

Title: Regional variations in semen quality of community-dwelling young men from Flanders are not paralleled by hormonal indices of testicular function.

Running title: Semen quality and indices of hormonal status

Authors: Willem Dhooge, Nicolas van Larebeke, Frank Comhaire, Jean-Marc Kaufman.

Institutions:

Department of Endocrinology, University Hospital Ghent, Belgium. (W Dhooge ir, Prof JM Kaufman MD, Prof F Comhaire MD); 185 De Pintelaan, B9000 Gent, Belgium.

Study Centre for Carcinogenesis and Primary Prevention of Cancer, Department of Radiotherapy, Nuclear Medicine and Experimental Cancerology, University Hospital Ghent, Belgium. (Prof N van Larebeke MD); 185 De Pintelaan, B9000 Gent, Belgium.

Correspondence to: Willem Dhooge, Department of Endocrinology, University Hospital Ghent, 185 De Pintelaan, B9000 Gent, Belgium.

E-mail: Willem.Dhooge@Ugent.be; TEL.: +32 9 240 3333; FAX.: +32 9 240 3897

Requests for reprints: Jean-Marc Kaufman, Department of Endocrinology, University Hospital Ghent, 185 De Pintelaan, B9000 Gent, Belgium.

E-mail: Jean.Kaufman@Ugent.be

Funding source:

The Flemish Environment and Health (Milieu en Gezondheid) Study was commissioned and financed by the Ministry of the Flemish Community (Brussels, Belgium). Willem Dhooge is supported by a fund from the Support Group Environment and Health, financed by the Flemish Government (Department of Science, Department of Public Health and Department of Environment, Brussels, Belgium).

1 Abstract

Background:

Epidemiological studies of sperm quality are hampered by problems such as low participation rates and poor comparability due to methodological differences in semen analysis. More objective sperm quality-related serum markers would facilitate worldwide comparison of male reproductive status. Our objectives were to investigate to what extent a set of hormonal indices of testicular function, previously established in clinical setting, can predict regional variations in seminal parameters in men from the general population.

Methods:

We recruited 101 men, aged 20-40 years, from two regions in Flanders, and assessed sperm parameters and serum reproductive hormones.

Results:

In one region compared to the other, participants had a lower sperm concentration (by 34%, $p=0.06$), total sperm count (by 41%, $p=0.02$) and sperm morphology (by 32%, $p<0.001$), which was paralleled by a significantly lower free testosterone (11%, $p=0.03$), while for total testosterone (T, 10%) and follicle stimulating hormone (FSH, 17%) the differences were non-significant (both $p=0.09$). There were no differences in inhibin B and the T/luteinizing hormone (LH) ratio, markers of testicular function. Receiver operating characteristic curve analysis demonstrated T/LH, inhibin B, and the inhibin B/FSH ratio to have significant discriminatory power between men with a sperm concentration below or above $13.5 \times 10^6/\text{mL}$.

Conclusions:

Regional variations in semen quality of community-dwelling individuals are not necessarily reflected in altered hormonal indices of testicular function and thus these markers, validated in clinical settings, fail to be substitutes for the traditional semen quality assessment in epidemiological population studies.

Key words: FSH, inhibin B, spermatogenic arrest, sperm quality, T/LH ratio.

Acknowledgements:

Fieldwork was done with the assistance of G Thys, C Van Turnhout, MP Lommaert, N Benoy, R Danckers, L Thijs and I Calders. We thank N De Clercq and S Stuyvaert for their expert technical help in semen and hormone analysis. We are grateful to D. De Bacquer for reviewing the statistics of the article and for providing helpful comments. The Flemish Environment and Health Study was commissioned and financed by the Ministry of the Flemish Community (Department of Science, Brussels, Belgium). Willem Dhooge is supported by a fund from the Support Group Environment and Health, financed by the Flemish Government (Department of Science, Department of Public Health and Department of Environment, Brussels, Belgium).

2 Introduction

Male sexual differentiation and development as well as male fertility and sexuality are under tight endocrine regulation by the hypothalamo-pituitary-testicular (HPT) axis. Any factor disturbing this axis may result in male gonadal dysfunction and reduced fertility. Conversely, primary testicular dysfunction tends to be reflected in altered hormonal parameters of HPT function. The determination of the latter hormonal parameters belongs to the standard clinical investigations for male infertility (Skakkebaek et al, 1994). Indeed, in non-obstructive azoospermia, high follicle stimulating hormone (FSH) values reflect drastic testicular failure (de Kretser et al, 1972). Similarly, in subfertile (or oligospermic) men, elevated FSH and, less distinctly, luteinizing hormone (LH) levels with decreasing sperm counts, indicate an adaptive gonadotropic reaction towards testicular impairment (Hunter et al, 1974; Purvis et al, 1975). Further, the testosterone on LH ratio (T/LH), a measure of Leydig cell responsiveness, seems to correlate positively with sperm concentration and negatively with FSH (Andersson et al, 2004a) and has been suggested to be an independent parameter of infertility in infertile men with normal FSH (Giagulli and Vermeulen, 1988). The testicular polypeptide hormone inhibin B, regulating the negative feedback of pituitary FSH secretion, reflects the functional state of the seminiferous epithelium, as was shown in normal men and men with testicular failure (Anawalt et al, 1996). The inhibin B on FSH ratio (Inh/FSH) has been suggested to be a better predictor of infertility than FSH or inhibin B separately (Andersson et al, 2004b). In contrast to the above described markers of testicular health, low to low-normal FSH and LH levels in conjunction with low testosterone suggests secondary hypogonadism (Nieschlag et al, 1999).

The complex interplay between environment, lifestyle and genetic factors, during foetal life as well as at adulthood hampers the assessment of their respective roles in male reproductive failure. Furthermore, epidemiological studies of sperm quality are rendered difficult by practical problems such as low participation rates and poor comparability of sperm analysis between laboratories (Cohn et al, 2002; Keel, 2004). In this respect, the question whether sperm quality has decreased worldwide with time in the past era of explosive growth of the

chemical industry is still a matter of debate (Carlsen et al, 1992; Sherins, 1995). In any case, regional differences in semen parameters, and/or in their decline seem to exist (Swan et al, 2000; Jørgensen et al, 2001, 2002). Whether this is paralleled by differences in reproductive endocrine status is a largely unanswered question (Jørgensen et al, 2002).

In 1999, as part of a feasibility study on biomarkers of environmental exposure and of health in humans (Flemish environment and health study, FLEHS) (Staessen et al, 2001), we recruited male adults of reproductive age living in a rural or urban region in Flanders. We observed regional differences in seminal parameters, and assessed whether these were paralleled by differences in reproductive endocrine status. Further, we examined the potential use of the above described hormonal indices of testicular function, previously validated mainly in cohorts of infertile or oligospermic men recruited through infertility clinics, in assessing the reproductive status of a sample of men from the general population. We restricted the participant's age range to 20-40 years to exclude, as much as possible, age-related endocrine alterations in the hypothalamo-pituitary-gonadal axis (Kaufman and Vermeulen, 2005). Further, we excluded men that might be occupationally exposed to testicular toxicants (Auger et al, 2001) or that commuted over long distances.

3 Materials and Methods

3.1 Geographical areas

We recruited male candidates from two regions in Flanders, the urban area of Antwerp and the rural area of Peer, during the summer period of 1999. A total of 2487 short questionnaires with an accompanying letter were mailed to young men, aged 20-40, selected randomly from the two municipal population registries. The response rate in Peer (28%) was slightly lower than that in Antwerp (33%). Of the 744 responders, 207 men were eligible after consideration of the following exclusion criteria: vasectomy, commuting over long distances, and having a job at risk to harmful conditions for male reproductive health (chemical industry, dry cleaning, industrial production, military airport). The candidates were contacted by telephone to check their willingness to participate and to provide further details regarding sperm sample collection and minimal hygienic conditions. They were asked to carefully report the hour of semen collection and the abstinence period, thereby stressing the wish to observe a 3 day abstinence period. A similar proportion of candidates in Antwerp (82%) and Peer (92%) were contacted before the number of 50 participants within each area was reached, whereby preference was given to non-smokers, and lifelong residents of the area. Overall, 60% of the contacted persons agreed to participate in the study. Reasons for non-participation were holiday (n=17), refusal (n=29) and personal reasons (n=20). Two currently smoking men were allowed to enter the study, one in each region. Due to a misclassification, an extra eligible person in Peer was included in the study (n=51). In neither areas, differences in age or having children were noted between those that were contacted and those that were not, or between those that entered the study and those that refused during the telephone conversation. A greater proportion of the eligible persons in Peer (72%) had proven fertility compared to those in Antwerp (44%), which was possibly related to the age difference in the two cohorts (34.5 and 31.6 years, respectively). All participants showed up at the investigation and gave their informed consent. The ethics committee of the Catholic University of Leuven approved the study.

3.2 Questionnaires, clinical measurements, semen collection and analysis.

Semen samples collected by masturbation were processed within 45 min at the site of investigation whereby a seminal smear was prepared by trained nurses, 2 mL of Hayem's anticoagulant and preservative solution was added, and the specimen was stored at 4°C. All samples reached the andrology department within 4 days of semen collection and technicians were unaware of their origin. Ejaculate volumes were estimated using a graduated pipette. A single technician scored sperm concentration in duplicate using disposable counting chambers (Cellvision, Heerhugowaard, The Netherlands), and sperm morphology on air dried Papanicolaou stained seminal smears. Using the morphology assessment method recommended by the 1992 World Health Organization (WHO) manual (World Health Organization, 1992), only spermatozoa with absolutely no defects were classified as normal. The total sperm count (TSC) was derived by multiplying the individual's sperm concentration and volume. The participants filled in a questionnaire to assess health, lifestyle, social class, use of tobacco and alcohol and intake of medicines, as previously reported for the adolescent cohort of the FLEHS project (Staessen et al, 2001). For every participant, weight and length were measured at the site of investigation and the body mass index (BMI) was calculated as kg/m^2 .

3.3 Hormone measurements

Venous blood was obtained and centrifuged at the site. Serum was aliquoted for determination of markers of hormone status and was stored at -20°C until analysis. Commercial immunoassays were used to determine serum levels of total testosterone (T) (Medgenix, Fleurus, Belgium), Luteinizing Hormone (LH) and Follicle Stimulating Hormone (FSH) (Roche Diagnostics, Vilvoorde Belgium), sex hormone binding globulin (SHBG; Orion Diagnostica, Espoo, Finland), total 17β -estradiol (E2) (Clinical Assay, DiaSorin s.r.l., Saluggia, Italy; adapted protocol with use of double amount of serum) and inhibin B (Serotec, Oxford, UK). The free fraction of T (fT) and E2 (fE2) was calculated from serum total T, SHBG, assuming a fixed albumin concentration using a validated equation (Vermeulen et al, 1999; Szulc et al, 2004). The intra- and interassay coefficients of variation for all assays were

less than 12%. For every individual the ratio of total testosterone on LH (T/LH) was calculated as nmol/IU, the ratio of T on E2 (T/E2) as pmol/pmol and the inhibin B on FSH ratio (inh/FSH) as ng/IU.

3.4 Statistical analyses and exclusions

Hormone levels are reported as molar units or international units (IU) expressed per liter serum except for inhibin B which is expressed as ng/L. One participant from Peer had a hormone and sperm profile suggestive of Klinefelter syndrome, and was excluded from the dataset, reducing the number in Peer to 50 participants. Normal data were described as average and standard deviation (SD). Data that were not normally distributed were described by median and interquartile range (IQR) or log-transformed and described by geometric mean (GM) and 95% confidence intervals (95%CI) in case of correction for confounding variables. We compared means and proportions across the two areas by a *Mann-Whitney U* test and Fisher's exact test, respectively. Next, applying analysis of covariance, potentially important covariates were forced into the models irrespective of statistical significance. We explored inter-hormone and hormone-sperm parameter relationships using scatter plots and Spearman's rank correlation (r). We subsequently built general linear models with inclusion of area of residence, an interaction term between the explanatory variable and the area of residence and known confounders. Odds ratio for a disorder was calculated using multiple logistic regression. Because of the non-linear relationship between sperm parameters and abstinence period, the latter was coded <2 days, between 2 and 4 days, between 4 and 6 days or higher, and corrected for in multiple linear regression. Age was included as a confounder during sperm parameter analyses (Kidd et al, 2001). Receiver Operating Characteristic (ROC) analysis was performed using Medcalc (version 8.1). All other statistical analyses were done with SPSS software (version 12.0).

4 Results

4.1 Characteristics

The characteristics of the population are presented in Table I. Participants from Peer were older and had a higher average BMI. On average, in Peer, blood was collected two hours later than in Antwerp. The groups were not different in education level. In both regions, a similar proportion of participants took medication, which had no influence on any hormone or sperm parameter, more than 80% of them felt “good” at the time of the investigation and none reported a “bad” health status. Participants with a history of serious illness or operation had a drastically lower total sperm count, before and after correction for age, completeness of the semen sample and abstinence period (GM (95%CI) 38.4 million/mL (20.3-73.0) vs 117.7 million/mL (95.1-145.8), $p=0.02$).

Sperm volume and total sperm count but not sperm concentration correlated positively with duration of abstinence (r for sperm volume: 0.35, $p<0.001$ and TSC 0.25, $p=0.01$), that was reported to be more than three days in 90% of the participants. In this relatively young population, age was not related to sperm concentration, sperm volume, total sperm count or sperm morphology. Time (minutes) between reported semen sample collection and first processing by a nurse was identical in both areas (37 min). Number of days between semen processing and laboratory investigations were significantly higher in Antwerp but this did not negatively influence measured sperm concentration, semen volume or TSC as the respective correlation coefficients were all small and non-significant. Finally, there were no differences in the reported completeness of the semen sample between the areas. Serum levels of E2, fE2, LH, FSH, SHBG and inhibin B were not related to age in contrast to T and fT (r : -0.32, $p=0.001$; -0.38, $p<0.001$ respectively). BMI correlated significantly with total testosterone (r : -0.38, $p<0.001$), but not with fT, E2, LH, FSH or any semen parameter. In a combined general linear model, testosterone declined by 3.2% (95%CI: 1.6%-4.7%, $p<0.001$), 3.4% (95%CI: 1.9%-4.9%, $p<0.001$) and 1.5% (95%CI: 0.6%-2.4%, $p=0.002$) with every later hour serum was sampled, every unit increase in BMI and every year of age respectively. The sperm and

hormone values of the two current smokers were well within the normal range of the total population.

4.2 Area differences in reproductive parameters

With regard to the threshold values of subfertility suggested by Guzick et al (2001), 16% of the men in Peer and 10% of the men in Antwerp (Fisher's exact test $p=0.6$) had a sperm concentration below <13.5 million cells/mL and 33% of the men in Peer vs 14% in Antwerp (Fisher's exact test $p=0.03$) had a sperm morphology $<9\%$ normal forms. The age adjusted odds ratio for men living in Peer to have a sperm morphology value below 9% was 3.0 (95%CI: 1.1-8.7; $p=0.04$). Differences in average values for hormone and sperm parameters, after correction for confounding variables, are presented in Table II. Uncorrected median (IQR) sperm concentration, sperm volume, total sperm count and sperm morphology in Peer vs Antwerp were respectively: 45.1 (19.5-76.9) vs 49.2 (30.9-90.4) $\times 10^6$ /mL ($p=0.18$), 2.8 (1.6-4.1) vs 3.1 (2.0-4.1) mL ($p=0.67$), 116.0 (49.2-177.7) vs 131.2 (90.2-250.0) $\times 10^6$ ($p=0.12$) and 12.0 (6.5-16.0) vs 18.0 (11.5-24.5) ($p<0.001$). Excluding those cases reporting a previous serious illness or operation, or including BMI as an extra confounding variable did not influence the statistics on regional differences in free testosterone, sperm morphology or total sperm count.

None of the participants had LH or SHBG levels outside the laboratory's reference range (1-12 IU/L and 11-71 nmol/L respectively) while in Antwerp, three persons had FSH values above 12 IU/L. In Peer, 4 persons had a free testosterone level below 0.21 nmol/L while 15 inhabitants from Peer and 5 from Antwerp had a total testosterone below the reference value of 11.1 nmol/L (odds ratio (95%CI): 3.86 (1.28-11.6); $p=0.02$). The latter area difference disappeared after correction for BMI, age and time of day blood was sampled ($p=0.5$).

In the whole population (i.e. both regions combined), sperm concentration and TSC were negatively related to FSH (r : -0.23, $p=0.02$; -0.20, $p=0.04$ respectively) and positively to inhibin B (r : 0.23, $p=0.02$ and 0.28, $p=0.005$) and the T/LH ratio (r : 0.25, $p=0.01$ and 0.20, $p=0.05$), but did not correlate with T, E2 or their free fractions. Further, inhibin B correlated negatively with FSH (r : -0.32, $p=0.001$) and positively with T (r : 0.24, $p=0.02$), that, together

with fT, correlated positively with LH (r : 0.25, $p=0.01$ and 0.24, $p=0.02$ respectively). Table III shows the correlation coefficients between hormones and between hormone and sperm parameters in the two populations. There was a general trend towards stronger correlations in Peer than in Antwerp. We further investigated a limited number of dose effect relationships using multiple linear regression including possible confounders (Table IV). Introducing area of residence and an interaction term between the latter parameter and the main explanatory variable did not reveal any relevant area differences in the investigated relationships.

Using ROC curve analysis, we determined the efficiency of FSH, inhibin B, Inh/FSH and T/LH ratio to discriminate between fertile and subfertile men according to the above described threshold values of subfertility. An area under the ROC curve (AUC) significantly larger than 0.5, and ideally close to one, indicates that the marker has the potential to discriminate between two groups. In categorizing men with a sperm concentration lower or higher than 13.5 million cells/mL, the AUC (95%CI) for FSH, Inhibin B, Inh/FSH and T/LH were 0.65 (0.55-0.74), $p=0.09$; 0.69 (0.59-0.78), $p=0.007$; 0.69 (0.59-0.78), $p=0.006$ and 0.72 (0.62-0.80), $p=0.001$ respectively. None of these hormone markers of testicular function had any significant power to discriminate between men with a sperm morphology lower or higher than 9% normal forms. In addition none of them could discriminate between men living in Peer or Antwerp, in contrast to testosterone, and, more importantly, free testosterone (AUC, 95%CI: 0.71 (0.61-0.80), $p<0.001$ and 0.74 (0.64-0.82), $p<0.001$ respectively).

5 Discussion

We investigated the reproductive capacity in a sample of community-dwelling young men from two regions in Flanders, 80 kilometer apart, with an alleged difference in patterns of lifestyle and environmental factors. We found significantly lower total sperm count and sperm morphology, that was paralleled by lower free testosterone levels in the rural area of Peer than in the industrialized city of Antwerp. However, FSH, inhibin B, Inh/FSH and T/LH, all considered to be markers of testicular function, were not different in the two populations.

Total sperm count and sperm concentration were more than 30% lower Peer than in Antwerp. The regional sperm quality measures were uniformly distributed and the observed differences could not be explained by outlying observations (data not shown). Although, in accordance with literature (Andersson et al, 2004b), inhibin B correlated significantly with both total sperm count and sperm concentration, its distribution in Peer overlapped completely with that in Antwerp (data not shown and Table II). The testicular polypeptidic heterodimer inhibin B is the physiologically relevant circulating inhibin form in men, constituting the afferent arm of the feedback loop to the pituitary, inhibiting the release of FSH (Hayes et al, 2001). At adulthood, normal inhibin B secretion is dependent on both Sertoli cell functioning and the presence of germ cells. It has been proposed that inhibin B may in fact be a joint product of Sertoli cells and germ cells of the perimeiotic pachytene to round spermatid stage, both in testes with normal spermatogenesis and with spermatogenic arrest (Andersson et al, 1998; Marchetti et al, 2003). Partial spermatogenic arrest at or beyond the stage of meiosis would thus lead to decreased sperm output, without significantly affecting inhibin B levels, as has been argued before (Andersson et al, 1998; von Eckardstein et al, 1999).

High T levels in the interstitial and seminal tubule of the testes are indispensable for succesfull spermatogenesis, (Sharpe, 1994). Total and free testosterone were slightly lower in Peer versus Antwerp. Although the peripheral testosterone level cannot be used to infer the intratesticular levels (Maddocks et al, 1993), it seems improbable that the slight decrease in serum testosterone levels in Peer is reflective of a drastically lower intratesticular

testosterone concentration that would be needed to, by itself, cause a more than 30% reduction in semen quality.

The importance of FSH action on sperm production has long been a matter of debate. The clinical and experimental evidence supports however a theoretical paradigm in which FSH exerts a stimulatory effect on both quantitative as well as qualitative aspects of the testosterone driven spermatogenesis (Sharpe 1994; Moudgal and Sairam, 1998; Nieschlag et al, 1999; Plant and Marshall, 2001). This hormonal synergism implies that relatively moderate alterations in testosterone and FSH can have profound effects on the spermatogenetic cycle, as has been suggested before for testosterone (Maddocks et al, 1993). In this respect it is important to note that FSH was 17% lower in Peer than in Antwerp, although this did not reach the $p=0.05$ level of significance. FSH has been described to affect spermiogenesis (Tesarik et al, 1998; Moudgal and Sairam, 1998), but how this relates to a possible effect on the traditional sperm morphology parameters remains unknown. Irrespective of the exact physiological mechanism, the observations of Skakkebaek et al (1973) of an increased level of spermatogenic arrest at the meiotic stage in infertile men presenting with both low sperm counts and poor morphology, compared to controls with good semen quality suggest that in this disorder both sperm quality measures go hand in hand, as was the case in the present study.

Jørgensen et al (2002) reported on an east-west gradient in better sperm quality in the Nordic Baltic area that was not paralleled by differences in FSH, LH or testicular hormones, including inhibin B and could not be explained by differences in confounding factors. However, the population that was recruited in these four countries were relatively young men, aged 18-19 years, an age category that might be less appropriate to detect possible subtle hormonal differences. In a group of men aged 20-35 years, Jensen et al (1997) found a significant regional difference in sperm concentration that was accompanied by a higher inhibin B and a lower FSH level in the men with better sperm quality, suggestive of differences in testicular functioning in the two cohorts. Our data do not indicate a testicular

factor causing the observed lower semen quality status in the population in Peer. The T/LH ratio, a marker of Leydig cell sensitivity (Andersson et al, 2004a) was similar in both cohorts. Further, an undefined testicular threat leading to impaired testosterone secretion logically would have resulted in increased compensatory LH levels, which was not seen in this study. Together with the decreased FSH levels in Peer compared to those in Antwerp, these findings however suggest the possibility of a partial spermatogenic arrest at the spermatid level related to altered pituitary function (Martin-du Pan and Campana, 1993; Morrow et al, 1986). It might be suggested that an interfering factor at the pituitary level would logically also have resulted in lower LH levels in Peer. However, differences in LH between groups are generally more difficult to assess due to the high intra-individual variation of LH secretion. We found a tendency towards stronger correlations (Table III) in the region with the poorer sperm characteristics. However, when further investigating a number of relevant relationships between hormones and semen parameters (Table IV), no significant regional differences in the dose-effect curves were seen. Nevertheless, our data are in line with those of Andersson et al (2004a) who reported similar relationships to be stronger in 357 idiopathic infertile men compared to 318 men with proven fertility. Uhler et al (2003) presented graphs from which strong associations of FSH with sperm concentration below 40 million/mL and sperm counts below 100 million/ejaculate can be deduced, with apparently less pronounced relationships at higher semen values in normal couples. Finally, our data are in accordance with those of Morrow et al (1986) who reported a positive relationship between testosterone and LH in people with mild hypospermatogenesis and germ cell arrest, which was different from those with severe forms of seminiferous tubule failure, a finding that has been corroborated by other studies (Aafjes et al, 1977; Giagulli and Vermeulen, 1988).

Jensen et al (1997) found in two samples from the general population that an FSH above 10IU/L combined with an inhibin B below 80 pg/mL had a 100% predictive power for having a sperm concentration below 20 mill/mL. Only 8 persons in our population had an FSH level above 10 IU/L, three of which were oligospermic while none of the participants had an inhibin

B below 80 pg/mL. However, Vernaeva et al (2002) concluded that inhibin B, alone or in combination with serum FSH, failed to predict the presence of sperm in men with non-obstructive azoospermia undergoing testicular sperm extraction. The Inh/FSH ratio has been described to be a better prognostic factor for infertility than either inhibin or FSH alone (Andersson et al, 2004b). Using ROC curve analysis we found that inhibin B and the Inh/FSH ratio indeed had some power to discriminate between people with a sperm concentration below 13.5 mill/mL which is a threshold of subfertility suggested by Guzick et al (2001). In addition, in our population T/LH was a better discriminator of subfertility than Inh/FSH. Although the existence of an association of lesions of the seminiferous tubules with Leydig cell dysfunction has been questioned (Ruder et al, 1974), T/LH, seems to correlate positively with sperm concentration and negatively with FSH (Andersson et al, 2004a), as was also found in this study. Further, this hormone ratio has been suggested to be an independent parameter of infertility in infertile men with normal FSH (Giagulli and Vermeulen, 1988). However, none of these markers of testicular function could discriminate between men living in Peer or Antwerp, in contrast to testosterone, and, more importantly, free testosterone.

We could not avoid the presence of area differences in a number of possibly important confounding variables in our study and corrected for these in regression analysis, where appropriate. Men from Peer were slightly older than those from Antwerp. Older age is associated with decreased levels of free testosterone not adequately compensated by a rise in LH secretion, but such age-related changes are usually seen mainly after the age of 50 years (Kaufman and Vermeulen, 2005). Moreover, the neuroendocrine regulation of FSH is better preserved, meaning that FSH generally rises with aging (Kaufman and Vermeulen, 2005), which contrasts with the lower FSH values in Peer.

It is thought that BMI lowers testosterone levels through an insulin mediated decrease in SHBG production (Kaufman and Vermeulen, 2005). However, although BMI was significantly higher in Peer than in Antwerp, SHBG levels were similar. Fat tissue possesses aromatase

activity, usually measured by the T/E2 ratio, and obese men have been described to suffer from a mild hypogonadotropic hypogonadism, possibly related to increased estrogen related negative pituitary feedback (Strain et al, 1982; Vermeulen et al, 1993). The T/E2 was lower in Peer which could be attributed to the decreased testosterone levels as estrogen levels did not differ. Further, BMI did not affect free testosterone or any sperm parameter, precluding BMI as a major explanatory factor for the observed testosterone and semen differences.

In Peer, the men were investigated on an average two hours later than in Antwerp. We accounted for differences in sampling hour using linear regression analyses, which is justifiable as for all hormones with diurnal secretory pattern, the decline during the time frame we operated, is generally monotonic linear (Ahokoski et al, 1998; Carlsen et al, 1999). Moreover, the decline in testosterone concentration of 3.2 % per hour is identical as the value reported by others (Andersson et al, 2004a).

Although not unusual for this type of study, an initial overall response rate of 30% may seem low, and thus our study is at risk of selection bias, reducing the representativeness of these sperm and hormone values for the total Belgian rural or urban population (Cohn et al, 2002). In particular, one cannot exclude the possibility that the observed differences are due to a differential recruitment in both areas. Nevertheless, if so, one would expect, in view of our results, an over-representation of infertile men in Peer (Larsen et al, 1998). Whereas this contrasts with the higher percentage of people from this rural area with proven fertility compared to Antwerp, it does not exclude that the responders in Peer had an increased awareness of impaired fertility and thus were more willing to participate in a study without incentive (Larsen et al, 1998; Muller et al, 2004). In any case, even if confounding factors have contributed to the observed regional semen quality differences, this does not alter the main conclusion that these differences are poorly reflected in the hormonal markers of testicular function.

It could be argued that the small sample size of the present study might compromise the development of relationships between semen parameters and hormonal indices. The 28%

regional difference in sperm concentration of 349 Danish men that was reported by Jensen et al (Jensen et al, 1997) was accompanied by a 28% lower average FSH level, and a 9% higher inhibin B in the 187 men with better sperm quality. In view of the more than 30% difference in semen quality we found, post-hoc analyses revealed that the present sample size of 50 men in each region yielded a statistical power above 80% to detect between-cohort differences of 10% in inhibin B and 30% in all other hormones measured including the T/LH ratio. Therefore, it would seem that in the present study meaningful differences in hormone values, if present, would most likely have been detected, further questioning the reliability of the investigated markers of testicular functioning as general predictors of spermatogenesis.

In conclusion, we found a more than 30% regional difference in sperm quality measures, that was accompanied by lower testosterone and FSH levels, but was not paralleled by altered inhibin B levels or T/LH and Inh/FSH ratios which are considered to reflect testicular health. These results indicate that serum indices of testicular function or spermatogenesis that have been validated for use in clinical settings do not seem reliable surrogates for semen analysis in population studies.

6 References

Aafjes JH, van der Vijver JC, Docter R, Schenck PE. Serum gonadotrophins, testosterone and spermatogenesis in subfertile men. *Acta Endocrinol (Copenh)*. 1977;86:651-658.

Ahokoski O, Virtanen A, Huupponen R, Scheinin H, Salminen E, Kairisto V, Irjala K. Biological day-to-day variation and daytime changes of testosterone, follitropin, lutropin and oestradiol-17beta in healthy men. *Clin Chem Lab Med*. 1998;36:485-491.

Anawalt BD, Bebb RA, Matsumoto AM, Groome NP, Illingworth PJ, McNeilly AS, Bremner WJ. Serum inhibin B levels reflect Sertoli cell function in normal men and men with testicular dysfunction. *J Clin Endocrinol Metab*. 1996;81:3341-3345.

Andersson AM, Jørgensen N, Frydelund-Larsen L, Rajpert-De Meyts E, Skakkebaek NE. Impaired Leydig cell function in infertile men: a study of 357 idiopathic infertile men and 318 proven fertile controls. *J Clin Endocrinol Metab*. 2004a;89:3161-3167.

Andersson AM, Müller J, Skakkebaek NE. Different roles of prepubertal and postpubertal germ cells and Sertoli cells in the regulation of serum inhibin B levels. *J Clin Endocrinol Metab*. 1998;83:4451-4458.

Andersson AM, Petersen JH, Jørgensen N, Jensen TK, Skakkebaek NE. Serum inhibin B and follicle-stimulating hormone levels as tools in the evaluation of infertile men: significance of adequate reference values from proven fertile men. *J Clin Endocrinol Metab*. 2004b;89:2873-2879.

Auger J, Eustache F, Andersen AG, Irvine DS, Jørgensen N, Skakkebaek NE, Suominen J, Toppari J, Vierula M, Jouannet P. Sperm morphological defects related to environment, lifestyle and medical history of 1001 male partners of pregnant women from four European cities. *Hum Reprod.* 2001;16:2710-2717.

Carlsen E, Giwercman A, Keiding N, Skakkebaek NE. Evidence for decreasing quality of semen during past 50 years. *BMJ.* 1992;305:609-613.

Carlsen E, Olsson C, Petersen JH, Andersson AM, Skakkebaek NE. Diurnal rhythm in serum levels of inhibin B in normal men: relation to testicular steroids and gonadotropins. *J Clin Endocrinol Metab.* 1999;84:1664-1669.

Cohn BA, Overstreet JW, Fogel RJ, Brazil CK, Baird DD, Cirillo PM. Epidemiologic studies of human semen quality: considerations for study design. *Am J Epidemiol.* 2002;155:664-671.

de Kretser DM, Burger HG, Fortune D, Hudson B, Long AR, Paulsen CA, Taft HP. Hormonal, histological and chromosomal studies in adult males with testicular disorders. *J Clin Endocrinol Metab.* 1972;35:392-401.

Giagulli VA and Vermeulen A. Leydig cell function in infertile men with idiopathic oligospermic infertility. *J Clin Endocrinol Metab.* 1988;66:62-67.

Guzick DS, Overstreet JW, Factor-Litvak P, Brazil CK, Nakajima ST, Coutifaris C, Carson SA, Cisneros P, Steinkampf MP, Hill JA, Xu D, Vogel DL. Sperm morphology, motility, and concentration in fertile and infertile men. *N Engl J Med.* 2001;345:1388-1393.

Hayes FJ, Pitteloud N, DeCruz S, Crowley WFJr, Boepple PA. Importance of inhibin B in the regulation of FSH secretion in the human male. *J Clin Endocrinol Metab.* 2001;86:5541-5546.

Hunter WM, Edmond P, Watson GS, McLean N. Plasma LH and FSH levels in subfertile men. *J Clin Endocrinol Metab.* 1974;39:740-749.

Jensen TK, Andersson AM, Hjollund NH, Scheike T, Kolstad H, Giwercman A, Henriksen TB, Ernst E, Bonde JP, Olsen J, McNeilly A, Groome NP, Skakkebaek NE. Inhibin B as a serum marker of spermatogenesis: correlation to differences in sperm concentration and follicle-stimulating hormone levels. A study of 349 Danish men. *J Clin Endocrinol Metab.* 1997;82:4059-4063.

Jørgensen N, Andersen AG, Eustache F, Irvine DS, Suominen J, Petersen JH, Andersen AN, Auger J, Cawood EH, Horte A, Jensen TK, Jouannet P, Keiding N, Vierula M, Toppari J, Skakkebaek NE. Regional differences in semen quality in Europe. *Hum Reprod.* 2001;16:1012-1019.

Jørgensen N, Carlsen E, Nermoen I, Punab M, Suominen J, Andersen AG, Andersson AM, Haugen TB, Horte A, Jensen TK, Magnus O, Petersen JH, Vierula M, Toppari J, Skakkebaek NE. East-West gradient in semen quality in the Nordic-Baltic area: a study of men from the general population in Denmark, Norway, Estonia and Finland. *Hum Reprod.* 2002;17:2199-2208.

Kaufman JM and Vermeulen A. The decline of androgen levels in elderly men and its clinical and therapeutic implications. *Endocr Rev.* 2005;26:833-876.

Keel BA. How reliable are results from the semen analysis? *Fertil Steril.* 2004;82:41-44.

Kidd SA, Eskenazi B, Wyrobek AJ. Effects of male age on semen quality and fertility: a review of the literature. *Fertil Steril.* 2001;75:237-248.

Larsen SB, Abell A, Bonde JP. Selection bias in occupational sperm studies. *Am J Epidemiol.* 1998;147:681-685.

Maddocks S, Hargreave TB, Reddie K, Fraser HM, Kerr JB, Sharpe RM. Intratesticular hormone levels and the route of secretion of hormones from the testis of the rat, guinea pig, monkey and human. *Int J Androl.* 1993;16:272-278.

Marchetti C, Hamdane M, Mitchell V, Mayo K, Devisme L, Rigot JM, Beauvillain JC, Hermand E, Defossez A. Immunolocalization of inhibin and activin alpha and betaB subunits and expression of corresponding messenger RNAs in the human adult testis. *Biol Reprod.* 2003;68:230-235.

Martin-du Pan RC and Campana A. Physiopathology of spermatogenic arrest. *Fertil Steril.* 1993;60:937-946.

Morrow AF, Baker HW, Burger HG. Different testosterone and LH relationships in infertile men. *J Androl.* 1986;7:310-315.

Moudgal NR and Sairam MR. Is there a true requirement for follicle stimulating hormone in promoting spermatogenesis and fertility in primates? *Hum Reprod.* 1998;13:916-919.

Muller A, De La Rochebrochard E, Labbé-Declèves C, Jouannet P, Bujan L, Mieuxset R, Le Lannou D, Guerin JF, Benchaib M, Slama R, Spira A. Selection bias in semen studies due to self-selection of volunteers. *Hum Reprod.* 2004;19:2838-2844.

Nieschlag E, Simoni M, Gromoll J, Weinbauer GF. Role of FSH in the regulation of spermatogenesis: clinical aspects. *Clin Endocrinol (Oxf)*. 1999;51:139-146.

Plant TM and Marshall GR. The functional significance of FSH in spermatogenesis and the control of its secretion in male primates. *Endocr Rev.* 2001;22:764-786.

Purvis K, Brenner PF, Landgren BM, Cekan Z, Diczfalusy E. Indices of gonadal function in the human male. I. Plasma levels of unconjugated steroids and gonadotrophins under normal and pathological conditions. *Clin Endocrinol (Oxf)*. 1975;4:237-246.

Ruder HJ, Loriaux DL, Sherins RJ, Lipsett MB. Leydig cell function in men with disorders of spermatogenesis. *J Clin Endocrinol Metab.* 1974;38:244-247.

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

Sherins RJ. Are semen quality and male fertility changing? *N Engl J Med.* 1995;332:327-328.

Skakkebaek NE, Bryant JI, Philip J. Studies on meiotic chromosomes in infertile men and controls with normal karyotypes. *J Reprod Fertil.* 1973;35:23-36.

Skakkebaek NE, Giwercman A, de Kretser D. Pathogenesis and management of male infertility. *Lancet*. 1994;343:1473-1479.

Staessen JA, Nawrot T, Den Hond E, Thijs L, Fagard R, Hoppenbrouwers K, Koppen G, Nelen V, Schoeters G, Vanderschueren D, Van Hecke E, Verschaeve L, Vlietinck R, Roels HA. Renal function, cytogenetic measurements, and sexual development in adolescents in relation to environmental pollutants: a feasibility study of biomarkers. *Lancet*. 2001;357:1660-1669.

Strain GW, Zumoff B, Kream J, Strain JJ, Deucher R, Rosenfeld RS, Levin J, Fukushima DK. Mild Hypogonadotropic hypogonadism in obese men. *Metabolism*. 1982;31:871-875.

Swan SH, Elkin EP, Fenster L. The question of declining sperm density revisited: an analysis of 101 studies published 1934-1996. *Environ Health Perspect*. 2000;108:961-966.

Szulc P, Claustrat B, Munoz F, Marchand F, Delmas PD. Assessment of the role of 17beta-oestradiol in bone metabolism in men: does the assay technique matter? The MINOS study. *Clin Endocrinol (Oxf)*. 2004;61:447-457.

Tesarik J, Guido M, Mendoza C, Greco E. Human spermatogenesis *in vitro*: respective effects of follicle-stimulating hormone and testosterone on meiosis, spermiogenesis, and Sertoli cell apoptosis. *J Clin Endocrinol Metab*. 1998;83:4467-4473.

Uhler ML, Zinaman MJ, Brown CC, Clegg ED. Relationship between sperm characteristics and hormonal parameters in normal couples. *Fertil Steril*. 2003;79:1535-1542.

Vermeulen A, Kaufman JM, Deslypere JP, Thomas G. Attenuated luteinizing hormone (LH) pulse amplitude but normal LH pulse frequency, and its relation to plasma androgens in hypogonadism of obese men. *J Clin Endocrinol Metab.* 1993;76:1140-1146.

Vermeulen A, Verdonck L, Kaufman JM. A critical evaluation of simple methods for the estimation of free testosterone in serum. *J Clin Endocrinol Metab.* 1999;84:3666-3672.

Vernaev V, Tournaye H, Schiettecatte J, Verheyen G, Van Steirteghem A, Devroey P. Serum inhibin B cannot predict testicular sperm retrieval in patients with non-obstructive azoospermia. *Hum Reprod.* 2002;17:971-976.

von Eckardstein S, Simoni M, Bergmann M, Weinbauer GF, Gassner P, Schepers AG, Nieschlag E. Serum inhibin B in combination with serum follicle-stimulating hormone (FSH) is a more sensitive marker than serum FSH alone for impaired spermatogenesis in men, but cannot predict the presence of sperm in testicular tissue samples. *J Clin Endocrinol Metab.* 1999;84:2496-2501.

World Health Organization (1992) WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction. Cambridge University Press, Cambridge, 94-98.

Table I: Characteristics of the population

	Peer (n=50)	Antwerp (n=50)	p-value*
Age, years	34.2 (4.8)	30.9 (6.3)	0.004
Anthropometrics			
Body weight, kg	80.5 (12.7)	79.7 (11.9)	0.77
Body height, cm	175.2 (4.7)	180.8 (7.2)	<0.001
BMI, kg/m ²	26.2 (3.8)	24.3 (3.1)	0.009
Self reported information			
Median residence period, years (IQR)	10.3 (5.2-25.8)	13.5 (6.8-24.8)	0.56
Education, n (%)			
workers	20 (44%)	15 (32%)	
middle class	22 (48%)	25 (53%)	
Educated professionals	4 (9%)	7 (15%)	0.43
Earlier serious illness or operation, n (%)	8 (16%)	3 (6%)	0.20
Taking medication, n (%)	16 (32%)	14 (28%)	0.83
Self appraisal of health status, n (%)			
good	43 (86%)	41 (82%)	
intermediate	7 (14%)	9 (18%)	0.79
Men having children	35 (68.6%)	22 (44%)	0.02
Men with complete semen sample, n (%)	46 (92%)	43 (86%)	0.53

Median abstinence period, days (range)	4 (1-15)	3.0 (0-7)	0.02
Median time between semen sample collection and laboratory analysis, days (range)	4 (3-7)	5.0 (4-12)	0.001
Blood collection time, hr:min	16:04 (3:10)	14:11 (3:31)	0.006

Data are mean (SD) unless stated otherwise. * p-value median: Mann Whitney U test; p value mean: T-test; categories p value: Fisher's Exact Test; † calculated for those that reported consuming alcohol; ‡ (n=84).

Table II: Cohort differences in semen parameters and hormone values, corrected for confounding variables.

Sperm Characteristics*	Peer	Antwerp	% difference	P value
Sperm concentration (x10 ⁶ /mL)	32.4 (24.0-43.8)	49.2 (36.5-66.5)	34.1	0.06
Sperm volume (mL)	2.5 (2.1-2.9)	2.8 (2.3-3.3)	10.7	0.40
Total sperm count (x 10 ⁶)	79.9 (59.1-108.0)	135.7 (100.4-183.4)	41.1	0.02
Sperm morphology (%)	11.9 (1.0)	17.5 (1.0)	32.0	<0.001
Hormone values and ratios				
Inhibin B (ng/L)†	230.7 (9.0)	225.9 (8.9)	0.8	0.72
FSH (IU/L)§	4.0 (3.5-4.7)	4.8 (4.2-5.6)	16.7	0.09
LH (IU/L) ‡	4.0 (3.6-4.5)	4.3 (3.8-4.8)	7.0	0.38
Testosterone (nmol/L)†	13.3 (12.3-14.4)	14.8 (13.6-16.0)	10.1	0.09
17β-Estradiol (pmol/L)†	65.1 (59.9-70.8)	70.3 (64.7-76.3)	7.4	0.22
SHBG (nmol/L)†	23.2 (20.8-25.8)	23.8 (21.4-26.4)	2.5	0.75
Free testosterone (nmol/L)†	0.31 (0.29-0.33)	0.35 (0.33-0.38)	11.4	0.03
Free 17β-estradiol (pmol/L)†	1.26 (1.15-1.37)	1.35 (1.24-1.48)	6.7	0.26
Inhibin B/FSH ratio (ng/IU)†	54.7 (45.1-66.2)	45.7 (37.8-55.2)	19.7	0.20
Testosterone/LH ratio (nmol/IU) †	3.3 (2.9-3.8)	3.5 (3.0-4.0)	5.7	0.62
Testosterone/Estradiol ratio†	218.8 (8.9)	224.9 (8.8)	2.7	0.65

Data are geometric mean (95%CI) or arithmetic mean (SEM), for each parameter, bottom rows are corrected values: * corrected for age, abstinence period (≤ 2 days, 2-4 days, 4-6 days and > 6 days) and reported completeness of semen sample collection; sperm

morphology was only corrected for age. † corrected for age, BMI and blood sampling time; § corrected for age and blood sampling time

Table III: Spearman's rank correlation coefficient (p-value) between semen parameters and hormone values in Peer and Antwerp.

	region	Sperm morph.*	Sperm conc.	TSC	LH	FSH
T	Peer	0.15 (0.31)	-0.04 (0.79)	0.01 (0.93)	0.29 (0.04)	0.17 (0.24)
	Antwerp	0.08 (0.60)	0.08 (0.56)	0.27 (0.05)	0.19 (0.20)	0.10 (0.49)
LH	Peer	-0.23 (0.11)	-0.32 (0.03)	-0.13 (0.37)		0.46 (<0.01)
	Antwerp	0.15 (0.29)	-0.09 (0.53)	-0.01 (0.94)		0.26 (0.07)
FSH	Peer	-0.07 (0.62)	-0.32 (0.02)	-0.34 (0.01)	0.46 (<0.01)	
	Antwerp	0.11 (0.43)	-0.21 (0.14)	-0.16 (0.26)	0.26 (0.07)	
Inhibin B	Peer	-0.03 (0.86)	0.26 (0.07)	0.30 (0.04)	-0.10 (0.48)	-0.44 (<0.01)
	Antwerp	-0.10 (0.48)	0.17 (0.25)	0.26 (0.07)	0.01 (0.94)	-0.27 (0.06)
T/LH ratio	Peer	0.32 (0.03)	0.32 (0.02)	0.16 (0.28)		-0.35 (0.01)
	Antwerp	-0.14 (0.35)	0.14 (0.34)	0.21 (0.15)		-0.16 (0.26)

*n=98. TSC= Total Sperm Count; LH= Luteinizing Hormone; FSH=Follicle Stimulating Hormone; T=Testosterone.

Table IV: Dose effect relationships between hormone levels or between hormone and sperm quality measures in the whole population.

Semen parameters*	Explanatory variable†	% change‡	95%C (%)	P
Sperm concentration	T/LH	57.0	13.4-117.2	0.007
Total sperm count	T/LH	46.8	6.2-103.0	0.02
Sperm morphology	T/LH	15.8	-3.9-39.5	0.12
Hormones§				
FSH	Inhibin B	-44.6	29.8-56.4	<0.001
Testosterone	LH	10.7	1.0-21.2	0.03
Inhibin B	Total sperm count	6.1	2.6-9.7	0.001
	Sperm concentration	5.6	2.1-9.2	0.002

*corrected for age; † T/LH: testosterone on luteinizing hormone (LH) ratio; ‡ percentage change in the dependent variable associated with a two fold increase in the independent variable; § corrected for BMI, age and sampling hour.