

## Time to pregnancy: a measure of reproductive function in either sex

Michael Joffe, Asclepios project (appendix)

### Abstract

**Introduction**—Growing evidence of reproductive effects associated with occupational and environmental agents has created the need for research with sensitive and well validated methods. There is a complex relation between manifest effects and underlying pathogenic processes. Conceptions will on average tend to be delayed in a population exposed to an agent that causes embryonic damage, an increase in germ cell mutations, or decreased fertility.

**Studying time to pregnancy**—Time to pregnancy can be used to measure the degree of delay in conceiving, across the whole continuum of biological fertility, in either men or women. The distribution of time to pregnancy largely reflects a sorting process, as the more fertile couples become progressively less well represented with the passage of time. The basic research strategy is comparison of the time to pregnancy within groups defined by their exposures, allowing for potential confounding factors relating not only to the study subject but also to his or her partner.

**Measurement and validity**—Prospective and retrospective methods are available, and each has strengths and weaknesses. Prospective studies have some theoretical advantages, but have unrepresentative populations and problems of feasibility and cost. Retrospective assessment of time to pregnancy is feasible with a short questionnaire, without intruding into sensitive areas of respondents' lives, with good validity at the group level, and without the necessity of large populations. Potential biases have been identified that can be minimised by careful design and analysis; the principal remaining problem is difficulty in obtaining exposure data retrospectively.

(*Occup Environ Med* 1997;54:289–295)

the effect could be recognised relatively easily because the particular defect, phocomelia, was vanishingly rare in the absence of the drug. It would be far more difficult to recognise an increase in a common abnormality, even if this were responsible for many cases.

Congenital malformations have enormous emotional and practical implications for those affected. However, there are many other ways in which reproductive processes can be disrupted.

There is growing evidence that occupational factors can adversely affect reproductive processes. This is relevant both because many women have now joined the workforce, and because of the possibility of male mediated effects. Much exploratory work has been done, with linkage of routine data sets in those countries where these exist, case control studies, and relatively insensitive indicators such as demographic fertility (number and timing of births).

At present, evidence for or against reproductive or genetic toxicity of specific agents still exists only for a few of those in common use—for example, ionising radiation, lead, and the soil fumigant dibromochloropropane (DBCP). There is also epidemiological evidence which refers to groups of agents or to broad occupational categories, such as exposure to organic solvents or work in the electronics industry. In general, the information available is seriously incomplete, and the lack of satisfactory evidence cannot be taken to indicate lack of effect.

Research is now urgently needed, with the more sensitive methods that have been developed in recent years. The purpose of this paper is to describe the use of the time taken to conceive, time to pregnancy, to assess reproductive function in either sex. The focus is mainly on short questionnaires for use in retrospective cohort studies. Before describing the method and outlining what is known about its validity, it is necessary to understand how time to pregnancy relates to other reproductive end points, and to the underlying biology.

### Biological dysfunction and manifest effect

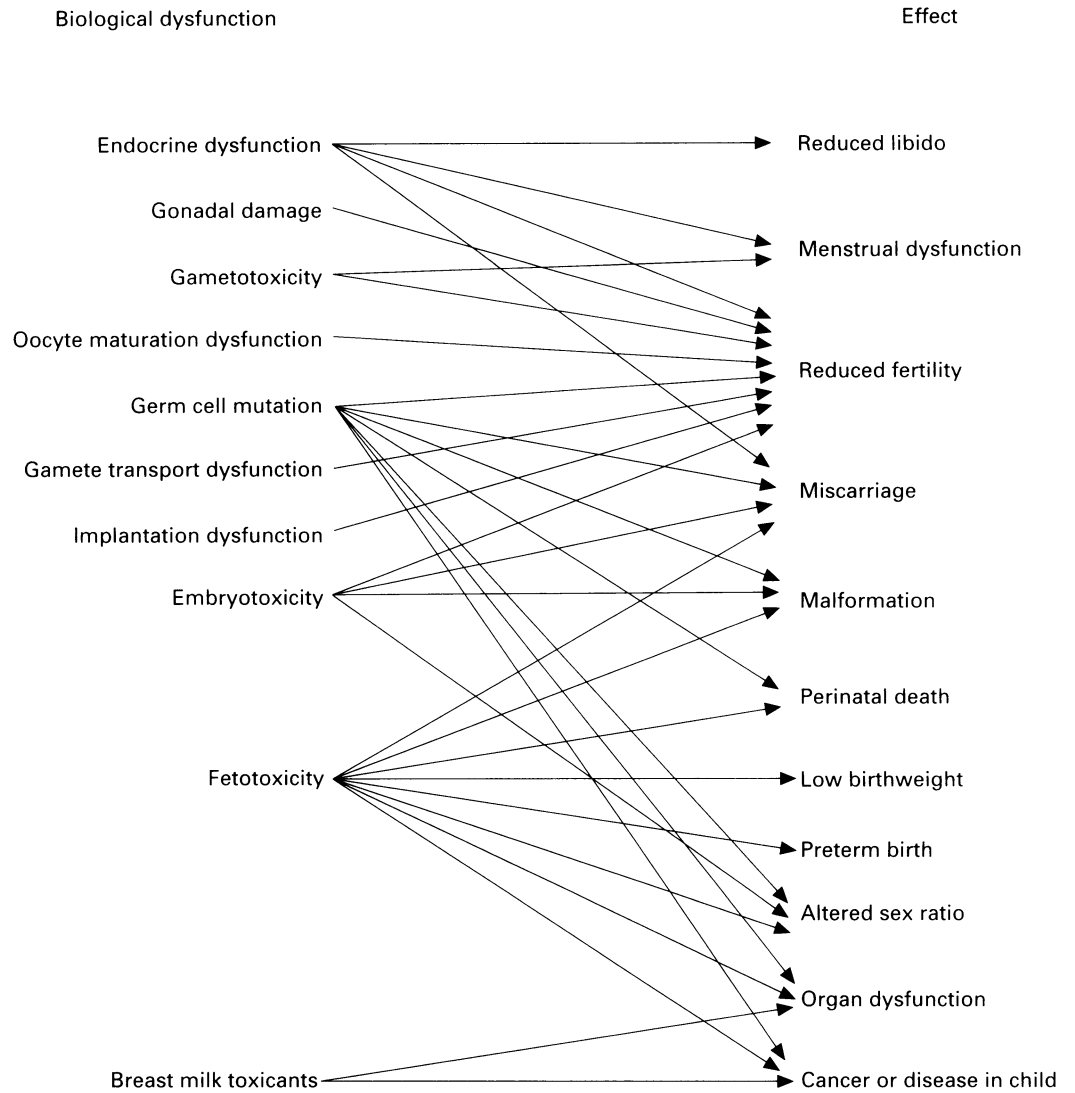
Several biological processes are involved in reproductive and genetic damage, and there are also several possible outcomes that could be found among those affected. The interrelation of these is not one to one: figure 1 shows an overview of the links between them.<sup>1</sup>

Keywords: research methods; questionnaires; time to pregnancy

### Why study time to pregnancy?

Over 30 years ago, the thalidomide tragedy showed that a chemical agent can cause devastating reproductive damage. In that instance,

Figure 1 Overview of possible relations between biological dysfunction and manifestation in reproductive outcomes (adapted with permission).<sup>1</sup>



A given pathogenic process can result in different manifestations. For example, an increase in germ cell mutations can cause various types of adverse health effects (fig 2): manifest genetic disease; increased frequency of heterozygous (carrier) status of recessive disorders; cancer, especially that occurring in childhood; and those congenital defects which have an important genetic component. Within each of these categories, many different clinical

conditions would be included. For example, a range of genetic diseases would occur, each one being rare; it would therefore be difficult to detect an increased risk in a particular population. Also, one would expect dominant effects causing fetal death, manifest as perinatal death or miscarriage, or embryonic death, which would be apparent only as a cycle during which conception did not occur, so that a clinically manifest pregnancy takes longer to

Figure 2 Possible consequences of germ cell mutation (adapted with permission).<sup>1</sup>

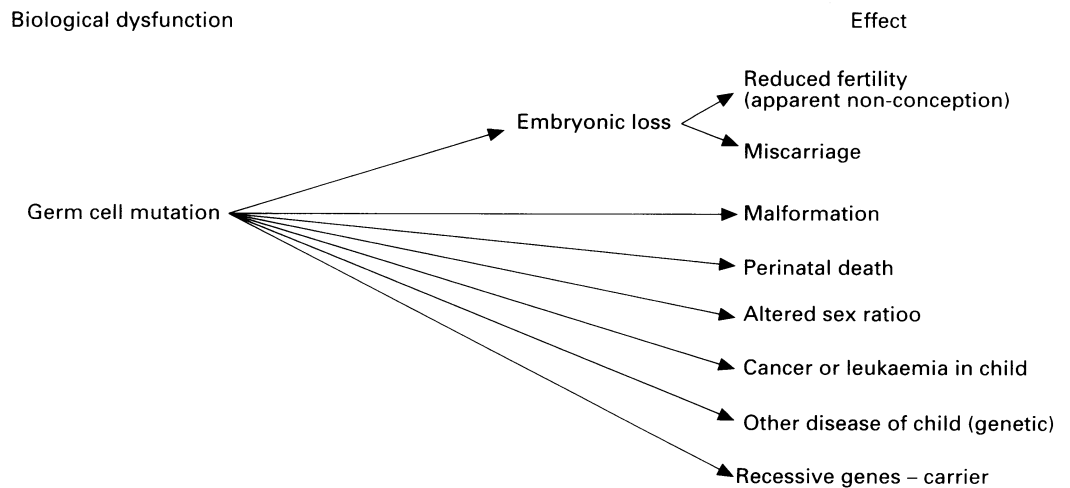
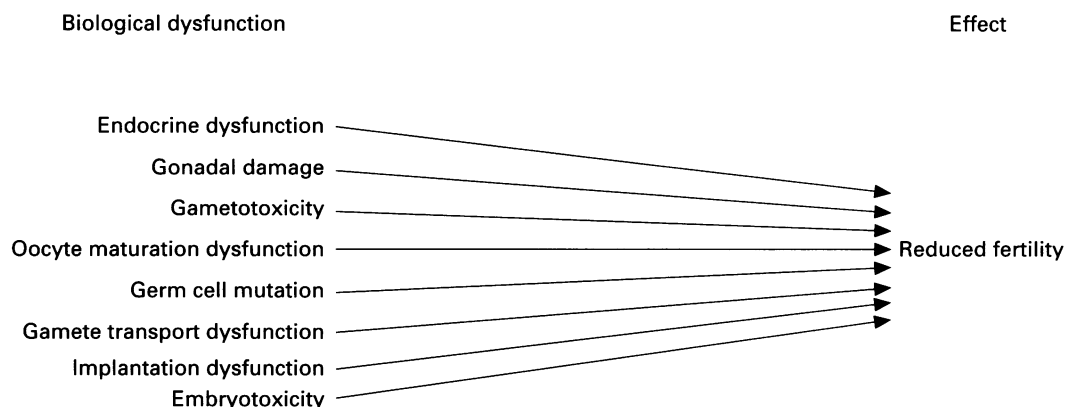


Figure 3 Possible antecedents of delayed conception (adapted with permission).<sup>1</sup>



conceive. At a group level, it would be manifest as a tendency for delay in conception in an exposed population.

On the other hand, a given manifestation can result from several types of biological dysfunction. For example, delay in conception may result not only from germ cell mutation but also from embryotoxicity (some embryotoxic agents are teratogens), as well as reduced biological fertility in a more literal sense, which can be due to toxic effects on the gamete, or gonad, or both, or to endocrine effects (fig 3); in some cases, a decrease in libido can also occur as a result of exposure to chemical agents. Most of these biological mechanisms can affect either sex.

#### The concept of time to pregnancy

The phrase "delay in conception" has been used to indicate that on average, the couples in a particular population take longer to produce a clinically recognisable pregnancy. This is always relative, a matter of comparing exposed and non-exposed conceptions, as there is no absolute standard.

It has long been recognised that human populations have a wide range of reproductive capacity: whereas some couples have a high probability of conception in a particular menstrual cycle, other couples have fertility that is low, or even zero (sterility). It is important to recognise that there is a smooth continuum between these extremes, and that categorisation into fertile and subfertile or infertile is necessarily arbitrary. As very fertile couples have a higher probability of conception within a few months of beginning unprotected intercourse, they are relatively unrepresented among couples who take a long time to conceive: thus, time to pregnancy measures a process that is partly a matter of sorting couples according to their degree of fertility, and partly a probabilistic occurrence for each couple, given their degree of fertility.

The most convenient way of representing time to pregnancy is as a cumulative percentage curve (fig 4).<sup>2</sup> The probability of conception in the first month is typically around 0.3, and gradually diminishes throughout the distribution; 50% of couples conceive within two months, and over 80% in six months; even after at least five years of non-conception have elapsed, the probability of conception is

greater than zero, so that the curve goes on rising.<sup>3</sup>

The practical importance of these findings is that it is much more statistically efficient to compare the time to pregnancy distributions of exposed and non-exposed populations than to compare proportions of those defined as clinically subfertile (no conception within for example, 12 months). Also, just one part of the curve may be affected, the classic case being recent use of the oral contraceptive pill which tends to delay conception in the early months, but does not affect the probability of conception in the long term.<sup>4</sup>

#### Study design

The natural design for studying time to pregnancy is a cohort study, populations being defined in terms of exposure status. The basic strategy is the comparison of the time to pregnancy distributions of exposed and non-exposed populations, or better, of subpopulations with differing degrees of exposure.

Research is typically aimed at studying only male mediated or female mediated effects, but the degree of fertility is a property of the couple, so potential confounding factors for both partners are needed. As in any epidemiological

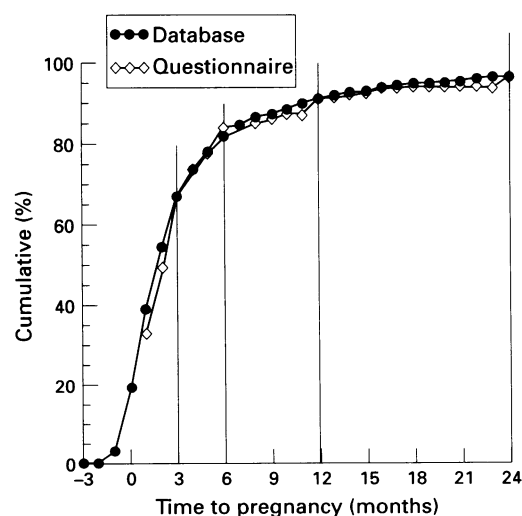


Figure 4 Comparison of time to pregnancy distributions recorded concurrently in the Oxford Family Planning Association study database and reported in a short self completion questionnaire, for 1392 matched pregnancies (taken with permission).<sup>2</sup>

study, every effort should be made to measure all variables that could be associated both with the outcome (time to pregnancy) and with the exposures of interest.

However, because the comparison is of populations, it is not essential to include information on every factor that could affect fertility. For example, in an individual case, conception could be delayed because illness or travel lead to an interruption in the sexual life of a couple. At the population level, this would only be important if an exposure group was especially predisposed to such events. The assumption can often be made that such factors are equally balanced between groups, or that their effect is small at the population level. The practical importance is that one does not have to acquire information at this level of detail, which is fortunate, as intimate questioning of this nature would not be appropriate for the occupational situation.

Similarly, although the frequency of intercourse is associated with the probability of conception, it is not essential to collect such data: it can often be assumed that its distribution is similar in the different exposure groups. In the special case of an agent that affects the libido, adjustment for frequency of intercourse would even obscure the adverse effect of the agent.

#### **Measuring time to pregnancy**

Time to pregnancy is a functional measure, corresponding not to any specific biological process, but rather to the final common path of conception. There are two approaches to its measurement, prospective and retrospective, corresponding to prospective and retrospective cohort study designs.

Conceptions that occurred in the presence of a normally effective method of contraception are excluded from consideration, so that time to pregnancy is only measured for that category of pregnancies that were conceived by couples who had embarked upon a period of unprotected intercourse. In principle, it is essential to base this distinction on actual behaviour and its biological implications, rather than on subjective aspects such as whether the couple was planning a pregnancy or trying to conceive.

In practice, this is especially problematic for prospective studies, as it affects recruitment: participating couples are likely to be those who are actively planning, so that those who have a more relaxed approach will be underrepresented. Retrospective studies have less of a problem in this respect, as the research population initially includes all pregnancies, and the decision is then which to exclude from the analysis, based on reported contraceptive behaviour; including knowledge of the exposure status of those excluded, any resulting bias can be evaluated. Also, the inherently blurred definitions involved can be explored statistically by carrying out a sensitivity analysis (described later).

The prospective method uses either a manifest biological outcome (birth of a baby),<sup>4</sup> or a

pregnancy test.<sup>5</sup> In the pregnancy test, a urinary human chorionic gonadotrophin (HCG) assay detects pregnancy within a few days after implantation of the fertilised ovum. Selection bias can arise, not only because those who tend to plan their pregnancies are favoured, as already mentioned, but also due to the need for participants to declare their intention to conceive, which they may be reluctant to do, especially in an occupational context. An additional problem, which is more fundamental, is that a sampling frame may not exist; however, this does not apply to occupational studies. Finally, problems of feasibility and cost limit the use of this approach, as large populations are required to generate enough pregnancies for study.

An early example of a prospective study was from the Oxford Family Planning Association contraceptive study, which showed that recent use of oral contraceptives was associated with an increased time to birth;<sup>4</sup> a subsequent report showed the effect of maternal smoking.<sup>6</sup> A more recent study in an occupational context, with a urinary pregnancy test, was concerned with physical activity.<sup>7</sup>

Retrospective questionnaire based studies have two important advantages. Firstly, many pregnancies are readily obtained (although only some of these are exposed pregnancies). Secondly, it is possible to study a historical period when exposures were considerably higher than at present; companies which have made improvements may be more motivated to take part in the research. A limitation is that exposure data may be difficult to obtain for earlier periods, and information on potential confounding factors is likely to be less accurate.

Some studies have been carried out with comparatively long duration of recall, of the order of 10 years or more. Others focus only on pregnancies conceived within a few years before the research, and so have advantages—such as the use of more detailed questioning, based on discrete menstrual cycles, and events such as illness or separation that occurred at a particular time; information on exposures and confounding can be expected to be more reliable than with longer recall. On the other hand, the advantages listed in the previous paragraph increase with the duration of recall.

An example of an occupational study that used relatively short recall is that of nitrous oxide exposure among dental assistants;<sup>8</sup> a study of pharmacy assistants focused on pregnancies occurring up to seven years previously;<sup>9</sup> and a study of exposure to pesticides among fruit growers used a 12 year period.<sup>10</sup>

In designing the questionnaire, it is advisable to focus on a single pregnancy, for two reasons: it is not statistically justifiable to include more than one pregnancy per couple, as these are not independent; and the more defined focus allows detailed questioning on potential confounding factors.

In the questionnaire that is being used in the Asclepius project, a major European collaborative study of male fertility in the occupational context, the most recent birth has been

chosen as the pregnancy of interest. The reason for restriction to births is that ascertainment of the occurrence and timing of births is clearer than in the case of miscarriages that avoid pregnancy recognition bias;<sup>11</sup> furthermore, the recall of time to pregnancy is almost certainly better when an actual child is involved. The most recent one is studied because the man is more likely to be in the same industry if this is chosen; in other words, it is the one most likely to be an exposed pregnancy.

Restriction of data collection to births raises the possibility previously mentioned, that an exposure may affect both fertility and the risk of miscarriage; pregnancies that miscarry tend to take slightly longer to conceive.<sup>12</sup> Summary data on all pregnancies is therefore also collected so that the miscarriage rate can also be estimated, as an additional outcome variable.

In the Asclepius questionnaire, exposure information is focused on the time 12 months before the baby was born, thus allowing nine months for the pregnancy and up to three months for spermatogenesis and spermiogenesis. This is not ideal, as exposures (and confounding factors) can vary during the waiting time before conception occurs. However, this will predominantly affect the relatively few couples with a long time to pregnancy, not the bulk of the population in each exposure group which is where the statistical power lies. Information on any other phases of infertility is also collected, which is important for those who have never conceived.

Potential confounding factors include age, smoking habit, and educational level of both parents, maternal occupation (for possible chemical exposures and occupational class), the last method of contraception (including breast feeding), and relevant medical conditions such as pelvic inflammatory disease. Maternal parity is another potential confounding factor, but it should not be automatically controlled for, as it reflects earlier fertility, which may be related to the exposure of interest.

From the practical point of view in an occupational context, it is vital to avoid questions about intimate or possibly stigmatised aspects of respondents' lives. Fortunately this is easy in time to pregnancy studies: asking about the duration of time before conception occurred has been found to be acceptable in a wide variety of different cultural settings. Frequency of intercourse as a possible confounding factor has been already discussed; no other information about sexual behaviour is relevant. The periods of respondents' lives that are not connected with a birth are irrelevant, as are contraceptive histories, casual or illicit sexual activity, deliberate termination of pregnancy, and any period of celibacy they may have had.

#### **Validity of retrospective questionnaire data**

In reproductive studies generally, it has been found that recall of characteristics relating to a pregnancy (and especially to an actual baby),

such as birthweight and gestational age, is remarkably accurate and durable.<sup>13</sup> This section outlines the evidence on recall of time to pregnancy.

In a 1987 study, male factory workers were found to be able to answer a questionnaire item on time to pregnancy, in the context of an interview study of reproductive effects of occupation, and the resulting distribution closely resembled that expected from prospective studies.<sup>14</sup> In the case of female workers, the predicted infertile worker effect<sup>15</sup> was found, that workplaces contain a relatively high proportion of women who have not produced a live birth. Apparently reasonable data were obtained with 10 or even 20 years of recall.

In an American study, the validity of recall obtained with a short, self completed questionnaire was compared with that from a detailed telephone interview, among 523 predominantly white dental assistants.<sup>16</sup> All the pregnancies had occurred within four years before the study. For most covariates, discrepancies were non-differential, the main effect of misclassification thus being to reduce statistical power. Accuracy was unrelated to duration of recall. Although menstrual cycles were used in the questioning, there was some evidence that women found it easier to think in months.

In a Dutch validity study, 132 participants in a previous prospective study of pregnancy failures were asked about the time to pregnancy; duration of recall was three to 20 months.<sup>17</sup> No systematic errors were found, but random errors were fewer with personal interview format than with telephone interview or self completed questionnaire. Reliability (test-retest) showed great stability of response after three to five weeks, in all formats. A question on phases of subfertility not related to a pregnancy was poorly understood.

A further validation study compared the values of time to pregnancy recorded essentially prospectively in the Oxford Family Planning Association contraceptive study with those reported several years later when members of that study were sent a short self completion questionnaire.<sup>2,18</sup> Figure 4 shows the findings, as a comparison at the group level. The curves are almost superimposed for most of the distribution, with a small degree of digit preference evident at 3, 6, 12, and 24 months.

At the individual level, considerable misclassification was found. This was greater for longer waiting times to pregnancy, confirming the view that research is best carried out as a comparison of whole distributions of time to pregnancy between population groups defined in terms of exposure characteristics, rather than as a comparison of the proportion of those defined as having clinical infertility.

The findings at group level are all the more remarkable, as the median recall time was 14 years (interquartile range 11 to 16 years); a similar graph confined to duration of recall above the median surprisingly showed an even closer correspondence between the two curves.<sup>2</sup>

The concept of time to pregnancy is thus

well understood by the general population; indeed it seems to form part of lay thinking. Typically a value for time to pregnancy can be obtained from over 90% of female respondents, and over 80% of male respondents.<sup>19</sup> It is best to ask the number of months as an open ended question, rather than to group the values—for example, 0–3, 4–6, etc.

If interviews are being used, it seems to be useful to provide potential respondents with advance warning of the topics that they will be asked about. If questionnaires are to be self completed in a male population, it is an advantage if the men can take the questionnaires home and complete them with their wives.

### Biases, analysis, and statistical power

There are several potential pitfalls in the design, statistical analysis, and interpretation of time to pregnancy studies.<sup>11</sup> Detailed discussion is beyond the scope of this paper. One example, pregnancy recognition bias, has already been mentioned. Another one, time trend bias, resulted in a spurious protective effect among dental assistants who wore gloves: time to pregnancy was apparently longer among assistants who did not wear gloves, because these women started trying to conceive before awareness of HIV was widespread.<sup>8</sup> Its occurrence depended on the study pregnancies being clustered in time, and would be much less of a problem in studies with long duration of recall.

Survival methods are appropriate for analysis, including cumulative percentage graphs (similar to fig 4, equivalent to Kaplan-Meier curves), log rank tests for bivariate analysis, and Cox regression for multivariate analysis. The Cox regression is ideal for comparing time to pregnancy distributions between populations with different exposures, allowing for potential confounding effects.

The criterion for including pregnancies in the basic analysis is that no effective form of contraception was used. In terms of the questionnaire, categories of birth control in use are: (a) not at all, (b) only sometimes, (c) nearly all the time, and (d) all the time; this means restricting analysis to (a) and (b). Also, to allow for the problem of blurred definitions already mentioned, a sensitivity analysis is needed, restricted to only (a).

Populations do not have to be very large. It has previously been estimated that to detect a 50% fall in the probability of conception with 80% power at the 5% level of significance would require only 55 exposed and 55 non-exposed pregnancies;<sup>20</sup> however, this assumed perfect information and no confounding. The degree of misclassification at the individual level in retrospective data means that stable estimates probably require more like 300 exposed pregnancies.

### Conclusion

Time to pregnancy is a sensitive measure of fertility in either sex, which can be measured

prospectively or retrospectively. Eligible pregnancies are defined in terms of objective criteria, relating to the use of contraception, not in terms of whether the couple was wanting, intending, or trying to conceive. Time to pregnancy is measured in menstrual cycles or months, and grouping into categories should be avoided, as should dichotomous analysis in terms of clinical infertility.

In retrospective cohort studies, with a suitably designed questionnaire, information can be obtained that is valid at the group level, and sample size requirements are such that many occupational and environmental exposures can be studied.

### Appendix

The Asclepius project is a European Union concerted action research project coordinated by The Steno Institute of Public Health, University of Aarhus, Denmark, with the following participants: Belgium, Gent (P Kiss, A Mahmoud, M Vanhoorne, H Verstraelen); Denmark, Aarhus (A Abell, JP Bonde, S Brixen Larsen, G Danscher, E Ernst, H Kolstad), Copenhagen (A Giwercman); England, London (A Dale, M Joffe, N Shah); Finland, Helsinki (M-L Lindbohm, H Taskinen, M Sallmen), Turku (J Lähdetie); France, Paris (P Jouannet, P Thonneau); Strasbourg (A Clavert); Germany, Erlangen (KH Schaller, W Zschiesche); Italy, Brescia (P Apostoli, S Porru), Milano (L Bisanti), Pietrasanta (L Lastrucci), Rome (M Spanò); Netherlands, Nijmegen (N Rocleveld, H Thuis, GA Zielhuis), Zeist (W de Kort); Poland, Lodz (K Sitarek). The Asclepius project is funded by the BIOMED programme of the European Union, and by several complementary national sources of funding. The questionnaire was developed and translated with the aid of a grant from the Health and Safety Executive. We thank an anonymous reviewer for helpful comments on an earlier draft.

- Joffe M. Epidemiology of occupational reproductive hazards: methodological aspects. *Rev Epidemiol Sante Publique* 1992;40:S17-25.
- Joffe M, Villard L, Li Z, Plowman R, Vessey M. A time to pregnancy questionnaire designed for long term recall: validity in Oxford, England. *J Epidemiol Community Health* 1995;49:314-9.
- Spira A. Epidemiology of human reproduction. *Hum Reprod* 1986;1:111-5.
- Vessey MP, Wright NH, McPherson K, Wiggins P. Fertility after stopping different methods of contraception. *BMJ* 1978;i:265-7.
- Canfield RE, O'Connor JF, Birken S, Krichevsky A, Wilcox AJ. Development of an assay for a biomarker of pregnancy and early fetal loss. *Environ Health Perspect* 1987;74:57-66.
- Howe G, Westhoff C, Vessey M, Yeates D. Effects of age, cigarette smoking, and other factors on fertility: findings in a large prospective study. *BMJ* 1985;290:1697-700.
- Florack EI, Zielhuis GA, Rolland R. The influence of occupational physical activity on the menstrual cycle and fecundability. *Epidemiology* 1994;5:14-8.
- Rowland A, Baird DD, Weinberg CR, Shore DL, Shy CM, Wilcox AJ. Reduced fertility among women employed as dental assistants exposed to high levels of nitrous oxide. *N Engl J Med* 1992;327:993-7.
- Schaumburg I, Olsen J. Time to pregnancy among Danish pharmacy assistants. *Scand J Work Environ Health* 1989;15:222-6.
- de Cock J, Westveer K, Heederick D, te Velde E, van Kooij R. Time to pregnancy and occupational exposure to pesticides in fruit growers in The Netherlands. *Occup Environ Med* 1994;51:693-9.
- Weinberg CR, Baird DD, Wilcox AJ. Sources of bias in studies of time to pregnancy. *Stat Med* 1994;13:671-81.
- Joffe M, Li Z. Association of time to pregnancy and the outcome of pregnancy. *Fertil Steril* 1994;62:71-5.
- Joffe M. Male- and female-mediated reproductive effects on occupation: the use of questionnaire methods. *J Occup Med* 1989;31:974-9.
- Joffe M. Feasibility of studying subfertility using retrospective self reports. *J Epidemiol Community Health* 1989;43:268-74.
- Joffe M. Biases in research on reproduction and women's work. *Int J Epidemiol* 1985;14:118-23.

- 16 Baird DD, Weinberg CR, Rowland AS. Reporting errors in time-to-pregnancy data collected with a short questionnaire: impact on power and estimate of fecundability ratios. *Am J Epidemiol* 1991;133:1282-90.
- 17 Zielhuis GA, Hulscher MEJL, Florack EIM. Validity and reliability of a questionnaire on fecundability. *Int J Epidemiol* 1992;21:1151-6.
- 18 Joffe M, Villard L, Li Z, Plowman R, Vessey M. Long-term recall of time-to-pregnancy. *Fertil Steril* 1993;60, 99-104.
- 19 Joffe M, Li Z. Male and female factors in fertility. *Am J Epidemiol* 1994;140:921-9.
- 20 Baird DD, Wilcox AJ, Weinberg CR. Use of time to pregnancy to study environmental exposures. *Am J Epidemiol* 1986;124:470-80.

## Vancouver style

All manuscripts submitted to *Occup Environ Med* should conform to the uniform requirements for manuscripts submitted to biomedical journals (known as the Vancouver style.)

*Occup Environ Med*, together with many other international biomedical journals, has agreed to accept articles prepared in accordance with the Vancouver style. The style (described in full in the *JAMA*<sup>1</sup>) is intended to standardise requirements for authors, and is the same as in this issue.

References should be numbered consecutively in the order in which they are first mentioned in the text by Arabic numerals on the line in square brackets on each occasion the reference is cited (Manson[1] confirmed other reports . . . [2][3][4][5] . . .). In future references to papers submitted to

*Occup Environ Med* should include: the names of all authors if there are seven or less or, if there are more, the first six followed by *et al*; the title of journal articles or book chapters; the titles of journals abbreviated according to the style of *Index Medicus*; and the first and final page numbers of the article or chapter. Titles not in *Index Medicus* should be given in full.

Examples of common forms of references are:

- 1 International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomed journals. *JAMA* 1993;269:2282-6.
- 2 Soter NA, Wasserman SI, Austen KF. Cold urticaria: release into the circulation of histamine and eosinophil chemotactic factor of anaphylaxis during cold challenge. *N Engl J Med* 1976;294:687-90.
- 3 Weinstein L, Swartz MN. Pathogenic properties of invading micro-organisms. In: Sodeman WA Jr, Sodeman WA, eds. *Pathologic physiology, mechanisms of disease*. Philadelphia: W B Saunders, 1974:457-72.