Selection bias in semen studies due to self-selection of volunteers

A. Muller1,7, E. De La Rochebrochard1, C. Labbé-Declèves2, P. Jouannet2, L. Bujan3, R. Mieusset4, D. Le Lannou5, J. F. Guerin6, M. Benchakib6, R. Slama1 and A. Spira1

1INSERM (National Institute For Health and Medical Research) U569 IFR69, 94276 Le Kremlin-Bicêtre, 2Service de Biologie de la Reproduction-CECOS, Hôpital Cochin, EA 1752 Université Paris 5, 75014 Paris, 3CECOS Midi-Pyrénées, Research Group ‘Human Fertility’ (E1 3694), 4Centre de Stérilité Masculine, Service d’Urologie Andrologie, Hôpital Paule de Viguier, 31059 Toulouse, 5Service de Biologie de la Reproduction-CECOS, Hôpital de l’Hôtel-Dieu, 35000 Rennes and 6Département de Médecine et Biologie de la Reproduction, Faculté de Médecine-CECOS, 69373 Lyon, France

7To whom correspondence should be addressed at: INSERM U569, 82 rue du Général Leclerc, 94276 Le Kremlin-Bicêtre cedex, France. E-mail: muller@vjf.inserm.fr

BACKGROUND: Reports of a secular decrease in semen quality remain controversial, particularly due to the possibility of selection bias. We aimed to describe the potential bias due to self-selection of volunteers in semen studies involving infertile men. METHODS: Using data from the French multicentre study REPRHOM, we compared the characteristics of the partners of pregnant women for three levels of participation: completion of a refusal questionnaire (n = 698), agreement to complete the study questionnaires only (n = 676) and agreement to complete the study questionnaires and give a semen sample (n = 331, 13% of the subjects approached). RESULTS: Poorly educated men refused more often to participate than highly educated men. Semen providers were more likely to have experienced unfavourable pregnancy outcomes (odds ratio 1.68, 95% confidence interval 1.14–2.49) compared with participants completing the questionnaires only. Time to pregnancy was similar for all participants. CONCLUSIONS: This study demonstrates the existence of selection bias in semen studies associated with fertility and socio-demographic characteristics of men. The results of semen analysis for this population sample cannot be extrapolated to the whole population from which the volunteers originate. More information is required on who participates, and participation rates should be reported in semen studies to make it possible to interpret the results correctly.

Key words: bias (epidemiology)/fecundity/participation/selection bias/semen studies

Introduction

A number of studies (Carlsen et al., 1992; Auger et al., 1995; Adamopoulos et al., 1996; Swan et al., 2000) have suggested that the quality of human sperm has decreased over the last 50 years. This secular decrease has been the source of considerable controversy, particularly given the low rates of participation in semen studies (Jouannet et al., 2001). These participation rates are rarely published even though transparency is necessary to interpret correctly the results obtained (Sandler, 2002). In studies based on voluntary participation, the declared participation rates are very low (one or two people for every 10 approached), raising questions about those who agree to participate. Handelsman (1997) suggested that voluntary participation of men in semen studies is linked to the self-selection of less fecund men and that this selection bias may be similar in amplitude to the secular decrease in sperm quality reported. A part of the observed decline could indeed be due to selection bias, particularly as not all studies have reported a secular decrease in semen quality (Bujan et al., 1996; Fisch et al., 1996; Andolz et al., 1999). These findings also raise the more general question of whether the results of semen studies based on voluntary participation can be extrapolated to the entire population from which the volunteers originate.

To date, few studies have dealt with selection bias in semen studies. Five studies (Handelsman et al., 1985; Handelsman, 1997; Larsen et al., 1998; Cohn et al., 2002; Lalos et al., 2003) have shown that men who volunteer for semen studies tend to be younger, to have experienced a long period of infecundity (measured as a time to pregnancy, TTP, of > 6 months; Larsen et al., 1998) and to have family members who have had problems conceiving a child.

The work presented here is based on data from a French cross-sectional multicentre study (Reproduction de l’Homme,
REPRHOM), the principal aim of which was to compare the fecundity (semen quality and TTP) of the partners of pregnant women in four French towns. The selection phenomena associated with voluntary participation were analysed at three levels: (i) refusal to participate in the study but completion of a short refusal questionnaire; (ii) completion of study questionnaires only; and (iii) completion of study questionnaires and collection of semen sample.

Materials and methods

Study population

The study sample consisted of the male partners of pregnant women attending the maternity units of the hospitals housing the four Centres for the Study and Conservation of Human Eggs and Sperm (CECOS) participating in the study. These centres were located in the French towns of Paris, Toulouse, Rennes and Lyons. Participants were recruited between April 2002 and December 2003. The inclusion criteria were: ‘natural’ pregnancy [rather than by means of artificial reproductive techniques (ARTs), with the use of ovulation inducers not considered to be an ART], male age between 20 and 45 years, born in continental France and living in the recruitment area for at least 1 year. The female partner had to be at least 14 weeks pregnant at the time of first contact with the interviewer.

Design

We used the same standardized recruitment protocol in each centre to minimize between-centre differences. There were no incentives at all to participate. An interviewer described the study to the pregnant women, checked that the couple was eligible and then asked them to participate (Figure 1). If the couple agreed to participate, three questionnaires (for the man, the woman and the mother of the man) were given to the woman. About 1 week later, the interviewer contacted the couple to ask the man to attend an appointment at the local CECOS for the collection of a semen sample. If the man refused, the interviewer tried to persuade the couple to complete and return the questionnaires. If the questionnaires were still not returned after several telephone calls, the interviewer attempted to get the man to complete a refusal questionnaire by telephone. This questionnaire asked about the man’s age, level of education and smoking habits. In all cases in which the man agreed to participate, he provided written consent to the medical team. The study protocol was approved by the French regulatory and consultative ethics bodies.

Data collection

Questionnaires. The questionnaires were completed at home by the male volunteers, their partners and their mothers (when alive). The principal subjects addressed were fecundity history and lifestyle 3 months before conception of the current pregnancy and during childhood. The longest questionnaire (the man’s questionnaire) took ~45 min to complete.

Semen collection. Each male volunteer who agreed to give a semen sample came to the local CECOS during his partner’s pregnancy or, at the latest, 1 month after delivery. The semen sample was obtained by masturbation after a recommended period of 2–10 days of sexual abstinence.

Definition of studied variables

Information was obtained for three groups of men: (i) men who refused to participate but who completed the refusal questionnaire; (ii) those who completed the study questionnaires only; and (iii) those who gave a semen sample and completed the study questionnaires.

The following information was recorded in all questionnaires (refusal and complete questionnaires): age of the man, his educational level, his smoking status and the study centre. Of men who refused to participate, 44% agreed to complete the refusal questionnaire. More detailed information was available on the completed questionnaires. Women were the first to be approached by the interviewers. As the study did not directly concern them, the vast majority of them (96%) agreed to ask their partner to participate. Therefore, we considered in this work that the choice to participate was the man’s decision.

‘Number of previous pregnancies’ included all pregnancies for which the man was the father (other than that for which the man had been included in the REPRHOM study) with either their

Figure 1. Method used to recruit the partners of pregnant women in the REPRHOM study, 2002–2003.
current partner or a previous one. ‘Unfavourable pregnancy outcomes’ included all spontaneous miscarriages, ectopic pregnancies, abortions for medical reasons and stillbirths resulting from every pregnancy for which the man was responsible. Men who had had cryptorchidism, inguinal hernia, varicocele, testicular torsion, testicular cancer, hypospadias, gonococcal infection, epididymitis or orchitis were considered to have a history of andrological disease. ‘Consultation for infecundity’ included declarations made by the men concerning their entire lifetime. TTP was estimated for the current pregnancy on the basis of declarations made by the pregnant women. It was not possible to estimate TTP for 40 couples, and 51 couples were using contraception at the time of conception. The statistical analyses of fecundity thus included 894 couples for whom TTP was known. Smoking habits were recorded for the 3 month period preceding conception. ‘Exposure to tobacco in utero’ was recorded on the questionnaire filled in by the volunteers’ mothers and concerned the period of intra-uterine development of the man.

Statistical analysis

We first compared the men who completed the refusal questionnaire with the group of participants (questionnaires only and questionnaires + semen sample) on the basis of all the information available in the refusal questionnaire. The adjusted odds ratios (ORs) for participation were calculated by logistic regression. Variables included in the model were man’s age, man’s education level, man’s smoking status and study centre.

We then compared the detailed information from the complete questionnaires for the two groups of participants (semen collection + questionnaires versus questionnaires only). The adjusted ORs associated with the characteristics of the men agreeing to provide a semen sample with respect to those completing questionnaires only were estimated by logistic regression. The variables taken into account in the logistic regression analysis were the man’s age, man’s education level, study centre, history of andrological disease, unfavourable pregnancy outcomes, consultation for infecundity, number of previous pregnancies, smoking status of the man, exposure to tobacco in utero and conception problems in the man’s family.

To complete this analysis, the TTP of the current pregnancy was studied using the discrete Cox model with a logistic link (Scheike and Jensen, 1997) on the total number of months of attempts, according to outcome of the attempt (failure/success). The fecundability ratio (FR) obtained is the ratio between the monthly probability of conceiving for a couple who participate in semen collection compared with that of a couple who only complete questionnaires. TTP was censored after 13 months, as couples not managing to conceive may change their behaviour or lifestyle in a way that might affect fertility after this time. The variables taken into account in the discrete Cox model were the same as in the previous logistic regression model plus frequency of sexual intercourse and history of gynaecological disease, variables known to be confounding factors. All ORs and FRs were estimated with 95% confidence intervals (CIs) using STATA SE 8.2* software (Stata Corporation, College Station, TX).

Results

Description of the whole sample

Among the 2581 eligible couples invited to participate, 676 (26%) agreed to complete the questionnaires only and 331 (13%) agreed both to complete the questionnaires and to provide a semen sample (Figure 1). In 22 cases (7% of semen samples), the questionnaires completed by men agreeing to semen collection were lost and therefore not taken into account by the coordinating team. Statistical analysis was therefore carried out for the 309 participants who agreed to give a semen sample for whom questionnaires were available. Of the 1574 men who refused participation, 44% agreed to complete the refusal questionnaire. The participation rates were similar in Paris, Toulouse and Rennes (Table I; 44, 44 and 46%, respectively, agreed to participate and 19, 13 and 15%, respectively, agreed to give a semen sample) but far lower in Lyons (26 and 5% agreement, respectively).

Table II summarizes the socio-demographic characteristics for each of the three groups and Table III describes the socio-demographic characteristics and characteristics associated with fertility only for men who completed detailed questionnaires. The mean age of the participants (men who at least completed the questionnaires plus men who completed questionnaires and gave a semen sample) was 32.6 years (4.4). A history of andrological disease was noted for 15% of participants and 20% had experienced an unfavourable pregnancy outcome.

Comparison of participants with those refusing to participate

Participants were less likely to be smokers (OR 0.72, 95% CI 0.57–0.91) and had a higher educational level than those who refused to participate (Table II).

Comparison of participants who gave semen sample with participants who only completed questionnaires

The men who provided semen had a higher educational level (test for trend: P < 0.001) than those who only completed questionnaires (Table IIIA). They were also more likely to have been exposed to tobacco in utero, even with a large confidence interval (OR 1.64, 95% CI 0.96–2.78). Their current and previous partners were more likely to have experienced unfavourable pregnancy outcomes (OR 1.68, 95% CI 1.14–2.49) than those of the men completing questionnaires only. The couples who agreed to provide a semen sample had a similar fecundability to those who only completed questionnaires (Table IIIIB).

Discussion

In the few semen studies that clearly reported participation rates, semen collection rates were between 13 and
Participation rates exceeding 20% are rare (Storgaard et al., 2003). Thus the REPRHOM study is within the generally observed range, with a semen collection rate of 13%. No differences in recruitment rates were observed between Paris, Toulouse and Rennes (Table I). However, the participation rate was lower in Lyons (slightly for the completion of questionnaires only, 21%, and particularly lower for semen collection, 5%). This variability may be explained by two points. First, it could result from a different study logistic management from the local team, as compared with the others. Secondly, there was

### Table II. Participation in the REPRHOM study among eligible couples approached according to characteristics available in the refusal questionnaire

<table>
<thead>
<tr>
<th>Age of the man</th>
<th>Total (n = 1705)</th>
<th>Refusal questionnaire (n = 698)</th>
<th>Participants in REPRHOM (n = 1007)</th>
<th>Adjusted OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>20–24 years</td>
<td>59 (4%)</td>
<td>31 (4%)</td>
<td>28 (3%)</td>
<td>0.35</td>
<td>0.12–1.03</td>
</tr>
<tr>
<td>25–29 years</td>
<td>403 (23%)</td>
<td>159 (23%)</td>
<td>244 (25%)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>30–34 years</td>
<td>756 (45%)</td>
<td>305 (44%)</td>
<td>451 (46%)</td>
<td>0.83</td>
<td>0.63–1.10</td>
</tr>
<tr>
<td>≥35 years</td>
<td>462 (28%)</td>
<td>203 (29%)</td>
<td>259 (26%)</td>
<td>0.75</td>
<td>0.55–1.02</td>
</tr>
</tbody>
</table>

### Table IIIA. Agreement to semen collection among volunteers who completed at least the questionnaires in the REPRHOM study, according to social and fecundity-linked characteristics

<table>
<thead>
<tr>
<th>Age of the man</th>
<th>Total (n = 1007)</th>
<th>Questionnaires only (n = 676)</th>
<th>Questionnaires + semen sample (n = 331)</th>
<th>Crude OR</th>
<th>95% CI</th>
<th>Adjusted OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>20–24 years</td>
<td>28 (3%)</td>
<td>24 (4%)</td>
<td>1 (1%)</td>
<td>0.35</td>
<td>0.12–1.03</td>
<td>0.56</td>
<td>0.18–1.75</td>
</tr>
<tr>
<td>25–29 years</td>
<td>244 (23%)</td>
<td>162 (24%)</td>
<td>82 (27%)</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30–34 years</td>
<td>451 (46%)</td>
<td>309 (46%)</td>
<td>142 (46%)</td>
<td>0.91</td>
<td>0.65–1.27</td>
<td>0.93</td>
<td>0.65–1.33</td>
</tr>
<tr>
<td>≥35 years</td>
<td>259 (26%)</td>
<td>178 (26%)</td>
<td>81 (26%)</td>
<td>0.90</td>
<td>0.62–1.32</td>
<td>0.90</td>
<td>0.59–1.37</td>
</tr>
</tbody>
</table>

OR = odds ratio; CI = confidence interval, adjusted, by using logistic regression, for man’s age, man’s educational level, man’s smoking status and study centre.
a cultural difference between the population approached in Lyons and in the others towns because it was mostly composed of an African community in which semen collection is taboo. The study topic was wrongly perceived and a vast majority of men refused to give their semen. This variability was taken into account in the model by including study centre, history of andrological disease, unfavourable pregnancy outcomes, consultation for infertility, number of previous pregnancies, man’s smoking status, exposure to tobacco in utero and conception problems in the man’s family.

Consultation by the man for infertility, at any time in his life, with his current partner or a previous one.

Adjusted, by using discrete Cox model, for man’s age, man’s educational level, study centre, history of andrological disease, unfavourable pregnancy outcomes, consultation for infertility, number of previous pregnancies, man’s smoking status, exposure to tobacco in utero, conception problems in the man’s family, frequency of sexual intercourse and history of gynaecological disease (history of genital inflammation, salpingitis, endometriosis, uterine fibroma, complications of appendicitis, diabetes, ovarian cysts or Chlamydia trachomatis infection in the current partner).

Table IIIB. Fecundability ratio (FR) for 894 pregnancies according to study participation

<table>
<thead>
<tr>
<th>Participation</th>
<th>Crude FR</th>
<th>95% CI</th>
<th>Adjusted FR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questionnaires only</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questionnaires</td>
<td>1.03</td>
<td>0.86–1.22</td>
<td>1.04</td>
<td>0.87–1.26</td>
</tr>
<tr>
<td>+ semen sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

OR = odds ratio; FR = fecundability ratio; CI = confidence interval.

*Adjusted, by using logistic regression, for man’s age, man’s educational level, study centre, history of andrological disease, unfavourable pregnancy outcomes, consultation for infertility, number of previous pregnancies, man’s smoking status, exposure to tobacco in utero and conception problems in the man’s family.

Consultation by the man for infertility, at any time in his life, with his current partner or a previous one.

*Adjusted, by using discrete Cox model, for man’s age, man’s educational level, study centre, history of andrological disease, unfavourable pregnancy outcomes, consultation for infertility, number of previous pregnancies, man’s smoking status, exposure to tobacco in utero, conception problems in the man’s family, frequency of sexual intercourse and history of gynaecological disease (history of genital inflammation, salpingitis, endometriosis, uterine fibroma, complications of appendicitis, diabetes, ovarian cysts or Chlamydia trachomatis infection in the current partner).

Were the participants who only completed questionnaires as fecund as those who provided a semen sample?

Men who donated sperm were no more likely to declare a family history of conception problems than were those who completed questionnaires only. Our initial hypothesis was that men who agreed to provide a semen sample were more likely to have a personal or a familial history of fecundity problems (Jouannet et al., 2001). This was shown in two studies of sperm donors in infertility clinics who donate out of a spirit of altruism (Handelsman et al., 1985), often because they have been sensitized to the issue as a result of sterility among their relatives (Lalos et al., 2003). However, the samples studied were very different: partners of pregnant women recruited for a fertility study on the one hand, and semen donors at infertility clinics on the other hand. The reasons for participating in a scientific research programme are undoubtedly different from the reasons for donating semen. As for those who refused to participate and participants, men with a low educational level were more likely to participate in questionnaires only than highly educated men who agreed more often to give a semen sample (Table IIIA: OR secondary school/further education 0.41, 95% CI 0.24–0.69).

The FR for the current pregnancy did not differ according to the level of participation (Table IIIB). In an occupational Danish study (Larsen et al., 1998), men who had experienced a long period of infertility were more likely than other men to agree to semen collection. There may be two reasons for these different results. First, participation tends to be higher in occupational studies because individuals are aware that they have impaired health, and may wish to explain their state of health in terms of professional exposure (Bonde et al., 1996). Secondly, the REPRHOM sample consisted of fecund...
couples (pregnant women and their partners), and the known effects of various factors on TTP have been shown to disappear or to be reversed if infecund couples are excluded. This is the case, for example, for maternal age (Juul et al., 2000).

The participants who provided a semen sample and those who completed the questionnaires only did not differ in terms of andrological history or male smoking status (Curtis et al., 1997; Kunzle et al., 2003). Female partners of men who donated semen were more likely to have experienced unfavourable pregnancy outcomes than were the partners of those who completed questionnaires only (OR 1.68, 95% CI 1.14–2.49). This suggests that these men were more likely to have experienced infecundity in the past than the others. Furthermore, exposure to tobacco in utero was slightly (but not significantly when adjusted) more frequent among men agreeing to provide semen than among the men who only completed the questionnaires (adjusted OR 1.64, 95% CI 0.96–2.78). Maternal smoking probably has a deleterious effect on sperm quality in the son (Storgaard et al., 2003; Jensen et al., 2004). It is therefore possible that sperm quality was lower in those who agreed to give a semen sample than in those who only completed the questionnaires: this cannot be tested as no semen sample was available for the latter.

In view of differences of unfavourable pregnancy outcomes and maternal smoking, sperm quality may be lower for participants providing a semen specimen than for men completing questionnaires only. Low semen quality is known to affect TTP (Bonde et al., 1998; Slama et al., 2002), but our analysis of TTP for this sample did not confirm the hypothesis that fecundity is different between participants who completed questionnaires only and men who provided a semen sample. It is possible that the semen quality variation studied was not sufficient to have a detectable effect on TTP. It may not be possible to extrapolate the distribution of sperm characteristics for the participants providing semen samples to the entire population of participants or to the whole of the population from which this sample originates. This element should be studied and considered in subsequent analyses of sperm quality. Given the similarity of fecundity in the two groups of participants, the analysis of TTP for all study participants should not be affected by selection bias in the REPRHOM sample.

In conclusion, our results show that men who volunteer to participate, even in part, in semen studies differ from those who refuse to participate (from which they originated), particularly in terms of educational level and smoking status. Moreover, men who agree to complete questionnaires only and men who agree to semen collection differ in terms of certain socio-demographic characteristics (educational level) and variables associated with fecundity (unfavourable pregnancy outcomes, exposure to tobacco in utero), despite having a similar FR (similar TTP). The results of semen analysis for this population sample cannot be extrapolated to the whole population from which the volunteers originate. As seen in our study, there may be certain differences in fertility between the refusal group and participants (in the REPRHOM sample, the refusal group could be less fecund than participants, but individuals who provided a semen sample could be less fecund than those who completed questionnaires only). Therefore, as selection bias could involve bilateral deviation in fertility characteristics, we recommend that the effect of this selection bias be systematically analysed for each sample in every new study. Above all, more information is required on participating men and men who refused to participate, and participation rates should be reported in semen studies to make it possible to interpret the results correctly.

Acknowledgements

The authors gratefully acknowledge Beatrice Ducot and Jean Bouyer for their critical review of the manuscript. This study was supported by contract QQLK-CT-1999-01422 from the European Union and grants from the French Health Department (DGS), the French Research Department, the French Environmental and Health Security Agency (AFSSE), the Total Fina Elf Group and the AGRICA Group.

References


Submitted on June 28, 2004; accepted on August 19, 2004