High frequency of sub-optimal semen quality in an unselected population of young men

A.G. Andersen1,5, T.K. Jensen2, E. Carlsen1, N. Jørgensen1, A.M. Andersson1, T. Krarup3, N. Keiding4 and N.E. Skakkebæk1

1Department of Growth and Reproduction, Rigshospitalet, section GR-5064, 9, Blegdamsvej, DK-2100 Copenhagen, 2Institute of Community Health, Department of Environmental Medicine, Odense University, 3Department of Urology, Aalborg Hospital and 4Department of Biostatistics, University of Copenhagen, Denmark

© European Society of Human Reproduction and Embryology

Material and methods

Group A

Young men from the general population were enrolled as they presented for the compulsory medical examination in two of the largest cities in Denmark (Copenhagen and Aalborg) in autumn 1996–spring 1997, and in Copenhagen during a second period in autumn 1997–spring 1998. Men who suffer from chronic diseases were rejected in advance, corresponding to 10–14% of the total population.

Introduction

The management of infertility problems has become an increasingly important part of health services during the past 20 years, at least in the industrialized countries. A substantial number of couples seek fertility treatment due to poor semen quality, and there is evidence in the literature that male reproductive function seems to have deteriorated considerably in the past 4–5 decades. In a meta-analysis, Carlsen et al. (1992) observed a significant decline in mean sperm concentration from 113 x 10^6/ml in 1940 to 66 x 10^6/ml in 1990, or 0.94 x 10^6/ml/year. This finding has been supported by other studies (Auger et al., 1995; Irvine et al., 1996; Swan et al., 1997), but it has also raised some controversy (Olsen et al., 1995; Fisch et al., 1996). Moreover, the trend in semen quality is probably not global, as the sperm counts of Finnish men (Suominen and Vierula, 1993) and possibly also of some American men (Fisch et al., 1996) have changed only little and have remained high. However, the incidences of other male reproductive abnormalities, such as testicular cancer (Adami et al., 1994; Forman and Møller, 1994), hypospadias, and cryptorchidism (Ansell et al., 1992) may also have increased in several Western countries during the corresponding period of time, and although the reason for these different changes is not known, they may very likely be inter-related (Skakkebæk et al., 1998). Thus, it is important that the possible causes of the deterioration of male fertility are investigated and attempts are made to prevent further decay.

So far clinical studies on semen quality have dealt with highly selected groups of men: volunteers enrolled after advertisement (Irvine et al., 1996; Paulsen et al., 1996; Lemcke et al., 1997), candidates for vasectomy (Sheriff, 1983; Fisch et al., 1996), semen donor candidates (Leto and Frensilli, 1995; Paulsen et al., 1996; Bujan et al., 1996; Swan et al., 1997), infertility patients (MacLeod and Wang, 1979; Bostofte et al., 1983; Ombelet al., 1996; Berling and Wölner-Hanssen, 1997; Andolz et al., 1999) and may therefore be flawed by selection bias. Studies of semen quality, reproductive hormones and testicular size in unselected populations of young males have not previously been reported.

Due to the military drafting system in Denmark, all 18 year old men are required to attend a compulsory medical examination to determine their fitness for military service. In collaboration with the military health board, a unique opportunity was given to study important parameters of male reproductive health, including semen quality and reproductive hormones, in a quite unbiased population of young Danish men.

Materials and methods

Group A

Young men from the general population were enrolled as they presented for the compulsory medical examination in two of the largest cities in Denmark (Copenhagen and Aalborg) in autumn 1996–spring 1997, and in Copenhagen during a second period in autumn 1997–spring 1998. Men who suffer from chronic diseases were rejected in advance, corresponding to 10–14% of the total population.
For a young man to participate in the present study he had to be born and raised in Denmark. During the autumn 1996–spring 1997 period 2166 young men fulfilled the inclusion criteria and 413 (19%) agreed to participate, and during the autumn 1997–spring 1998 period 1862 young men fulfilled the inclusion criteria and 295 (16%) agreed to participate. Physical examination, blood sampling and semen sample delivery took place at hospitals in Copenhagen and Aalborg, and study subjects received approximately £40 for their participation.

**Group B**
The relatively low participation rates obtained in group A might lead to selection bias. Therefore it was investigated whether the participants in group A were self-selected with regard to reproductive health or whether the low participation rates simply reflected a low motivation to participate in the rather demanding examination programme. Consequently, participants were enrolled according to a much less demanding protocol, which was expected to result in a higher participation rate and which would enable comparison of reproductive hormone concentrations as a measure of reproductive health. Subjects in this group (group B) were enrolled when they reported for their medical examination at the military health board during a 2 week period (November 10–21, 1997). Inclusion criteria were the same as in group A. In this group the only request was for a blood sample drawn immediately after the medical examination was completed. Blood samples were drawn in a mobile laboratory placed just outside the building in which the military health board was located. Study subjects received approximately £30 for their participation. Altogether 195 men (79%) agreed to participate. Subsequently 12 of the participants from group B agreed to participate according to protocol A and were thus included in the 295 participants in the 1997–98 cohort of group A instead of in group B. Serum concentrations of follicle stimulating hormone (FSH), inhibin B, luteinizing hormone (LH) and testosterone were compared between group A and group B.

**Semen analysis**
In group A each man provided a semen sample that was obtained by masturbation into a wide mouthed plastic container in a room close to the semen laboratory. The period of abstinence was recorded and the semen sample was analysed according to the World Health Organization’s guidelines (WHO, 1992) modified in accordance with Jørgensen et al. (1997). A minor change in the procedure of estimating the semen volume was introduced between the two group A cohorts. During the study of the first cohort, the semen volumes were estimated by weighing the plastic containers with semen samples and subsequently subtracting a standard weight of 16.1 g corresponding to a previously calculated mean weight of an empty plastic container. However, when it subsequently became clear that the weight of plastic containers varied from batch to batch (~0.5 g), all plastic containers used during the study of the second cohort were weighed before being handed out, and in this cohort the semen volumes were calculated as the weight difference between the filled plastic container and the empty container.

A previous study found inter-observer variability in semen analysis between different laboratories (Jørgensen et al., 1997), especially concerning motility assessments. In the present study only three technicians from our laboratory were involved in the semen analyses. Our technicians performed weekly intra-laboratory control by examining the same semen samples blindly and the between-technician variation in all semen parameters was less than 10% throughout the study period.

**Hormone analysis**
The concentrations of FSH, LH and testosterone were determined by time-resolved-immunofluorometric assay (DELFIA®, Wallac, Turku, Finland). Intra- and interassay coefficients of variation were below 8% in all DELFIA® assays. Inhibin B was determined using a specific enzyme immunometric assay, as described by Groome et al. (1996), with a detection limit of 18 pg/ml and intra- and interassay coefficients of 15 and 18% respectively.

**Physical examination**
All physical examinations were performed by three physicians (two in Copenhagen and one in Aalborg). Tanner stage of pubic hair and genital development, testicular volumes (determined by use of a Prader orchidometer), and the possible presence of a varicocele, a hydrocele or any genital malformation was recorded.

**Questionnaire**
Participants in group A each completed a questionnaire that was returned to the physician at the time of the physical examination. The questionnaire included information on previous history of fertility and/or genital diseases as well as other previous or current diseases. In addition, information was obtained on lifestyle factors such as smoking and drinking habits, as well as education and occupation. For participants in group B and for all non-participants, information on place and year of birth as well as education and occupation was obtained from the military health board.

**Statistical analysis**
Between-group differences in hormone concentrations, sperm motility, and period of abstinence were tested by the non-parametric Mann–Whitney test. The significance of between-group differences in sperm concentrations, semen volumes, and total sperm counts was tested by analysis of covariance controlling for period of abstinence. Analysis of covariance was performed on cubic root transformed values, as the cubic root transformation gave a normal distribution (sperm concentration and total sperm count) or near normal distribution (semen volume) of these parameters. Correlations between testsis volume, semen parameters and hormone concentrations were tested by Spearman’s rank correlation. All statistical calculations were performed in the statistics package for social sciences (SPSS) for windows, release 7.5.2.

**Results**
Results of the semen analyses are summarized for the 1996–97 cohort and for the 1997–98 cohort of group A separately, and for the two group A cohorts combined in Table I. In Table II the results of the hormone analyses are summarized for the two cohorts of group A separately, for the group A cohorts combined and for group B. The distribution of sperm concentrations (Figure 1, upper panel) was smoothly decreasing, with no indication of clearly delineated groups. In the lower panels of Figure 1, the concentrations of inhibin B and FSH are stratified according to sperm concentrations. Results of the physical examinations and questionnaires are summarized in Table III.

The place and year of birth did not differ between participants (groups A and B) and non-participants.

After controlling for period of abstinence, no difference was found in sperm concentration between men enrolled in 1996–97 and 1997–98 (see Table I). A minor difference in semen volume, and therefore in total sperm count, was observed between the two cohorts. However, this difference...
was most likely due to a slight change in the method of estimating the volume, as described in the methods section, which is supported by the fact that there were no differences in sperm motility or morphology. Furthermore, no differences were found in the concentrations of any of the reproductive hormones between the 1996–97 and 1997–98 cohorts except for a minor difference in FSH concentration (Table II), a difference which is of no clinical relevance and may be explained by the long-term (1–2 year) variation in the assay. In addition, the FSH concentrations in group B and the second cohort of group A, which were analysed during the same period, were concordant whereas the FSH concentrations in the first cohort of group A, which was analysed 1.5 years earlier, showed a small (and similar) difference with regard to both group B and the second cohort of group A.

When the two group A cohorts were combined, no significant differences in the concentrations of FSH or any of the other reproductive hormones were observed between men in group A and men in group B (see Table II).

As the two cohorts of group A were very similar with regard to both semen parameters and reproductive hormones, they are described and discussed combined in the following. Among all men in group A 25 and 48% respectively had sperm concentrations below 20 10^6/ml and 40 10^6/ml. A total of 17% had less than 30% morphologically normal spermatozoa and 15% had less than 50% motile spermatozoa which are defined as normal reference values by WHO (WHO, 1992). WHO recommends a period of abstinence of at least 48 h (WHO, 1992). A total of 521 participants in group A had a period of abstinence above 48 h and
Table III. Age, clinical findings and self-reported information (when entering the study) on fertility and reproductive health diseases in young Danish men

<table>
<thead>
<tr>
<th></th>
<th>Mean (SD)</th>
<th>Median (5–95 percentile)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>19.4 (1.3)</td>
<td>19.0 (18.5–21.8)</td>
</tr>
<tr>
<td>Clinical findings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>23.0 (3.2)</td>
<td>22.5 (19.0–28.9)</td>
</tr>
<tr>
<td>Testicular volume (ml)²</td>
<td>20.0 (4.1)</td>
<td>20.0 (13.5–26.5)</td>
</tr>
<tr>
<td>Fully virilized (%)</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>Varicocele (%)</td>
<td>10.0</td>
<td></td>
</tr>
<tr>
<td>Self-reported information on fertility and previous diseases (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Having caused pregnancy/pregnancies</td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td>Unprotected intercourse without achieving pregnancy</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>Cryptorchidism, spontaneous descensus</td>
<td>8.8</td>
<td></td>
</tr>
<tr>
<td>Cryptorchidism, treated</td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td>Varicocele</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>Testicular torsion</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>Inguinal hernia</td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td>Epididymitis</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>Gonorrhea</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Chlamydia</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>Parotitis in adulthood</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>Testicular cancer</td>
<td>0.0</td>
<td></td>
</tr>
</tbody>
</table>

BMI = body mass index.
²Mean of left and right testis, measured by use of Prader orchidometer.
³Men who were at least in Tanner stage 5 of pubic hair development.
⁴Men who had experienced repeated unprotected intercourse for at least a year without achieving a pregnancy.
⁵Men who were born with one or two undescended testicles, but whose testicle(s) descended without treatment.
⁶Men previously treated for cryptorchidism, either by hormonal treatment, surgery or both.
⁷Men previously diagnosed as having a varicocele, irrespective of whether treated or not.
⁸Men previously diagnosed as having an inguinal hernia, irrespective of whether treated or not.
⁹All these men were also included in the group who had a clinical varicocele.

The following conditions may affect semen quality: current or previously treated varicocele, current or previously treated cryptorchidism, torsion of one or both testicles, inguinal hernia, epididymitis, genital infections, parotitis in adulthood, and testicular cancer. Altogether 29% of the men in group A had currently or had previously had one or more of these diseases (Table III). The median sperm concentration (44 × 10⁶/ml) and the median total sperm count (121.6 × 10⁹) among the men with no present or previous diseases in the reproductive organs were higher than among the men who had diseases in the reproductive organs (34 ×10⁶/ml and 74.5 × 10⁹). However, even among the men who had no diseases in the reproductive organs and had a period of abstinence above 48 h (n = 381), 18 and 40% respectively had sperm concentrations below 20 and 40 ×10⁶/ml.

The correlations between testicular volume, semen quality and reproductive hormones are summarized in Table IV.
Since season of collection has been linked to sperm count (Levine et al., 1988), semen quality was also examined by month of collection. All except three men delivered the semen samples during the autumn, winter, and spring seasons. No differences were observed in sperm concentrations or total sperm counts between these three seasons.

### Discussion

Surprisingly low sperm concentrations were found in this population of 18–20 year old men: the median sperm concentration of 41 × 10⁶/ml was significantly lower than the values found in previous national and international studies (Carlsen et al., 1992; Bonde et al., 1998a). The distribution of sperm concentrations (Figure 1) did not support a hypothesis of a clearly delineated dyspermic group, rather the decrease seemed to be a general phenomenon. Initially, the study was planned to last only 1 year, but due to the finding of unexpectedly low sperm counts it was decided to extend the study with a second cohort the following year, and also to add the alternative protocol (group B) in order to validate the findings. The resulting material of semen samples from more than 700 men and concentrations of reproductive hormones from almost 900 men appears to be the most comprehensive study of reproductive health in males from the general population that has ever been published.

An important question to consider is whether selection bias may play a role, since less than 20% of the potential study populations participated in the semen studies. However, the place and year of birth of the participants as well as their educational status did not differ from the non-participants. Furthermore, it is believed that the participants are representative of the general population of 18–20 year old men in Denmark for the following reasons: (i) the men had essentially no prior knowledge of their own fertility potential and therefore this is unlikely to have affected their motivation to participate; (ii) no differences in sperm counts and biomarkers of spermatogenesis (FSH and inhibin B) were found between the 1996–97 and 1997–98 cohorts; (iii) the serum concentrations of the reproductive hormones FSH, inhibin B, LH and testosterone found in those men who delivered semen samples did not differ from the concentrations found in the participants in group B, who only had a blood sample drawn and who it is believed represent the general population due to the high participation rate (79%). As serum concentrations of both FSH and inhibin B have previously been shown to reflect spermatogenesis (Jensen et al., 1997), this suggests that the spermatogenic potential observed in the present semen quality study can be considered to reflect the status in the general population of young men.

It might be argued that this group of young men does not represent mature adults. However, as many as 97% were fully virilized with normal adult pubic hair (at least Tanner stage 5), and they had testicular sizes within the normal range. Furthermore, in a previous study (Nielsen et al., 1986) the first occurrence of spermatozoa (spermarche) in the urine was found at a median age of 13.4 years and a median Tanner stage of 2.5. Other studies have shown that adult concentrations of reproductive hormones are established before the age of 18 years (Andersson et al., 1997). In addition, preliminary results from a follow-up study of 150 of the men from the 1996–97 cohort, including semen analysis every 3 months, do not indicate any improvement in sperm concentration within a year (unpublished data). Thus, immaturity of the volunteers does not seem to explain the results.

The frequencies of current or previous genital diseases, including present varicocele (10%) and cryptorchidism with spontaneous descensus (8.8%), correspond well with published rates in the general Danish male population (Østergaard, 1971; Blom, 1984; Møller et al., 1996). This indicates that the poor semen quality is not due to a high rate of men with genital diseases.

A crucial question is whether the findings have relevance with respect to fertility. According to the WHO guidelines (WHO, 1992) a sperm concentration below 20 × 10⁶/ml is abnormal. However, these guidelines are not based on studies of fertility but are arbitrarily defined by a committee of international experts. Previous reference ranges for normal men were considerably more strict and delineated the lower range for sperm concentration at 60 × 10⁶/ml (MacLeod, 1946), but also these reference ranges were arbitrarily defined. A recent study related semen quality to fecundity (waiting time to pregnancy of a couple) (Bonde et al., 1998b) and a decreasing time to pregnancy was found with increasing sperm concentration up to approximately 40 × 10⁶/ml. It raises great concern that 40% of the men in the present study who had a
period of ejaculation abstinence above 48 h, and who had never suffered from any genital diseases, fell below this value and were thus in the range of suboptimal semen quality. Even using the current WHO guidelines, a major fraction of these individuals (18%) would be classified as having oligozoospermia. In addition, a substantial number of the participants had a relatively high number of abnormal forms of spermatozoa as well as reduced motility. It is well known, though, that great inter-observer variability in sperm motility and morphology exists between different laboratories (Jørgensen et al., 1997). In contrast, determination of sperm concentration based on the conventional haemocytometer, which was used in the present study, is a reliable method with acceptable intra- and inter-individual variation (Jørgensen et al., 1997).

It is difficult to explain the low values of sperm count in young Danish men. A short period of abstinence may contribute to a low sperm count, but even when excluding those with a period of abstinence below 48 h, more than 40% of the men had a sperm concentration below 40 ×10^6/ml. A significant correlation was also found between reproductive hormones, sperm count and testicular volume, indicating that low sperm counts were due to intrinsic biological factors rather than a short period of abstinence.

Seasonal variations in sperm concentrations, with the lowest values during the summer and peak values during late winter and spring months have previously been reported (Levine et al., 1988). Since the semen samples in the present study were collected during autumn, winter and spring seasons, the low sperm counts could not be ascribed to seasonal variation.

It remains to be seen whether these findings are generally applicable to populations of young men in the industrialized countries. Denmark seems to have relatively high rates of male reproductive abnormalities, including cryptorchidism (Ansell et al., 1992) and testicular cancer (Adami et al., 1994; Forman and Møller, 1994). The latter disease is possibly of fetal origin, and it has been postulated that an increased exposure to factors that inhibit the development of the fetal germ cells may be important for the pathogenesis of spermatogenic disorders and testicular cancer (Sharpe and Skakkebæk, 1993). According to this hypothesis, changes in semen quality and frequency of testicular cancer should be related to year of birth. Previous studies (Auger et al., 1995; Irvine et al., 1996; Bergström et al., 1996; Bonde et al., 1998a) have in fact indicated that this might be the case, and the birth cohort effect of testicular cancer seems to be stronger than the time trend effect (Bergström et al., 1996). Furthermore, in a Danish study including 1196 men from 10 cross-sectional occupational studies (Bonde et al., 1998a), the median sperm concentration fell from 63 ×10^6/ml among men born from 1937–49 to 52 ×10^6/ml among men born from 1970–74. Sperm concentrations decreased by 1.24% per year after control for confounders. The present cohort born from 1976–79 seems to fit into this pattern with a median concentration of 41 ×10^6/ml. In this connection it is interesting to compare the present findings with recent data on semen quality in men who were born approximately 10 years before the present group of participants. Firstly, in this group of men with a mean age of 28.2 years (SD 3.0) and with no knowledge of their fertility, the median sperm concentration was significantly higher (52 versus 41 ×10^6/ml) and secondly, the concentration of inhibin B, a biomarker of spermatogenesis, was also significantly higher (median 202 versus 184 pg/ml) (Jensen et al., 1997).

In conclusion, surprisingly low sperm counts were found in this study of young men. In fact, more than 40% of young adult Danish men have sperm counts below 40 ×10^6/ml, which according to a recent study (Bonde et al., 1998b) is associated with decreased fertility. Thus, the findings indicate that a high proportion of young Danish men may experience reduced fertility. Possible causes for this high frequency of suboptimal semen quality are obscure and need to be explored. Also it remains to be seen whether these findings apply to young male populations in other industrialized countries. An international collaboration to address this question has been initiated.

Acknowledgements
This study was supported by grants from the European Union Biomed 2 programme (contract number BMH4-CT96-0314), the Danish Medical Research Council (grant number 9700909 and 9700833), and the A.P. Møller and wife Chastine McKinney Møllers foundation. We are indebted to support from the physicians at the military health board Axel Elsborg and Peer Fischer. Technicians from the laboratory are greatly acknowledged for performing the semen and hormone analyses.

References
A.G. Andersen et al.


Received on August 20, 1999; accepted on October 26, 1999