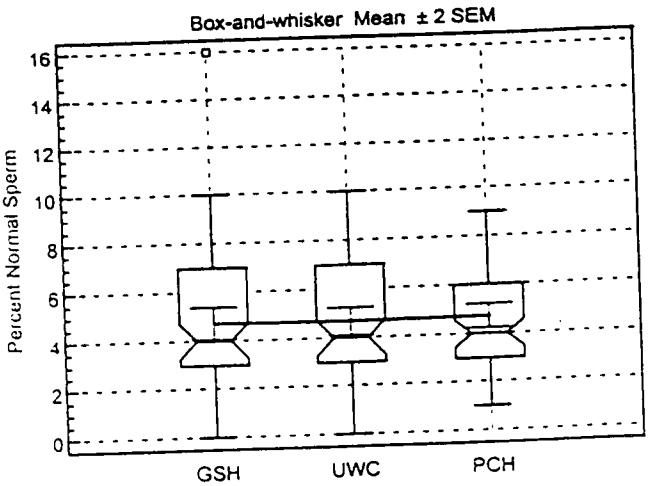
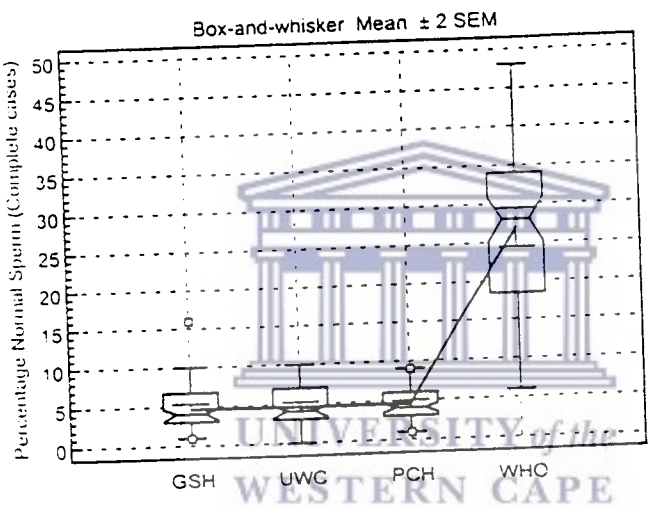


Fig. 3.2.18: Show the correlation between each of the three technicians for TSC criteria and one technician for WHO criteria

3.2.19



3.2.20



Figs: 3.2.19 - 3.2.20 are notched box and whisker plots summaring the data of 77 proven fathers sperm head

3.2.5 Summary statistics for sperm morphology of 77 proven fathers

Figs 3.2.19 - 3.2.20 are notched box and whisker plots summarizing the data of 77 proven fathers. Table 3. 2. 1 Shows summary data for three technicians scoring sperm smears according to TSC. It can be seen that the averages for the proven fathers are almost identical for the three different laboratories and the 95% confidence intervals overlap. From this data it appears that the average normal sperm morphology for these proven fathers are 4.9%. This is surprising in view of the fact that this value represents the lowest infertility class according to TSC.

Fig. 3. 2. 17 combines the data from Fig. 3. 2. 18 with the average data of WHO. On the basis of this comparison it appears that an average TSC score of 5% corresponds to a WHO score of 28% for WHO. The cut off value for subfertile according to WHO is 30% while it is <14% for TSC. It therefore appears that the sperm morphology data for proven fathers according to WHO make more biological sense than the TSC data.

QUANTITATIVE IMAGE ANALYSIS OF THE HUMAN SPERM

3.3 Quantitative Image Analysis of the human sperm head

3.3.1 General aspects

Sections 3.1 and 3.2 showed that despite the problems indicated, normal/abnormal sperm can be described and selected with relative ease by the trained technician. It was, however, indicated under 3.2 that scoring of the same human sperm smears by different technicians may show considerable variation. One of the reasons for this is the subjective nature of the scoring methods. The main aim of this part of the research was to analyze normal and abnormal sperm from patients visiting the routine andrology laboratory at Groote Schuur Hospital by means of quantitative image analysis. The advantage of this approach is that the data produced in this way may be used for quantitative and automated computer sperm morphology analysis. Normal sperm of thirty donors from Groote Schuur hospital as well as normal sperm of 77 proven fathers have also been used for comparisons in this study.

3.3.2 Key to presentation of results

Table 3.3.1 summarizes the various sperm forms analyzed as well as the donor/patient types. Table 3.3.2 summarizes the morphometric parameters analyzed for each sperm type listed in Table 3.3.1. The abbreviations for the above parameters are furthermore indicated in Table 3.3.3 to serve as legend for future tables and figures. Each of these parameters were determined by means of the FIPS image analysis system as described under Materials and Methods 2.3.5. Table 3.3.4 summarizes the best cut-off points among the nine different sperm head types, of human semen for seven morphometric parameters.

Statistical analysis has been performed as indicated under Materials and Methods. For the sake of simplicity the presentation of these results are presented as tables, Box- and-Whisker Plots (See Fig 3.3.2 which illustrates the various features of the Box- and-Whisker Plot used throughout this study), and Receiver Operating Characteristic curves (ROC-curves) which each time illustrates the cut-off point between normal and each abnormal sperm type for each of four morphometric parameters as indicated in Table 3.3.5 a - 3.3.5 d. Table 3.3.6 summarizes ROC -curve analysis showing cut-off values of normal sperm heads with a Sensitivity and Specificity level of at least 70%.

Table 3.3.7 (a, b) shows sperm which clearly fall outside and inside the range for each parameter with a Sensitivity **or** Specificity / Sensitivity **and** Specificity of at least 70% but usually higher.

3.3.3 Quantitative morphometric analysis of the human sperm head

Summary statistics of seven morphometric parameters of normal and abnormal sperm are indicated in Table 3.3.2. These results are furthermore graphically presented as Box-and-Whisker plots (Figs.3.3.1).

From Table 3.3.1 and Fig.3.3.1, it can be seen that there are statistically significant difference between normal and the various abnormal sperm types for most sperm head parameters. Normal sperm heads of patients differed significantly ($P< 0.0001$ - ANOVA and $P<0.05$ for pairwise comparisons using the Student-Newman-Keuls test) from all abnormal sperm heads for sperm head length, area and Avfer. Flame and big heads differ significantly from normal heads (patients) for all seven parameters. Small heads differ significantly from normal heads for all parameters except for width (Minfer). The parameter which showed the least significant differences among different sperm forms was sperm head shape. Here, normal sperm heads of patients differed significantly from AH, BH and TH ($P<0.0001$-ANOVA and $P<0.05$ for pairwise comparisons - Student-Newman-Keuls test).

The most surprising finding of this investigation is the statistical significant differences that exist among normal sperm (NH) from patients attending a fertility clinic and normal sperm from donors (DO) and proven fathers (PF). Table 3.3.2 and Table 3.3.4 show that the NH group differs significantly from PF in terms of Perimeter and from DO in terms of Area. On the basis of these results it appears that sperm from both PF and DO are larger than normal sperm from patients visiting the fertility clinic. It should be emphasized that the shape of these three types of normal sperm are similar (See Fig.3.3.1 (i) but they differ in size as can be seen in Tables 3.3.1 and 3.3.5; Figs 3.3.1(b) (e.g. Area). It is unlikely that differences in size of these three types of "normal sperm" can be readily distinguished by means of subjective microscopic evaluation.

It was indicated that these results may be useful in computer automated image analysis of sperm morphology. ROC-curves represent a useful means of establishing the cut-off points

between normal and abnormal forms. Tables 3.3.5 a -d and Fig 3.3.3 are ROC-curve analyses which show the cut-off points between NH and each abnormal sperm type. In Table 3.3.6 the cut-off value is indicated in bold when the sensitivity and specificity are above 70%. Figs. 3.3.3 represents a visual presentation of the ROC-curves and emphasize the area under the ROC-curve. Tables 3.3.5 a and b represent sperm heads which clearly fall outside and inside the range for each parameter with a sensitivity and specificity of at least 70% but usually higher. A perfect cut-off point is one which shows 100% sensitivity and 100% specificity and the area under the ROC-curve is 1 (the maximum for this feature). The best cut-off points were observed for sperm head length (Maxfer) and Perimeter (Peri).

Tables 3.3.7a and 3.3.7b were constructed on the basis of ROC-curve analysis and present a summary of ideal cut-off points that can be used in automated computer analysis to distinguish between normal and abnormal sperm. In Table 3.3.7(a) the perimeter of 10.71 - 14.30 and area of 8.98 - 9.88 are good discriminants for FH, BG; and AH, DH, , FH, TH, RH BH, respectively; Minimum ferret of 2.73 - 2.96 and Maximum ferret of 3.78 - 4.94 are also good discriminants for AH, PF, FH, RF; and SH and FH. Good discriminants for AH, DH, SH, FH, TH, BH, RH; and BH, TH; AH, BH; are Average ferret, Aspect ratio and Shape factor for the cut-off points of 3.77 - 3.90, 0.52 - 0.68 and 0.64 - 0.74 respectively as compared to the normal sperm head. For all these comparisons the sensitivity **or** specificity was at least 70%.

Table 3. 3. 7 (b) represents the sperm heads that fall outside and inside the selected range for the normal sperm head for each parameter with a sensitivity **and** specificity of at least 70% or higher. Good discriminants for BH and DH, FH and SH, TH, FH, respectively are perimeter, area and length respectively and good discriminants for FH and SH, FH, RH and BH respectively are width, average ferret and shape factor respectively as compared to the normal sperm head. The poor discriminant is aspect ratio. When using a high level of specificity and sensitivity (Fig. 3.3.7b) it was not possible to distinguish between normal sperm heads and amorphous sperm heads for any sperm head parameter.

3.3 Quantitative Image Analysis of sperm head

3.3.1: Morphometric parameters of sperm head.

Table 3.3.1 indicates various sperm forms and donor / patients types.

Normal Head Amorphous Head Big Head Double Head Flame Head Round Head Small Head Tapered Head	Sperm smears of patients visiting the andrology clinic at Groote Schuur Hospital for semen evaluation.
Donor Head	Refer to only normal sperm from sperm smears of 30 healthy male volunteers aged between 25 - 35 years who regularly supplied ejaculates at Groote-Schuur andrology clinic and are rated fertile by all WHO parameters and Tygerberg Strict Criteria for morphology.
Proven fathers sperm head	Refer to only normal sperm of fathers who had one or two children in the past 6-18 months from the Centre for Fertility Studies at Pretoria Central Hospital.

From Table **3.3.1**, it is clear that morphometric characteristics have been analysed of normal sperm from patients with potential infertility problems, normal sperm from apparently fertile donors by WHO criteria and normal sperm from proven fathers. Morphometric parameters of all abnormal sperm types have only been evaluated in patients visiting an andrology clinic because of potential fertility problems.

Table 3.3.2 Population averages and ± SD for seven morphometric parameters of normal and abnormal sperm-head indicated in Table 3.3.1.

Sperm head form and donor types	Perimeter (μm)	Area (μm^2)	Length(max ferret) (μm)	Width (min. ferret) (μm)	Average ferret (μm)	Aspect ratio	Shape factor
Amorphous heads	11.90 ± 0.94*	7.73 ± 1.15 *	4.32 ± 0.39 *	2.50 ± 0.22	3.44±0.24*	0.58 ±0.06	0.69 ± 0.06*
Big Head	18.36 ± 1.80*	14.01±3.48*	6.47 ± 0.54*	3.19 ± 0.74*	4.89±0.40*	0.49 ±0.12*	0.52 ± 0.10*
Double Head	11.90 ± 1.80*	14.75 ±3.48*	5.26 ± 0.79*	3.22 ± 0.68*	4.30 ± 0.59 *	0.61 ± 0.10	0.72 ± 0.16 *
Donor Head	12.03 ± 1.09*	8.61 ± 2.09 *	4.49 ± 0.41	2.57 ± 0.24	3.55 ±0.28	0.58 ±0.10	0.67 ± 0.70
Flame Head	9.13 ± 0.72*	4.65 ± 0.89 *	3.21 ± 0.40 *	2.08 ± 0.28*	2.68±0.24*	0.66 ± 0.07 *	0.70 ± 0.10*
Normal Head	12.41 ± 1.27	8.08 ± 1.16	4.42 ± 0.38	2.56 ± 0.23	3.52 ± 0.24	0.58 ±0.12	0.66 ± 0.07
Proven Fathers Head	13.32 ± 1.81*	9.17 ± 0.91 *	5.02 ± 0.47*	2.61 ± 0.26	3.77±0.07*	0.52 ± 0.07 *	0.67 ± 0.12
Round Head	14.07 ± 2.10*	10.25 ± 2.79*	5.25 ± 0.46*	2.71 ± 0.70	4.00±0.14*	0.52 ± 0.14*	0.65 ± 0.11
Small Head	-	9.28 ± 0.50 *	3.47±0.19*	2.78 ±0.13*	2.88±0.03*	0.64 ± 0.03*	0.71 ± 0.06 *
Tapered head	14.33 ±1.52*	9.98 ±1.53*	5.55 ±0.45*	2.66 ±0.95*	4.09 ±0.33*	0.48 ±0.10*	0.63 ± 0.11*

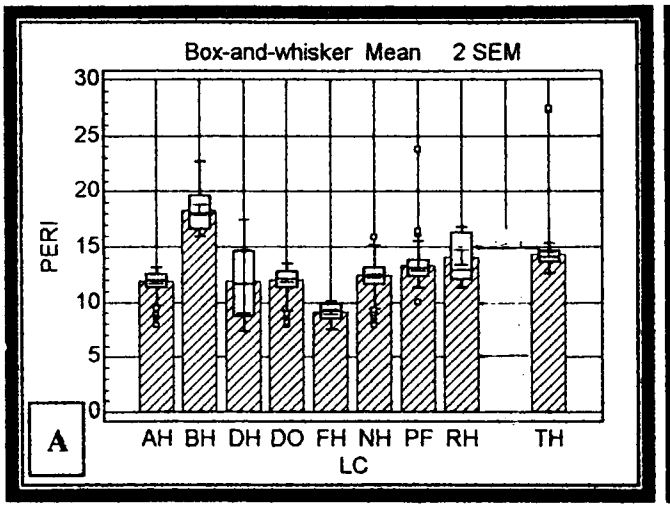
- ***denotes a statistically significant difference and $P < 0.05$ in columns**

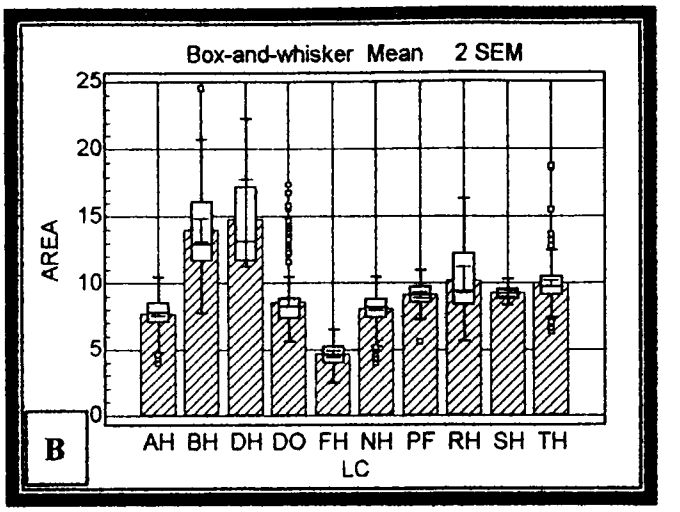
Table 3.3.3 of abbreviations that serves as legend for all the tables and figures:

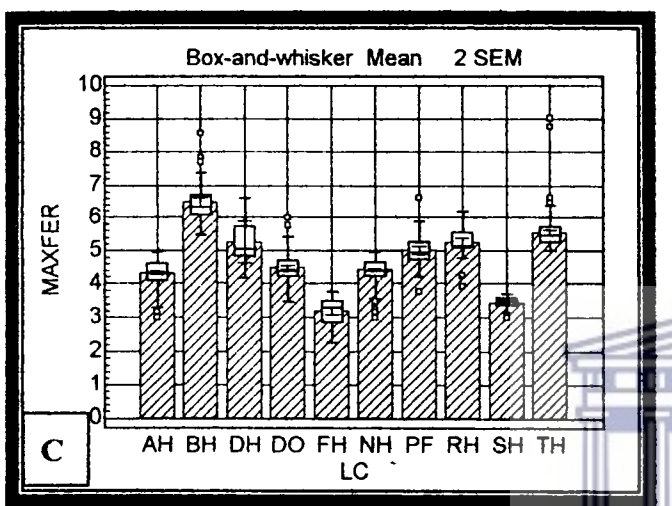
PARAMETER	ABREVIATIONS USED IN TABLES	ACTUAL TYPE OR MEASUREMENT
PERIMETER	PERI	Represents the total circumference of the sperm head as measured in µm
AREA	AREA	Represents the total surface area of the sperm head as measured in μm^2
MAXIMUM FERRET	MAXFER	Represents the length of the sperm head as measured in µm
MINIMUM FERRET	MINFER	Represents the width of the sperm head as measured in µm
AVERAGE FERRET	AVEFER	Represents the average between the length and the width of the sperm head as measured in µm
ASPECT RATIO	ASPECT	Represents the ratio of the perimeter and the area and the value is between 0 and 1
SHAPE FACTOR	SHAPE	Represents the regularity of shape and the value is between 0 and 1

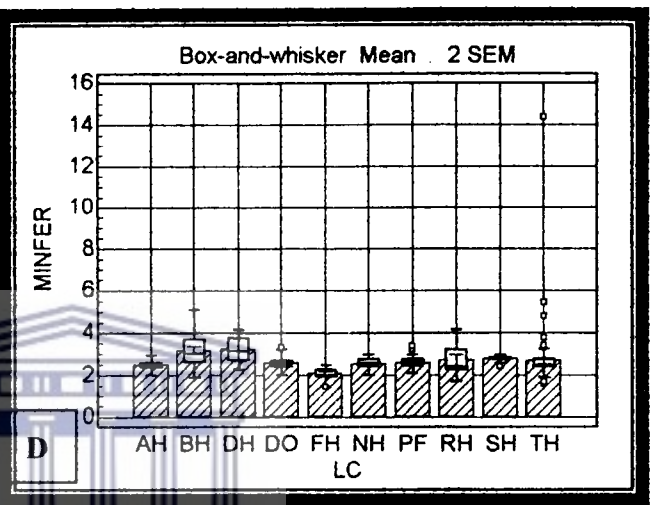
Table 3. 3. 4 : ROC-curve analysis showing cut-off points among nine different sperm head forms compared to normal sperm head of humans for seven morphometric parameters.

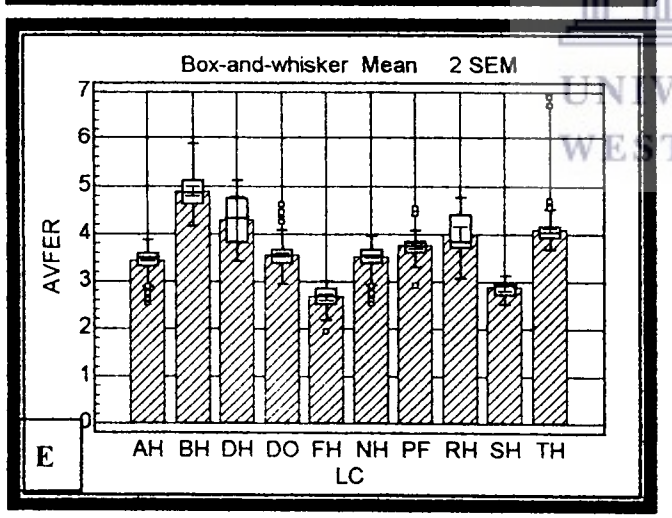
SPERM HEAD PARAMETER	AH	DH	DO	PF	SH	TH	FH	BH	RH
PERIMETER	>13.22	>10.55	>13.26	<=13.00	-	<=13.03	>10.17	<=15.93	<=15.19
AREA	>8.55	<=10.50	<=10.46	<=8.28	<=8.45	<=8.98	>6.00	<=9.88	<=9.36
MINFER	<=2.03	<=2.81	>2.75	<=2.50	<=2.73	<=2.96	>2.35	<=2.96	>2.27
MAXFER	>4.41	<=4.94	<=4.98	<=4.79	>3.72	<=4.98	>3.77	<=4.98	<=4.98
AVEFER	>3.59	<=3.90	<=3.11	<=3.60	>3.15	<=3.77	>3.03	<=3.99	<=3.68
ASPECT	<=0.62	<=0.68	>0.54	>0.52	<=0.60	>0.52	<=0.66	>0.49	>0.47
SHAPE	<=0.61	<=0.73	<=0.58	<=0.74	<=0.71	>0.64	<=0.69	>0.61	>0.64

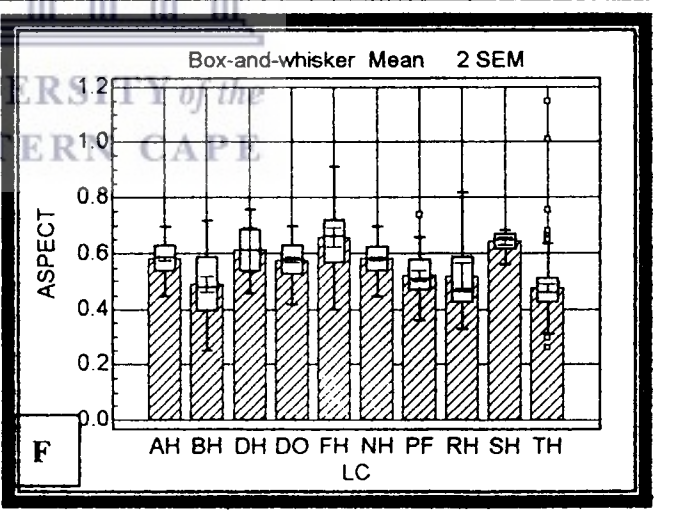


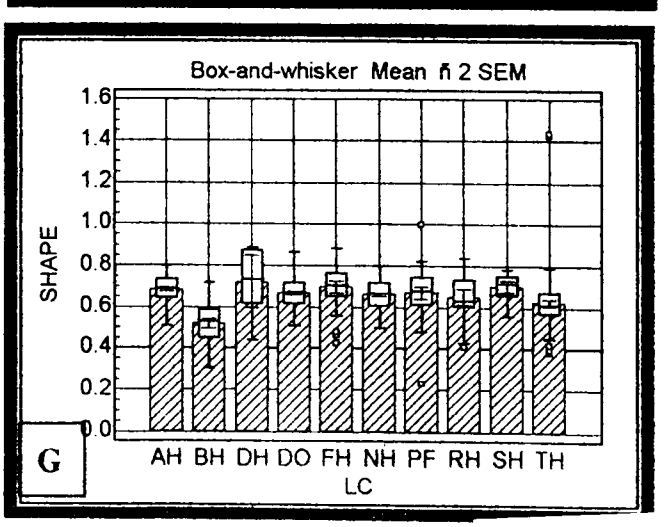




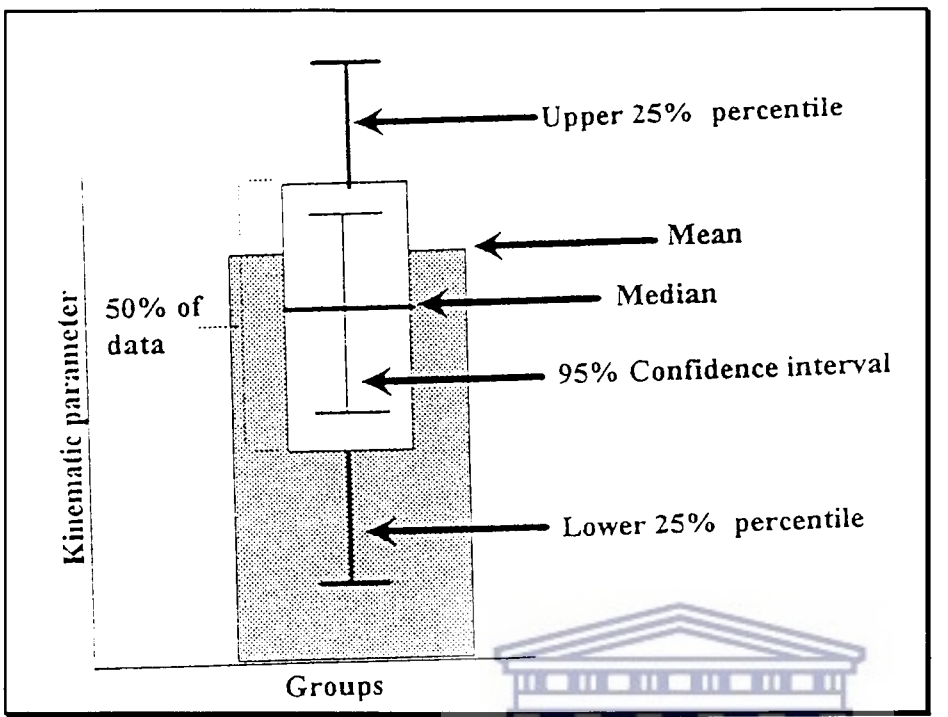








Figs. 3.3.1 A - G illustrate the Box - and - Whiskers plots of human sperm head types for seven parameters



Legend:

Fig. 3.3.2 This T - thin bar represents 95% CI (2± SEM) (From the M.Sc thesis of Mdhluli, C, 1998, UWC)

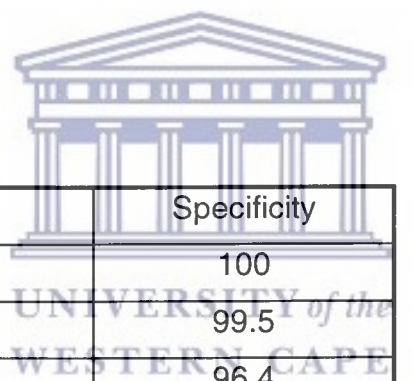
Table 3.3.5 a: Receiver Operating Characteristic curve data indicating best cut-off point (with asterisk in bold) with sensitivity and specificity level indicated for perimeter of normal sperm head (Positive group)(as measured with FIPS) compared to perimeters of small, tapered, flame and big sperm heads (Negative groups) of humans. In this example the best cut-off values of the following perimeters together with their sensitivity and specificity have been highlighted: tapered, flame and big sperm heads.

ROC CURVE

VARIABLE = PERI

Area under the ROC curve = 0.897

A : TAPERED HEADED SPERM



Criterion	Sensitivity	Specificity
< 7.967	0	100
<=12.69	61.1	99.5
<=12.991	70.1	96.4
<=13.032 *	**71.2**	**96.4**
<=13.038	71.2	95.8

ROC-CURVE

VARIABLE = PERIMETER

Area under the ROC curve = 0.987

B: FLAME HEADED SPERM

Criterion	Sensitivity	Specificity
> 10.005	97.1	91.1
> 10.094	96.9	91.1
> 10.119	96.7	95.6
> 10.138	96.5	97.8
> 10.167 *	**96.2**	**100**
> 15.926	0	100

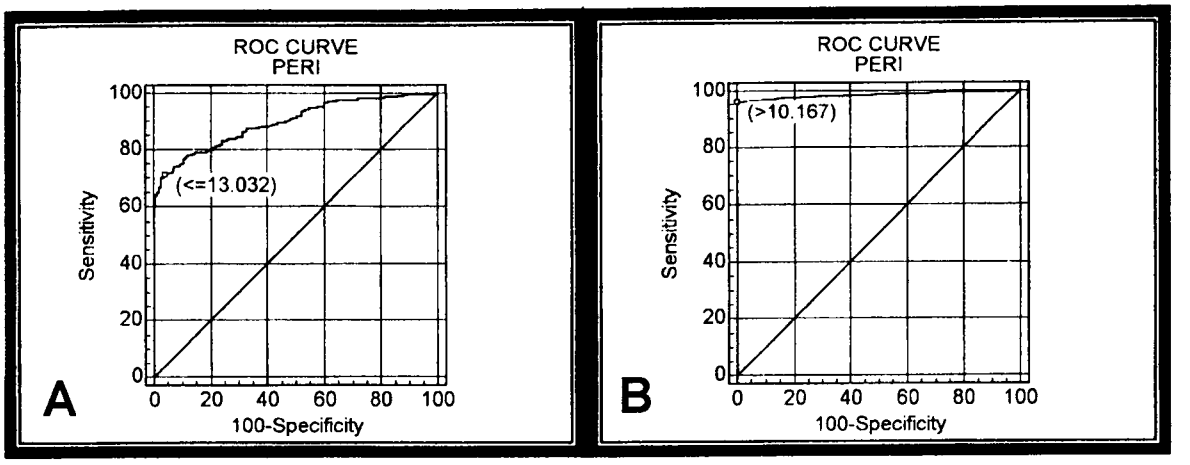
ROC CURVE

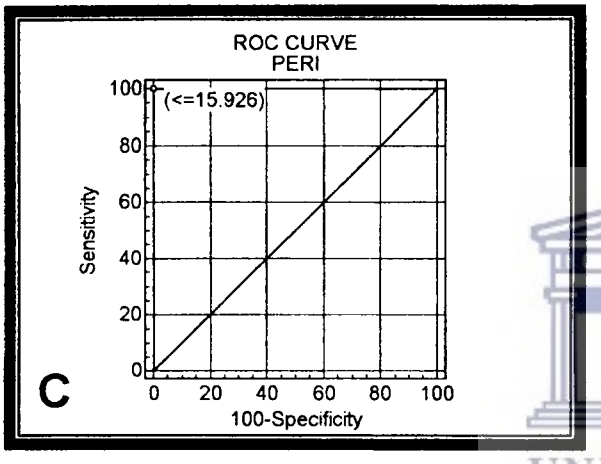
VARIABLE = PERIMETER

Area under the ROC curve = 1.000

C: BIG HEADED SPERM

Criterion	Sensitivity	Specificity
>7.967	0.0	100.0
<=15.926*	**100.0**	**100.0**
<=22.766	100.0	0.0





Legend:

Figs 3.3.3: These are ROC-curves of sperm head for perimeter showing the cut-off points of

A) Tapered headed sperm

B) Flame headed sperm

C) Big headed sperm

Table 3.3.5 b: Receiver Operating Characteristic curve data indicating best cut-off points (with asterisk in bold) with sensitivity and specificity level indicated for area of normal sperm head (Positive group) (as measured with FIPS) compared to area of double, small, tapered, flame and big sperm heads (Negative group) of human sperm. In this example , the best cut-off values of the following areas together with their sensitivity and specificity have been highlighted : double, tapered, flame and big sperm heads.

ROC CURVE

VARIABLE = AREA

Area under the ROC curve = 1.000

A. DOUBLE HEADED SPERM

Criterion	Sensitivity	Specificity
< 3.931	0	100
<=10.495 *	**100**	**100**
<=22.38	100	0

ROC CURVE

VARIABLE = AREA

Area under the ROC curve = 0.880

B. TAPERED HEADED SPERM

Criterion	Sensitivity	Specificity
<=8.872	74.1	84.9
<=8.958	78.1	83.9
<=8.976 *	**79.4**	**83.9**
<=9.01	80.1	82.3

ROC CURVE

VARIABLE = AREA

Area under the ROC curve = 0.983

C : FLAME HEADED SPERM

Criterion	Sensitivity	Specificity
>=2.467	100	0
> 3.931	99.8	22.2
> 5.987	93.8	95.6
> 5.995 *	**93.6**	**97.8**
> 6.549	88.9	97.8
> 10.495	0	100

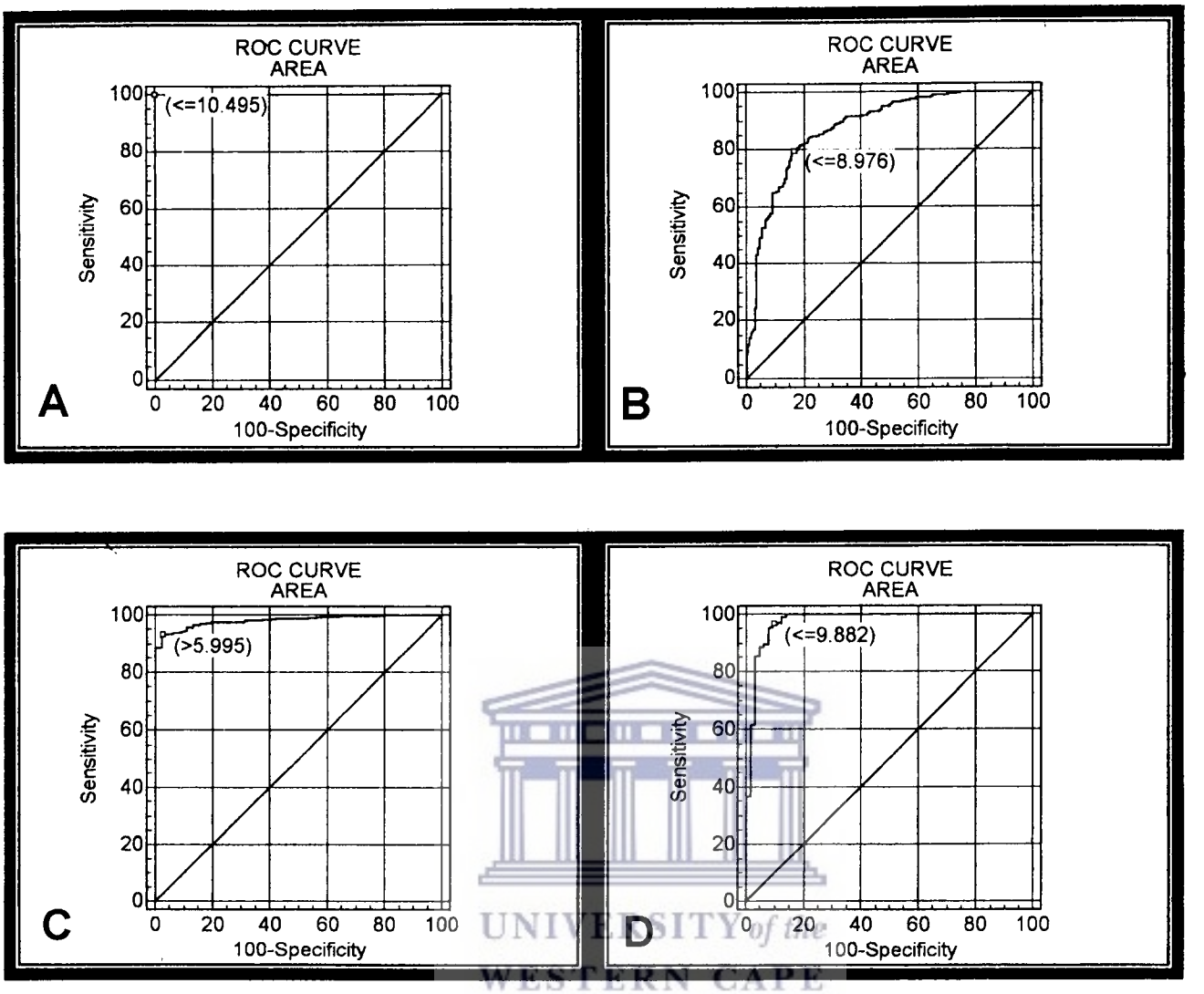
ROC CURVE

VARIABLE = AREA

Area under the ROC curve = 0.977

D : BIG HEADED SPERM

Criterion	Sensitivity	Specificity
<=9.788	95.4	92.3
<=9.846	95.8	90.8
<=9.882 *	**96.9**	**90.8**
<=9.918	96.9	89.2
<=24.488	100	0



Legend:

Figs 3.3.4: These are ROC-curves of sperm head for area showing cut-off points of

A) Double headed sperm
B) Tapered headed sperm
C) Flame headed sperm
D) Big headed sperm

Table 3.3.5 c: Receiver Operating Characteristic curve data indicating best cut-off point (with asterisk in bold) with sensitivity and specificity level indicated for maximum ferret of normal sperm head (Positive group)(as measured with FIPS) compared to maximum ferret of double, proven fathers, tapered, flame, big and round sperm heads (Negative groups) of humans. In this example , the best cut-off values of the following areas together with their sensitivity and specificity have been highlighted : double, proven fathers, tapered, flame, big and round sperm heads.

ROC CURVE

VARIABLE = MAXFER

Area under the ROC curve = 0.840

A . DOUBLE HEADED SPERM

Criterion	Sensitivity	Specificity
< 3.015	0	100
<=4.104	19.2	100
<=4.67	67.9	75
<=4.94 *	**91.4**	**75**
<=4.97	91.4	50
<=4.975	100	50
<=6.63	100	0

ROC CURVE

VARIABLE = MAXFER

Area under the ROC curve = 0.837

B. PROVEN FATHER HEADED SPERM

Criterion	Sensitivity	Specificity
< 3.015	0	100
<=3.769	6.4	100
<=4.636	67.9	81.7
<=4.673	75.2	75
<=4.788 *	**78.5**	**75**
<=4.975	100	43.3
<=6.63	100	0

ROC CURVE

VARIABLE = MAXFER

Area under the ROC curve = 1.000

C . TAPERED HEADED SPERM

Criterion	Sensitivity	Specificity
< 3.015	0	100
<=4.975 *	**100**	**100**
<=9.0433	100	0

ROC CURVE

VARIABLE = MAXFER

Area under the ROC curve = 0.985

D FLAME HEADED SPERM

Criterion	Sensitivity	Specificity
> 3.572	96.5	84.4
> 3.618	96.2	88.9
> 3.648	95.6	91.1
> 3.724	94.2	93.3
> 3.769 *	**93.6**	**100**
> 4.975	0	100

ROC CURVE

VARIABLE = MAXFER

Area under the ROC curve = 1.000

E . BIG HEADED SPERM

Criterion	Sensitivity	Specificity
< 3.015	0	100
<=4.975 *	**100**	**100**
<=8.593	100	0

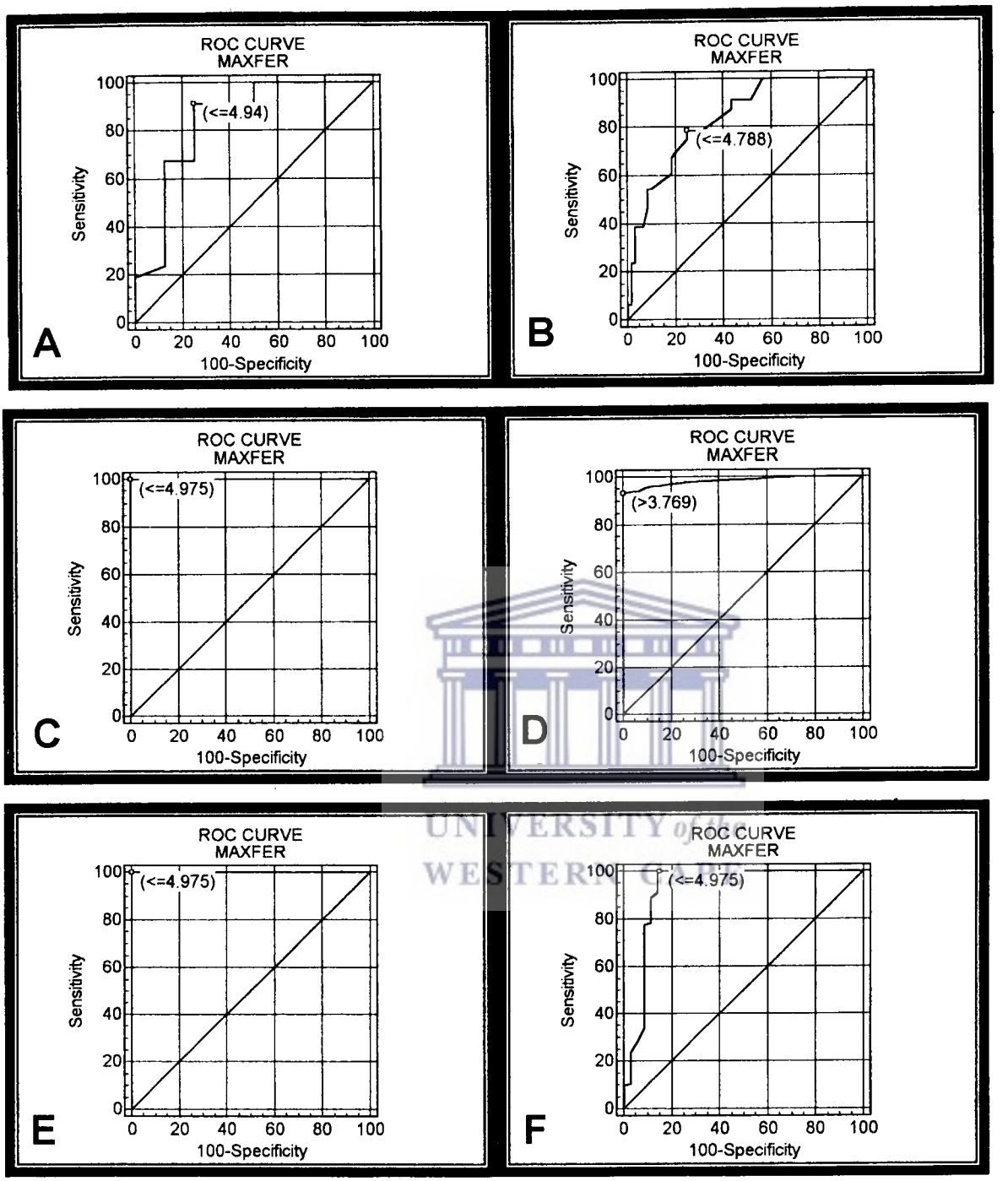
ROC CURVE

VARIABLE = MAXFER

Area under the ROC curve = 0.924

F . ROUND HEADED SPERM

Criterion	Sensitivity	Specificity
< 3.015	0	100
<=4.221	27.9	94.3
<=4.256	33.8	91.4
<=4.712	77.7	91.4
<=4.94	91.4	85.7
<=4.975 *	**100**	**85.7**
<=6.181	100	0



Legend:

Figs. 3.3.5: These are ROC-curves of sperm head for maximum ferret showing cut-off points of

A) Double headed sperm
B) Proven fathers headed sperm
C) Tapered headed sperm
D) Flame headed sperm
E) Big headed sperm
F) Round headed sperm

Table 3.3.5 d: Receiver Operating Characteristic curve data indicating best cut- off point (with asterisk in bold) with sensitivity and specificity level indicated for average ferret of a normal sperm head (Positive group), (as measured with FIPS) compared to average ferret of small, tapered, flame, big and round sperm heads (Negative groups) of human sperm. In this example , the best cut-off values of the following areas together with their sensitivity and specificity have been highlighted :small, tapered, flame, big and round sperm heads.

ROC CURVE

VARIABLE = AVFER

Area under the ROC curve = 0.972

A . SMALL HEADED SPERM

Criterion	Sensitivity	Specificity
> 2.933	97.3	55
> 2.942	97.1	55
> 3.094	93.6	95
> 3.141	92	95
> 3.151 *	**91.6**	**100**
> 3.986	0	100

ROC CURVE

VARIABLE = AVFER

Area under the ROC curve = 0.982

B . TAPERED HEADED SPERM

Criterion	Sensitivity	Specificity
<=3.735	85.4	97.9
<=3.751	87.4	97.4
<=3.768 *	**88.9**	**96.4**
<=3.948	98.7	71.9

ROC CURVE

VARIABLE = AVFER

Area under the ROC curve = 0.988

C . FLAME HEADED SPERM

Criterion	Sensitivity	Specificity
> 2.866	98.2	75.6
> 2.876	98	75.6
> 2.895	97.8	77.8
> 2.974	96.5	93.3
> 3.028 *	**94.7**	**100**
> 3.986	0	100

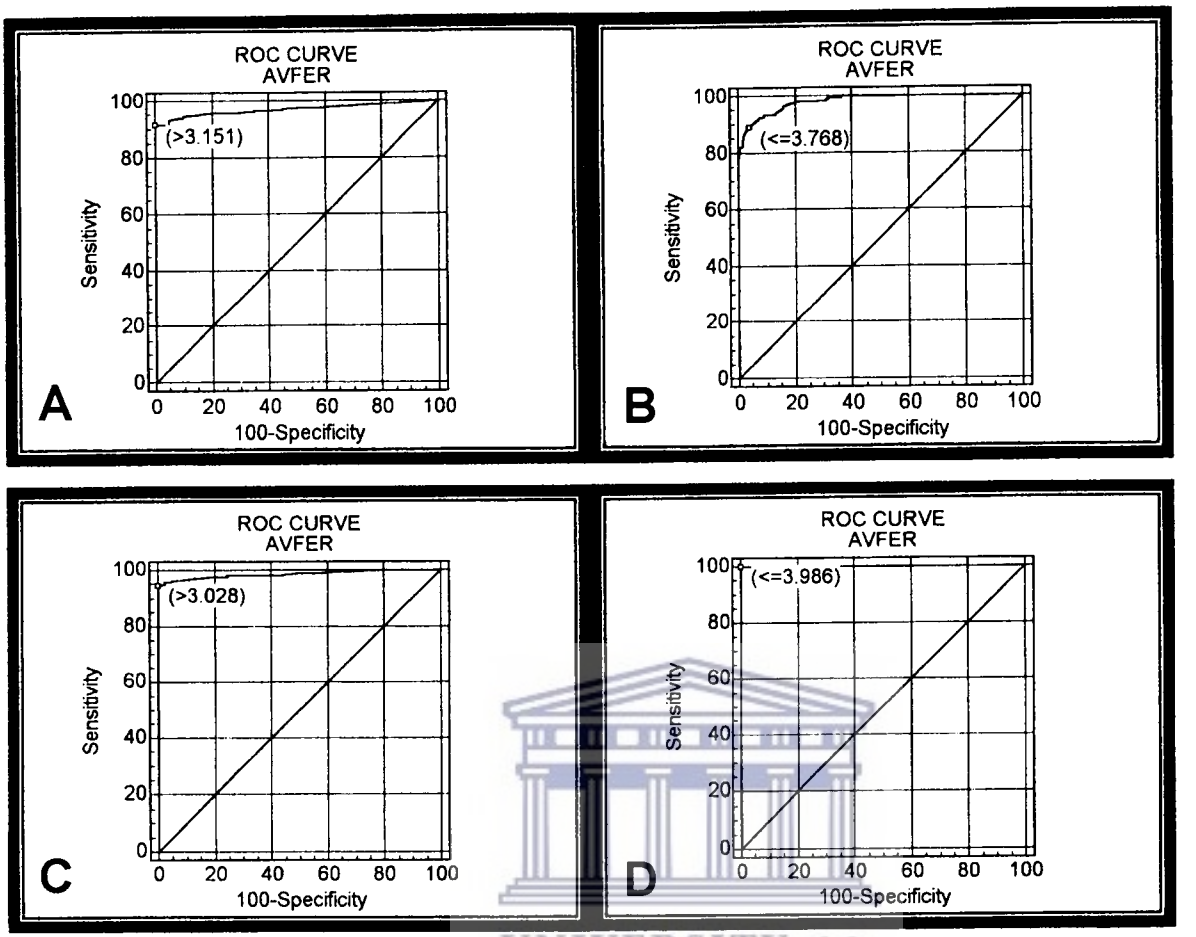
ROC CURVE

VARIABLE = AVFER

Area under the ROC curve = 1.000

D . BIG HEADED SPERM

Criterion	Sensitivity	Specificity
< 2.552	0	100
<=3.986 *	**100**	**100**
<=5.884	100	0



Legend:

Fig. 3.3.6: These ROC-curves of sperm heads fall average ferret showing cut-off points of

A) Small headed sperm

B) Tapered headed sperm

C) Flame headed sperm

D) Big headed sperm

Table 3.3.6: Represents the ROC-curve analysis which shows the cut-off values of the nine different sperm head types compared to the normal sperm head of humans: Sensitivity and Specificity levels above 70% are indicated in bold to emphasize best cut-off points for different parameters.

		SPERM HEAD PARAMETER						
SPERM HEAD TYPE		Peri	Area	Minfer	Maxfer	Avefer	Aspect	Shape
AH	Cut-off Value	>13.22	>8.55	<=2.03	>4.41	>3.59	<=0.62	<=0.61
	Area Under ROC-curve	0.62	0.60	0.50	0.58	0.60	0.51	0.59
	Sensitivity	24.1	39.4	1.3	49.6	44.7	71.2	25.9
	Specificity	100.0	74.2	98.7	62.5	70.6	32.1	88.0
DH	Cut-off Value	>10.55	<=10.50	<=2.81	<=4.94	<=3.90	<=0.68	<=0.73
	Area Under ROC-curve	0.56	1.00	0.79	0.84	0.89	0.61	0.66
	Sensitivity	93.6	**100.0**	**85.4**	**91.4**	**98.2**	94.2	76.8
	Specificity	50.0	**100.0**	**75.0**	**75.0**	**75.0**	37.5	62.5
DO	Cut-off Value	>13.26	<=10.46	>2.75	<=4.98	<=3.11	>0.54	<=0.58
	Area Under ROC-curve	0.57	0.52	0.50	0.52	0.50	0.52	0.52
	Sensitivity	22.3	99.8	27.4	100.0	7.1	77.0	19.7
	Specificity	99.5	10.5	76.8	5.8	98.4	30.0	86.8
PF	Cut-off Value	<=13.00	<=8.28	<=2.50	<=4.79	<=3.60	>0.52	<=0.74
	Area Under ROC-curve	0.67	0.78	0.52	0.84	0.79	0.75	0.55
	Sensitivity	70.1	51.3	37.8	**78.5**	56.4	81.2	83.0
	Specificity	55.0	93.3	75.9	**75.0**	88.3	58.3	31.7
SH	Cut-off Value	-	<=8.4	<=2.73	>3.72	>3.15	<=0.60	<=0.71
	Area Under ROC-curve	-	0.83	0.79	0.98	0.97	0.81	0.67
	Sensitivity	-	56.9	**78.8**	**94.2**	**91.6**	60.8	69.7
	Specificity	-	100.0	**85.0**	**100.0**	**100.0**	95.0	60.0

TH	Cut-off Value	<=13.03	<=8.98	<=2.96	<=4.98	<=3.77	>0.52	>0.64
	Area Under ROC-curve	0.90	0.88	0.50	1.0	0.98	0.87	0.65
	Sensitivity	**71.2**	**79.4**	98.2	**100.0**	**88.9**	**84.5**	61.9
	Specificity	**96.4**	**83.9**	12.5	**100.0**	**96.4**	**76.0**	62.5
FH	Cut-off Value	>10.17	>6.0	>2.35	>3.77	>3.03	<=0.66	<=0.69
	Area Under ROC-curve	0.99	0.98	0.91	0.99	0.99	0.72	0.64
	Sensitivity	**96.2**	**93.6**	**97.5**	**93.6**	**94.7**	89.4	60.2
	Specificity	**100.0**	**97.8**	**88.0**	**100.0**	**100.0**	51.1	64.4
BH	Cut-off Value	<=15.93	<=9.88	<=2.96	<=4.98	<=3.9	>0.49	>0.61
	Area Under ROC-curve	1.0	0.98	0.76	1.0	1.0	0.74	0.88
	Sensitivity	**100.0**	**96.9**	98.2	**100.0**	**100.0**	93.1	**74.8**
	Specificity	**100.0**	**90.8**	53.8	**100.0**	**100.0**	55.4	**86.2**
RH	Cut-off Value	<=15.19	<=9.36	>2.27	<=4.96	<=3.6	>0.47	>0.64
	Area Under ROC-curve	0.70	0.73	0.54	0.92	0.85	0.71	0.53
	Sensitivity	99.8	89.8	86.5	**100.0**	**74.3**	98.2	63.5
	Specificity	42.9	51.4	42.9	**85.7**	**85.9**	54.3	54.3

Table 3.3.7(a): Sperm head which clearly fall outside and inside the range for each parameter with a Sensitivity **or** Specificity of at least 70% but usually higher. DO and PF are excluded as good discriminants.

PARAMETER	SELECTED RANGE FOR NORMAL SPERM HEAD AS DETERMINED BY ROC	SPERM HEAD THAT FALL OUTSIDE THE RANGE (GOOD DISCRIMINANT)	SPERM HEAD THAT FALL WITHIN THE RANGE (POOR OR NO DISCRIMINANT)
PERI	10.71 - 14.30	DH, FH, BH and RH	AH, DO, PF, and TH
AREA	8.98 - 9.88	AH,DH, SH, and FH,	TH, BH and RH
MAXFER	3.78 - 4.94	DH, SH, TH, FH, BH and RH	AH, DH and PF
MINFER	2.73 - 2.96	AH, FH and RH	DH, DO, SH, TH and BH
AVEFER	3.77 - 3.90	AH, SH, FH, BH and RH	DH and TH
ASPECT	0.52 - 0.68	BH and RH	AH, DH, DO, PF, SH, TH and FH
SHAPE	0.64 - 0.74	BH	AH, DH, PF, SH, TH, FH, and RH

Table 3.3.7(a): Sperm head which clearly fall outside and inside the range for each parameter with a Sensitivity **or** Specificity of at least 70% but usually higher. DO and PF are excluded as good discriminants.

PARAMETER	SELECTED RANGE FOR NORMAL SPERM HEAD AS DETERMINED BY ROC	SPERM HEAD THAT FALL OUTSIDE THE RANGE (GOOD DISCRIMINANT)	SPERM HEAD THAT FALL WITHIN THE RANGE (POOR OR NO DISCRIMINANT)
PERI	10.71 - 14.30	DH, FH, BH and RH	AH, DO, PF, and TH
AREA	8.98 - 9.88	AH,DH, SH, and FH,	TH, BH and RH
MAXFER	3.78 - 4.94	DH, SH, TH, FH, BH and RH	AH, DH and PF
MINFER	2.73 - 2.96	AH, FH and RH	DH, DO, SH, TH and BH
AVEFER	3.77 - 3.90	AH, SH, FH, BH and RH	DH and TH
ASPECT	0.52 - 0.68	BH and RH	AH, DH, DO, PF, SH, TH and FH
SHAPE	0.64 - 0.74	BH	AH, DH, PF, SH, TH, FH, and RH

3. 4 Quantitative image analysis of the human sperm acrosome

3.4.1 General aspects

The main aim of this part of the research was to analyze normal and abnormal sperm acrosome from patients visiting the routine andrology laboratory at Groote Schuur Hospital by means of quantitative image analysis. This may be used in for quantitative and automated computer sperm acrosome morphology analysis. Normal sperm of thirty donors from Groote Schuur hospital as well as normal sperm of 77 proven fathers have also been used for comparisons in this study.

3.4.2 Key to presentation of results

Table 3.4.1 summarizes the various sperm acrosome types analyzed as well as the donor/patient types. Table 3.4.2 summarizes the morphometric parameters analyzed for each sperm acrosome type listed in Table 3.4.1. The abbreviations for the above parameters are furthermore indicated in Table 3.3.3 to serve as legend for future tables and figures. Each of these parameters were determined by means of FIPS image analysis system as determined under Materials and Methods (page 23).

Statistical analysis has been performed as indicated under 2.3.6 (Materials and Methods). For the sake of simplicity the presentation of these results are presented in tables, Box-and-Whiskers plots (see Fig.3.4.1 which illustrates the various features of the Box-and-whiskers plots used throughout this study), and Receiver Operating Characteristic curves (ROC-curve) which each time illustrates the cut-off point between normal and each abnormal sperm acrosome type for each of seven morphometric parameters as indicated in Table 3.4.3.a - 3.4.3 d. Table 3.4.4 summarizes the best cut-off points for acrosomes among the nine different types of sperm for seven morphometric parameters. Receiver Operating Characteristic curves (ROC-curve) which shows the cut-off values of normal sperm acrosomes and Sensitivity and Specificity above 70 are indicated in Table 3.4.5. Table 3.4.6 shows sperm acrosome which clearly fall outside / inside the range for each parameter with a Sensitivity / Specificity of at least 70% but usually higher.

3.4.3 Quantitative morphometric analysis of the human sperm acrosome

Summary statistics of seven morphometric parameters of normal and abnormal sperm acrosomes are indicated in Table 3.4.2. These results are furthermore graphically presented as Box-and-Whisker plots (Fig. 3.4.1). The Amorphous sperm acrosome (AH) and Normal sperm acrosome (NH), show similarities in terms of Peri, Area, Maxfer, Minfer, Avefer and Aspect (see figs. 3.4.1 a, b, c, d, e, and f). Flame sperm acrosomes and big head sperm acrosomes differ statistically significantly from normal sperm acrosomes for all seven parameters ($P< 0.05$ -ANOVA). Round sperm acrosomes differ statistically significantly from Tapered sperm acrosomes for all parameters except for width (minfer) and aspect ratio.

It is however surprising that proven father's sperm acrosomes differ from normal sperm acrosomes for all parameters except the aspect ratio. On the basis of these results it appears that the sperm acrosome from NH and DO are larger than proven fathers sperm acrosomes, with the exception of the aspect ratio and the area (Table 3.4.3 a-d and Fig.3.4.1). The acrosomes of the Double headed (DH) sperm differ from all other sperm acrosomes for most parameters ($P< 0.05$).

It was indicated previously that these results may be useful in computer automated image analysis of sperm morphology. ROC-curves represent a useful means of establishing the cut-off points between normal and abnormal. Tables 3.4.4 a-d and Figs 3.4.3 are ROC-curve analyses which show the cut-off points between NH and each abnormal sperm type. In Tables 3.4.5 the cut-off value is indicated in bold and information is also supplied on the sensitivity and specificity for each cut-off point for each parameter. Fig.3.4.3 represents a visual presentation of the ROC-curves and emphasizes the area under the ROC-curve. Table 3.3.6 (a,b) represent sperm acrosomes which clearly fall outside and inside the range for each parameter with a sensitivity **or/and** specificity of at least 70% but usually higher. A perfect cut-off point is one which shows 100% sensitivity and 100% specificity and the area under the ROC-curve is 1 (the maximum for this feature). The best cut-off points were observed in Maxfer.

Table 3. 4. 6 (a) represents sperm accrosomes which clearly fall outside and inside the selected range for normal sperm acrosome for each parameter with a sensitivity **or** specificity of at least 70% but usually higher as determined by ROC-curve analysis. Good discriminants for DH, BH; SH, BH; and DH, BH respectivily as compared to the normal sperm acrosome, are Peri, Area and Maxfer and with the cut-off points of 4.04 - 8.79, 1.82 - 2.74 and 2.41 - 2.66 respectively. The following cut-off points: 1.55 - 1.67, 2.03 - 3.00, 0.52 - 0.58 and 0.49 - 0.52 are good discriminants for AH, TH, BH, RH; TH, FH, BH; DH, TH, FH, BH RH; and SH, FH respectively as compared to the normal sperm acrosome for width, avefer, aspect ratio and shape factor.

Table 3. 4. 6 (b) represents sperm acrosomes which fall outside and inside the selected range for each parameter with a sensitivity and specificity of at least 70% but usually higher. Good discriminants for DH, RH; TH as compared to the normal sperm acrosome are perimeter, area, length respectively, with the cut-off points of 4.05 - 8.79, 1.82 - 2.74 and 2.41 - 2.66 respectively. While good discriminants for TH, RH; and TH, BH are width and avefer, with the cut-off points of 1.55 - 1.67 and 2.03 - 3.00. The poor discriminants are aspect ratio and shape factor.

3.4: Morphometric parameters of sperm acrosome.

Table 3.4.1: Various sperm forms and donor / patients types.

Normal acrosome Amorphous acrosome Big acrosome Double acrosome Flame acrosome Round acrosome Small acrosome Tapered acrosome	Sperm smears of patients visiting the andrology clinic at Groote Schuur Hospital for semen evaluation.
Donor acrosome	Refer to only normal sperm acrosome from sperm smears of 30 healthy male volunteers aged between 25 - 35 years who regularly supplied ejaculates at Groote-Schuur andrology clinic and are rated fertile by all WHO parameters and Tygerberg Strict Criteria for morphology.
Proven fathers sperm acrosome	Refer to only normal sperm acrosomes of fathers who had one or two children in the past 6-18 months from the centre for fertility studies at Pretoria Central Hospital

From Table 3.4.1, it is clear that morphometric characteristics have been analysed of normal sperm-acrosome from patients with potential infertility problems, normal sperm-

acrosome from apparently fertile donors by WHO criteria and normal sperm-acrosome from proven fathers. Morphometric parameters of all abnormal sperm post-acrosome types have only been evaluated in patients visiting an andrology clinic because of potential fertility problems.

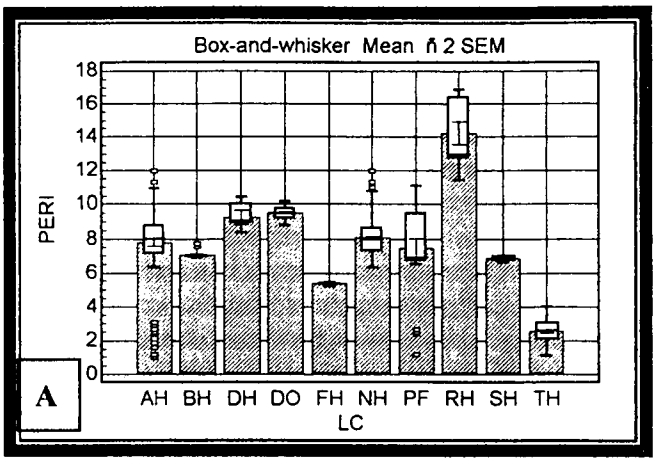
Table 3.4.2: Population averages and ± SD for seven morphometric parameters of normal and abnormal sperm - acrosome as indicated in Table 3.4.1.

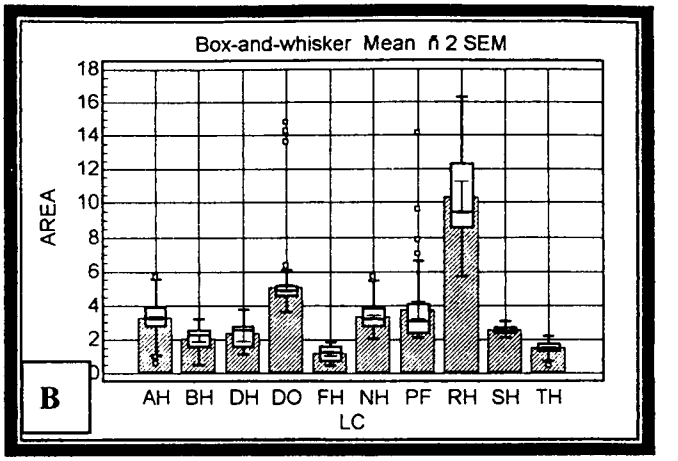
Sperm acrosome form and donor types	Perimeter (μm)	Area (μm^2)	Length(max ferret) (μm)	Width (min. ferret) (μm)	Average (ferret) (μm)	Aspect ratio	Shape factor
Amorphous heads	7.61 ± 1.83*	3.29 ± 0.91	2.89 ± 0.45*	1.86 ± 0.32*	2.37 ± 0.26	0.61 ± 0.09	0.65 ± 0.09
Big Head	11.00 ± 3.93*	6.61 ± 4.70*	4.14 ± 1.51*	2.16 ± 1.07*	3.15 ± 1.23*	0.53 ± 0.17*	0.59 ± 0.16*
Double Head	9.24 ± 0.77	2.34 ± 0.85*	3.31 ± 0.81*	1.43 ± 0.39*	2.37 ± 0.31	0.47 ± 0.18 *	0.36 ± 0.15 *
Donor Head	9.49 ± 0.37*	4.85 ± 0.47*	3.41 ± 0.27*	2.24 ± 0.22*	2.83 ± 0.20*	0.66 ± 0.07*	0.68 ± 0.06*
Flame Head	5.37 ± 0.07*	1.17 ± 0.44 *	2.18 ± 0.39*	0.94 ± 0.29*	1.56 ± 0.19*	0.45 ± 0.19*	0.51 ± 0.19 *
Normal Head	8.06 ± 0.99	3.34 ± 0.77	2.96 ± 0.38	1.76 ± 0.21	2.36 ± 0.25	0.60 ± 0.09	0.65 ± 0.08
Proven Fathers Head	7.46 ± 2.61*	3.56 ± 1.57	2.87 ± 0.56	1.69 ± 0.35	2.28 ± 0.42	0.59 ± 0.10	0.61 ± 0.09 *
Round Head	7.07 ± 0.36*	2.14 ± 0.76*	2.58 ± 0.44*	1.38 ± 0.33*	1.98 ± 0*	0.55 ± 0.16*	0.53 ± 0.17*
Small Head	6.87 ± 0.13*	2.53 ± 0.29*	2.57 ± 0.19*	1.57 ± 0.16*	2.07± 0.11*	0.62 ± 0.09	0.67 ± 0.06
Tapered Head	2.61 ± 0.70*	1.51 ± 0.32*	2.1 ± 0.21*	2.82 ± 0.45*	2.50 ± 0.31*	0.58 ± 0.15	0.54 ± 0.17*

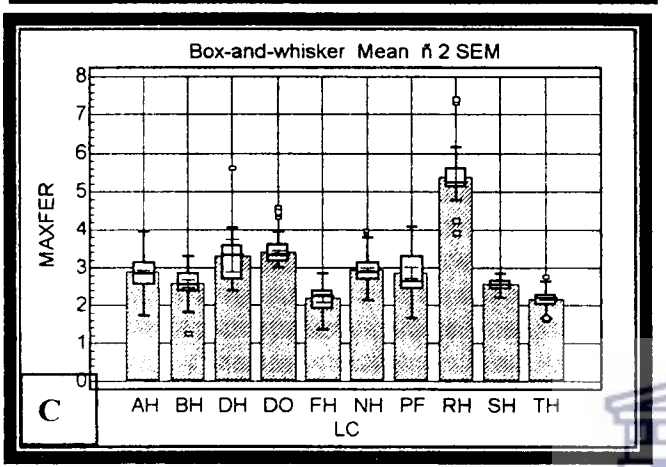
*** denotes a statistically difference and $p< 0.05$ in columns**

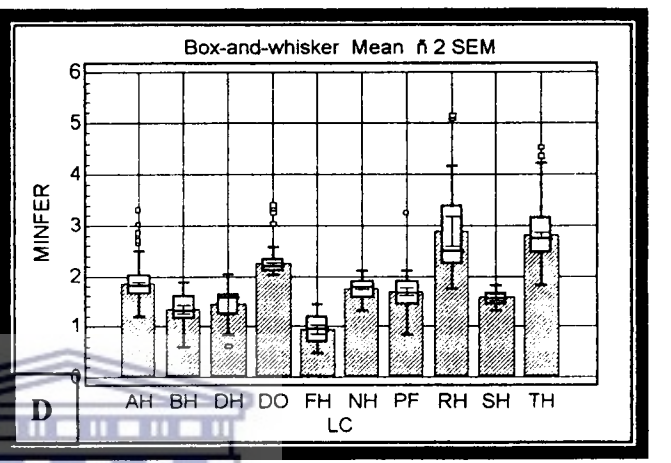
Table 3.4.3 : ROC-curve analysis showing cut-off points among 9 different sperm acrosome forms compared to normal sperm head of humans for seven morphometric parameters:

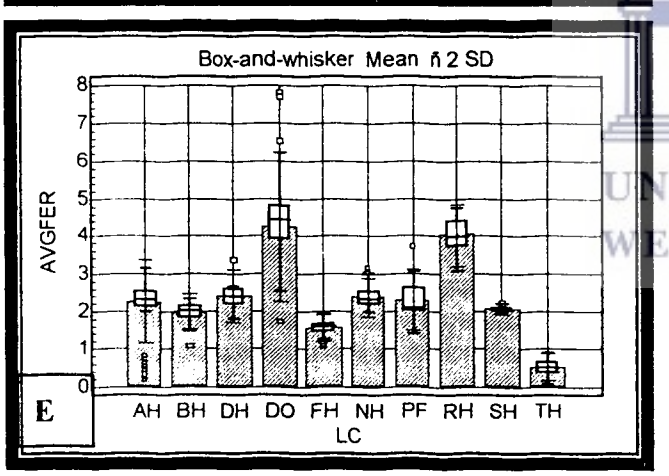
SPERM ACROSOME PARAMETER	AH	DH	DO	PF	SH	TH	FH	BH	RH
PERIMETER	>6.63	<=8.93	<=8.79	>6.98	>7.04	>4.05	>7.11	<=11.23	>5.48
AREA	>1.82	>2.74	<=4.17	>2.60	>3.08	>2.20	>2.59	<=5.67	>1.83
MINFER	<=1.68	>1.55	<=1.97	>1.44	>1.67	<=2.12	>1.67	<=2.12	>1.36
MAXFER	>2.66	<=3.32	<=3.02	>2.66	>2.66	>2.50	>2.51	<=3.80	>2.41
AVEFER	>2.03	<=2.59	<=2.54	>2.08	>2.23	>0.9	>1.93	<=3.05	>2.16
ASPECT	<=0.58	>0.45	<=0.58	>0.51	<=0.62	>0.49	>0.46	>0.49	>0.49
SHAPE	>0.52	>0.49	<=0.61	>0.65	<=0.59	>0.50	>0.55	>0.49	>0.49

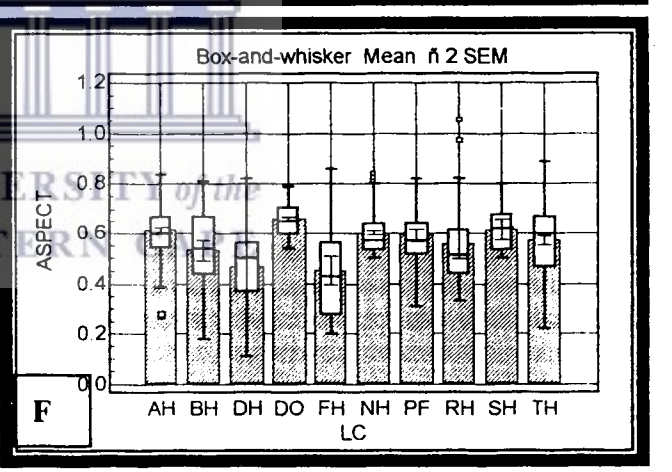


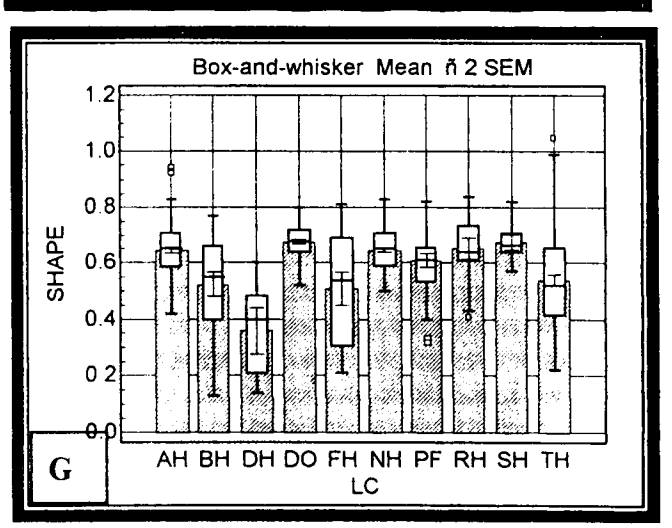












Figs. 3.4.1 A - G illustrate the Box - and - Whiskers plots of human sperm acrosome types for seven parameters.

Table 3.4.4 a: Receiver Operating Characteristic curve data indicating best cut- off point (with asterisk in bold) with sensitivity and specificity level indicated for perimeter of a normal sperm acrosome (Positive group), (as measured with FIPS) compared to perimeter of double, donor, tapered, flame, big and round sperm acrosomes (Negative groups) of human sperm. In this example , the best cut-off values of the following areas together with their sensitivity and specificity have been highlighted : double, donor, tapered, flame, and round sperm acrosomes.

ROC CURVE

VARIABLE = PERI

Area under the ROC curve = 0.790

A : DOUBLE HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
< 1.14	0	100
<=1.14	0	93.3
<=8.36	65	93.3
<=8.93 *	**81.2**	**80**
<=10.43	98	0
<=11.98	100	0

ROC CURVE

VARIABLE = PERI

Area under the ROC curve = 0.905

B : DONOR HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
< 6.034	0	100
<=8.79 *	**78.5**	**100**
<=8.82	78.5	99.5
<=8.83	79	99.5
<=8.99	82.5	88.8
<=10.18	96.9	0
<=11.98	100	0

ROC CURVE

VARIABLE = PERI

Area under the ROC curve = 1.000

C : TAPERED HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
>=1.12	100	0
> 4.05 *	**100**	**100**
> 11.98	0	100

ROC CURVE

VARIABLE = PERI

Area under the ROC curve = 0.812

D : FLAME HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
>=5.22	100	0
> 7.08	81.6	72.2
> 7.11 *	**81.2**	**88.9**
> 7.63	61.3	88.9
> 11.98	0	100

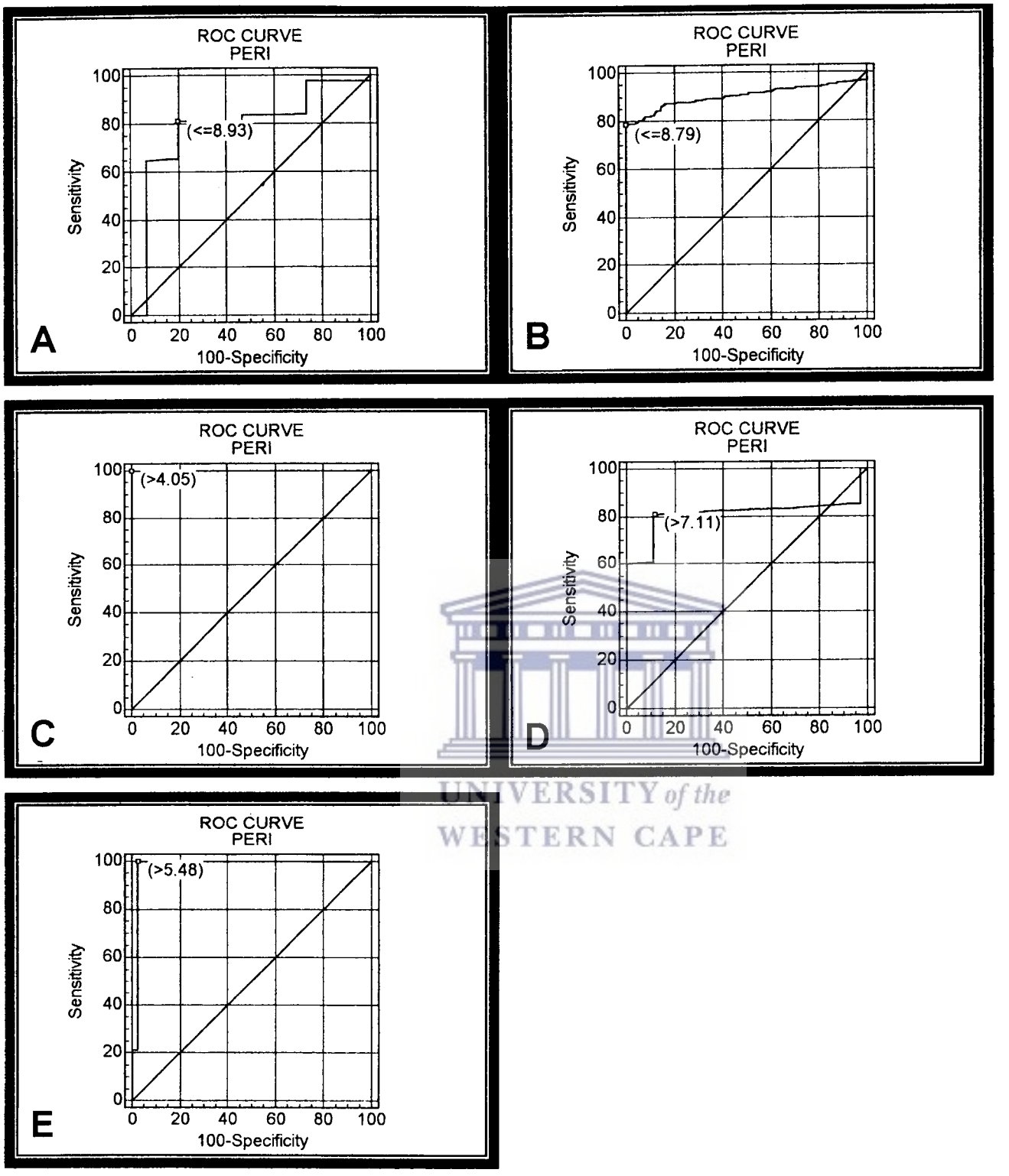
ROC CURVE

VARIABLE = PERI

Area under the ROC curve = 0.983

E : ROUND HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
>=5.22	100	0
> 5.48 *	**100**	**97.8**
> 8.79	21.5	97.8
> 11.98	0	100



Legend:

Figs. 3.4.3: These are ROC-curves of sperm acrosome for perimeter showing cut-off points of

A) Double headed sperm acrosome
B) Donor headed sperm acrosome
C) Tapered headed sperm acrosome
D) Flame headed sperm acrosome
E) Round headed sperm acrosome

Table 3.4.4 b: Receiver Operating Characteristic curve data indicating best cut-off points (with asterisk in bold) with sensitivity and specificity level indicated for area of normal sperm acrosome (Positive group) (as measured with FIPS) compared to area of double, donor, tapered, flame, big and round sperm acrosomes (Negative groups) of human sperm. In this example , the best cut-off values of the following areas together with their sensitivity and specificity have been highlighted :double, donor, tapered, flame and round sperm acrosomes.

ROC CURVE

VARIABLE = AREA

Area under the ROC curve = 0.737

A : DOUBLE HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
>=1.19	100	0
> 2.73	75	53.3
> 2.74 *	**74.3**	**73.3**
> 3.37	43.4	73.3
> 3.79	28.8	93.3
> 6.66	0	100

ROC CURVE

VARIABLE = AREA

Area under the ROC curve = 0.941

B : DONOR HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
< 2.06	0	100
<=4.08	85.4	94.7
<=4.17 *	**85.8**	**94.7**
<=4.19	85.8	93.1
<=4.2	86.1	91.5
<=5.74	100	3.2
<=6.4	100	0

ROC CURVE

VARIABLE = AREA

Area under the ROC curve = 1.000

C : TAPERED HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
>=0.48	100	0
> 2.12	98.7	99.3
> 2.2 *	**98.2**	**100**
> 5.74	0	100

ROC CURVE

VARIABLE = AREA

Area under the ROC curve = 0.869

D : FLAME HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
>=0.5	100	0
> 2.58	83.8	69.4
> 2.59 *	**83.6**	**72.2**
> 3.23	49.3	100
> 5.74	0	100

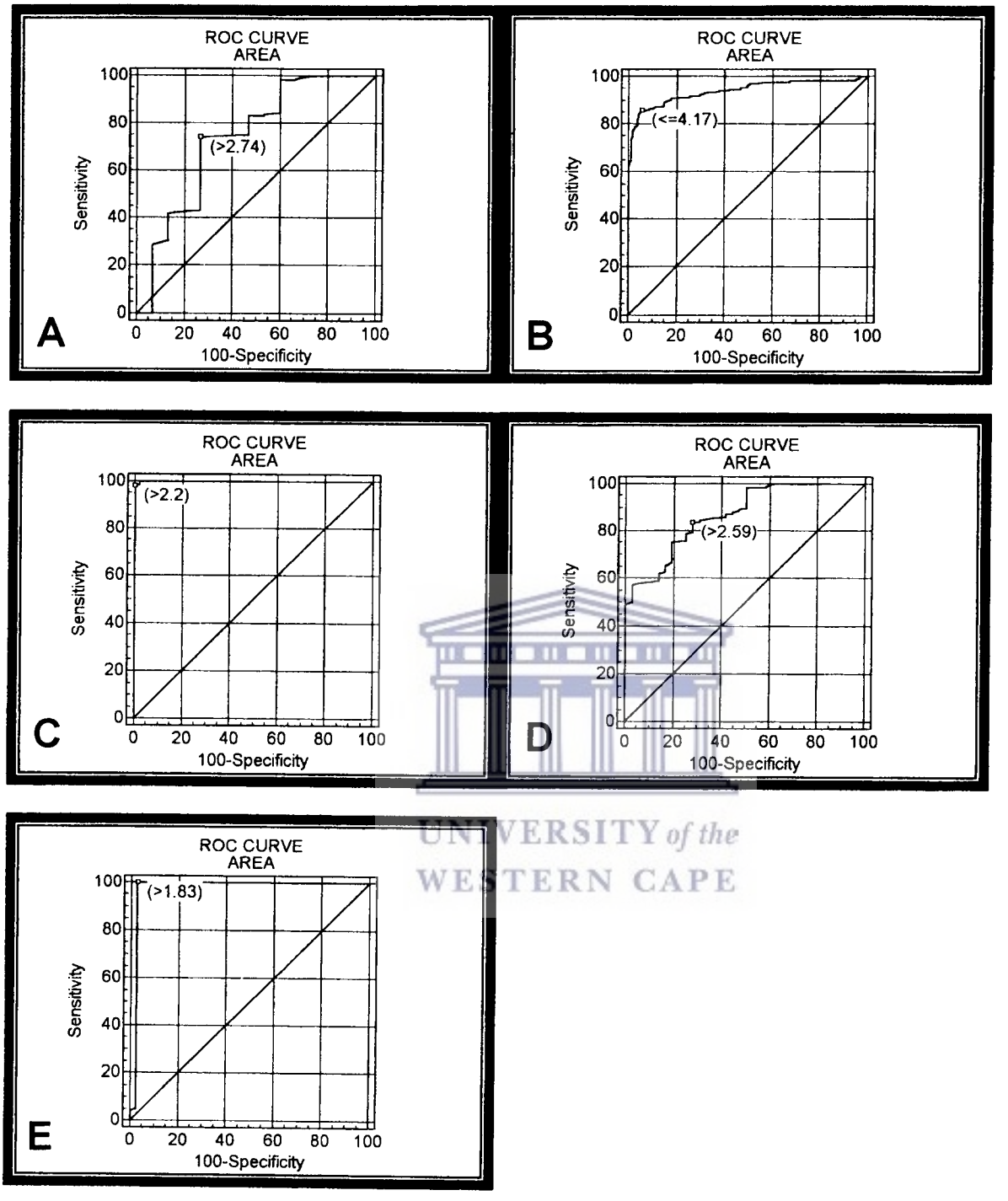
ROC CURVE

VARIABLE = AREA

Area under the ROC curve = 0.979

E : ROUND HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
>=0.46	100	0
> 1.83 *	**100**	**97.8**
> 4.75	5.3	97.8
> 5.74	0	100



Legend:

Figs. 3.4.4: These are ROC-curves of sperm acrosome for area showing cut-off points of

A) Double headed sperm acrosome
B) Donor headed sperm acrosome
C) Tapered headed sperm acrosome
D) Flame headed sperm acrosome
E) Round headed sperm acrosome

Table 3.4.4 c: Receiver Operating Characteristic curve data indicating best cut-off point (with asterisk in bold) with sensitivity and specificity level indicated for maximum ferret of normal sperm acrosome (Positive group)(as measured with FIPS) compared to maximum ferret of donor, small, tapered, flame, big and round sperm acrosomes (Negative groups) of human sperm. In this example , the best cut-off values of the following areas together with their sensitivity and specificity have been highlighted : small, tapered, flame, big and round sperm acrosomes.

ROC CURVE

VARIABLE = MAXFER

Area under the ROC curve = 0.833

A : SMALL HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
>=2.13	100	0
> 2.58	81	68.4
> 2.66 *	**75.9**	**78.9**
> 2.86	51.3	100
> 3.95	0	100

ROC CURVE

VARIABLE = MAXFER

Area under the ROC curve = 0.983

B. TAPERED HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
>=1.62	100	0
> 2.43	92.5	88.7
> 2.5 *	**92.5**	**97**
> 2.51	88.3	97
> 2.74	64.2	99.7
> 2.75	64.2	100
> 3.95	0	100

ROC CURVE

VARIABLE = MAXFER

Area under the ROC curve = 0.663

C : BIG HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
< 1.96	0	100
<=3.17	78.8	54.7
<=3.8 *	**98.9**	**54.7**
<=3.92	99.8	53.1
<=3.95	100	53.1
<=7.41	100	0

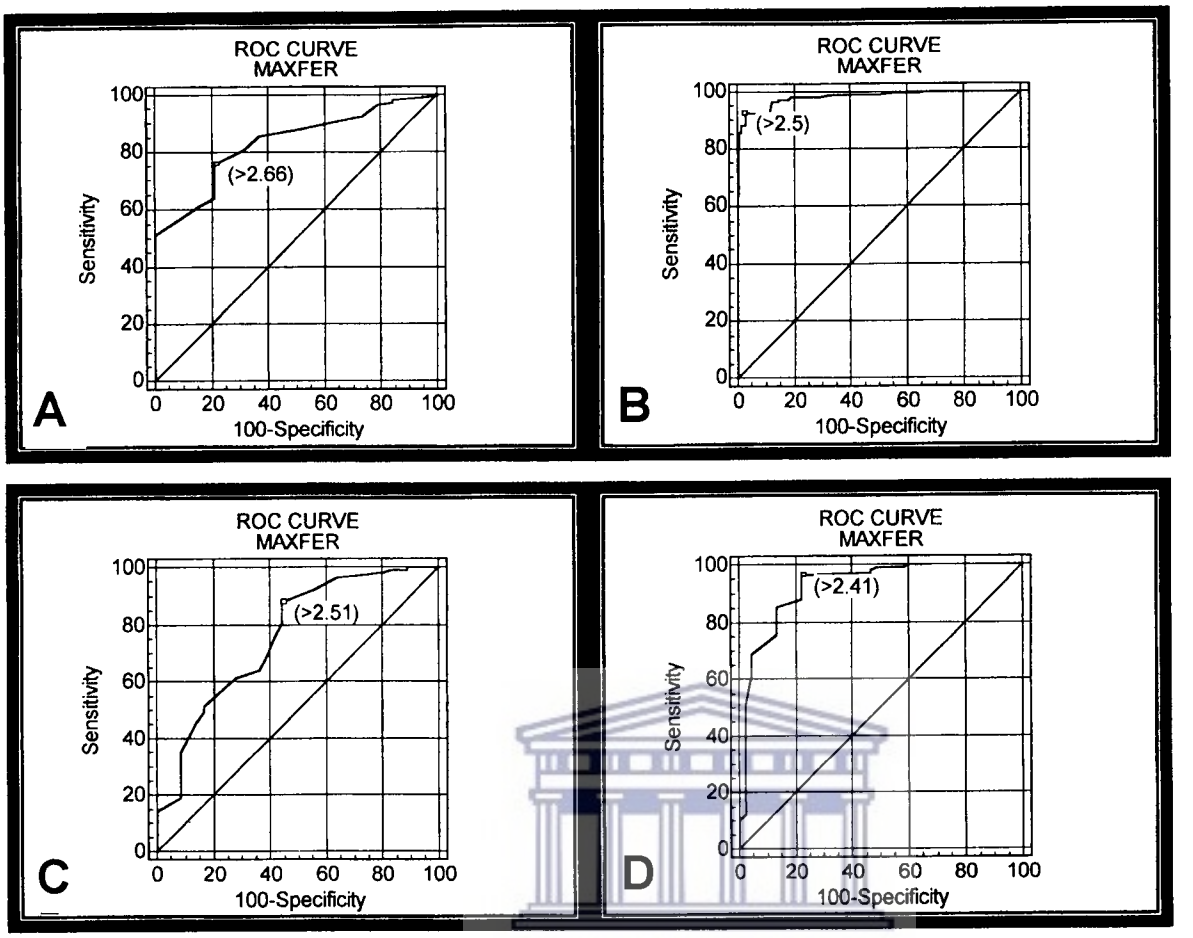
ROC CURVE

VARIABLE = MAXFER

Area under the ROC curve = 0.926

D : ROUND HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
>=1.36	100	0
> 2.11	100	40
> 2.13	99.8	40
> 2.36	97.3	53.3
> 2.41 *	**96.5**	**77.8**
> 2.51	88.3	77.8
> 2.56	85.6	86.7
> 3.47	10.4	100
> 3.95	0	100



Legend:

Figs. 3.4.5: These are ROC-curve of sperm acrosome for maximum ferret showing cut-off points of

A) Small headed sperm acrosome

B) Tapered headed sperm acrosome

C) Flame headed sperm acrosome

D) Round headed sperm acrosome

Table 3.4.4 d: Receiver Operating Characteristic curve data indicating best cut-off point (with asterisk in bold) with sensitivity and specificity level indicated for average ferret of normal sperm acrosome (Positive group)(as measured with FIPS) compared to the average ferret of proven fathers, small, tapered, flame and big sperm acrosomes (Negative groups) of human sperm. In this example , the best cut-off values of the following areas together with their sensitivity and specificity have been highlighted : tapered, flame and big sperm acrosomes.

ROC CURVE

VARIABLE = AVGFER

Area under the ROC curve = 1.000

B: TAPERED HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
>=0.092	100	0
> 0.902 *	**100**	**100**
> 3.16	0	100

ROC CURVE

VARIABLE = AVGFER

Area under the ROC curve = 1.000

B : FLAME HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
>=1.044	100	0
> 1.926 *	**100**	**100**
> 3.16	0	100

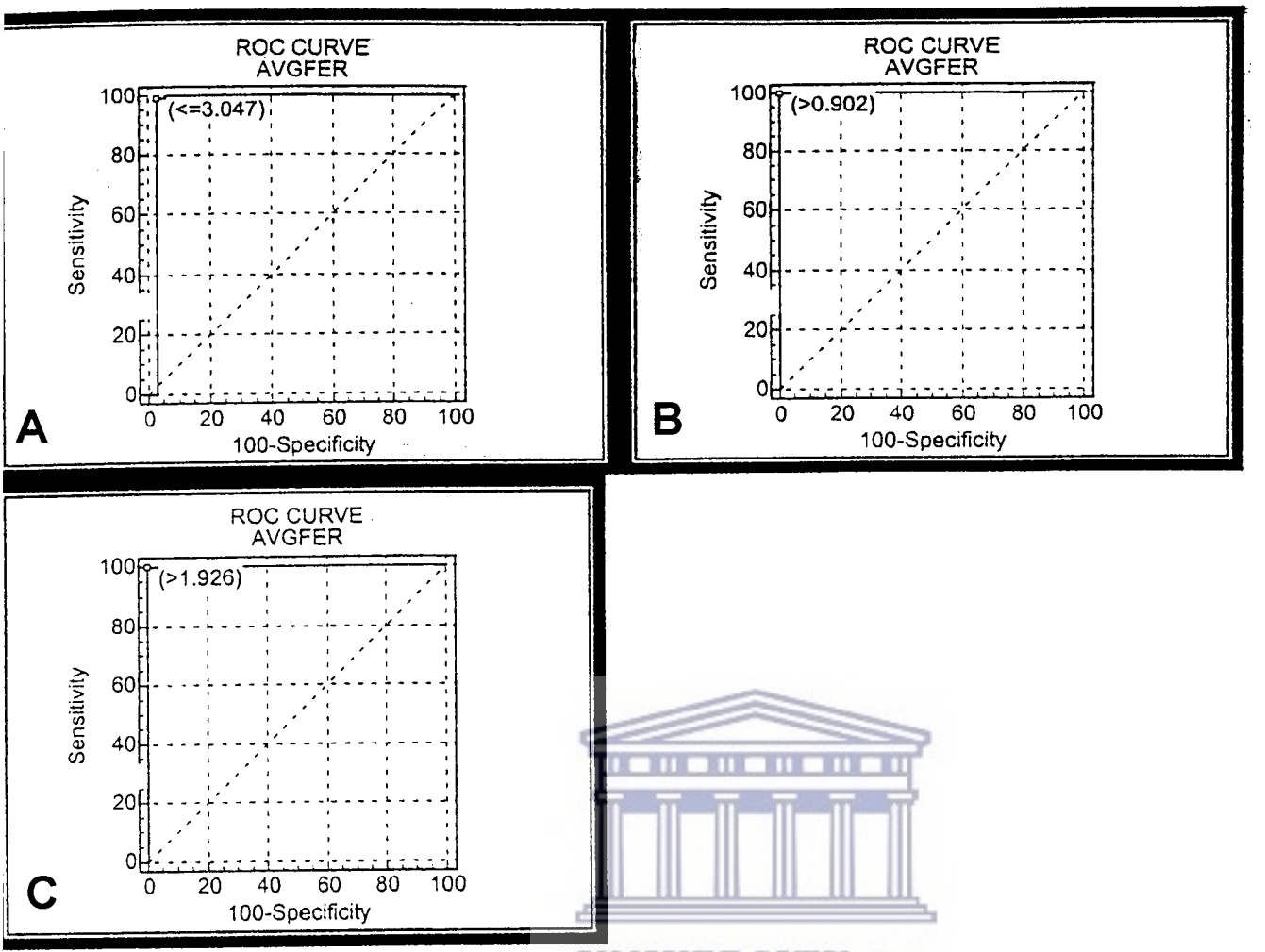
ROC CURVE

VARIABLE = AVGFER

Area under the ROC curve = 0.972

C : BIG HEADED SPERM ACROSOME

Criterion	Sensitivity	Specificity
< 1.504	0	100
<=1.504	0	97.2
<=3.047 *	**99.1**	**97.2**
<=3.092	99.1	94.4
<=3.16	100	94.4
<=4.774	100	0



Legend:

Figs. 3.4.6: These are ROC-curves of sperm acrosome for average ferret showing cut-off points of

A) Tapered headed sperm acrosome

B) Flame headed sperm acrosome

C) Big headed sperm acrosome

Table 3.4.5: The ROC-curve analysis which shows the cut-off values of the nine different sperm acrosomes compared to the normal sperm acrosome of humans: Sensitivity and Specificity levels of at least 70% are indicated in bold to emphasize best cut-off points for different parameters.

		SPERM ACROSOME PARAMETER						
SPERM ACROSOME TYPE		Peri	Area	Minfer	Maxfer	Avefer	Aspect	Shape
AH	Cut-off Value	>6.63	>1.82	<=1.68	>2.66	>2.03	<=0.58	>0.52
	Area Under ROC-curve	0.50	0.50	0.57	0.55	0.52	0.55	0.50
	Sensitivity	97.3	100.0	42.9	75.9	95.8	54.2	90.9
	Specificity	9.7	7.3	37.0	35.0	12.3	54.0	12.3
DH	Cut-off Value	<=8.93	>2.74	>1.55	<=3.32	<=2.59	>0.45	>0.49
	Area Under ROC-curve	0.79	0.74	0.78	0.63	0.52	0.77	0.97
	Sensitivity	**81.2**	**74.3**	**74.3**	85.6	81.0	100.0	**100.0**
	Specificity	**80.0**	**73.3**	**73.3**	53.3	33.3	53.3	**80.0**
DO	Cut-off Value	<=8.79	<=4.71	<=1.97	<=3.02	<=3.61	<=0.58	<=0.61
	Area Under ROC-curve	0.91	0.94	0.97	0.84	0.97	0.73	0.62
	Sensitivity	**78.5**	**85.8**	**81.9**	64.8	**100.0**	54.2	40.7
	Specificity	**100.0**	**94.7**	**100.0**	97.3	**81.0**	88.8	84.0
PF	Cut-off Value	>6.98	>2.60	>1.44	>2.66	>2.06	>0.51	>0.65
	Area Under ROC-curve	0.60	0.55	0.59	0.58	0.62	0.53	0.59
	Sensitivity	86.1	83.4	85.2	75.9	91.4	85.8	45.1
	Specificity	61.7	45.0	33.3	51.7	46.7	26.7	73.7

SH	Cut-off Value	>7.04	>3.08	>1.67	>2.66	>2.23	<=0.62	<=0.59
	Area Under ROC-curve	0.89	0.83	0.75	0.55	0.88	0.55	0.61
	Sensitivity	**83.6**	55.8	65.3	75.9	65.9	67.7	31.2
	Specificity	**100.0**	100.0	73.7	35.0	100.0	52.6	94.7
TH	Cut-off Value	>4.05	>2.20	<=2.12	>2.50	>0.90	>0.49	>0.50
	Area Under ROC-curve	1.0	1.0	0.99	0.98	1.0	0.49	0.73
	Sensitivity	**100.0**	**98.2**	**100.0**	**92.5**	**100.0**	100.0	97.8
	Specificity	**100.0**	**100.0**	**95.0**	**97.0**	**100.0**	2.3	48.3
FH	Cut-off Value	>7.11	>2.59	>1.67	>2.51	>1.93	>0.46	>0.55
	Area Under ROC-curve	0.81	0.87	0.82	0.76	1.0	0.55	0.69
	Sensitivity	**81.2**	**83.6**	65.3	88.3	**100.0**	100.0	83.2
	Specificity	**88.9**	**72.2**	80.6	55.6	**100.0**	27.8	52.8
BH	Cut-off Value	<=11.23	<=5.67	<=2.12	<=3.80	<=3.05	>0.49	>0.49
	Area Under ROC-curve	0.60	0.59	0.57	0.66	0.97	0.67	0.58
	Sensitivity	99.8	99.8		98.9	**99.1**	100.0	100.0
	Specificity	54.7	54.7	46.9	54.7	**97.2**	51.6	26.6
RH	Cut-off Value	>5.48	>1.83	>1.36	>2.41	>2.16	>0.49	>0.49
	Area Under ROC-curve	0.98	0.98	0.97	0.93	0.88	0.95	0.66
	Sensitivity	**100.0**	**100.0**	**96.0**	**96.5**	**78.0**	100.0	100.0
	Specificity	**97.8**	**97.8**	**91.1**	**77.8**	**78.5**	57.8	44.4

Table 3.4.6(a): Sperm acrosome which clearly fall outside and inside the range for each parameter with a Sensitivity **or** Specificity of at least 70% but usually higher. DO and PF are excluded as good discriminants.

PARAMETER	SELECTED RANGE FOR NORMAL SPERM ACROSOME AS DETERMINED BY ROC	SPERM ACROSOME THAT FALL OUTSIDE THE RANGE (GOOD DISCRIMINANT)	SPERM ACROSOME THAT FALL WITHIN THE RANGE (POOR OR NO DISCRIMINANT)
PERI	4.05 - 8.79	DH and BH	AH, DO, SH, PF, TH, FH and RH
AREA	1.82 - 2.74	SH and BH	AH, DH, PF, TH, FH and RH
MAXFER	2.41 - 2.66	DH and BH	AH, PF, SH, TH, RH and RH
MINFER	1.55 - 1.67	AH, TH, BH and RH	DH, SH and FH
AVEFER	2.03 - 3.00	TH, FH and BH	AH, DH, DO, PF, SH and RH
ASPECT	0.52 - 0.58	DH, TH, FH, BH and RH	PF, SH and FH
SHAPE	0.49 - 0.52	SH and FH,	AH, DH, TH, BH and RH

Table 3.4.6(b): Sperm acrosome which clearly fall outside and inside the range for each parameter with a Sensitivity **and** Specificity of at least 70% but usually higher. DO and PF are excluded as good discriminants.

PARAMETER	SELECTED RANGE FOR SPERM ACROSOME AS DETERMINED BY ROC	SPERM ACROSOME THAT FALL OUTSIDE THE RANGE (GOOD DISCRIMINANT)	SPERM ACROSOME THAT FALL WITHIN THE RANGE (POOR OR NO DISCRIMINANT)
PERI	4.05 and > 8.79	DH and RH	DO, SH, TH and FH
AREA	1.82 - 2.74	DO and RH	DH, TH and FH
MAXFER	2.41 - 2.66	TH	RH
MINFER	1.55 - 1.67	DO, TH and RH	DH
AVEFER	2.03 - 3.00	DO, TH and BH	FH and RH
ASPECT	0.52 - 0.58	-	-
SHAPE	0.49 - 0.52	-	DH

CHAPTER 4

DISCUSSION

4.1 Discussion on different microscopic techniques

4.1.1 General

In this study different microscopic techniques have been used to study and compare human sperm morphology. Bright field microscopy of Papanicolaou (PAP)stained smears, Normaski differential interference microscopy, Scanning electron microscopy and Laser Scanning Confocal Microscopy were employed.

4.1.2 Light Microscopy

The findings of the present study have potentially important implications for our pathophysiology understanding of the human sperm. A number of studies have shown that there are significant correlations between male fertility and sperm morphology (Aitken *et al.* 1983, 1988, Albertson *et al.*).

A positive correlation has also been reported between the incidence of normal sperm morphology in semen and the ability of the spermatozoa to fertilize human oocytes *in vitro* (Jeyendran *et al.*,1986). However, it is not difficult to envision how morphologic abnormalities of the sperm head organelles could lead to fertilization failure.

The Papanicolaou stained smears of sperm morphology has been discussed widely in the past, and different criteria exist throughout the world to asses sperm morphology using PAP smears (Glover et al., 1990). In this study, Papanicolaou staining was used because the procedure was found to be affordable, accessible and fast although the preparative

artifacts occurring with it are unavoidable. The fine details of the staining and/or fixation technique can influence the fine details of sperm morphology (Harasymowzcx *et al.* 1976). This was also confirmed by Meschede *et al.* (1993), who showed the influence of three different preparatory techniques on the results of human sperm morphology. These authors showed that when Papanicolaou (Pap), Shorr stain (Sho) and Wet preparatory methods were used artifacts were evident and this distorted results. Davis *et al.* (1993) demonstrated that there are significant differences in the percentage of sperm currently digitized when different cell staining techniques were used. They found the PAP method the worst approach and the new GZIN stain was the best. With the exception of the CellSoft measurement the tabulated sperm head dimensions show agreement between systems, irrespective of the staining procedure used. CellSoft dimensions of Wang *et al.* (1991b) are larger than those of other investigators, but those of other authors are comparable. However, they emphasized that Papanicolaou staining is still the best.

Enginsu *et al.* (1991) emphasized that slides stained with PAP and Diff-Quik have almost the same quality, enabling the morphological status of spermatozoa to be evaluated in detail. However, the latter has advantages in having a clear background of the stained smears. They advocated that Diff-Quik staining is a time-saving procedure compared to other staining techniques. It makes use of quality-controlled commercially prepared reagents and is easy to handle (Kruger *et al.*, 1987 b). Nevertheless, even though PAP is better according to this study the background is blurred and at times the smears are understained or overstained so that visualizing them is difficult. Although, there are problems encountered, using the PAP staining for morphometric analysis it appears to be the best routine method. Katz *et al.* (1986), in their laboratory used the Papanicolaou

staining technique for human spermatozoa because of its optical advantages, and because it is commonly used by other laboratories in evaluating human semen. This was also echoed by Meschede *et al.* (1993) who indicated that for practical purposes the PAP method seems the most suited, due to its widespread use and the fact that the current criteria for normal morphology have been elaborated based on this method (Eliasson, 1977).

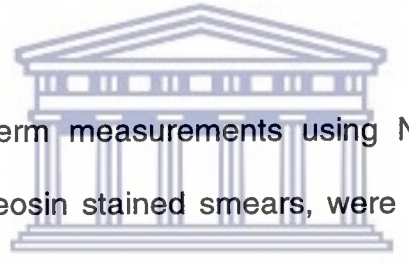
Garret *et al.* (1995), stated that irrespective of the staining procedure used the results of the morphometric parameters agree between systems. Parameters dependent on the stain intensity have been demonstrated to be important features for the discrimination of some abnormal categories (Perez-Sanchez *et al.* 1994) who further added that those parameters can only be assessed accurately using image analysis techniques, because differences in stain uptake are more difficult to appreciate by visual assessment than are differences in size or shape.

Katz *et al.* (1986) discovered that most laboratories use smears in which the cells were dried and stained and such procedures dehydrate the sperm and shrinkage is likely to occur. There was usually a reduction in the dimensions of human sperm when assessed with Papanicolaou. Meschede *et al.* (1993) acknowledged that owing to fixation of the specimens, the midpiece on PAP and Shorr slides were shrunken and on average more slender. The effects induced by fixation and staining has been described previously for sperm head by Katz *et al.* (1986). This was emphasized by Wang *et al.* (1991b) that Shorr stain caused less shrinkage of the spermatozoa compared with the Papanicolaou procedure. The results of the present study show that all abnormal sperm forms can be

clearly recognized when using Bright field microscopy in conjunction with Papanicolaou stained smears.

4.1.3 Nomarski Differential Interference Microscopy (NDIM)

During the last decade approximately150 articles appeared on various aspects of sperm morphology as visualized by NDIM viz., Althouse *et al.*(1995) on boar sperm; van der Horst *et al.* (1991) on ferret sperm. However, none of these refer to a detailed analysis of normal/abnormal human sperm. This is surprising in view of the fact that most laboratories performing intracytoplasmic sperm injection (ICSI) use NDIM or equivalent such as Hoffman optics.



The only in depth comparison of sperm measurements using NDIM and bright field microscopy of unstained and nigrosin/eosin stained smears, were performed by van der Horst *et al.* (1991) on sperm of three ferret species. In the latter investigation measurements between the two techniques showed a high level of concordance. However, NDIM was superior to Bright field microscopy in revealing details of the acrosome. These authors could furthermore perform a detailed morphometric study on ferret sperm by using NDIM in conjunction with the Quantimet image analysis system. In the current investigation normal and all abnormal sperm forms could be visualized with great clarity by using NDIM optics. The 3D nature of NDIM sperm images furthermore assists in assessing sperm form. It is important that in future studies NDIM of human sperm morphology be compared in a controlled study with current routine methods such as PAP particularly in relation to the percentage normal sperm.

NDIM furthermore has the advantage that no staining is required and sperm in seminal plasma or glutaraldehyde fixed sperm can be visualized instantaneously after collection of

a semen sample.

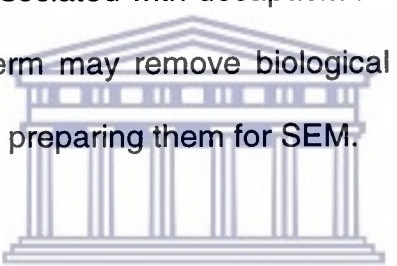
The major disadvantage of the NDIM technique is that protein coatings of human seminal plasma sometimes obscure sperm surface details. This was particularly evident for the acrosome and is in contrast with the findings of van der Horst *et al.* (1991) on ferret sperm. However, this can be overcome by washing the sperm in a suitable buffer.

4.1.4 Scanning Electron Microscopy (SEM)

Bright field microscopy and SEM of sperm have been described by Domagala *et al.* (1982), Mikel *et al.* (1980) and van Horst *et al.* (1991). Van der Horst *et al.* (1989) developed a new technique for SEM of sperm and employed Scanning electron microscopy to study human sperm morphology. In the analysis of semen this should be more useful than transmission electron microscopy, which is too time consuming and expensive a technique to be of general use in the clinical laboratory.

Few spermatologists have compared sperm morphology using Scanning electron microscopy with sperm measurements using Bright field microscopy. Van der Horst *et al.* (1991) compared SEM and Bright field microscopy measurements of sperm of three ferret species. They concluded that measurements such as length and width of the sperm head was on an average 25% smaller when using SEM than when using Bright field microscopy. This dramatic difference between Bright field microscopy and SEM in other cells than sperm have also been reported by other investigators according to van der Horst *et al.* (1991). The predictive value of sperm morphology by both methods needs to be analysed for human sperm.

In this investigation sperm studied by means of SEM can be seen with clarity but protein-like material often cover the posterior part of the sperm heads. After processing some protein-like material remain attached to the sperm and obscures surface detail. Van der Horst *et al.* (1989), demonstrated that the material could be removed by prolonged fixation with osmium tetroxide. While sperm morphology can be better visualized after etching with osmium tetroxide this protein-like material may be an important physiological-structure component of the sperm surface. The latter authors further emphasized that the protein coat may represent an inhibitory coat associated with decapacitation and fertilization. Dott (1969) pointed out that washing of sperm may remove biological substances inherently associated with the sperm surface when preparing them for SEM.



Although scanning electron microscopy gives good discrimination between normal and abnormal sperm, it remains a time-consuming and expensive procedure and cannot be recommended for routine laboratory purposes. However, most of the sperm forms that are seen with Papanicolaou staining can be readily recognized using Scanning electron microscopy. Therefore, on a routine basis SEM would not be an alternative for Bright field microscopy of Papanicolaou stained sperm smears as it does not add more information on the identification of normal/abnormal sperm and is too time consuming.

4.1.5 Laser Scanning Confocal Microscopy (LSCM)

Since the last five years about 60 research articles appeared on various aspects of sperm structure and function by making use of LSCM. Most of these studies deal with animal sperm and none with human sperm morphology in terms of normal/abnormal forms.

Many investigators such as Wang *et al.*(1991b) and George *et al.* (1996) have used Confocal Laser Scanning Microscopy to study aspects of human oocytes but not spermatozoa.

The current study represents the first application of LSCM to evaluate normal/abnormal human sperm morphology. In this study the LSCM observations of spermatozoa were from infertile/fertile men. Although the laser scanning confocal microscope is a high cost instrument, its use in and research in the field of andrology is important since the 3D structure can be elucidated with greater clarity than most other techniques such as SEM. One of the greatest advantages of this technique is that laser optical sections can be made and used to construct 3D images. These 3D images can be rotated by using specific software. It is therefore possible to view sperm from theoretically all angles. Laser Scanning Confocal Microscopy in this investigation was useful in showing that only a few degrees rotation may change the classification of a sperm from normal to abnormal. Because of the semi-rounded nature of the head of the human sperm it may well be possible that spermatozoa may attach to a glass surface (when making sperm smears) in various positions. It is particularly normal and borderline amorphous which may present problems in this context and lead to over or under estimation of the percentage normal sperm. On the basis of this information a quantitative study of normal/abnormal sperm by means of LSCM should be performed in future studies. At the time that this research was performed, the software for quantitative measurements was not in place. Currently, software is available which will make detailed measurements by means of LSCM possible and it is parameters such as sperm head volume that may assist greatly to distinguish among the different sperm forms.

4.2 Comparison of technician assessment of sperm morphology on the basis of Tygerberg Strict Criteria and WHO

Numerous studies have reported significant variation in the percentage normal human sperm when specimens were analyzed by different technicians and laboratories (Zaini *et al.* 1995; Baker *et al.* 1987; Dunphy *et al.* 1989; Menkveld *et al* , 1990; Menkveld *et al.* 1991). This was confirmed in the present study. Three technicians from three different laboratories scored 43% of patients in different fertility classes when Tygerberg Strict Criteria was used (the same sperm smears from the same 77 patients were evaluated by all three technicians).

Several factors are responsible for this technical variation, including differences in the methods used to prepare and stain specimens (Davis *et al.* 1993; Katz *et al.* 1986) and differences in proficiency among technicians (Zaini *et al.* 1995; Baker *et al.* 1987; Dunphy *et al.* 1987; Menkveld *et al.* 1991).

The variation in the percentage of normal sperm reported in this study is similar to the variation reported in other studies where multiple microscopic slides of the same specimens were sent to several laboratories for analysis (Baker *et al.*, 1987, Dunphy *et al.*, 1989). The present study also demonstrated that there was a high correlation between two technicians i.e GSH/UWC and this was highly significant but the correlations between UWC/PCH and GSH/PCH were lower than for UWC/GSH. These results show that while two technicians may score within relatively narrow limits a third technician may score differently and indicate that it will be difficult to accept the accuracy of scoring the percentage normal sperm on an international basis.

The main difference between studies of morphology evaluation is in the application of suitable criteria for evaluation. Enginsu *et al.* (1991), indicated that when morphology was evaluated according to the guidelines of WHO, there was a correlation between normal morphology and fertilization *in vitro* (r =0.282), though the correlation was worse than using Strict Criteria (r =0.555, $P< 0.001$). The mean percentage of spermatozoa with normal morphology are < 50% suggested by the WHO (1987). Liu et al., (1988, 1989 a, b) found a good correlation between the percentage of normal morphology and fertilization *in vitro.* Their studies gave lower percentages than was suggested by the WHO. These findings and the results of this study indicate disagreement in the percentages of normal morphology because it appears that a 2% TSC score coincides with a 6% WHO score, a 4% TSC score coincides with a 22% WHO score and a 5% TSC score coincides with a 30% WHO score. These results are in agreement with Neuwinger *et al.* (1990) who illustrated that the interlaboratory variability of the results of semen analysis is not without consequence for the infertility patients. Based on the semen sample, a patient may be classified normal by WHO and subfertile/infertile by a TSC laboratory. Whereas the patient be accepted for artificial insemination or in an *in vitro* program by WHO, could be excluded by TSC. Furthermore, a 3 to 5% TSC range coincides with a 6 to 30% range for WHO. It therefore appears that the 3 to 5% TSC range possibly reflects a wide range of differential fertilities when compared to WHO and it has been shown that the accuracy of scoring among three technicians in different laboratories on a consistent basis exceeds 2%. On this basis the TSC system can not be considered sufficiently reliable to be accepted internationally to categorize patients in different fertility classes.

There are many clinicians who believe that standard sperm morphological criteria as

determined by the WHO may be one of the most reliable indices of male fertility potential (Check *et al.* 1992).

Better agreement of the results obtained by different laboratories can only be achieved by further standardization of the methods as well as by strict internal and external quality control. It is also important to appreciate that national standards for minimum levels of technologist training and proficiency need to be established, along with a national program for andrology and *in vitro* laboratory accreditation. Only when such measures are widely adopted will the accuracy and precision of andrology assays approach the standards of excellence that are now expected clinically. However, without a reliable and objective quantitative and automated method, this may never be achieved.

4.3 Quantitative Image Analysis of the human sperm head

Image analysis provides the key to quantification in the assessment of sperm morphology. Digital image processing enables quantification and, together with computer based automation, has the potential to eliminate the biases, subjectivity, and consequent lack of intra-technician and inter-technician reproducibility inherent in conventional manual assessment (Garrett *et al.*, 1995). Despite deficiencies in the subjective visual assessment, many groups have reported on the correlation between the percentage of sperm with normal morphology and fertility (Liu *et al.* 1989; Mahadevan *et al.* 1984), thus providing additional incentive for the development of an automated objective approach to morphology assessment.

Perez-Sanchez *et al.* (1994) have mentioned that the assignment of some of the abnormal

spermatozoa to a distinct category based on visual assessment is extremely difficult, and in those cases morphometrics is a useful tool. These authors demonstrated clearly that in some studies a set of four to six morphometric parameters were estimated for the whole sperm population without distinguishing between normal and abnormal forms (Jagoe *et al.* 1986; Katz *et al.* 1986; Turner *et al.*1989).

In the present study quantitative sperm morphology was performed on the donor/ presumably fertile/subfertile patients attending the Groote-Schuur Hospital as well as on proven fathers from Pretoria Central Hospital. Table 3.3.2 demonstrates population averages and standard deviations of the ten sperm head forms as indicated previously. Garrett *et al.* (1995) published a paper indicating all authors that have researched sperm morphometric parameter measurements by means of either fully automated or semi-automated image analysis systems. All of those are in agreement in terms of normal sperm morphometrics with the current study except Wang et al. (1991) and Perez-Sanchez *et al.* (1994) who show considerably larger sperm measurements (Table 4.1). In the present study results for most parameters of normal sperm as indicated in Table 3.3.2 and Table 4.1 fall within a narrow range when compared to those of Garrett *et al.* (1995), Katz *et al.* (1986), Jeyendran *et al.* (1986), Schrader *et al.* (1990) and Davis *et al.* (1992) for normal sperm. Most of the sperm morphometric results of the current study differ by less than 5% from those of the latter authors above.

Table 4.1 furthermore elucidates important aspects relating to normal sperm morphometrics when sperm were stained by different methods. The sperm head length of patients as determined by Garrett *et al.* (1995) and in the current study is almost identical (4.43µm and 4.42µm respectively) despite the fact that the former used Shorr stain and

the current study used the PAP stain. This is in agreement with Garrett *et al.* (1995) who stated that morphometric measurements of sperm agree between systems irrespective of the staining procedure employed. The measurements of sperm head width, sperm head perimeter and sperm surface area in the current investigation are also within an acceptable range when compared to the investigators listed in Table 4.1. However, the average value for sperm head width (Minfer) in the current investigation represents the lowest value when compared to all other investigators (Table 4.1). The measurement of this parameter using the FIPS system is based on the smallest distance from the centre of the sperm head (centroid) to the sperm head periphery as determined by 360 measurements (each time at a different angle from the centroid and is referred to as minimum ferret). It therefore represents a very accurate means of determining minimum width. It is possible that the width meaurement as determined by the authors above is based on maximum width and may account for this discrepancy. Furthermore, the current study made use of a combination of semi-automatic and fully automated procedures whereas the other investigators listed in Table 4.1 used either fully automated or fully semi-automated procedures. In this study sperm head thresholding was performed in the semi-automated mode. While this represents a time consuming procedure it ensures quality control since the sperm head is accurately framed (thresholded) for measurement. The actual measurement of seven morphometric parameters were then performed in the fully automated mode.

The current study furthermore represents the first investigation comparing normal sperm morphometrics of patients attending a fertility clinic, presumably fertile donors and proven fathers. Although sperm morphometrics of patients and donors were not significantly

different for most sperm head parameters, proven fathers differed from patients for several parameters (See Table 3.2.2). In the current study the sperm of proven fathers were larger than those of patients and donors (Table 4.1). Some investigators listed in Table 4.1 indicated no differences in most sperm morphometric measurements when patients were compared to donors. In the current investigation normal sperm heads from donors and normal sperm heads from patients were for example not significantly different in length (Table 3.3.2). However, none of the other investigators listed in Table 4.1 studied sperm representing proven fathers only and may account for the discrepancy in the comparative results. Frequency distribution histograms for head length have been constructed for normal sperm heads of patients, proven fathers and donors (Fig. 4.1). Fig. 4.1 emphasizes that the main frequency of sperm head length of proven fathers are shifted to the right (4.5 - 5.5µm) when compared to patients (4.2 - 4.6µm) and donors (4.0 - 5.0 µm). Furthermore, the two main frequencies for normal sperm heads of patients are 4.2 to 4.4 µm and 4.8 to 5.0 µm. This suggests that patients have two main populations of normal sperm heads of which one is "shorter" and the other is "longer". The population of longer normal sperm heads of patients may be equivalent to normal sperm heads of proven fathers who seem to have a more constant sperm length. The results of the current investigation on sperm heads of proven fathers are unique and has important implications in clinical spermatology. With a larger sample size of proven fathers, the head dimensions of their normal sperm could be used in future studies as a standard for normal sperm.

There is less agreement between Davis *et al.* (1992), Perez-Sanchez *et al.* (1994), and the author's results on morphometric measurements of abnormal sperm. However, the measurements for normal sperm heads indicated by Perez-Sanchez *et al.* (1995) fall

outside the range and are much larger than those of most other investigators (Table 4.1). It appears that most of their measurements for abnormal sperm heads are also larger than those reported in the current study. Davis *et al.* (1992) only give measurements for tapered and amorphous sperm head forms and based their findings on only three patients per category. Despite this, the latter authors' results and that of the current investigation differ by less than 5% for sperm head length (4.13 and 4.32μm respectively) and sperm head width (2.72 and 2.5μm) for amorphous sperm on an average basis. Tapered forms in the two studies differed greatly.

Some investigators showed that four common parameters (length, width, area and perimeter) were sufficient to differentiate normal from abnormal sperm. Although as many as seven morphometric parameters were used in this study to differentiate normal and abnormal sperm head dimensions not all are good discriminators. In the current study sperm head length (Maxfer) showed the best level of discrimination shape factor showed the poorest levels of discrimination.

ROC-curve analysis showing cut-off points among the nine sperm head forms compared to normal sperm heads were indicated in Table 3.3.4. Unfortunately no other research papers have used this novel method to distinguish among different sperm forms with a high level of sensitivity and specificity. The data presented here provides a good quantitative basis for future automatic objective analysis of the percentage normal sperm of a patient. Table 3.3.7a and b furthermore show that a combination of cut-off points for the different morphometric parameters will provide the best discrimination between normal and each abnormal sperm head form and allow inferences with fertility/infertility and IVF data. It was unfortunately not possible to discriminate between normal sperm heads (patients) and

amorphous sperm heads with a high level of sensitivity and specificity (above 70%) but only when lower levels of either sensitivity or specificity were employed. However, when ROC-curve analysis was performed on normal sperm heads of proven fathers and amorphous sperm heads, cut-off points with a high level of specificity and sensitivity were obtained for sperm head length and Avefer (average of maximum length and minimum width) (Fig. 4.2). This emphasize the need for a larger sample size (proven fathers) in order to use the normal sperm of proven fathers as a standard for comparison and ROC-curve analysis.

The current investigation appears to be the first and only study that quantified the human sperm acrosome by means of image analysis for normal and abnormal sperm in terms of seven morphometric parameters. It was interesting to note that sperm heads with the largest lengths also had the largest acrosome lengths and *vice versa*. Table 4.2 was constructed to indicate this relationship and to indicate which percentage cover of the total sperm length was contributed by the acrosome. Table 4.2 indicates that the range extends from 38% (acrosomes of tapered heads) to 75% (acrosomes of donor heads). Normal and amorphous sperm acrosomes covered 67% of the length of the sperm head. This quantitative information is in agreement with both WHO and TSC which appear to define a normal acrosome as one which covers approximately 40 - 70% of the human sperm head length. Surprisingly, the value for the acrosomes of proven fathers was only 57% of head length. However, the actual acrosome length of proven fathers and patients did not differ significantly.

ROC-curve analysis has also been performed for acrosomes and the cut-off points produced in this context can be used in conjunction with the sperm head cut-off points in future sperm morphometry studies and should provide a broader scientific basis to distinguish among normal and abnormal sperm forms.

Recent literature on computer-assisted assessment of human sperm morphology indicate both advantages and disadvantages in the clinical environment. Davis *et al.* (1992) indicated that the CFH automated sperm morphometry instrument exceeds the accuracy and precision of most manual approaches. However, Wang *et al.* (1991a) concluded that the results of the zona-free hamster egg penetration test could be predicted using manual assessment of sperm morphology and computer-assisted morphometric analysis did not add further information. Wang *et al.* (1991b) furtermore indicated that there was no advantage of the morphologizer that they used over the manual method in sperm morphology classification. These authors did agree that the morphometric parameters of spermatozoa have to be defined.

Steigerwald and Krause (1998) developed the latest sperm morphometric system using neuronal network software. Their system did not reveal significant differences of the percentage of normal forms when compared to the direct microscopical estimation. However, nearly all the classes of abnormal sperm heads were estimated by the two methods as being significantly different. They furthermore indicated that the morphometric system of analysis took twice as long as the manual method.

The literature cited above seems to agree that computer-aided methods of sperm

morphometry have the advantage of laboratory quality control but there does not currently appear to be major advantages above manual assessment. The reasons for this is not surprising. Firstly, the field of sperm morphometry is relatively young and most of the basic research has only been performed during the last decade. Secondly, there is very poor conformity in terms of the methodology of computer-aided methods and the algorithms employed. It is erroneous that there is currently probably less conformity in computer-aided morphometry than in the manual assessment of sperm morphology. Davis *et al.* (1992) indicated that with improvements in sperm recognition and type of classification algorithms, computer-aided methods could significantly improve the reliability of morphology assays in clinical and research laboratories. It should be emphasized that unless there can be agreement on an international "gold standard" in sperm morphometry measurements (sperm preparation, staining, equipment, cut-off points and algorithms), no system will provide a better option than current manual/direct microscopic assessment.

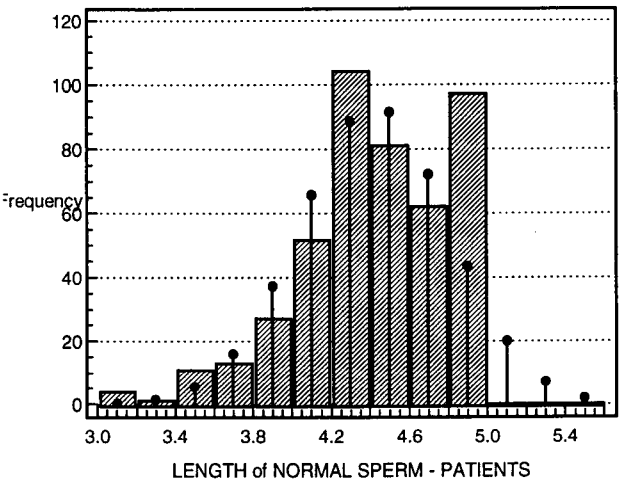
The current study has contributed much in the above context by providing detailed measurements of seven parameters of all the main sperm head types and their acrosomes in semen. In addition cut-off points with a high level of sensitivity and specificity have been defined by means of ROC-curve analysis for normal versus abnormal sperm head types and acrosomes. This approach differs to a large extent from previous studies in providing a fully quantitative standard. Most of the investigations cited made use of a method whereby the computer was "trained" to recognize typical examples of normal and abnormal sperm and therefore lack empirically derived quantitative cut-off points among the different sperm types. It was unfortunately out of the realm of this investigation to use the above

quantitative information in the context of an automatic morphometric system. However, the FIPS system used in this study with its library of macros could be employed in future studies as an automated sperm morphometry system.

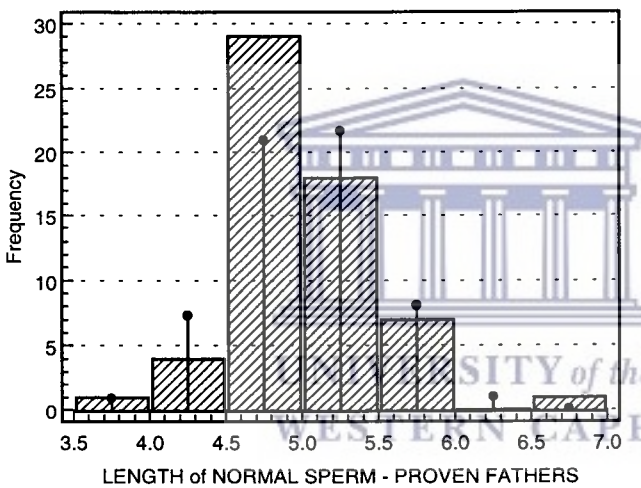
Table 4.1: Results for the current investigation and summary of the published data for sperm head morphometric parameters. The tabulated sperm head dimensions show agreement between systems, irrespective of the staining procedure used except for Wang *et al.* (1991) and Perez-Sanchez *et al.*(1994).

	This investigation (1998)	**This investigation (1998)**	**This investigation (1998)**	**Garrett et al. (1995)**	**Garrett et al. (1995)**	**Schrader et al. (1990)**	**Schmassmann et al. (1979)**	**Jeyendran et al. (1986)**	**Katz et al. (1986)**	**Wang et al. (1991)**	**Perez-Sanchez *et al.* (1994)**	**Davis et al. (1992)**
Stains	PAP	PAP	PAP	Shorr	Shorr	PAP*	HEM*	PAP	PAP	Shorr	Hema-color	HEM
Sample source	Donor	Proven fathers	Patient	Patient	Donor	Donor	Patient	Patient	Donor	Normal Sperm	Normal Sperm	Normal Sperm
(no.)	30	77	298	58	23	45	20	81	30	50	186	9
L(µm)	4.49	5.02	4.42	4.43	4.35	4.53	4.70	4.37	4.37	6.32	5.31	4.20
W(µm)	2.57	2.61	2.56	2.85	2.89	2.85	3.00	2.79	2.83	3.93	3.92	2.74
A(μm^2)	8.61	9.17	8.08	9.64	9.60	8.72	10.30	9.52	9.79	18.30	14.39	8.50
P(µm)	12.03	13.32	12.41	12.5	12.4	12.7	12.30	11.60	11.70	21.40	-	11.40

A



B



C

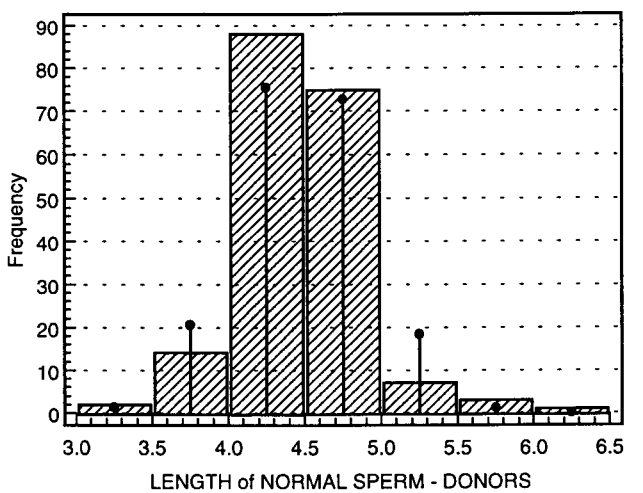


Fig. 4.1: Frequency distribution of the Length (Maxfer) of normal sperm from patients (A), from proven fathers (B), and from donors (C).

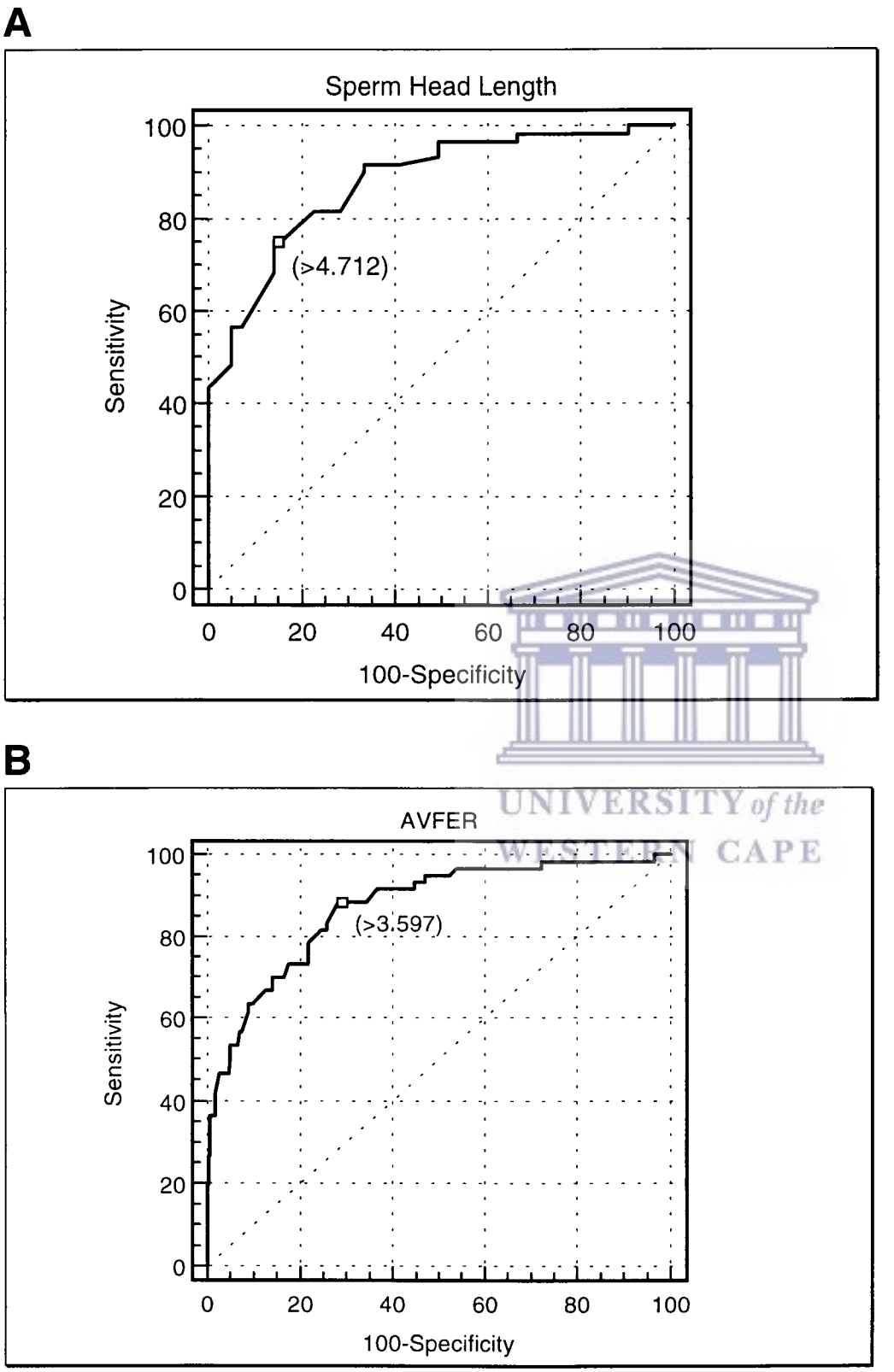


Fig. 4.2: ROC-curves showing cut-off points between sperm head length (A) andAvefer (B) for normal sperm of Proven fathers (positive group) and amorphous sperm of patients (negative group). In both A and B cut off points with high sensitivities and specificities are evident.

Table 4.2: The percentage contribution of the acrosome length to the sperm head length.

Sperm morphology types	Head length (:m)	Acrosome length (:m)	Percentage of acrosome length to sperm head length (%)
Amorphous	4.32	2.89	67
Double head	6.47	3.31	63
Donor (Normal)	4.49	3.41	75
Flame	3.21	2.18	67
Patient (Normal)	4.42	2.96	67
Proven father (Normal)	5.02	2.87	57
Round	5.25	2.58	49
Small	3.47	2.57	74
Tapered	5.55	2.1	38
Big	6.47	4.14	64

5. Summary

In the first part of this investigation, four techniques for microscopic evaluation of human sperm morphology, viz., bright field microscopy, Normaski Differential Interference Microscopy (NDIM), Scanning Electron Microscopy (SEM) and Laser Scanning Confocal Microscopy (LSCM) were compared. Bright field microscopy of Papanicolaou stained sperm smears appears to be an acceptable method to distinguish among the various sperm forms. However, Laser Scanning Confocal Microscopy allows a more accurate distinction between normal and abnormal sperm than any of the other three techniques used. Because sperm may adhere to the surface of a glass-slide in various positions, microscopic techniques that do not allow rotation of the image may produce misleading results. Unfortunately this latter technique is too expensive and tedious for routine morphological analysis.

Nomarski Differential Interference Microscopy (NDIM) shows great promise in the assessment of sperm morphology. A major advantage of NDIN is that live or glutaraldehyde-fixed, but unstained, sperm are visualized instantaneously. This technique is a good candidate for routine application in the clinical andrology laboratory.

In the second part of this investigation, the reliability/repeatability of scoring among three technicians from three different laboratories using the Tygerberg Strict Criteria (TSC) and bright field microscopy of PAP stained smears was tested. There was a better correlation for Groote-Schuur Hospital (GSH) and University of the Western Cape (UWC), but a

significantly lower correlation coefficient between GSH / (PCH) and UWC / PCH. A 3-5% TSC range of abnormal sperm corresponded to a 6-30% WHO range. The high variability of these results suggests that the TSC may not be useful as an international standard for sperm morphology assessment. It was surprising that 77 proven fathers had an average normal sperm morphology of less than 5% when assessed by means of Strict Criteria.

In the third part of the investigation normal and abnormal sperm dimensions were quantified by quantitative image analysis. The Flexible Image Processing System was used to establish quantitative criteria that could assist to automate computer analysis of normal and abnormal sperm forms. ROC-curve analysis was used to establish cut-off points among normal and abnormal sperm heads and sperm acrosomes for seven parameters. The best parameters for discriminating between normal and abnormal sperm were sperm head length (Maxfer) and the average for sperm head length and width (Avefer). The poorest discriminators were aspect ratio and shape factor. The results of this investigation suggest a combination of sperm dimensions that discriminate quantitatively and objectively among normal and abnormal human sperm heads with a high level of specificity and sensitivity.

A unique finding of this investigation is that normal sperm from proven fathers are significantly larger than normal sperm from patients. The various parameters established for the sperm head dimensions of proven fathers may in the future be used a standard for "normal fertile sperm".

6. REFERENCES

Aitken, R. J., Warner, P., Best, F. S. M., Templeton, A. A., Djahanbakhch, O. and Lees, M. M. (1983). The predictability of subnormal penetrating capacity of semen in cases of unexaplained infertility. Int. J. Androl., **6:** 212 - 220.

Aitken, R.J. (1988). Assessment of sperm function for IVF. Hum Reprod., **3:** 89

Albertsen, P. C., Chang, T. S. K., Vindivich, D., Robinson, J. C. and Smyth, J. W. (1983). A critical method of evaluating tests for male infertility. J. Urol., **130:** 467 - 475.

Algren, M., Borstrom, K. and Malmqvist, R. (1974). Sperm transport and survival in women with special reference to the fallopian tube. In: Sperm transport, survival and fertilizing ability in vertebrates. Eds ESE Hafez and CG Thibault. INSERM, Paris, **26:** 183-206.

Althouse, G. C., Bruns, K. A., Evans, L. E., Hopkins, S. M. and Hsu, W. H. (1995). A simple technique for the purification of plasma membranes from ejaculated boar spermatozoa. Reprod. Fertil. Dev., **7(2):**197-210

Appleton, T. C. and Fisher, S. B. (1984). Morphology and X-Ray Microprobe Analysis of Spermatazoa from Fertile and Infertile Men in in Vitro Fertilization. Journal of in vitro Fertilization and Embroy Transfer, **1(3):**188 - 203.

Asch, R. H. (1976). Laparascopic recovery of sperm from peritoneum, in patients with negative or poor Sims-Huhner test. Fertil. Steril., **27:** 1111-1114.

Baker, H. W. G. and Clarke, G. N. (1987) Sperm morphology: consistency of assessment of the same sperm by different observer. Clinical Reproduction and Fertility, **5:** 37-43.

Berenyi, M. and Corradi, Gy. (1982). A simple staining method for the morphologic classification of spermatozoa. International Urology and Nephrology, **14:** 184-188.

Cannon, S. B., Veazy, J. M.and Jackson, R. S. (1978) . Epidemic kepone poisioning in chemical workers. Am. J. Epidemiol., **107 :** 529-537.

Comhaire, F. and Vermeulen, L. (1995) Human semen analysis. Hum. Reprod., **1:**343-362.

Chia, S-E., Ong, C. N., Lee, S. T. and Tsakok, F. H. M. (1992) Blood concentrations of lead, cadmium, mercury, zinc and copper and human semen parameters. Arch. Androl., **29:**177-183.

Check, J. H., Bollendorf, A., Press, M. and Bleu, T. (1992). Standard sperm morphology as a predictor of male fertility potential. Archives of Andrology. **28:** 184-188.

Chong, A. P., Walters, C. A. and Weinrieb, S. (1983). The neglected laboratory test. J. Androl., **4:** 280.

David, G., Bisson, J. P., Czyglik, F., Jouannet, P. and Gernigon, C. (1975). Anomalies morphologiques du spermatozoide humain. 1) propositin pour un systme de classification. J. Gynecol. Obstet. Reprod., **1:** 17.

David, G., Severs, C. and Jouannet, P. (1981). Kinematics of human spermatozoa. Gamete Res., **4:** 83 - 86.

Davis, R. O. and Gravance, C. G. (1993). Standardization of specimen preparation, staining, and sampling methods improves automated sperm-head morphometry analysis, Fertility and Sterility. Fertil Steril., **55:** 412 - 417.

Davis, R. O., Bain, D. E., Siemers, R. J., Thal, D. M., Andrew, J. B. and Gravance, O. G. (1992). Accuracy and precision of the Cellform- Human automated sperm morphometry instrument. Fertility and Sterility. **58:** 763- 769.

Davis, R. O., Bain, D., Obasaju, M. F., Andrwe, J. B. and Katz, D. F. (1990). Computerized morphometric analysis and classification of human spermatozoa.

J. Urol., **11:** 26.

Dott, H. W. (1969). Preliminary examination of bull, ram, and rabbit spermatozoa with the stereoscan electron microscope. J. Reprod. Fertil., **18(1):** 133 - 134.

Domagala, W. and Kotanska, K. (1982). Method of preparation of the maternal obtained by fine - needle aspiration biopsy for examination by Scanning electron microscopy. Patol - Pol., **33(3-4):** 201 - 204.

Dunphy, B. C., Kay, R., Barrat, C. L. R. and Cooke, I. D. (1989) Quality control during the conventional analysis of semen, an essential exercise. Journal of Andrology. **10:** 378-385.

Duijn, C. van, Jr.(1975) Menstruation of spermatozoa. Biblphy. Reprod., **25:**121.

Enginsn, M. E., Dumoulin, J. C. M., Pieters, M. H. E., Bras, M,, Evers, J. L. H. and Geraedts, J. P. M. (1991). Evaluation of human sperm morphology using strict criteria after Diff- Quik staining: correlation of morphology with fertilization *in vitro* Human Reproduction, **6:** 854 - 858.

Eliasson, R. (1971a). Standards for investigation of human semen. Andrologia. **3:** 49-64.

Eliasson, R. (1971b). Analysis of semen. In: The Testis (eds H. Burger and D. de Kretser), pp. 381 - 399. Raven Press, New York.

Eliasson, R. (1977) Supravital staining of human spermatozoa. Fertil Steril., **28:**1257

Evans, H. J., Fletcher, J., Torrance, M. and Hargreave. T. B. (1981). Sperm abnormalities and cigarette smoking. Lancet **1:** 627 - 629.

Falk,H. C.and Kaufman, S. A. (1950) What constitutes a normal semen? Fertil Steril., **1:** 489 - 503.

Frank, P., McNamee, R., Hannafort, P. C. and Kay, C. R. (1994). Effect of changes in maternal smoking habits in early pregnancy of infant birthweight. Br. J. Gen,. Pract., **44(379):** 57 - 59.

Fredricsso, B. (1979). Morphologic evaluation of spermatozoa in different laboratories. Andrologia. **11:** 57-61.

Fredricsson, B. and Bjork, G. (1977) Morphology of postcoital spermatozoa in the cervical secretion and its clinical significance. Fertil Steril., **28:**841

Freund, M. (1996a). Standards for the rating of human sperm morphology: a cooperative study. International Journal of Fertility, **11:** 97-118.

Freund, M. (1996b). A cooperative stidy on the rating of human sperm morphology. Proc. Vth. Int. Congr. Fertil. Steril., Stockholm, 556 - 561.

Freund, M. (1968). Semen analysis. In progress in fertility, ed Behrman, S.J. and Kistner, R.W. First edition, chapter 25, pp 593-627, London.

Garrett, C. and Baker, G. H. W. (1995). A new fully automated system for the morphometric analysis of human sperm heads. Fertility and Sterility, **63(6):** 1306-1317.

Gravance, C. G. and Davis, R. O. (1995). Automated sperm morphology analysis (ASMA) lin the rabbit. Journal of Andrology, **16:** 88-93.

Gamblin, T. C. and Williams, R. C. Jr. (1995). Determination of microtubule polarity *in vitro* by the use of video-enhanced differential - interference contrast light microscopy and Chlamydomonas flagellar axonemal pieces. J. Biol. Chem., **270(31):** 18539 -18542

George, J. D., Fail, P. A., Grizzle, T. B. and Heindel, J. J. (1996). Nitrofurazone: reproductive assessment by continuous breeding in Swiss mice. Fundam. Appl. Toxicol., **34(1):** 56-66

Glover, T. D., Barrat, C. L. R., Tyler, J. P. P. and Hennessey, J. F. (1990). Human Male Fertility and Semen Analysis, pp 132 - 138. Academic Press, San Diego, CA.

Haidl, G. and Schill, W-B.(1993) Sperm morphology in fertile men. Arch. Androl., **31:**153-157.

Hall, J. A., Fishel, S. B., Timson, J. A., Dowell, K. and Klentzeris, L. D. (1995) Human sperm morphology evaluation pre - and post - Percoll gradient centrifugation. Hum. Reprod., 1**0:**342-346.

Harasymouycz, J., Ball, L. and Siedel, G. E. (1976). Evaluation of bovine spermatozoal morphologic features after staining or fixation. Am J Vet Res., **37:** 1053-1057.

Hargreave, T. B. and Nilsson, S. (1983). Seminology. In male infertility, ed, Hargreave, T.B. Chapter 4, pp 56-74. Berlin, Heidelberg, New York, London, Paris, Tokyo: Springer-Verlag.

Hellinga, G. (1976). Clinical Andrology. William Heine-mann Medical Books, London.

Hendricks, S. (1996). Structural and functional aspects of sperm after transit through the fallopian tube of the sheep. M. Sc. Thesis. University of the Western Cape.

Hofmann, N. and Haider, S. G. (1985). Neue Ergebnisse morphologischer Diagnostik der Spermatogenesestnrunge. Gynakologe, **18:** 70-80

Holt, W. V., Moore, H. D. M. and Hillier, S. G. (1985). Computer-assisted measurement of sperm swimming speed in human semen: correlation of results with the *in vitro* fertilization assay. Fertil. Steril., **44:** 112 - 119.

Hotchkiss, R. S., Brunner, E. K. and Grenley, P. (1938). Semen analysis of two hundred fertile men. Am. J. Med. Sex., **196:** 362-384.

Jagoe, J. P., Washbrook, N. P. and Hudson, E. A. (1986). Morphology of spermatozoa using semiautomatic image analysis. J. Clin. Pathol., **39:** 1347-1352.

Jeyendran, R. S., Schrader, S. M., van der Ven, H. H., Burg, J., Perez- Pelaez, M. and Al- Hasani, *et al.* (1986). Association of the in - vitro fertilizing capacity of human spermatozoa with sperm morphology as assessed by three classification systems. Hum Reprod., **1:** 305-308.

Jequior, A. M. and Ukombe, E. B. (1983). Errors inherent in the performance of a routine semen analysis. Br. J. Urol., **55:** 434.

Katz, D. F., Diet, L. and Overstreet, J. W. (1982) Differences in the movement of morphologically normal and human seminal spermatozoa. Biol. Reprod., **26:** 566-570.

Katz, D. F., Overstreet, J. W., Samuels, S. J., Niswander, P. W., Bloom, T. D. and Lewis, E. L. (1986). Morphometric analysis of spermatozoa in the assessment of human male fertility. J Androl., **7:** 203- 210.

Katz, D. F., Drobnis, E. S. and Overstreet, J. W. (1989). Factors regulating sperm migration through the female reproductive tract and oocyte vestments. Gamete Res., **22:** 443-469.

Katz, D. F., Morales, P., Samuels, S. J. and Overstreet, J. W. (1990). Mechanisms of filtration of morphologically abnormal human sperm by cervical mucus. Fertil Steril., **54:** 513.

Keefe, D., Tran, P., Pellegrini, C. and Oldenbourg, R. (1997). Nanomolar concentrations of nocodazole alter microtubule dynamics instability *in vivo* and *in vitro.* Mol. Biol. Cell., **8(6):** 973 - 985.

Krause, W. (995). Computer - assisted semen analysis systems: Comparison with routine evaluation and prognostic value in male fertility and assisted reproduction. Hum. Reprod., 10 suppl. **1:** 60 - 66.

Kruger, T., Menkveld, R., Stander, F. S. H., Lombard, C. J., Van der Merve, J. P., van Zyl, J. A. and Smith, K. (1986). Sperm morphologic features as a prognostic factor in in vitro fertilization. Fertility and Sterility. **46:** 1118.

Kruger, T. ., Haque, D., Acosta, A. A., Pleban, P., Swanson, R. J., Simmons, K. F., Matta, J. F., Morshedi, M. and Oehninger, S. (1988b). Correction between sperm morphology, acrosin and fertilization in an IVF program. Arch Androl., **20:** 237-241.

Kruger, T. F., Acosta, A. A., Simmons, K. F., Swanson, R. J., Matta, J. F. and Oehninger, S. (1988a) Predictive value of abnormal sperm morphology in *in vitro* fertilization. Fertil Steril., **49:**112-117.

Kruger, T. F., Acosta, A. A, Menkveld, R. and Oehninger, S. (1990) Basic semen analysis: Clinical importance of morphology. In Acosta, A.A., Kruger,T.F., Swanson, R.T., van Zly, J.A., Ackerman,S.B.and Menkveld, R. (eds), Human spermatozoa in assisted reproduction. Williams & Williams, Baltimore, M.D. pp 176-181.

Kruger, T. F., Du Toit, T. C., Franken, D. R., Acosta, A. A., Oehninger, S. C., Menkveld, R. and Lombard, C. J. (1993). A new computerized method of reading sperm morphology (strict criteria) is as efficient as technician reading. Fertil Steril., **59:**202-209.

Knuth, U. A., Neuwinger, J. and Nieschlag, E. (1989). Bias of routine semen analysis by uncontrolled changes in laboratory environment - detection by long term sampling of monthly means for quality control. Int. J. Androl., **12:** 375.

Lancranjan, I., Popescu, H. I.and Gavanescu, O. (1975). Reproductive ability of workmen occupationally exposed to lead. Arch. Environ. Health., **30:**396-401.

Lane-Roberts, C., Sharman, A. and Walker, K. (1939). Sterility and impaired fertility, Pathonogenesis, Diagnosis and treatment. Page 93. London. Hamish Hamilton.

Liu, D. Y. and Baker, H. W. G. (1988). The propotion of human sperm with poor morphology but normal intact acrosomes detected with Psium sativum agglutinin correlates with fertilization *in vitro.* Fertil Steril., **50:** 288.

Liu, D. Y. and Baker, H. W. G. (1992). Morphology of spermatozoa bound to the zona pelluda of human oocytes that failed to fertilize *in vitro.* J. Reprod Fertil., **94:** 71 - 84.

Liu, D. Y., Clarke, G. N., Lopata, A., Johnston, W. I. J. and Baker, H. W. G. (1989a). A sperm zona pelluda binding test and *in vitro* fertilization. Fertil. Steril., **52:** 281 - 287.

Liu, D. Y., Lopata, A., Johnston, W. I. J. and Baker, H. W. G. (1989b). Human zona-pellucida binding, sperm characteristics and *in vitro* fertilization. Hum. Reprod., **4:** 696 - 701.

Mahadevan, M. M. and Trounson, A. O. (1984). The influence of seminal characteristics on the success rate of human *in vitro* fertilization. Fertil Steril., **42:** 400.

Macleod, J. and Gold, R. Z. (1951). The male factor in fertility and infertility.1V. Sperm morphology in fertile and infertile marriage. Fertil Steril., **2:** 394-414.

MacLeod, J. and Gold, K. Z. (1953). The factor in fertility and infertility. V1 semen quality and certain other factors in relation to ease of conception. Fertil Steril., **4:**10-33.

MacLeod, I. C., Irvine, D. S., Masterton, A., Taylor, A. and Templeton, A. A. (1974). Assessment of the conventional criteria os semen quality by computer assisted image analysis: evaluation of the Hamilton-Thorn analysis in the context of a service andrology laboratory. Hum. Reprod., **9:** 310-319.

MacLeod, E. (1974). Hospital volunteers. N - J. Nurs - J., **67(12):** 25

Mdhluli, C. (1998). The triterpene, oleanolic acid, induces sterility in male Wistar rats. M Sc Thesis. University of the Western Cape.

Menkveld, R., Stander, F. S. H., Kotze, T. J. W., Kruger, T. F., and Van Zyl, J. A. (1990). The evaluation of morphological characteristics of human spermatozoa according to stricter criteria. Human reproduction, **5:** 586-592.

Menkveld, R., Franken, D. R., Kruger, T. F., Oehninger, S and Hodgen, G. D. (1991). Sperm selection capacity of the human zona pellucida. Mol. Reprod. Dev.,

30(4): 346 -352.

Menkveld, R., Lacquet. F.M., Kruger,T.F., Lombard, C.J., Sanchez-Sarmiento, C.A. and de Villiers, A. (1997). Effects of different staining and washing procedures on the results of human sperm morphology evaluation by manual and computerised methods. Andrologia, **29:** 1-7.

Meschede, D., Keck, C., Zander, M., Cooper, T. G., Yeung, C.- H. and Nieschlag, E. (1993). Influence of three different preparation techiques on the results of human sperm morphology analysis. Internationa Journal of andrology, **16:** 362-369.

Mikel, U. V, and Johnson, F. B. (1980). A simple method for study of the same cells by light and scanning electron microscopy. Acta Cytol., **24:** 252 - 254.

Moench, G. L. and Holt, H. (1930). Sperm morphology in relation to fertility. Am J. Obst & Gynec., **22:**199 - 210.

Moench, G. L. and Holt, H. (1932). Biometrical studies of head lengths of human spermatozoa. J. lab. Clin. Med., **17:** 297.

Mortimer, D., Leslie, E. E., Kelly, R. W. and Templeton, A. A. (1982). Morphological selection of human spermatozoa in vivo and in-vitro. J. Reprod. Fert., **51:** 99-104.

Mortimer, D., Leslie, E. E., Kelly, R. W. and Templeton, A. A. (1982). Morphological selection of human spermatozoa *in vivo* and *in vitro.* J. Reprod. Fertil., **64:**391-399.

Mortimer, D., Shu, M. A. and Tan, R. (1986) Standardization and quality control of sperm concentration and sperm motility counts in semen analysis. Hum. Reprod., p 299.

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

O'Brien *et al* (1997). How calcium causes microtubule depolymerization. Cell Motil. Cytoskeleton, **36(2):** 125-135

Oehninger, S., Acosta, A. A., Morshedi, M., Veek, L., Swanson, R. J., Simmons, K. and Rosenwaks, Z. (1988). Corrective measures and pregnancy outcome in in-vitro fertilization in patients with severe sperm morphology abnormalities. Fertility and Sterility, **50:** 283-287.

Oehninger, S. C., Acosta, R., Morshedi, M., Philput, C., Swanson, R. J. and Acosta, A. A. (1990). Relationship between morphology and motion characteristics of human spermatozoa in semen and in swim - up sperm fraction. J. Androl.,
11: 446-453.

Olsen, G. W., Lanham, J. M. and Bodher, K. M. (1990). Determinants of spermatogenesia recovery among workers exposed to 1,2-dibromo-3-chloropropane. J. Occup. Med., **32(10):** 979-984.

Olivaris, D. and Guercilena, S. (1985). Morphological selection of human spermatozoa in cervical mucus in vivo. Andrologia, **17(5):** 508-512.

Page, E. W. and Houlding, F. (1951). The clinical interpretation of 1000 semen analysis among application for sterility studies. Fertil Steril., 2:140-151.

Papanicolaou, G. N. (1942). A new procedure for staining vaginal smears. Science, 438-441.

Perez-Sanchez, F., de Monserrat, J. J. and Soler, C. (1994). Morphometic analysis of human sperm morphology. International journal of andrology, **17:** 248-255.

Perrey, G., Glezerman, M. and Insler, V. (1977). Selection filtration of abnormal spermatozoa by cervical mucus *in vitro.* In Insler, V. and Bettondorf, G. (eds), *The uterine Cervical in Reproduction. Stuttgart:* George Thieme Publishers, Stuttgart, Gremany, p.118.

Qureshi, M. S. A., Pennington, J. H., Goldsmith, H. J. and Cox, P. E. (1972). Cyclophosphamide therapy. Fertil Steril. Lancer, **ii:** 1209 - 1211.

Ragni, G., Pietro, R., Bestetti, O., De Lauretis, L., Olivaris, D. and Guercilena, S. (1985). Morphological selection of human spermatozoa in cervical mucus in vivo. Andrologia,**17(5):** 508-512.

Salisbury, G. W. and Van Denmark, N. L. (1961). Physiology of reproduction and Artificial Insemenation by Cattle. WH Freeman and Co., London.

Schmassmann, A., Mikuz, G., Bartsch, G. and Rohr, H. (1979). Quantification of human sperm morphology and mortility by means of semi automatic image analysis systems. Micro Acta., **82:** 163- 178.

Schmassmann , A., Mikuz, G., Bartsch, G. and **Rohr, H.** (1982). Objective and reproducible methods for evaluating sperm morphology. Eur Urol., **8:** 274- 279.

Schrader, S. M., Turner, T. W. and Simon, S. D. (1990). Longitudinal study of semen quality of unexposed worker: sperm head morphometry. Journal of Andrology, **11:** 32- 39.

Schoonjans, F. (1996). MedCalc, MedCalc Software. Mariakerke. Belgium.

Sofikitis, N. V., Miyagawa, I., Zavos, P. M., Toda, T., Lino, A. and Terakawa, N. (1994). Confocal scanning laser microscopy of morphometric human sperm parameters: correlation with acrosin profiles and fertilizing capacity. Fertil. Steril., **62:** 376 - 386.

Steigerwald and Krause (1998). A new CASA system using a neuronal network. Andrologia, **30:** 23 - 27. (Original paper not seen. Abstract cited by J. Cummins on SpermMail).

Sun, J. G., Jurisicova, A. and Casper, R. F (1977). Detection of deoxyribonucleic acid fragmentation in human sperm: correction with fertilization *in vitro.* Biol. Reprod., **56:** 602-607.

Sujan, S., Danezis, J. and Sobrerro, A. J. (1963). Semen migration and cervical mucus studies in individual cycles. J. Reprod. Fertil., **6:** 87.

Turner, T. W., Schrader, S. M., Perez-Pelaez, M., Karuhn, R.F., van der Ven, H. H. and Jeyendran, R. S. (1989). Morphometric and volumetric comparison of human spermatozoa. Arch. Androl., **23:** 201 - 206.

van der Horst, G., Curry, P. T., Kitchin, R. M., Burgess, W., Thorpe, E. T., Kwiatkowski, D., Parker, M. and Atherton, R. W. (1991). Quantitative light and scanning electron microscopy of ferret sperm. Mol. Reprod. Dev., **30(3):** 232-240

van der Horst, G., Kitchin, R. M., Curry, P. T. and Atherton, R. W. (1989). The use of membrane filters and osmium tetroxide etching in the preparative of sperm for scanning electron microscopy. J. Electron Microsc. Tech., **12:** 65-70.

van Duijn, C.Jr., ven Voorst, C. and Hellinga, G.(1972). Precesion measurements of dimensions, shape and mass density of spermatozoon heads in normal and subfertile human males. Eur. J. Obstel Gynecol., **2:** 37-54.

Vine, M. F. (1996). Smoking and male reproduction: a review. Int. J. Androl., **19:**323-337.

Viczian, M. (1969). Ergebrisse von spermauntersuchungen bei zigarettenraunchen. Z. Haut Geschlechts Kr., **44:** 183 - 187.

Wang, M. D., Leung, A., Tsoi W-L., Leung, J., J. Ng V. and Lee, K. F. and Steven, Y. W. (1991a). Computer- assisted assessment of human sperm morphology: usefulness in predicting fertilizing capacity of human spermatozoa. Fertility and Sterility, **55 :** 989 - 1993.

Wang, C., Leung, A., Tsoi, W- L., Leung, J. Ng V. and Lee, K. F. *et al.* (1991b). Computer- assisted assessment of human sperm morphology: comparison with visual assessment. Fertility and Sterility, **55:** 938- 988.

William, W. W. and Savage, A. (1925). Observation on the seminal micropathology of bulls. Cornell Vet., **15:** 353 - 375.

Williams, W. W. (1946). Trans. Am. Soc. Study of Sterility, pp139.

Whorton, M., Krauss, R. M., Marshall, S. and Milby, T. H. (1977). Infertility in male pesticide workers. Lancet, **2:** 1259-1261.

Whorton, M. D. and Meyer, C. R. (1981). Sperm count results from 861 American chemical / agricultural workers from 14 separate studies. Fertility and Sterility, **35:** 46-53.

World Health Organization (1980) WHO Laboratory Manual for the Examination of human semen and sperm - cervical mucus Interaction, Press Concern, Singapore.

World Health Organization (1987). Laboratory manual for the examination of human semen and semen-cervical mucus interaction, 1st ed. Cambridge: Cambridge University Press. 1 - 67.

World Health Organization (1987) WHO Laboratory Manual for the Examination of human semen and sperm - cervical mucus Interaction, 2nd ed. Cambridge University Press, Cambridge.

World Health Organization. (1993). WHO laboratory manual for the examination of human semen and semen-cervical mucus interaction, 3rd ed. New York: Cambridge University Press.

Yang, Y. S., Chen, S.U., Ho, H. N., Chen, H. F., Chao, K. H., Lin, H. R., Huang, S. C. and Lee, T.Y. (1995). Correlation between sperm morphology using Strict Criteria in original semen and swim- -up inseminate and human *in vitro* fertilization. Arc. Androll., **34:** 105 - 113.

Zaini, A., Jennings, M. G., and Baker, H. W. G. (1985). Are conventional sperm morphology and motility assessment of predictive value in subfertile men? International Journal of Andrology. **8:** 427-435.

7 ACKNOWLEDGEMENT

I would like to express my sincere gratitude to the following, whose help, advice and encouragement greatly facilitated the compilation of this investigation :

My promoter , Professor van der Horst for his advice , guidance , dedication and for constructive criticism and support during the course of this investigation, and his assistance with statistics was greatly appreciated.

My co-promoter Professor Bornman, of (Centre of Fertility Studies, University of Pretoria), in whose laboratories the foundation of this research was laid, for her supervision and also providing me with slides of proven fathers sperm,

The andrology staff of H.F Verwoerd Hospital , who assisted me in many ways with staining and routine semen analysis during the course of this study, especially, Ms. E. Pienaar, whose help, advice and encouragement and scoring the proven fathers slides for me were greatly valued and highly appreciated.

And the following institutions, colleague, technical assistants, students and friends played a vital role with various forms of assistance and I am grateful to them:

Mr. P.J. Kempen, Mr. C. North and Mrs L. White for their support and supervision for the entire year in their laboratory, at Groote Schuur Hospital with the routine semen analysis,

of great importance, was Mr. P.J. Kempen who read my morphological slide of proven father sperm.

Dr. Ian Harper, Mr. Keith Williams of the Experimental Biology Group, Medical Research Council for their assistance with the Confocal Laser Scanning Microscopy ; Dr. A. Grobblaar who helped with the taking of photographs for the second time.

Mr. S. Allie of the University of Western Cape for his technical assistance and the motivational support, Ms S. Hendricks, Mrs N. Hattingh and Mr. C Mdhluli for their assistance with FIPS or with the typing of the many results, Mrs. S, Kaskar and Mr. Q. Terhoven who assisted with fixation of the semen samples for the scanning electron microscopy. Mrs J. Africa, Mrs A. Pretorius and Mr. R. Larney for the typing of the manuscript. Bernhard van der Horst for the statistical computerized input.

My late friend Feziwe Zibi, Tsitsi de Langen for their editorial work, Norry Mntonintshi and Itumeleng Qhuenya for their moral support, Mr. W. de Langen for printing the thesis and Anna Gangevoort, Thuli Ncube and Tseleng Masasa for their wonderful accommodation for the last one and half years, thank you so much.

Once more to my wonderful husband, thank you for his emotional and telephonic support. Thanks a million once more.

To my brothers, sisters, and to all my friends for their encouragement, love, support and prayers without which this investigation could not have been successfully completed. Thank you for bringing this investigation to a wonderful finish.

The following institutions rendered financial support and without their assistance this investigation would not have been possible. The FRD and Medical Research Council.

Finally but not least my friend, Dr. Tom. Pisarri of Creighton University, Nebraska for his advice, encouragement and support are greatly appreciated.

The author would like to name each individual (and they are many), who contributed to the successful completion of this investigation, but is unable due to space limitation. A collective "THANK YOU "is extended.