Is there a declining trend in ovarian function among infertility clinic patients?

A. Sobek Jr1,4, E. Tkadlec2,3, B. Hladíková1, and A. Sobek1

1Fertimed, Infertility Centre, 77200 Olomouc, Czech Republic 2Department of Ecology and Environmental Sciences, Palacky University Olomouc, 77200 Olomouc, Czech Republic 3Institute of Vertebrate Biology, Academy of Sciences of the Czech Republic, 675 02 Studenec, Czech Republic

4Correspondence address. E-mail: ales.sobek@seznam.cz

BACKGROUND: There is a growing body of evidence that testicular function has decreased rapidly over the last 50 years. However, much less is known about corresponding trends in ovarian function. Herein, we examine the temporal changes in ovarian function in a large sample of infertile patients from the Czech Republic over a period of 14 years.

METHODS: In a retrospective study, we analysed a large body of data from women, 20–40 years of age, undergoing IVF/ICSI treatment between 1995 and 2008. We defined ovarian function using five variables: basal FSH level, estradiol (E2) level on the day of HCG administration, dose of gonadotrophins used for ovarian stimulation, number of retrieved oocytes and dose of gonadotrophins per oocyte. Controlling simultaneously for temporal changes in patient age and stimulation protocol, we applied generalized additive models to describe the temporal trends.

RESULTS: During the study period the mean age of the study population increased by 2.7 years. Whereas the basal FSH and gonadotrophin dose did not change over time, the E2 level and oocyte retrieval declined, and the dose of FSH per oocyte increased during the study period.

CONCLUSIONS: The results are indicative of a small, but detectable decrease in ovarian function over a period of 14 years, which is not causally related to the ageing population.

Key words: ovarian function / decrease / infertility / IVF

Introduction

Changes in lifestyle in the industrialized world have been accompanied by a remarkable decline in fertility rates (Pearce et al., 1999; Jensen et al., 2004). There is little doubt that demographic shifts in modern societies are grounds for such a decline (Lutz, 2006). Although lifespan is increasing, fertility rates have declined to unprecedented low levels with a contemporaneous increase in age at first childbirth. As a result, the population age structure is getting older, including women who are treated for infertility (van den Bergh et al., 2005; Maheswari et al., 2008; Sobek et al., 2008). The adverse effect of increased age on fertility is well-documented for both women (Menken et al., 1986; van Noord-Zaadstra et al., 1991; Sobek et al., 2008) and men (Carlson et al., 1992; Adamopoulos et al., 1996). The poor quality of semen and the older age of women planning pregnancies are therefore reasonably considered to be factors contributing to the declining reproductive rates in the population.

Although there has been interest in documenting a steady decline in testicular function for a long time (Nelson and Bunge, 1974), the corresponding temporal changes in ovarian function have rarely been investigated, and when studied, the observations have been indirect by analysing data on declining fertility rates (Jensen et al., 2008). This may be partly because of the difficulties in defining a direct marker for ovarian function similar to semen quality in males, and partly because of continuous changes in stimulation protocols precluding direct comparisons with earlier data. Population aging and delaying reproduction to later ages are additional factors masking time trends in ovarian performance. Consequently, we have no information on ongoing changes in ovarian function independent of those caused by age.

We examined the temporal change in ovarian response to infertility treatment by analysing data on five ovarian variables measured over a period of 14 years. Although the study period may not seem sufficiently long relative to the generation interval, the Czech population...
has undergone a dramatic demographic change in that period characterized by a 5-year shift in maternal age at first birth. Controlling statistically for the concomitant effect of age and stimulation protocol type, we present evidence which is at least suggestive of such an ongoing decline in the ability of women to conceive.

Materials and Methods

Data

In a retrospective study, we analysed data on IVF/ICSI cycles obtained from patients in a central and northern Moravian region of the Czech Republic between 1995 and 2008. During the 14 year period, there have been remarkable demographic changes in this region. The average age of the population increased from 36.8 to 40.3 years and the mean age of the mother at the time of first childbirth rose from 22.7 to 27.3 years (Statistical Yearbooks published annually by the Czech Statistical Office). Early in the period, the total fertility rate (the average number of children that would be born to a woman over her lifetime) decreased slightly from 1.278 in 1995 to 1.133 in 1999 and then slowly increased to 1.438 in 2007. Life expectancy for women steadily increased over the period from 76.6 years in 1995 to 79.9 years in 2007.

As ovarian function cannot be assessed directly by measuring one parameter, we performed separate analyses with five operational variables, which collectively define ovarian function: (i) basal FSH on Day 3 of the menstrual cycle (FSH), (ii) the dose of gonadotrophins used for ovarian stimulation, (iii) estradiol levels on the day of HCG administration (E2), (iv) the number of collected oocytes and (v) the ratio of gonadotrophin dose per one oocyte recovered derived from variables (ii) and (iv).

In each patient, we recorded the age on the day of oocyte collection and measured the five variables as given above. The serum FSH concentrations were measured using commercially available immunoradiometric assay (IRMA) kits (Immunotech, Praha, CZ). The FSH assay was standardized against the WHO second international standard 94/632 reference material. The sensitivity of the FSH assay was 0.2 IU/l and the intra- and inter-assay CVs were 2.6 and 6.3%, respectively. Estradiol concentrations were measured using an immunnoassay (Advia Centaur; Bayer, Leverkusen, DE). The sensitivity of the E2 assay was 37 pmol/l and the intra- and inter-assay CVs were 9.7 and 10.2%, respectively.

All patients came from the same centre and region and were treated by the three clinicians following the same clinical and therapeutic guidelines. There have been few changes in the concept of stimulation protocols, but none were crucial. A long agonist protocol was primarily used (51.7–61.7%) during the entire period, with the exception of during 2000 and 2001, when a short agonist protocol was preferred (long agonist protocol was 32–34%). The protocol with antagonists of GnRH has been used since 2002 in about 15–40% of the cycles in low responders and patients >35 years of age. Recombinant FSH (rFSH) rather than urinary FSH (uFSH) has increasingly been used since 1998, rising from 8.4 to 50% of the patients. The mean age over the study period was 29.7 years (SD, 4.24), ranging from 20 to 40 years. The sample did not include oocyte donation cycles.

Statistical analysis

To avoid issues of non-independence of data, we analysed only the first cycle from each of the 3658 patients. Because biological functions are inherently non-linear, we analysed the variation in response variables by fitting generalized additive models (GAMs), including the continuous variables [age and time (both measured in years)] as predictors. As the type of stimulation protocol could affect the E2 response, gonadotrophin dose, number of oocytes and dose per oocyte, we first homogenised the sample by discarding patients stimulated by protocol with antagonist GnRH. The remaining 2760 patients were stimulated by four types of protocol: (i) long protocol combined with rFSH (LR), (ii) long protocol combined with uFSH (LU), (iii) short protocol combined with rFSH (SR) and (iv) short protocol combined with uFSH (SU). To account for variations associated with different protocols, we included in the GAMs a third categorical variable with four levels (LR, LU, SR, SU). Hence, the concomitant effects of age and protocol variations that could influence the estimate of temporal trend were statistically removed. By excluding patients stimulated by antagonists, we removed data from patients with the poorest ovarian function in the sample who were treated by a different protocol type early in the study period. Thus, this adjustment is conservative in terms of obtaining evidence for ovarian deterioration. Because of incomplete data records for some patients, we had 2751, 2675, 2710 and 2703 cycles for analysis of gonadotrophin dose, E2, oocyte number and dose per oocyte, respectively. In GAMs, we assumed a normal error distribution with identity link function for FSH, gonadotrophin dose and E2. This was justified because they were all roughly normally distributed. The skew was small and had negligible effects on the results as proven by analysing residuals for the models that fitted log-transformed data or assumed a gamma error distribution. To keep the results and interpretation as simple as possible, we assumed normal error distribution. The oocyte number was modelled as a Poisson variable using a log link function. The gonadotrophin dose per one oocyte was log-transformed prior to the statistical analysis replacing the zero values for the oocyte number by 0.5, and then treated as a normal variable.

Since the effects of age and time can interact well with each other in biological systems, we always checked for both additive (age + time + protocol) and interactive (age * time + protocol) model structures. The best model structure was selected according to the lowest value of the Akaike information criterion (AIC), with a difference >4 considered significant, as recommended by Burnham and Anderson (2002). Additive model structures, containing a substantially lower AIC, were always superior to interactive ones, indicating a consistent response over time in all patients irrespective of age. For fitting GAMs, we used package mgcv (Wood, 2006) implemented in the statistical software, R (R Development Core Team, 2008), which estimates the effective degree of smoothness (edf) of model terms as part of fitting. Model residuals were checked for using function gam.check in R. To obtain an estimate of an overall trend over the whole period, we subsequently replaced the smoothing term for time by a linear parametric function, which allowed us to calculate the mean annual change for each variable on the basis of an estimated regression slope.

Results

The age of patients increased steadily over the study period from 28.3 years (SE, 0.449) in 1995 to 31.0 years (SE, 0.227) in 2008; the slope of a trend from the linear model was 0.225 (95% CI, 0.188–0.263). The best GAM structure to explain variation in basal FSH level on Day 3 contained both age and time (the smoothing term for time: edf = 6.1, F = 12.0, P < 0.001; the difference in AIC with respect to the interactive model structure variant was 13.1). Whereas the increase in the basal FSH with age was linear (Fig. 1a), the temporal trend was highly non-linear, with peak values around 2002 (Fig. 1b). When we replaced the smoothing term in the best GAM by the linear term, the overall trend was increasing but without statistical significance (slope = 0.017, 95% CI, −0.005 to 0.039).

The dose of gonadotrophin increased with the age of patients (smoothing term for age: edf = 4.29, F = 32.0, P < 0.001), but
remained, on average, the same during the study period (the additive model structure, smoothing term for time: edf = 1 remained, on average, the same during the study period (the additive model structure: age (edf = 2.37, F = 74.6, P < 0.001) and time (edf = 6.58, F = 8.9, P < 0.001; Fig. 5), the ratio increasing annually on an original scale by about 4 IU of FSH per oocyte (slope of a linear trend on a log scale = 0.0241, t = 5.13, P < 0.001).

**Figure 1** Partial residuals of basal FSH (IU/l) as modelled by fitting the generalized additive model containing smoothing terms for age (a) and time (b).

The dashed lines indicate 95% confidence intervals. The points are partial residuals of FSH for a smooth term. They are obtained by dropping the term concerned from the model, whereas leaving all other variables fixed.

**Figure 2** Partial residuals of gonadotrophin dose (IU) as modelled by fitting the generalized additive model containing smoothing term for age (a) and time (b).

The partial residuals are not shown as their range is too large compared with estimated effects. The dashed lines indicate 95% confidence intervals.

Accounting simultaneously for the effect of age and protocol type, we also checked for temporal trends in gonadotrophin dose needed for the recovery of one oocyte. As revealed by the best GAM, which included additive effects of age, time and the parametric term for protocol, the dose increased with both age (edf = 2.37, F = 74.6, P < 0.001) and time (edf = 6.58, F = 8.9, P < 0.001; Fig. 5), the ratio increasing annually on an original scale by about 4 IU of FSH per oocyte (slope of a linear trend on a log scale = 0.0241, t = 5.13, P < 0.001).

**Discussion**

Unlike efforts to monitor changes in testicular function over time, evidence on ongoing trends in ovarian function is lacking. In contrast to testicular function for which semen quality serves as a reliable performance marker, studying temporal trends in the function of ovaries is much more difficult. We chose five ovarian parameters and examined their temporal variation in a sample of Czech women treated for infertility over a period of 14 years. Controlling for the concomitant effects of changing age and stimulation protocol type, we found evidence for temporal deterioration in three of these variables. Specifically, the hCG day estradiol and the number of recovered oocytes decreased, whereas the gonadotrophin dose required per one oocyte increased over time. Together, these results are indicative of a small, but detectable decrease in ovarian function over the period of 14 years, which is not causally related to ageing. However, it should be emphasised that it remains to be shown how the results obtained among infertility clinic patients apply to women in a population at large and whether these 14-year observations represent real long-term population processes reflecting the modern shift in human life history towards a longer life span.

Female fertility usually begins to decline after 30 years of age and progressively leads to sterility after 41 years of age (Wood, 1989; van Noord-Zaadstra et al., 1991; te Velde and Pearson, 2002). Consequently, the response of ovaries to treatment tends to become weaker in older women (Goverde et al., 2005). The average age of women in our group has increased by 2.7 years over a period of 14 years. This corresponds to the data from a Swiss national register,
predictive value (Broekmans et al., 2008). Assess ovarian function, but there is no real consensus about their levels and antral follicle count have been proposed in order to for IVF (Sobek et al., 2008). In our patients, there has been an overall tendency in the basal FSH to increase with age, but the temporal trend was highly non-linear and complex, with peak around 2002, suggesting that additionally to age, other unknown variables were operating during the study period (e.g. the cohort effect).

The amount of gonadotrophin used for stimulation remained the same during the studied period, increasing only with the age of the patients. The source and quality of FSH has changed during the study period, which could have become a potential source of bias if we had not controlled for this effect. As expected, the activity of rFSH has been shown to be higher than that of uFSH (Fisch et al., 1995; Daya and Gunby, 2007). Unlike gonadotrophin dose, however, E2 and the number of oocytes recovered decreased and gonadotrophin dose per oocyte increased during the study period. Hence, with the same dose, lower response has been achieved. These three variables can thus be considered as the measures of ovarian responsiveness to treatment by gonadotrophins, with the number of collected oocytes further demonstrating ovarian capacity (Lawson et al., 2003). As age and stimulation protocol were both controlled for, the observed decline in ovarian sensitivity over time had to be caused by different factors. The factors which are known to have a negative impact on the ovarian function include smoking (Cooper and Moley, 2008; Soares and Melo, 2008), obesity (van der Steeg, 2008), autoimmune diseases (Alper and Garner, 1985; Wheatcroft et al., 1997) and alcohol consumption (Kinney et al., 2007). These factors, with the exception of smoking, are relatively rare in this patient population and could therefore explain only a limited proportion of the observed effect. Although complex hormonal machinery interacting with environmental factors is likely involved in these processes (Vitzthum, 2008), we are unable to provide any plausible proximate mechanism explaining these trends. At present, the only explanation we can offer is based on an evolutionary argument bearing on the life-history trade-off between survival and reproduction, a pervasive feature in animals and plants (Bell and Koufopanou, 1986). The fact that natural selection is unable to maximize both of these processes at once has become a major tenet of modern biological theory, widely accepted by most life-historians and evolutionary biologists. If human demographic schedules for fertility and survival now change in favour of a longer life span, then some costs on the reproductive side can be expected. There is a huge bulk of evidence from yeast and plants to birds and mammals, including humans (Westendorp and Kirkwood, 1998; Thomas et al., 2000), supporting this view (see Stearns, 1992; Roff, 2002 for a review).

An important question is whether the changes observed in a group of infertility patients are an unbiased reflection of the population at large. Clearly, the patients are not a random sample drawn from a population. However, what we emphasise here is not the parameter values, but the changes over time. If the infertility patients come from one tail of the ovarian function distribution, which moves as a whole towards lower values (because of largely additive effects of older age and time-specific factors), then changes in patients could mirror quite closely those in a general population. However, there

---

**Figure 4** Partial residuals of the number of oocytes recovered as modelled by fitting the generalized additive model containing a smoothing term for age (a) and time (b).

The partial residuals are not shown as their range is too large when compared with the estimated effects. The dashed lines indicate 95% confidence intervals. Note that the number of oocytes was modelled as a Poisson variable, so the effects are plotted on a log scale.

---

**Figure 5** Partial residuals of gonadotrophin dose (IU) per one oocyte recovered as modelled by fitting the generalized additive model containing a smoothing term for age (a) and time (b).

The partial residuals are not shown as their range is too large when compared with the estimated effects. The dashed lines indicate 95% confidence intervals. Note that the effects are plotted on a log scale because the ratio was log-transformed prior to the analysis.

---

in which an increase of 2 years during a period of 10 years was noted (van den Bergh et al., 2005). The increase in the average age did not have a significant impact on the success of IVF treatment in our case, but it did influence the outcome measures of ovarian sensitivity to the treatment. These data emphasize the prevailing demographic trend characterised by progressive aging of the population at large is, itself, an important variable affecting conception rates in women treated for IVF (Sobek et al., 2008).

Several markers, such as basal FSH and estradiol levels, inhibin B levels and antral follicle count have been proposed in order to assess ovarian function, but there is no real consensus about their predictive value (Broekmans et al., 2006). Moreover some of the markers, including anti-Müllerian hormone, are relatively new and we had little experience with them early in the study. Basal FSH measured on Day 3 of the menstrual cycle remains the most common predictor of ovarian response and is used in most IVF units as an indicator of ovarian responsiveness (Maheshwari et al., 2006). Though its efficiency as a routine test has been questioned (van Montfrans et al., 2000; Wolff and Taylor, 2004), its indication value has been supported by a bulk of evidence from several studies (Muasher et al., 1988; Scott et al., 1989; Lawson et al., 2003).
Is there a declining trend in ovarian function?

is no doubt that a stronger inference would be made by measuring changes directly in a random sample from the population.

The declining trends in ovarian performance were observed on a relatively small time scale, amounting to about one-half of a generation time. However, we believe that the dynamic pattern observed in demographic processes is important for detection of changing human physiology as well. Observing large opposite shifts in birth and death rates at the population level might afford a better chance of detecting the corresponding change at the individual level. The dramatic changes in the Czech demography, brought about by political change during the 1990s, have created such an opportunity. Hence, it may be difficult to capture corresponding trends in demographically stable countries within the same time frame. On the other hand, the observed trends in patients from the Czech population are fairly small. Consequently, it seems wise not to overstate the significance of our findings until we have further evidence that these short-time trends persist in the long run, analogous to the trend observed in males.

Using the five parameters related to the function of ovaries, we demonstrated a statistically significant decrease in ovarian sensitivity to the treatment by gonadotrophins over the period of 14 years. These results are novel, suggesting that the deteriorating tendency in reproductive performance, observed in males, is likely to be present in females as well. Unlike Jensen et al. (2008), who demonstrated the declining trends in female fertility indirectly through analysis of the data from registers on birth and ascribed this decline to deteriorating male reproductive health, we have provided direct evidence for a temporal decline in ovarian performance. We have shown that from a life-history perspective, these findings are not surprising if we take into account the ongoing changes in human demography. On the other hand, it becomes increasingly clear that obtaining more convincing evidence will not be a simple task. It will necessitate systematic monitoring of ovarian function over long periods of time using generally accepted markers and unified methodological approaches. In this respect, developing a common scheme for monitoring ovarian function is highly desirable. It could also be instrumental in interpretation of evidence if future research could include not only patients with fertility problems but also women from a healthy part of the society. This could provide new insights into the interplay between the demographic processes taking place at the level of population and the physiologic processes governing the development and reproductive life of an individual.

Funding

E.T. was funded by grant MSM6198959212.

References


Scott RT, Toner JP, Muasher SJ. Follicle-stimulating hormone levels on cycle day 3 are predictive of in vitro fertilization outcome. Fertil Steril 1989;51:651–654.


Submitted on July 9, 2009; resubmitted on September 24, 2009; accepted on September 28, 2009.