

## Feasibility of recruiting a birth cohort through the Internet: the experience of the NINFEA cohort

Lorenzo Richiardi · Iacopo Baussano · Loredana Vizzini · Jeroen Douwes · Neil Pearce · Franco Merletti

Received: 16 May 2007 / Accepted: 2 October 2007 / Published online: 23 October 2007  
© Springer Science+Business Media B.V. 2007

**Abstract** The NINFEA cohort is an Italian birth cohort aiming at recruiting pregnant women through the Internet and following-up their children. To understand whether Internet-based recruitment was feasible we started a pilot in July 2005 targeted to pregnant women visiting the hospitals of the city of Turin (900,000 inhabitants), where we advertised the study. For this purpose we constructed a website ([www.progettoninfea.it](http://www.progettoninfea.it)), with on-line questionnaires to be completed during pregnancy and at 6 and 18 months after delivery. Participants' characteristics were compared with those of women giving birth in Turin, which are routinely released as individual anonymous records. We also compared complete with partial respondents. We also carried out a survey of 122 women giving birth in the main Turin obstetric hospital to estimate the proportion of pregnant women with access to the Internet and awareness of the NINFEA cohort. By December 2006, we had recruited 670 women. Participation was associated with being older, a university graduate, primiparous and born in Italy. Complete response ( $n = 633$ ) was associated with being primiparous and participation after the first trimester. In the survey, 66% (95% confidence interval: 56–74%;  $n = 80$ ) of the women had access to the Internet

and 42% (33–51%;  $n = 51$ ) were aware of the study; 6.5% (2.9–12.5%;  $n = 8$ ) had participated in the NINFEA cohort. Our study indicates that recruitment of an Internet-based birth cohort is feasible. As with many other types of birth cohort study, the participants are a self-selected sample. To minimise selection bias analyses should therefore be limited to internal comparisons.

**Keywords** Longitudinal study · Internet · Bias · Birth cohort

### Introduction

Although the Internet is becoming increasingly accessible in most countries of the world, and on-line surveys are becoming popular in social science research, this tool has been little used to date in health research [1]. Some studies have used the Internet to administer web-based questionnaires to pre-selected groups of individuals with known contact information, such as students [2], health professionals [3], patients [4] and healthy subjects sampled from the general population [5–7]. Intervention programs, such as those for smoking cessation and weight reduction, have been offered on-line [8, 9]. A web-based questionnaire has been used for the 10-year follow-up interview of a general population cohort of 50,000 Swedish women recruited in 1991/1992 [10]. Participation through a web-based questionnaire has been offered to subjects enrolled in a large US cohort study of military personnel [11]. Recently, some studies have posted surveys on websites, recruiting participants whose contact information was unknown before participation [12–14]. The results of several of these studies are encouraging, suggesting that the use of the

L. Richiardi (✉) · I. Baussano · L. Vizzini · F. Merletti  
Cancer Epidemiology Unit, CeRMS and Center for Oncologic  
Prevention, University of Turin, V Santena 7, 10126 Turin, Italy  
e-mail: [lorenzo.richiardi@unito.it](mailto:lorenzo.richiardi@unito.it)

I. Baussano  
Division of Epidemiology, Public Health and Primary Care,  
Imperial College London, London, UK

J. Douwes · N. Pearce  
Centre for Public Health Research, Massey University  
Wellington Campus, Wellington, New Zealand

Internet is well-accepted by the participants and is an efficient tool for collecting epidemiological data [15].

Birth cohort studies are an essential tool for life-course epidemiology and are becoming increasingly common with the growing recognition that exposures occurring early in life, from the foetal period to early childhood, may affect health status in adulthood [16–18]. An Internet-based birth cohort study may have considerable efficiency advantages in settings in which a “traditional” birth cohort study is not possible because of cost and/or organisational restrictions. We therefore carried out a pilot study (the NINFEA cohort) to assess the feasibility of using the Internet for recruitment of pregnant women and follow-up of their children in a birth cohort study.

## Methods

### The cohort

The NINFEA cohort is an Internet-based birth cohort established in Italy with the aim of investigating the effects of early life exposures on the health of newborns, infants, adolescents and adults. Starting from July 2005, a pilot study to test the feasibility of the Internet recruitment was carried out in the city of Turin (900,000 inhabitants) which includes three public general hospitals with Departments of Obstetrics, each with 800–1,000 newborns per year, and a large public obstetric and gynaecological hospital with about 10,000 newborns per year. The results of the pilot study will provide valuable information for the extension of the cohort to the Italian population, and the establishment of a similar birth cohort study in New Zealand.

The source population of the NINFEA cohort includes all babies born to pregnant women who have enough knowledge of the Italian language and the use of the Internet to complete the online questionnaires. Members of the cohort are babies born to women in this source population who become aware of the study and volunteer to participate through registration on the study website ([www.progettoninfea.it](http://www.progettoninfea.it)). The study was approved by the local Ethical Committee.

For the pilot phase, we advertised the study only in the city of Turin, limiting the source population primarily to pregnant women visiting Turin hospitals, although a small number of women ( $n = 38$ ) heard about the study from friends or colleagues. The study was advertised with the following methods: (1) posters at the main hospitals of the city; (2) leaflets enclosed with the results of laboratory tests and ultrasounds carried out for the prenatal screening; and (3) leaflets distributed at the pre-delivery classes after giving a brief overview of the study. Some gynaecologists also actively distributed leaflets during medical

examinations at hospitals, although we favoured methods of advertisement that did not require additional work for the health care personnel. Also, since in the pilot study we aimed at limiting the advertising to the city of Turin we did not post-information on national web-sites relevant to pregnant women. Registered women were asked to complete three Internet-based questionnaires: during pregnancy and 6 and 18 months after delivery. Moreover, they were asked to give consent for passive follow-up of the child and the mother, using linkage with records of health-related databases, such as the Hospital Discharge Registry and Cancer Registries.

The baseline questionnaire included 16 sections and about 450 items, with questions on demographic characteristics, lifestyle, environmental and occupational exposures, reproductive and medical history, drugs and supplements and some specific questions about the use of the Internet. At the end of the questionnaire the participants were asked if they would have donated a sample of urine, nail, saliva or blood if requested. Women were asked to submit each section of the questionnaire before moving on to the next section. Apart from the contact information and the date of birth, no answer was mandatory.

### Group comparisons and exposure information

Participants' characteristics were compared with those of women giving birth in the city of Turin during the same period. Comparison data were obtained from the Piedmont Birth Register which was established in 2002 [19]. Since information included in the Birth Registry is collected at delivery, whereas information for our cohort was collected during pregnancy, we limited the comparisons to demographic characteristics, which are not time-dependent (Table 1). To estimate the maternal age at delivery for the participants in the NINFEA cohort, we calculated the expected date of birth by adding 40 weeks from the date of conception, either estimated from the date of the last menstruation or as confirmed by ultrasound. Since the recruitment period for the NINFEA cohort was July 2005 to December 2006, participants were expected to deliver during 2005–2007, whereas the comparison was with the total population of women giving birth in Turin in the first semester of 2005, which was the most recent period for which data from the Birth Registry were available. However, changes in maternal age, parity, place of birth and maternal educational level in the population would have been relatively small over this time period.

We also compared the women who completed the entire questionnaire with those who were partial respondents; the comparisons were made by age, parity, place of birth (Italy versus non-Italy), ever smoking before pregnancy and

**Table 1** Selected characteristics of women participating in the NINFEA cohort compared with women giving birth in the city of Turin

Characteristics	NINFEA cohort ( <i>n</i> = 670)		Turin ( <i>n</i> = 5,959)		<i>P</i> -value
	Numbers	(%)	Numbers	(%)	
Maternal age (years)					
≤29	121	18.1	1,541	25.9	<0.001
30–32	187	27.9	1,352	22.7	
33–35	195	29.1	1,436	24.1	
36+	167	24.9	1,630	27.3	
Number of previous pregnancies					
0	412	61.5	2,730	45.8	<0.001
1+	258	38.5	3,229	54.2	
Place of birth					
Italy	648	96.7	4,618	77.5	<0.001
Non-Italy	22	3.3	1,341	22.5	
Educational level					
Undergraduate	286	42.7	5,003	84.0	<0.001
Graduate	384	57.3	956	16.0	
	Mean	SD	Mean	SD	
Maternal age (years)	33.3	3.8	32.5	5.0	<0.001

NINFEA cohort, Italy, 2005–2006

*SD* standard deviation

gestational age at participation. Age was categorised in two categories using the median (33 years) as the cut-off, whereas gestational age was categorised in trimesters.

#### Ad hoc survey

During the pilot study, the source population was babies born to pregnant women visiting Turin hospitals who had enough knowledge of the Italian language and the use of the Internet to complete in the online questionnaires. To better understand the process of self-selection from this source population we conducted an ad hoc survey within women who were hospitalised for reasons related to the delivery at the main public gynaecologic and obstetric hospital of the city. The survey was intended to estimate: (1) the proportion of pregnant women who were aware of the study, (2) the proportion of women having access to the Internet (defined as having the possibility of using the net at least twice a month) and (3) the proportion of women who, being aware of the study and having access to the Internet, actually participated in the study. For this survey, we interviewed, in a given week of July 2006, ward by ward, 122 women using an anonymous face-to-face interview.

#### Statistical analysis

We tested the differences between participants in the NINFEA cohort and women giving birth in Turin using univariate tests—chi-square tests for categorical variables

and *t*-tests for maternal age [20]. Prevalence odds ratios (PORs) and 95% confidence intervals (CI) of participation and partial response (compared with full response) were estimated using unconditional logistic regression.

Finally, we used information obtained from the 122 women participating in the ad hoc survey to estimate the proportions (and 95% CI) of pregnant women with access to the Internet, awareness of the existence of the study, and participation. PORs of having access to the Internet were estimated by age and place of birth.

#### Results

We recruited 676 women from July 2005 to December 2006, with an average of 9.4 women per week. Another 11 women registered but completed only a small portion of the questionnaire before “logging off”. After exclusion of women with missing values for age and/or educational level, 670 subjects remained for the analysis. As summarised in Table 1, participants were older, of a higher educational level, more often primiparous and more often born in Italy compared with all women delivering in Turin in the first semester of 2005. Multivariate analyses revealed interactions of maternal age with educational level and place of birth. Therefore, we estimated the probability of participation by maternal age stratified by educational level and limited to women born in Italy, as the number of members of the NINFEA cohort born outside Italy was too small to carry out stratified analyses. As reported in Table 2, within each stratum of educational level, the

participation rate was higher among younger mothers, especially among graduate women.

Most of the women completed the questionnaire either at home (65.4%) or at work (31.3%); 94.5% of the women completed all 16 sections of the questionnaires, whereas 98.4% completed the first eight. Responses to the specific items varied from 100 to <10%. For example, all women completed the questions on smoking, and 82.8% answered all of the 19 questions on food items. Selected characteristics of partial and complete respondents are reported and compared in Table 3. Multiparous women, ever smokers, and those born outside Italy were more likely to be partial respondents, although the associations for the latter two characteristics did not reach statistical significance. Women participating during the first trimester of the pregnancy were also more likely to be partial respondents.

Table 4 reports the findings for the questions on hypothetical donation of biological samples. Similar proportions of women (about 80%) would have given a hypothetical donation of a sample of urine, nails and saliva, whereas a smaller proportion (64.2%) would have donated a blood sample. The proportion who would have agreed to a hypothetical donation of a sample of saliva taken from the child for the purpose of genetic analyses was relatively high (71.4%).

The ad hoc survey indicated that 66% (95% CI: 56–74%;  $n = 80$ ) of the women delivering at the main hospital of the city of Turin had access to the Internet and 42% (95% CI: 33–51%;  $n = 51$ ) were aware of the NINFEA cohort. The PORs of having access to the Internet were 2.08 (95% CI: 0.90–4.79) for women aged 33 years or more compared with women aged <33 years, and 14.4 (95% CI: 3.00–69.1) for being born in Italy compared with being born outside of Italy. Overall, 8 of the 122 interviewed women had participated in the NINFEA cohort, corresponding to 22% (95% CI: 10–38%) of the 37 women who were both aware of the study and had access to the

Internet. The most common reasons given for non-participation were lack of interest in the study ( $n = 15$ ), lack of time ( $n = 6$ ) and problems in using the Internet ( $n = 6$ ). No woman reported that they did not participate for fear of revealing personal information.

## Discussion

The findings of this study show that it is feasible to recruit a birth cohort through the Internet in a population where 60–70% of pregnant women have access to the Internet. The response rates could increase if the study were better advertised; furthermore, the source population and numbers recruited could increase if, as expected, the proportion of individuals having access to the Internet increases rapidly over time.

The use of the Internet to recruit a birth cohort offers some key advantages. First, recruitment procedures are simple. Once the website has been constructed, researchers only need to focus on advertising the study. This implies that recruitment may continue for several years and that the study population and period of recruitment and follow-up can be extended, facilitating the recruitment of a large cohort. Second, in multicentre studies there is no need for each centre to have advanced research infrastructures and full-time researchers as is required for a “traditional” birth cohort. Third, having all the members of the cohort with access to the Internet means that it is possible to use email as a further tool for follow-up, thus making the procedures more efficient, and possibly reducing losses to follow-up.

These advantages should be balanced with the limitations of the Internet-based approach. Selection bias may occur as participants in our cohort were different from the general population of pregnant women both because they had access to the Internet and because they were self-selected volunteers. This may explain the strong

**Table 2** Prevalence odds ratio of participation in the NINFEA cohort for maternal age stratified by educational level

Maternal age (years)	Educational level											
	Undergraduate						Graduate					
	NINFEA cohort ( $n = 277$ )		Turin ( $n = 3,773$ )		POR	95% CI	NINFEA cohort ( $n = 371$ )		Turin ( $n = 845$ )		POR	95% CI
	Numbers	(%)	Numbers	(%)			Numbers	(%)	Numbers	(%)		
≤29	69	24.9	811	21.5	1.20	0.84–1.72	46	12.4	54	6.4	2.46	1.54–3.93
30–32	71	25.6	965	25.6	1.08	0.76–1.53	112	30.2	148	17.5	2.34	1.66–3.31
33–35	70	25.3	936	24.8	1.14	0.81–1.61	120	32.3	293	34.7	1.46	1.06–2.00
36+	67	24.2	1,061	28.1	1.00		93	25.1	350	41.4	1.00	

Women born in Italy. NINFEA cohort, Italy, 2005–2006

POR prevalence odds ratio adjusted for number of previous pregnancies, CI confidence intervals

**Table 3** Prevalence odds ratios of complete response for selected characteristics

Characteristics	Complete respondents ( <i>n</i> = 633)		Partial respondents ( <i>n</i> = 37)		Crude POR	Adjusted POR (95% CI)
	Numbers	(%)	Numbers	(%)		
Maternal age (years)						
<33	289	(45.7)	19	(51.4)	1.00	
33+	344	(54.3)	18	(48.6)	1.26	1.57 (0.78–3.20)
Number of previous pregnancies						
0	398	(62.9)	14	(37.8)	1.00	
1+	235	(37.1)	23	(62.2)	0.36*	0.33 (0.16–0.69)
Place of birth						
Italy	614	(97.0)	34	(91.9)	1.00	
Non-Italy	19	(3.0)	3	(8.1)	0.35	0.34 (0.09–1.28)
Educational level						
Undergraduate	268	(42.3)	18	(48.6)	1.00	
Graduate	365	(57.7)	19	(51.4)	1.29	0.98 (0.49–1.96)
Gestational age at participation (trimesters)						
I	89	(14.0)	12	(32.4)	1.00	
II	234	(37.0)	9	(24.3)	3.51*	2.98 (1.19–7.49)
III	310	(49.0)	16	(43.3)	2.61	2.03 (0.90–4.58)
Smoking habit before pregnancy						
Never	448	(70.8)	21	(56.8)	1.00	
Ever	185	(29.2)	16	(43.2)	0.54	0.61 (0.30–1.24)

NINFEA cohort, Italy, 2005–2006

POR prevalence odds ratio; adjusted POR, adjusted for all variables listed in the table, CI confidence intervals

\**P* < 0.05

association that we found between educational level and participation in the NINFEA cohort, as well as the interaction between maternal age and educational level, as women with a higher education tend to give birth at older ages, and younger people, within each stratum of socioeconomic status, may have more frequently access to the Internet. The implication is that the study findings cannot be simply extrapolated to the overall population of pregnant women. However, selection does not necessarily produce a bias when the focus of the analyses is on internal comparisons. In a web-based health survey in 47,859 Swedish women [10], 41% of them responded to the web-based questionnaire and 31% to a paper questionnaire. Socioeconomic status and being a non-smoker were predictors of participation in the web-based questionnaire, but the associations were not strong, and they applied to both the web and paper questionnaires [10]. Furthermore, some of the observed differences between participants and non-participants in the current study may also be related to the methods used to advertise the study. For example, the distribution of leaflets during the pre-delivery classes is probably part of the reason for the overrepresentation of primiparous women in our cohort.

In fact, members of any cohort are almost always a selected group of the source population, irrespective of the

method used for the recruitment. For example, about 30% of the study population was enrolled in the Danish National Birth Cohort study, which recruited 100,000 pregnant women from 1997 to 2002 [21]. Lack of recruitment was equally due to non-collaboration of family doctors and women declining to take part in the study. The participants were estimated to include a higher proportion of women aged 25–35 years, multiparous, of normal weight and non-smokers or former smokers. However, the estimated associations of some selected risk factors with adverse perinatal events, such as smoking and foetal growth or body mass index and stillbirth, were only weakly biased by these differences in participation rates. This was not unexpected since, as noted above, non-participation or volunteer bias in cohort studies will not necessarily produce a bias in internal comparisons unless participation is jointly differential with regards to both the exposure and the outcome [22, 23].

The validity and completeness of the data collected by various methods is another important issue. The proportion of partial respondents in our study was small and similar to that observed in the web-based health survey of 47,859 Swedish women mentioned above [24]. It has been observed that responses to a web version of a questionnaire are usually more complete than the responses to a printed

**Table 4** Numbers and proportions of women who would give a hypothetical donation of a biological sample

Biological sample	<i>n</i>	%	(95% CI)
<b>Mother's sample</b>			
Urine for analyses on the metabolism			
Yes	507	80.6	(77.5–83.7)
No	122	19.4	
No answer	10	–	
Nail for the assessment of environmental exposures			
Yes	512	80.8	(77.7–83.8)
No	122	19.2	
No answer	5	–	
Saliva for genetic analyses			
Yes	502	79.4	(76.3–82.6)
No	130	20.6	
No answer	7	–	
Blood for genetic analyses			
Yes	406	64.2	(60.5–68.0)
No	226	35.8	
No answer	7	–	
<b>Child's sample</b>			
Saliva for genetic analyses			
Yes	446	71.4	(67.8–74.9)
No	179	28.6	
No answer	14	–	

NINFEA cohort, Italy, 2005–2006

CI confidence intervals

questionnaire [6]. Furthermore, the quality of data collected through the web may be higher than in postal questionnaires as there is the possibility of including immediate checks for impossible or inconsistent answers. In addition, data are entered by the respondents themselves and there is no possibility of mistakes due to the researchers' data entry. In our study there were no mandatory answers, which resulted in a certain degree of non-response for some questions. However, mandatory answers can be introduced for specific exposures which are judged to be of fundamental importance for the study.

Another possible limitation of recruitment through the Internet is the lack of contact with the participants, which makes the collection of biological samples more difficult to achieve. However, our findings suggest that, after recruitment has occurred, it is possible to ask the women to mail biological samples, such as a sample of saliva. For example, an 80% compliance from smokers recruited through the Internet has been found in a study involving the collection of saliva to carry out DNA and cotinine analyses [25].

Finally, the use of the Internet raises concerns about privacy. In this respect, it is reassuring that unwillingness to reveal personal data was never mentioned as a cause of non-participation in our survey. Furthermore, the “anonymity” of an Internet-based questionnaire may yield

higher response rates, and more accurate responses, for questions of a personal nature, than is achieved with face-to-face interviews or even postal questionnaires [13].

The present study was, to our knowledge, the first attempt to recruit a birth cohort through the Internet. We found that this approach was feasible and well-accepted by the potential participants. As with all recruitment methods, using the Internet has advantages and limitations which should be carefully evaluated. We suggest that in countries and centres in which a “traditional” cohort is difficult to establish (either for financial or logistical reasons) and access to the Internet is widespread, the use of the Internet can be an efficient alternative to traditional methods of recruitment and follow-up in birth cohort studies. Where this is feasible, even with relatively low response rates, larger study sizes, more complete follow-up, and greater international collaboration can be more easily achieved.

**Acknowledgements** The study was supported by Compagnia di San Paolo/FIRMS, the Piedmont Region and the Ministry for University and Research. Methods used in this study have been developed within the framework of projects supported by the Italian Association for Cancer Research. The Centre for Public Health Research is supported by a Programme Grant from the Health Research Council of New Zealand. We thank Fabio Saccona, Silvano Stralla and Monica Michelis for their IT support and Olga Sassu for her support to the coordination of the study. We are also grateful to health personnel involved in the advertisement of the study.

## References

1. Ekman A, Litton JE. New times, new needs; e-epidemiology. *Eur J Epidemiol* 2007;22:285–92.
2. Baer A, Saroiu S, Koutsky LA. Obtaining sensitive data through the Web: an example of design and methods. *Epidemiology* 2002;13:640–45.
3. Braithwaite D, Emery J, De Lusignan S, Sutton S. Using the Internet to conduct surveys of health professionals: a valid alternative? *Fam Pract* 2003;20:545–51.
4. Wolters FL, van Zeijl G, Sijbrandij J, et al. Internet-based data inclusion in a population-based European collaborative follow-up study of inflammatory bowel disease patients: description of methods used and analysis of factors influencing response rates. *World J Gastroenterol* 2005;11:7152–8.
5. Link MW, Mokdad AH. Alternative modes for health surveillance surveys: an experiment with web, mail, and telephone. *Epidemiology* 2005;16:701–4.
6. Balter KA, Balter O, Fondell E, Lagerros YT. Web-based and mailed questionnaires: a comparison of response rates and compliance. *Epidemiology* 2005;16:577–9.
7. Silver RC, Holman EA, McIntosh DN, Poulin M, Gil-Rivas V. Nationwide longitudinal study of psychological responses to September 11. *JAMA* 2002;288:1235–44.
8. Wang J, Etter JF. Administering an effective health intervention for smoking cessation online: the international users of Stop-Tabac. *Prev Med* 2004;39:962–8.
9. Tate DF, Jackvony EH, Wing RR. Effects of Internet behavioral counseling on weight loss in adults at risk for type 2 diabetes: a randomized trial. *JAMA* 2003;289:1833–6.
10. Ekman A, Dickman PW, Klint A, Weiderpass E, Litton JE. Feasibility of using web-based questionnaires in large population-based epidemiological studies. *Eur J Epidemiol* 2006;21:103–11.
11. Smith B, Smith TC, Gray GC, Ryan MA. When Epidemiology meets the Internet: web-based surveys in the millennium cohort study. *Am J Epidemiol* 2007. In press.
12. Marquet RL, Bartelds AI, van Noort SP, et al. Internet-based monitoring of influenza-like illness (ILI) in the general population of the Netherlands during the 2003–2004 influenza season. *BMC Public Health* 2006;6:242.
13. Ross MW, Mansson SA, Daneback K, Cooper A, Tikkanen R. Biases in internet sexual health samples: comparison of an internet sexuality survey and a national sexual health survey in Sweden. *Soc Sci Med* 2005;61:245–52.
14. Ritter P, Lorig K, Laurent D, Matthews K. Internet versus mailed questionnaires: a randomized comparison. *J Med Internet Res* 2004;6:e29.
15. Gosling SD, Vazire S, Srivastava S, John OP. Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *Am Psychol* 2004;59:93–104.
16. Kogevinas M, Andersen AM, Olsen J. Collaboration is needed to co-ordinate European birth cohort studies. *Int J Epidemiol* 2004;33:1172–3.
17. Eaton W. The logic for a conception-to-death cohort study. *Ann Epidemiol* 2002;12:445–51.
18. Frank J, Di Ruggiero E, McInnes RR, Kramer M, Gagnon F. Large life-course cohorts for characterizing genetic and environmental contributions: the need for more thoughtful designs. *Epidemiology* 2006;17:595–8.
19. Boldrini R, Di Cesare M, Pennazza F. Certificato di assistenza al parto (CeDAP). Analisi dell'evento nascita—Anno 2002. Roma: Ministero della Salute, Dipartimento della Qualità, Direzione Generale Sistema Informativo, Ufficio di Direzione Statistica; 2004.
20. Armitage P, Berry G, Matthews JNS. *Statistical methods in medical research*, 4th ed. Oxford: Blackwell Science; 2002.
21. Nohr EA, Frydenberg M, Henriksen TB, Olsen J. Does low participation in cohort studies induce bias? *Epidemiology* 2006;17:413–8.
22. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;15:615–25.
23. Greenland S. Response and follow-up bias in cohort studies. *Am J Epidemiol* 1977;106:184–7.
24. Ekman A, Klint A, Dickman PW, Adami HO, Litton JE. Optimizing the design of web-based questionnaires—experience from a population-based study among 50,000 women. *Eur J Epidemiol* 2007;22:293–300.
25. Etter JF, Neidhart E, Bertrand S, Malafosse A, Bertrand D. Collecting saliva by mail for genetic and cotinine analyses in participants recruited through the Internet. *Eur J Epidemiol* 2005;20:833–8.