Correlation of testicular sperm extraction with morphological, biophysical and endocrine profiles in men with azoospermia due to primary gonadal failure

U.I.O.Ezehi, H.D.M.Moore1,2,3 and I.D.Cooke1

1University Departments of Obstetrics and Gynaecology, Jessop Hospital For Women, Leavygreave Road, Sheffield S3 7RE and 2Department of Molecular Biology and Biotechnology, University of Sheffield S10 2TN, UK
3To whom correspondence should be addressed at: University Departments of Obstetrics and Gynaecology, Jessop Hospital For Women, Leavygreave Road, Sheffield S3 7RE, UK

To identify the predictive factors for testicular sperm extraction (TESE) and to understand the pathology associated with TESE, we carried out a prospective study in 40 consecutive men with azoospermia due to primary gonadal failure. The main outcome measure was the retrieval of at least one testicular spermatozoon. Endocrine and biophysical profiles, testicular histology, Johnsen score and testicular spermatids were used as predictors of sperm extraction. Spermatogenesis was quantified with the Johnsen score. A variable pattern of spermatogenesis was common, being present in 20 (50%) patients. Visualisation of testicular spermatids on testicular histology showed a strong association with TESE. (P < 0.0001). Statistically significant differences were detected in plasma follicle stimulating hormone (FSH) and testicular volume between patients with hypospermatogenesis and Sertoli cell-only or maturation arrest. There were no significant differences in Johnsen score, biophysical and endocrine profiles between the groups with successful and failed TESE. However, a statistically significant trend occurred with changes in histological pattern [χ2 for trend, P = 0.001; Pearson’s coefficient (r) = 0.6], Johnsen score (P = 0.022; r = 0.5), testicular volume (P = 0.01; r = 0.5) and plasma FSH concentrations (P = 0.044; r = 0.4), albeit to a limited degree. Differences in the interpretation of histological patterns with different assessors was observed. The type of occupation or risk factors for azoospermia showed no association with testicular pathology or TESE. Variable histological patterns in different tubules in the same individual may explain the poor correlation of TESE with endocrine and biophysical profiles, Johnsen score and histological pattern. Differences in the amount of tissue used for TESE and histopathology, and misinterpretation of testicular histology rather than failure to quantify spermatogenesis may explain the poor correlation between histological patterns and TESE. Testicular spermatids predicted TESE. However, considerable overlap in values means that no single variable can provide a perfect discrimination between the groups with successful and failed TESE.

Key words: azoospermia/sperm extraction/testis pathology

Introduction

The use of testicular biopsy started more than 50 years ago as a common diagnostic procedure to distinguish obstructive from non–obstructive azoospermia (Charny, 1940). In recent years, its use has focused on extracting testicular spermatozoa for intracytoplasmic sperm injection (ICSI) of the oocyte for treating men with non-obstructive azoospermia (Silber et al., 1996a). Nevertheless, testicular biopsy is an invasive procedure that may be associated with significant potential complications (Beierdorffer and Schirren, 1979). Testicular sperm extraction (TESE) is a laborious process successful in 30–70% of patients with non-obstructive azoospermia (Jow et al., 1993; Tournaye et al., 1996a, 1997). The need to restrict testicular biopsy to those with a high chance of yielding testicular spermatozoa is therefore widely acknowledged. Little is known about the features associated with TESE. The role of follicle stimulating hormone (FSH) as a predictor of the status of spermatogenesis is now uncertain (Hauser et al., 1995; Martin-du-Pan and Bischof, 1995). Although there are non-invasive means of testicular biopsy such as needle aspiration techniques, which provide suitable materials for analysis of testicular histology (Malilidis and Baker, 1993; Craft et al., 1997), the testicular histology patterns associated with TESE in men with non-obstructive azoospermia is debatable. There is speculation as to whether the controversy is due to misinterpretation of histology (Hovatta, 1996) or differences in the number of biopsies taken for histopathology and sperm extraction (Silber, 1996). Studies correlating testicular histology with sperm count in men found that a qualitative description of testicular biopsy was too subjective and vague for the evaluation of spermatogenesis in oligozoospermic men (Johnsen, 1970; Zukerman et al., 1978). It is possible that the failure to quantify testicular pathology in azoospermic men may be responsible for the controversy observed.

Several methods have been used for quantitative analysis of spermatogenesis in oligozoospermic and azoospermic men in the past. Semiquantitative methods including cell counts/unit tubular wall length (Clermont, 1963; Steinberger et al., 1973) or cell count/Sertoli cell (Skakkebaek and Heller, 1973) have been found unsatisfactory because of their very poor correlation with sperm count and because their use was time-consuming. These gave rise to fully quantitative methods including sperm count/mature testicular spermatids (Silber and Rodriguez-Rigau, 1981) and a scoring system based on the percentage of tubules with mature spermatids or arrest of spermatogenesis at primary spermatocyte level (Holstein and Schirren, 1983) but both methods are unsuitable for patients with late maturation arrest. These difficulties made the Johnsen...
score, in which each tubule was scored on a scale of 1–10 according to the presence or absence of the main cell types arranged in order of maturity, the most popular clinical method of quantifying spermatogenesis. Although DNA cytometry of testicular biopsy has become a more recent, rapid method of quantifying spermatogenesis, it is expensive and does not count diploid cells.

We postulated that the biophysical and endocrine variables associated with spermatogenesis can predict TESE and that differing opinions in the interpretation of testicular histology in patients with non-obstructive azoospermia are due to failure to quantify spermatogenesis. The aim of this study was to determine the predictive factors and pathology associated with TESE for ICSI of the oocyte.

Materials and methods

Study population
We studied 40 consecutive patients with non-obstructive azoospermia who underwent testicular biopsy at the Jessop Hospital for Women in Sheffield, UK with a view ultimately to undergoing ICSI of oocytes documented in every patient. Each patient underwent physical examination which included an evaluation of secondary sexual characteristics, examination of the penis, vasa deferentia, epididymides and a rectal examination to exclude prostatic pathology. The testicular history of urological operations and exposure to gonadotoxins were documented in every patient. Each patient underwent physical examination which included an evaluation of secondary sexual characteristics, examination of the penis, vasa deferentia, epididymides and a rectal examination to exclude prostatic pathology. The testicular volume was estimated with a Prader orchidometer. Data on height and weight were used to calculate the body mass index (kg/m²).

Histopathology
Testicular tissue (~5–10 mg) was prepared for histology. Semi-thin paraffin wax testicular tissue sections (4 µm thick) fixed in Bouin’s solution were dewaxed and rehydrated by transfer through graded alcohol/xylene, stained with haematoxylin and eosin, and were examined under a light microscope at ×100–1000 magnification using standard techniques. Testicular histology was classified into hypospermatogenesis (reduction in the degree of normal spermatic- genic cells), maturation arrest (an absence of the later stages of spermatogenesis), Sertoli cell-only (the absence of germ cells in the seminiferous tubules), and tubular sclerosis (no germ cell or Sertoli cell present in the tubules). Isolated seminiferous tubules with few spermatids in the field of seminiferous tubules that were otherwise maturation arrest, Sertoli cell-only pattern or tubular sclerosis were classified as focal spermatogenesis. In addition, between 100 and 200 tubules were examined per slide; each slide was scored using the Johnsen score (Johnsen, 1970), whereby seminiferous tubules are scored on a scale of 1–10, with those with maximum activity (at least five or more spermatozoa in the lumen) scored as 10. In order to obtain a mean index, the number of tubules recorded at each score was multiplied by the score and the sum of all multiplications was divided by the total number of tubules recorded. Where bilateral testicular biopsy

<table>
<thead>
<tr>
<th>Table I. Eligibility of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not eligible</td>
</tr>
<tr>
<td>Obstructive azoospermia</td>
</tr>
<tr>
<td>Vasectomy</td>
</tr>
<tr>
<td>Vasectomy reversal</td>
</tr>
<tr>
<td>Sexually transmitted diseases</td>
</tr>
<tr>
<td>Swollen testis</td>
</tr>
<tr>
<td>Distended epididymis</td>
</tr>
<tr>
<td>Absence of vasa deferentia/epididymides</td>
</tr>
<tr>
<td>Semen pH &lt; 7</td>
</tr>
<tr>
<td>Presence of sperm agglutinins</td>
</tr>
<tr>
<td>Retrograde ejaculation</td>
</tr>
<tr>
<td>Low semen volume (&lt;1 ml)</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
</tr>
<tr>
<td>Retroperitoneal lymphadenectomy</td>
</tr>
<tr>
<td>Bladder neck surgery</td>
</tr>
<tr>
<td>Spinal injury</td>
</tr>
<tr>
<td>Hypogonadotrophism</td>
</tr>
<tr>
<td>Kallmann’s syndrome</td>
</tr>
<tr>
<td>Endocrine disorder</td>
</tr>
<tr>
<td>Low FSH/LH</td>
</tr>
</tbody>
</table>

| Eligible                          |
| Testicular failure                |
| Idiopathic                       |
| Chemotherapy                      |
| Radiotherapy                      |
| Malignant disease                 |
| Cryptorchidism                    |
| Orchidopexy                       |
| Torsion of the testis             |
| Mumps orchitis                    |
| Abnormal karyotype                |
| Testicular atrophy (<12 ml)       |

FSH = follicle stimulating hormone; LH = luteinizing hormone.
was performed the average of the scores for both testes was found. The slides were read by three different assessors unaware of the results of the testicular sperm retrieval.

**Main outcome measures**

The endocrine profile (plasma FSH, LH and testosterone concentrations), biophysical profile (patients’ age, paternal and maternal ages at birth, occupation, risk factors for azoospermia, height, weight, body mass index, and the sum of volume of both testes), whether or not testicular spermatids were visualized, Johnsen score and testicular histological patterns (hypospermatogenesis, maturation arrest, Sertoli cell-only, tubular sclerosis and focal spermatogenesis) were used as predictors of retrieval of at least one testicular spermatozoon in consecutive patients.

**Blinding of variables and end-points**

The persons performing the semen analysis, history evaluation, TESE and clinical evaluation were blinded from each other. Each of these procedures was performed by the same person on each occasion.

**Statistical analysis**

Analysis of continuous variables with regards to TESE were compared with the Mann–Whitney U-test. The difference between means was determined by P-values with confidence intervals; confidence intervals excluding 0 for means were considered to be significant. \( \chi^2 \)-Test for trend was used to compare proportions stratified by histological patterns, Johnsen score, testicular volume and plasma FSH concentration. The Kruskal–Wallis test was used for the comparison of the testicular histological types with biophysical and endocrine variables. \( \chi^2 \)-Test or Fisher’s exact test (with Yates’ correction) was used for comparison of categorical variables as appropriate. Two-sided P-values < 0.05 were considered significant.

**Results**

Patients with gonadal failure were selected on clinical criteria (Table I). All the patients showed histological evidence of testicular failure. The median age, height, testicular size, plasma FSH concentrations of all the patients studied were 34 years, 154 cm, 32 ml, and 18 IU/l, respectively. The mean weight (± SD) of tissue taken for histopathology was 585 (± 20) mg and for sperm extraction 479 (± 12.1) mg (\( P > 0.05 \)). TESE was successful in 28 out of the 40 patients studied, giving a sperm retrieval rate of 70%.

**Testicular pathology**

All the biopsies were considered suitable for analysis. The relative frequencies of the testicular histological patterns were hypospermatogenesis in 10 (25%) patients, maturation arrest in seven (18%), Sertoli cell-only pattern in 12 (30%), and focal spermatogenesis occurred in 11 (28%) patients, with Sertoli cell-only the commonest abnormality found. None of the patients had tubular sclerosis as an isolated abnormality. Tubular sclerosis co-existed with Sertoli cell-only syndrome pattern in six patients, three patients showed a mixed histological pattern of Sertoli cell-only and maturation arrest and focal areas of normal spermatogenesis co-existed with those of maturation arrest or Sertoli cell-only in 11 patients, giving a variable pattern of spermatogenesis rate of 50%. There was no evidence of malignancy in any of the specimens.

A block of spermatogenesis at the level of spermiogenesis was not encountered in any of the patients. In those with hypospermatogenesis, round spermatids were always present in association with the long spermatids and a prolonged search of the entire slides always showed a few spermatozoa. Table II examines the relation of testicular spermatids to TESE. A statistically significant difference was found between the group with and without testicular spermatids visualized (\( \chi^2 = 10.999; P = 0.0009 \)). However, the sensitivity was 71%, specificity 92%, positive predictive value 95% and negative predictive value 58% for predicting the likelihood of TESE, giving a false positive rate of 5% and false negative rate of 42%.

Table III examines the variation in the interpretation of histological patterns between different assessors. The variation in the interpretation of histological patterns with different assessors. The variation in the interpretation of histological patterns with different assessors.

<table>
<thead>
<tr>
<th>Testicular spermatids</th>
<th>Proportion of patients with testicular spermatozoa (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent (n = 19)</td>
<td>8 (42)</td>
</tr>
<tr>
<td>Present (n = 21)</td>
<td>20 (95)</td>
</tr>
</tbody>
</table>

\( \chi^2 \)-Test: \( P < 0.0009 \).

Values in parentheses are percentages.

Table III. The variation in the interpretation of histological patterns with different assessors

<table>
<thead>
<tr>
<th>Testicular spermatids</th>
<th>Proportion of patients with testicular spermatozoa (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent (n = 19)</td>
<td>8 (42)</td>
</tr>
<tr>
<td>Present (n = 21)</td>
<td>20 (95)</td>
</tr>
</tbody>
</table>

\( \chi^2 \)-Test: \( P < 0.0009 \).

Values in parentheses are percentages.

**Discussion among the assessors resolved any discrepancies in order to reach a final diagnosis.**

There was no statistically significant difference in the median Johnsen scores between the groups with successful and failed TESE (Table IV). There is considerable overlap in Johnsen score in both the groups with successful and failed TESE (Figure 1). None of the patients had a Johnsen score of <2, which is in keeping with the fact that no patients had tubular sclerosis as an isolated abnormality.

**The relation of biophysical and endocrine profile to TESE**

The median patient’s age, maternal or paternal age at birth, height, weight and body mass index were similar in both the group with successful and failed TESE (Table IV). The estimated 95% CI of the median difference in testicular size ranged from 0 to 18 (Table IV). In addition, there was

![Table II. Testicular sperm retrieval in relation to testicular spermatids](image-url)

<table>
<thead>
<tr>
<th>Testicular spermatids</th>
<th>Proportion of patients with testicular spermatozoa (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent (n = 19)</td>
<td>8 (42)</td>
</tr>
<tr>
<td>Present (n = 21)</td>
<td>20 (95)</td>
</tr>
</tbody>
</table>

\( \chi^2 \)-Test: \( P < 0.0009 \).

Values in parentheses are percentages.
The relation of testicular histology to biophysical and endocrine profiles

Statistically significant differences were found between the pattern of histology and testicular volume (Kruskal–Wallis test: T = 13.23; P = 0.02) (Table V). The testicular volume of the group with hypospermatogenesis was significantly larger than that of the group with focal spermatogenesis (P = 0.008), Sertoli cell-only syndrome (P = 0.0002) or maturation arrest (P = 0.01). There was a significant association between the plasma FSH concentrations and the pattern of histology (Kruskal–Wallis test: T = 14.34; P = 0.0025). The mean plasma FSH was significantly higher in the group with Sertoli cell-only (P = 0.0005) or maturation arrest (P = 0.0016) compared with the group with hypospermatogenesis. Similarly, the group with focal spermatogenesis had significantly lower plasma FSH concentrations compared with the group with Sertoli cell-only (P = 0.008) or maturation arrest (P = 0.016). No statistically significant difference was found between testicular volume and plasma FSH concentrations on the one hand and other histological patterns. Ten (100%) patients with the histological diagnosis of hypospermatogenesis and nine out of 11 (82%) with focal spermatogenesis, four out of seven (57%) patients with maturation arrest and five out of 11 (46%) patients with Sertoli cell-only had normal testicular volumes (>24 ml). One (8%) patient with Sertoli cell-only pattern, one (14%) with maturation arrest, five (46%) patients focal spermatogenesis and eight (80%) with hypospermatogenesis had normal plasma FSH concentration. No testicular volume or plasma FSH concentration was typical of any histological pattern. There was considerable overlap in testicular volume and plasma FSH concentration between the groups of patients with different testicular histology patterns (Figure 2).

Risk factors for azoosperma

The patients’ occupation and risk factors for azoosperma were considered in relation to histological patterns and TESE. Because of the small number of patients in each category of risk factors (mumps orchitis, chemotherapy or radiotherapy, undescended testis, abnormal karyotype or testicular trauma), pooled risk factors for azoosperma were considered. For similar reasons the patients’ occupation was classified into engineering (steel working, machining, structural and building engineering, electrical working, warehouse working, carpentry, window fitting, labouring, warehouse working) and non-engineering groups (driver, teacher, civil servant, sales assistant, shop-keeper, bank manager, policeman, private detective, unemployed).

Of the 19 patients with identifiable risk factors for azoosper-
Table V. Biophysical and endocrine profiles correlated with testicular histological patterns

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hypospermatogenesis</th>
<th>Focal Maturation</th>
<th>Sertoli cell-only arrest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients’ age^a (years)</td>
<td>36 (11)</td>
<td>32 (5)</td>
<td>38 (11)</td>
</tr>
<tr>
<td>Paternal age^a (years)</td>
<td>28 (8)</td>
<td>32 (7)</td>
<td>30 (4)</td>
</tr>
<tr>
<td>Maternal age (years)</td>
<td>27 (7)</td>
<td>28 (8)</td>
<td>26 (4)</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>176 (10)</td>
<td>108 (85)</td>
<td>155 (69)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>78 (15)</td>
<td>79 (11)</td>
<td>77 (8)</td>
</tr>
<tr>
<td>Body mass index (kg/m^2)</td>
<td>25 (3)</td>
<td>31 (12)^c</td>
<td>29 (14)^d</td>
</tr>
<tr>
<td>FSH (IU/l)</td>
<td>8 (6)^b</td>
<td>14 (13)^c</td>
<td>26 (11)^d</td>
</tr>
<tr>
<td>LH (IU/l)</td>
<td>6 (5)</td>
<td>7 (6)</td>
<td>11 (11)</td>
</tr>
<tr>
<td>Testosterone (nmol/l)</td>
<td>16 (5)</td>
<td>17 (5)</td>
<td>13 (5)</td>
</tr>
</tbody>
</table>

Values are means (SD). Kruskal–Wallis test was used for analysis.

^a Represents the age at the birth of the patient.

^b,c,d Represent significant difference in comparison between the groups. P < 0.05 was considered statistically significant.

LH = luteinizing hormone; FSH = follicle stimulating hormone.

TESE was successful in 12 out of 17 (71%) patients who did engineering-related work compared with 16 out of 23 (70%) who did non-engineering-related work ($\chi^2 = 0$, P = 1.00). In the group who did engineering-related work, testicular histology showed hypospermatogenesis in three patients, Sertoli cell-only in seven, maturation arrest in two, and focal spermatogenesis in five compared with seven, five, and six who did non-engineering work, respectively ($\chi^2 = 2.465$, P = 0.48). The patients’ occupation was not associated with TESE.

An abnormal karyotype was recorded in eight (20%) patients who had testicular biopsy: Klinefelter’s syndrome ($n = 4$), macroscopic deletion of the distal end of the Y chromosome ($n = 2$) and chromosome translocation ($n = 2$). In contrast to patients with Klinefelter’s syndrome testicular spermatozoa were retrieved from two patients with translocation and one of the two patients with Y chromosome deletion. Having an abnormal karyotype therefore does not exclude the possibility of successful sperm retrieval. Patients with an abnormal karyotype showed no typical histological pattern. One patient with a chromosome translocation of 3/4 showed hypospermatogenesis and the other, with satellite deletion of chromosome 9, showed focal spermatogenesis. Different histological patterns were seen in patients with Klinefelter’s syndrome: one showed Sertoli cell-only, another Sertoli cell-only with tubular sclerosis, and the third showed a maturation arrest histological pattern. One patient with Y deletion showed Sertoli cell-only and the other maturation arrest.

Trends in TESE and changes in testicular volume, plasma FSH concentrations, Johnsen score and histological pattern

A statistically significant trend occurred between TESE and changes in the pattern of histology [Pearson’s $r^2$ for trend = 13.62, $P = 0.001$, Pearson’s coefficient ($r$) = 0.6]. Johnsen score (Pearson’s $r^2$ for trend = 9.63, $P = 0.022$, $r = 0.4$), testicular size (Pearson’s $r^2$ for linear trend = 9.171, $P = 0.01$, $r = 0.5$) and plasma FSH concentrations (Pearson’s $r^2$ for linear trend = 6.72, $P = 0.035$, $r = -0.4$), albeit to a limited degree (Table VI). Testicular spermatozoa were not recovered from any of the seven patients with maturation arrest. No testicular spermatozoa were retrieved from four
patients with very severe testicular atrophy (the sum of testicular volume of both testes ranged from 1 to 12 ml). These patients had Klinefelter's syndrome.

**Discussion**

This prospective study investigated the relation of testicular pathology and the current clinical parameters associated with spermatogenesis to TESE in patients with azoospermia due to defects in spermatogenesis. A poor correlation was found between the endocrine and biophysical profiles, histological patterns and Johnsen score with TESE.

Very few studies have addressed the relationship between testicular pathology and sperm extraction. Jow et al. (1993) and Tournaire et al. (1996a) correlated the testicular histological pattern with the outcome of TESE. Our finding, that the histological pattern best associated with successful sperm retrieval is hypospermatogenesis, is consistent with the findings of other studies. Jow et al. (1993) reported successful TESE in seven out of 11 (54%) patients with hypospermatogenesis, two out of nine (33%) patients with maturation arrest and none of nine patients with Sertoli cell-only pattern among 29 azoospermic men with azoospermia due to gonadal failure. Tournaire et al. (1996a) studied 124 patients with both obstructive and non-obstructive azoospermia and found that the number of testicular spermatozoa retrieved decreases as the histological pattern changes from normal to maturation arrest and Sertoli cell-only. In our study, testicular spermatozoa were retrieved in all the patients with hypospermatogenesis, 91% of patients with focal spermatogenesis, none of the patients with maturation arrest and 67% of patients with Sertoli cell-only pattern. Nevertheless, there are a number of other differences between this and previous studies. First, the previous studies were retrospective. Second, we explored the relationship between testicular spermatids and sperm extraction. Third, the Johnsen score was used to investigate whether failure to quantify spermatogenesis is responsible for the disparity between the result of histopathology and that of sperm extraction. Fourth, we used a comparable amount of testicular tissue and number of testicular biopsies for histopathology and sperm extraction. Fifth, the relationship between the biomedical, anthropometrical and endocrine parameters associated with spermatogenesis and TESE were explored. Finally, the trends in testicular volume, plasma FSH concentrations, Johnsen score and histological pattern with regards to TESE were investigated.

Our results (sensitivity 71%, specificity 92%, positive predictive value 95% and negative predictive value 58%) are also consistent with two other studies that demonstrated a good relationship between visualizing testicular spermatids and TESE. Mulhall et al. (1997) studied 30 patients with non-obstructive azoospermia and found that seven out of 14 patients (50%) with germ cell aplasia, six out of eight (75%) patients with maturation arrest or hypospermatogenesis and all the eight patients with late spermatids seen on histological analysis had spermatozoa retrieved from their testes, giving a sensitivity of 38%, a specificity of 100%, a positive predictive value of 100% and a negative predictive value of 41% for predicting the probability of TESE. In contrast to our study, their study classified hypospermatogenesis with maturation arrest. Like our study, their study was prospective and the finding of late spermatids was highly predictive of TESE and their absence led to a high false positive rate. A very recent retrospective study evaluating the distribution of spermatogenesis in the testiciles of 45 azoospermic men found that 22 of 26 (85%) men with mature spermatids had successful sperm retrieval (Silber et al., 1997). This rate is similar to the prediction rate of 95% observed in our study. However, their study differs from our study in some respects. Some of their cases with spermatids were classified as incomplete Sertoli cell-only or maturation arrest whereas these were termed focal spermatogenesis in our study. In addition, successful sperm retrieval was observed in only one of 19 (5%) patients without spermatids seen on prior testicular biopsy, a false negative rate of only 5% (sensitivity 96%, specificity 82%, positive predictive value 85% and negative predictive value 95%) which differs from our rate of 42% and 58% reported by Mulhall et al. (1997). The reason for discordance in the false negative results remains unexplained. That mature spermatids are predictive of TESE is not surprising. This is because maturation arrest is usually rare (Soderstrom and Souminen, 1980; Johnson et al., 1980; Silber et al., 1997), apoptosis has reduced remarkably and cell divisions have ceased at the level of spermiogenesis (Steinberger and Tjioe, 1968; Silber and Rodriguez-Rigau, 1981), so that some testicular spermatozoa would invariably be produced once spermatogenesis has progressed to this stage. This observation also concurs with our findings and those of others that spermatogenic maturation arrest occurs in meiosis rather than in spermiogenesis (Silber and Rodriguez-Rigau, 1981; Silber et al., 1996b; Mulhall et al., 1997) as well as the findings of previous studies that demonstrated excellent correlations between the mean number of mature spermatids per seminiferous tubule and ejaculated sperm counts (Silber and Rodriguez-Rigau, 1981; Silber et al., 1996b; Mulhall et al., 1997).

---

**Table VI.** Trends in testicular sperm retrieval in relation to testicular volume, plasma follicle stimulating hormone (FSH) concentration, Johnsen score and histological patterns

<table>
<thead>
<tr>
<th>Variables</th>
<th>Success</th>
<th>Failure</th>
<th>P-value</th>
<th>Pearson’s r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasma FSH (IU/l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0–12 (n = 12)</td>
<td>13 (93)</td>
<td>1 (7)</td>
<td>0.035</td>
<td>0.4</td>
</tr>
<tr>
<td>12.1–24 (n = 15)</td>
<td>10 (67)</td>
<td>5 (33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;24 (n = 11)</td>
<td>5 (46)</td>
<td>6 (54)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Testicular volume (ml)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0–12. (n = 3)</td>
<td>0 (0)</td>
<td>3 (100)</td>
<td>0.010</td>
<td>0.5</td>
</tr>
<tr>
<td>12.1–24 (n = 10)</td>
<td>6 (60)</td>
<td>4 (40)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;24 (n = 27)</td>
<td>22 (82)</td>
<td>5 (18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Histological patterns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypospermatogenesis (n = 10)</td>
<td>10 (100)</td>
<td>0 (0)</td>
<td>0.001</td>
<td>0.6</td>
</tr>
<tr>
<td>Focal spermatogenesis (n = 11)</td>
<td>10 (91)</td>
<td>1 (9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maturation arrest or Sertoli</td>
<td>8 (24)</td>
<td>11 (58)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cell-only (n = 19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johnsen score</td>
<td></td>
<td></td>
<td>0.022</td>
<td>0.4</td>
</tr>
<tr>
<td>8–10 (7)</td>
<td>7 (100)</td>
<td>0 (0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6–7.9 (10)</td>
<td>9 (90)</td>
<td>1 (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3–5.9 (10)</td>
<td>4 (40)</td>
<td>6 (60)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–2.9 (13)</td>
<td>8 (62)</td>
<td>5 (38)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Values in parentheses are percentages.

FSH = follicle stimulating hormone.

---

**Testicular sperm extraction**
et al (1996b) or testicular spermatozoa extracted (Mulhall et al., 1997; Silber et al., 1996a). However, failure to visualize the testicular spermatids on histological analysis does not exclude the possibility of sperm production since foci of sperm extraction could be occurring in some area of the testis not included in the specimen for histopathology.

Focal areas of spermatogenesis co-existing with Sertoli cell-only histological patterns were first noted by Levin (1979) in 6% of 102 patients undergoing testicular biopsies for oligozoospermia, obstructive azoospermia and azoospermia due to primary gonadal failure. Later, Silber et al. (1995) reported similar findings in four patients in relation to TESE, but the proportion of patients with azoospermia due to primary gonadal failure with variable spermatogenesis remains unknown. In this study, focal areas of spermatogenesis co-existing with either maturation arrest or Sertoli cell-only pattern were found in 28% of the patients who underwent bilateral testicular biopsy. In addition, 22% of the patients had areas of Sertoli cell-only in addition to maturation arrest or tubular sclerosis. The frequency of variable spermatogenesis of 50% in this study suggests that a variable histological pattern is common. The tubule-to-tubule variability runs counter to the orthodoxy that spermatogenesis in the testis is uniform and that a piece of testicular tissue as small as 10 mg is representative of the entire testis (Steinberger and Tjoe, 1968; Johnsen et al., 1980). While this may be the case in normal or oligozoospermic men, the findings of this study suggest that this is not the case in men with azoospermia due to primary gonadal failure. Opinions differ with regard to how the seminiferous tubules with active spermatogenesis are distributed in the testis. Tournaye et al. (1996) had speculated that these tubules are distributed in a patchy manner in the testis since sperm extraction was successful in those with maturation arrest and Sertoli cell-only. In contrast, Silber et al. (1997) have recently proposed a homogeneous distribution of these tubules on the basis of the high predictive value of spermatids for predicting the likelihood of TESE. The low negative predictive value of spermatids for predicting the likelihood of TESE observed in this study and others (Mulhall et al., 1997) supports a more patchy distribution.

The cause of tubule-to-tubule variation is unknown. A number of factors may be responsible. One can speculate that the degeneration of cells in the seminiferous tubules occurs in a progressive fashion in such a way that the tubules lose cells in a particular order. Such a progressive degeneration of cells has been reported previously (Johnsen, 1967). That some of our patients showed Sertoli cell-only with focal spermatogenesis without obvious areas of maturation arrest casts doubt on such a hypothesis. To understand exactly the distribution of seminiferous tubules with active spermatogenesis and the mechanism of tubular degeneration in the testis of azoospermic men will entail serial sampling of the entire testes obtained at autopsy. Desquamation of Sertoli cells or different spermatogenic cells in the seminiferous tubules during histological preparation may offer another explanation, but the fact that all our histological slides were suitable for analysis, and that tissues were collected by an atraumatic technique and fixed in Bouin’s solution makes this possibility unlikely. It is not related to the predisposing factors of azoospermia as no relationship was found between the underlying cause of azoospermia or patient’s occupation and TESE. Perhaps the variable spermatogenesis could be a reflection of variable expression of the same gene or the existence of multiple genes involved in spermatogenesis (Reijo et al., 1995). This study confirms that misinterpretation of testicular histology accounts for the discrepancy between the result of TESE and that of histology in some cases. Review of the histology by three different assessors and the extensive review of slides required to estimate the Johnsen score reduces the possibility of misinterpretation of the histology in our patients. Although a comparable amount of testicular tissue was taken for both TESE and histopathology and the testicular tissues for both procedures were taken through the same testicular incision, the fact that only a relatively tiny amount of testicular tissue (4 µm thick) was actually used for slide preparation for histological examination compared to an average of 485 mg of tissue used for sperm extraction could explain the disparity between histology and the outcome of TESE. This is because the tissues for sperm extraction are more likely to contain a greater number of foci of spermatogenesis. In addition, mincing of therapeutic testicular biopsy allows observation of each isolated testicular cell. In contrast, the diagnostic testicular biopsy does not involve dispersion of cells. The data presented here do not support the hypothesis that the lack of correlation between testicular histological patterns and TESE was caused by failure to quantify testicular histology. This is because of considerable overlap in the value of Johnsen’s indices in both groups with successful and failed TESE. The result is that a patient with a high Johnsen score may have successful TESE but a low Johnsen score does not exclude the possibility of TESE. The reason for the overlap is unclear. One can speculate that the variability in the nature of spermatogenesis between one tubule and another in the same patient may be so extreme that Johnsen’s scoring system is not adequate to offset the impact of such variability. Because of the small sample size, the data were presented in four groups instead of 10 but the classification was still in agreement with the principles of Johnsen (1970): presence of spermatozoa scores 10, 9 or 8; presence of spermatids 7 or 6; presence of spermatocytes or spermatagonia 5, 4, 3; presence of only Sertoli cells scores 2. It is possible that using a greater number of patients than used in this study to correlate Johnsen’s scores with TESE, biophysical and endocrine variables may provide a better insight into the relation of these variables with the Johnsen score. However, the continuation of this study will provide the opportunity to address this question in future.

Because our study was prospective, all the important clinical data were available. Despite the fact that we considered variables that have an effect on spermatogenesis none of the biomedical, anthropometric and endocrine variables in the groups had any effect on TESE. Age-related changes in testicular function, semen analysis and endocrine profile have been reported (Matthieu et al., 1995; Wide, 1995; Haidl et al., 1996). Previous studies demonstrated a relationship between mitochondrial (mt) DNA mutation, reactive oxygen species and fertility (Aitken, 1994). Cummins et al. (1994) recently
drew attention to the relationship between premature ageing of the testis, mtDNA mutation and generation of reactive oxygen species. A recent report by Keefe et al. (1995) showed that the ovaries from older women harbour mutated mitochondrial DNA. The age of the patient as well as paternal and maternal ages at birth of the patient may therefore affect spermatogenesis and TESE. The young age of patients who participated in this study, and the age of their parents at the time of their birth coupled with the fact that fertility in the male, unlike the female, persists into old age, may account for the poor correlation between TESE and the patient’s age or the age of their parents at the time of their birth. There is a relationship between height and male infertility although the reason for this is not clear. For example, Klinefelter’s syndrome has a well-defined association with height. The fact that there were only four patients with Klinefelter’s syndrome in this study, as well as the fact that the mean height of the patients who participated in this study was within normal limits, may explain the lack of correlation between TESE and height. The considerable overlap in FSH and testicular volume between the groups confirms the findings of other studies which found that foci of testicular spermatozoa could still be present in the testis despite raised plasma FSH concentration and testicular atrophy (Hauser et al., 1995). Nevertheless, the frequency of raised plasma FSH concentration and testicular atrophy in azoospermic patients with active spermatogenesis is not known (Hauser et al., 1995). This study shows that testicular spermatozoa could be retrieved in 58% of azoospermic patients with raised plasma FSH concentration (>12 IU/l) and 46% of those with testicular atrophy (<24 ml) respectively.

In contrast to other published studies (Tournaye et al., 1997), this study found that the rate of TESE significantly declines with the deterioration in testicular pathology (Pearson’s coefficient, r = 0.5), testicular volume (r = 0.5) and increase in plasma FSH concentration (r = 0.3), albeit to a limited degree. This may reflect the fact that the group with hypospermatogenesis and focal spermatogenesis had significantly larger testicular volume and lower plasma FSH concentrations compared with the group with maturation arrest and Sertoli cell-only histological patterns. This study shows that the best chance of TESE is when the histology shows hypospermatogenesis and the worst when the testicular volume is in the range of 1–12 ml. This is attributed to the fact that patients with hypospermatogenesis had relatively abundant number of spermatozoa in their testicular tissues while those with severe degrees of testicular atrophy (1–12 ml) all had Klinefelter’s syndrome which usually is incompatible with normal spermatogenesis. Some authors, however, have reported successful TESE, pregnancy and birth in a few patients with mosaic and non-mosaic forms of Klinefelter’s syndrome (Tournaye et al., 1996b; Palermo et al., 1998). The difference in testosterone concentrations between the groups may reflect the severe concomitant impairment of Leydig function in the patients with failed TESE. On the other hand, peripheral serum testosterone levels do not reflect delicate changes in Leydig cell secretory function (Steinberger et al., 1973). The very poor association of LH relative to FSH concentrations with TESE reflects the differential sensitivity of FSH and LH to gonadal steroids. The negative feedback of testosterone is more pronounced on FSH than on LH (Martin-du-Pan and Bischof, 1995).

The classification of spermiogenic arrest has become important because the outcome of microspermatid injection of oocyte is better in those with incomplete spermiogenic arrest (round, enlongating and enlongated visualized) compared to those with complete spermiogenic arrest (no development beyond the round spermatids stage) (Vanderzwalmen et al., 1997; Fishel et al., 1997; Sousa et al., 1998; Amer et al., 1998). It will therefore be of interest in future to relate complete and spermiogenic arrest observed in diagnostic biopsy with similar situations in therapeutic biopsy. However, like others (Steinberger and Tjioe, 1968; Soderstrom and Souminen, 1980; Silber and Rodriguez-Rigau, 1981; Silber et al., 1996a,b) we did not find round spermatids in the absence of long spermatids in our therapeutic biopsy. We did not classify our therapeutic biopsy at the time of this research, partly because microspermatid injection of oocytes is banned in the UK and partly because the importance of this classification was not realized at the time of this research.

A number of methods such as the use of erythrocyte lysing buffer, in-vitro culture of testicular tissues and the use of enzymatic procedures have been developed to improve the efficiency of sperm recovery from therapeutic testicular biopsy. The benefit of a combination of procedures is reflected in our high sperm retrieval rate of 70%. With a mechanical separation technique alone, the efficiency of sperm recovery is limited by damaged cells, free nuclei and erythrocytes contaminating the spermatozoa. The addition of erythrocyte lysing buffer allows clearer identification of spermatozoa because the main obstacle to finding spermatozoa in the therapeutic testicular tissue is the abundance of erythrocytes (Nagy et al., 1997) while in-vitro culture permits migration of spermatozoa from the Sertoli cells and debris and confirms their viability (Zu et al., 1995). However, there remains unsubstantiated concern that the sperm viability could be compromised during the in-vitro culture. We did not use proteolytic enzymes in our TESE procedure because of our concern that they may modify sperm membrane proteins, although recent reports suggest that they may be safe and may improve the efficiency of sperm retrieval (Crabbe et al., 1997).

In conclusion, we found a poor correlation between Johnsen score, biophysical and endocrine profiles with TESE data. This may be due to varying histological patterns not only between patients but also between different tubules in the same patient. The difference in the amount of testicular tissue used for sperm extraction and histopathology and misinterpretation of testicular histological patterns rather than failure to quantify spermatogenesis may account for the discrepancy between histological patterns and TESE. The chance of sperm retrieval increases with less impairment of testicular function, the larger the testes and the lower the plasma FSH concentrations. Visualization of testicular spermatids showed a very strong association with TESE. However, considerable overlap in these variables between the groups occurred, so that no single variable can provide a perfect discrimination between those with successful and failed TESE.
Acknowledgements

The authors are indebted to Mr Nicholas A. Taub (MSc) of the Department of Epidemiology and Public Health, University of Leices-
ter, Leicester, UK, for reviewing the statistics, and the staff of the
University Research Clinic for their help with running the
azoospermia clinic.

References

Dev., 6, 19–24.

Amer, M., Soliman, E., El-Sadek, M. et al. (1998) IS is complete spermiogenesis

Beierdorfer, H., and Schirren, C. (1979) Peculiarities and side effects of


Chil lion, M., Casals, T., Gimenez, J. et al. (1994) Analysis of the CFTR gene
con firms the high genetic heterogeneity of the Spanish population: 43
mutations account for only 78% of CF chromosomes. Hum. Genet., 93,

Anat., 112, 35–45.


Craft, I., Tsirigotis, G., Courtauld, E. et al. (1997) Testicular biopsy as an
alternative to biopsy for the assessment of spermatogenesis. Hum Reprod.,
12, 1483–1487.

human male infertility: links with ageing, mitochondria genetics, and

Fishel, S., Green, S. and Hunter, A. (1997) Human fertilisation with round

Reprod., 11, 558–560.

Hausser, R., Temple-Smith, P.D., Southwick, G.J. et al. (1995) Fertility in


Johnsen, S.G. (1970) Testicular biopsy score count—a method for registra-
tion of spermatogenesis in human testes: normal values and results in 335
hypogonadal males. Hormones, 1, 2–25.

Johnsen, S.G. (1967) The mechanisms involved in testicular degeneration in


biopsy specimens. J. Androl., 14, 94–98.

deletions in oocytes and reproductive ageing in women. Fertil. Steril., 64,
577–584.

Levy, H.S. (1979) Testicular biopsy in the study of male infertility. Its current
usefulness, histologic techniques, and prospects for the future. Hum.
Pathol., 10, 569–584.


hormone in infertile men. Is increased plasma FSH always due to damaged

Matthieu, C., Echochard, R., Bied, V. et al. (1995) Cumulative conception rate
following intrauterine artificial insemination with husband’s spermatozoa:

Mulhall, J.P., Burgess, C.M., Cunningham, D. et al. (1997) The presence of
mature sperm in testicular parenchyma of men with non obstructive

procedure for testicular biopsy specimens offers more efficient recovery:

Palermo, G.D., Schlegel, P.N., Sills, E.S. et al. (1998) Births after intracytoplas-

Reijo, R., Lee, T., Salo, P. et al. (1995) Diverse spermatogenic defects in
humans caused by Y chromosome deletions encompassing a novel RNA-

Skakkebæk, N.E. and Helli, C.G. (1973) Quantification of human seminiferous
epithelium. I. Histological studies in twenty-one fertile men with normal

Steinberger, E. and Tjoer, D.Y. (1968) A method for quantitative analysis of


biopsy: determination of partial obstruction and prediction of sperm count


resulting from testicular sperm extraction and intracytoplasmic sperm
injection for azoospermia due to maturation arrest. Fertil. Steril., 66,
110–117.

Silber, S.J., Nagy, Z., Doeyere, P. et al. (1996b) Distribution of spermatogenesis
in the testicles of azoospermic men: the presence or absence of spermatids

104, 476–84.


Steinberger, E., Root, A., Fisher, M. and Smith, K.D. (1973) The role of
Metab., 37, 746–751.

factors for successful testicular sperm recovery in azoospermic patients?
Hum. Reprod., 14, 80–86.

Tournaye, H.J., Liu, J., Nagy, Z. et al. (1996a) Correlation between testicular
histology and outcome after intracytoplasmic sperm injection using testicular

Tournaye, H., Staessen, C. and Liebaers, I. (1996b) Testicular sperm recovery
in nine 47 XXX Klinefelter's patients. Hum Reprod., 11, 1650–1653.

injection of spermatids retrieved from testicular tissue: influence of testicular
pathology, type of selected spermatids and oocyte activation. Hum. Reprod.,
12, 1203–1213.


FSH, LH and FSH: 11. Relationship to sex and age. Acta Endocrinol.,
109, 190–197.

World Health Organization (1992) Laboratory Manual for the Examination of
Human Semen and Cervical Mucus Interaction, 3rd edn. Cambridge
University Press, New York.

analysis of the seminiferous epithelium in human testicle biopsies and the