Smoking habits among pregnant Danish women: reliability of information recorded after delivery

Ulrik Kesmodel, Sjurður Fröði Olsen

Abstract

Study objective—To compare recall of smoking habits during pregnancy 0.5–3 years after delivery across groups defined by recall time (5 six month periods) and pregnancy outcome (pre-eclampsia, pregnancy induced hypertension, intrauterine growth retardation, preterm or post-term delivery compared with controls).

Design—Case-control nested in cohort study.

Setting and participants—A subsample of 503 women from a cohort of 6347 women established between 1989 and 1991 in Aarhus University Hospital.

Main results—Measures of agreement between concurrent and retrospective data on smoking status varied between 0.93 and 1.0 (sensitivity), 0.90 and 0.98 (specificity), and 0.79 and 0.98 (κ). Spearman’s correlation coefficients for number of cigarettes smoked/day varied between 0.87 and 0.97; mean differences were all close to zero. Accuracy of recall tended to diminish with increasing alcohol intake, particularly among women smoking ≥10 cigarettes/day.

Conclusions—Information on smoking habits could be accurately obtained retrospectively independent of recall time and pregnancy outcomes studied here. Accuracy diminished with increasing alcohol intake, particularly among heavy smokers.

It is well documented that smoking in pregnancy is associated with reduced fetal growth, abruptio placentae and placenta previa, perinatal mortality, and other ill health. When the complications studied are rare, the case-control design with retrospective assessment of exposure offers a very cost effective epidemiological design. A major weakness inherent in this design, however, is that spurious associations may be created between such retrospectively collected exposure data and pregnancy outcome caused by deterioration of recall over time and differential misclassification between cases and controls.

The extent to which such misclassification takes place when smoking in pregnancy is assessed after delivery has been studied under various conditions, but because of inconsistent results further exploration and documentation is warranted.

In Danish women in whom smoking habits in pregnancy had been assessed concurrently and retrospectively we attempted to examine to what extent recall depended on recall time and pregnancy outcome. Observed differences were also related to baseline variables such as social status and alcohol intake.

Methods

STUDY POPULATION AND SELECTION PROCEDURE

Data used in this study were from an earlier published case-control nested in cohort study. Subjects were a subsample from a cohort of 9434 pregnant women established during an approximate two year period between 1989 and 1991 in Aarhus University Hospital, covering a well defined geographical area. Approximately 99% of women in the area comply with the antenatal care programme. The associations between certain dietary factors in pregnancy and the risk of pre-eclampsia, pregnancy induced hypertension (PIH), intrauterine growth retardation (IUGR), preterm and post-term delivery were examined by expanding the cohort data base by a questionnaire sent to selected cases and controls. Apart from dietary information, information about smoking in pregnancy was obtained. As smoking information was also part of the cohort data base, we could compare retrospective and concurrent information on smoking as presented here.

Women with language problems (461), fetal death before 28 completed weeks of gestation (120), or multiple pregnancies (259), women delivering outside of a predetermined 24 month period (590), as well as women on whom information on baseline data required for the selection procedure was missing (1657) were excluded leaving a total of 6347 women as sampling frame from which cases and controls were selected for the study.

For the case-control study, all 43 pre-eclamptic women and all 136 women with PIH were selected, whereas samples of 182 women of 252 with IUGR, 153 of 221 delivering preterm, and 189 of 673 delivering post-term were obtained by random sampling within each group. The control group sample were to be at least the size of the largest of the case groups. A control group of 256 women was sampled from the entire cohort ensuring that they were evenly distributed over and covering all months of the period corresponding to the recall time in the case groups. For the purpose of the present analyses women eligible for more than one case group were included in all case groups for which they were eligible. Thus, the total number of women sampled amounts to less than the sum of women in the case and control groups. A total of 1011 women were sampled for the study.
Table 1 Measures of agreement for smoking status (smoking or not) are proportion of agreement ($P_o$), sensitivity (sens), specificity (spef), and $\kappa$, and for number of cigarettes smoked Spearman’s correlation coefficient ($r_v$) and mean difference ($d$). Comparison of concurrent and retrospective information on smoking habits in 2nd half of pregnancy: subdivision according to length of recall period. (Case and control groups were aggregated). Sensitivity and specificity were estimated using concurrent information as a reference.

<table>
<thead>
<tr>
<th>Smoking status (yes/no)</th>
<th>Number</th>
<th>$P_o$</th>
<th>Sens</th>
<th>Spef</th>
<th>$\kappa$</th>
<th>$r_v$</th>
<th>$d$</th>
<th>95% CI*</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1 year</td>
<td>30</td>
<td>0.97</td>
<td>1.00</td>
<td>0.95</td>
<td>0.92</td>
<td>0.96</td>
<td>−0.10</td>
<td>−0.38, 0.18</td>
</tr>
<tr>
<td>1–1.5 years</td>
<td>124</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.92</td>
<td>0.95</td>
<td>0.06</td>
<td>−0.19, 0.32</td>
</tr>
<tr>
<td>1.5–2 years</td>
<td>118</td>
<td>0.97</td>
<td>0.95</td>
<td>0.97</td>
<td>0.93</td>
<td>0.96</td>
<td>−0.10</td>
<td>−0.50, 0.29</td>
</tr>
<tr>
<td>2–2.5 years</td>
<td>132</td>
<td>0.95</td>
<td>0.97</td>
<td>0.94</td>
<td>0.87</td>
<td>0.88</td>
<td>−0.01</td>
<td>−0.54, 0.53</td>
</tr>
<tr>
<td>2.5–3 years</td>
<td>99</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.90</td>
<td>0.93</td>
<td>−0.31</td>
<td>−0.88, −0.13</td>
</tr>
</tbody>
</table>

*95% confidence intervals for mean difference.

Pre-eclampsia was defined as diastolic blood pressure >90 mm Hg with proteinuria >0.3 g/l or >> using dipsticks; PIH as diastolic blood pressure >90 mm Hg without proteinuria; IUGR as birth weight below the 5th centile for gestational age; preterm delivery as gestational age <259 days; and post-term delivery as gestational age >294 days.

For pre-eclampsia and PIH all case reports were checked manually to determine the time of diagnosis. All pre-eclamptic women and all but five women with PIH were diagnosed after 30th week of gestation.

Results

Of the 1011 women invited to participate in the retrospective study 764 (76%) returned the questionnaire with useful dietary information. Five hundred and three women (50%) of the study base had answered all questions on smoking habits in both questionnaires. Response rates for each subgroup were 60% (pre-eclampsia), 57% (PIH), 39% (IUGR), 40% (preterm delivery), 54% (post-term delivery), and 65% (controls).

Proportions of smokers among respondents in each subgroup (concurrent information) were 19% (pre-eclampsia), 9% (PIH), 50% (IUGR), 40% (preterm delivery), 22% (post-term delivery), and 26% (controls). For comparison, proportions of smokers in each subgroup in the cohort data base (concurrent information) were 14% (pre-eclampsia), 11% (PIH), 55% (IUGR), 42% (preterm delivery), 26% (post-term delivery), and 30% (controls).

Table 1 presents agreement between concurrent and retrospective information on smoking according to length of recall period. No substantial differences were seen in agreement as a function of recall interval, although mean differences might suggest a modest tendency towards retrospective overreporting of smoking habits. Thus, in women with the largest recall interval of 2.5–3 years, the mean difference amounted to −0.51 (95% CI: −0.88, −0.13). Apart from that the general impression was that of a high stability over time.

Table 2 presents agreement between concurrent and retrospective information on smoking according to pregnancy outcome. No noteworthy differences were seen in agreement between groups representing cases and controls.

We attempted to differentiate between various levels of smoking to evaluate whether the amount smoked influenced reliability of recall but could not detect any systematic trend in

---

**STATISTICAL METHODS**

Various statistical measurements have been recommended for assessing validity and reliability of epidemiological tools. Following Bellach we present the results of various statistics to reflect the stability of our results: Reproducibility of binary questions (yes/no) was assessed by proportion of agreement ($P_o$) among pairwise comparison in a two by two table, sensitivity, specificity, and $\kappa$. Sensitivity and specificity were estimated using concurrent information as a reference. Reproducibility of continuous variables (number of cigarettes per day) was assessed by Spearman’s correlation coefficient ($r_v$) and the mean difference ($d$). Difference was defined as concurrent measurement—retrospective measurement.

To account for potential effect modification and confounding all analyses were stratified by maternal social status. Subsequently we performed multiple linear regression analyses using the mean difference as dependent variable. Models included as independent variables all the above covariates as well as either case status or recall time (five categories in table 1 or as a continuous variable). Regression analyses were performed for all women together as well as for various levels of smoking.

The study design was approved by the regional ethics committee.

---

**DATA COLLECTION METHODS**

Concurrent information on smoking habits was obtained from the cohort data base and derived from a questionnaire filled in at 30th week of gestation by women attending routine antenatal care. Retrospective information from cases and controls was obtained during the autumn of 1992 from a mailed self administered questionnaire. A second set of the questionnaire was mailed in early 1993 to all non-respondents.

There were slight differences in the questioning on the two occasions: Concurrent questions related to 30th week of gestation, whereas retrospective questions related to 2nd half of pregnancy. This was done to facilitate information retrieval and thus avoid reporting error. Questions on smoking habits were otherwise similarly phrased and related to smoking status (smoking yes/no) as well as amount smoked (number of cigarettes per day).

The concurrent questionnaire focused primarily on psychosocial factors, education and work conditions, the retrospective questionnaire on dietary habits during pregnancy. Thus, both questionnaires focused mainly on factors other than smoking habits.

Information on covariates was obtained from the cohort data base and included maternal alcohol intake at 30th week of gestation (<1, 1–2, 3–14 drinks/week), maternal height (five categories), maternal prepregnancy weight (five categories), parity (four categories), and maternal social status as defined by the Social Research Institute in Copenhagen, (originally five categories, of which categories 1–3 were combined because of small numbers in categories 1 and 3).
Table 2. Measures of agreement for smoking status (smoking or not) are proportion of agreement ($P_x$), sensitivity ($s$), specificity ($sp$), $\kappa$, and for number of cigarettes smoked Spearman’s correlation coefficient ($r_s$) and mean difference ($d$). Comparison of concurrent and retrospective information on smoking habits in 2nd half of pregnancy: subdivision according to case and control groups. Sensitivity and specificity were estimated using concurrent information as a reference.

<table>
<thead>
<tr>
<th>Smoking status (yes/no)</th>
<th>Number of cigarettes smoked</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$P_x$</td>
</tr>
<tr>
<td>Controls</td>
<td>167</td>
</tr>
<tr>
<td>Pre-eclampsia</td>
<td>26</td>
</tr>
<tr>
<td>PIH</td>
<td>78</td>
</tr>
<tr>
<td>IUGR</td>
<td>99</td>
</tr>
<tr>
<td>Preterm</td>
<td>89</td>
</tr>
<tr>
<td>Post-term</td>
<td>102</td>
</tr>
</tbody>
</table>

*95% confidence intervals for mean difference.

Discussion

In these data Danish women seem to be able to accurately recall their smoking habits during pregnancy independent of recall time and pregnancy outcome. Recall, however, seems to diminish with increasing alcohol intake, particularly among women smoking >10 cigarettes/day.

Previous investigations studying possible deterioration of recall over time have shown that smoking status (smoking or not) among non-pregnant women is fairly accurately recalled, whereas recall of the number of cigarettes smoked in pregnancy seems to diminish with time as does recall of other life style factors such as dietary information. In our study the tendency towards overreporting of number of cigarettes smoked after a recall period of 2.5–3 years was less than one cigarette on average, and no time trend was noted over the recall period of 0.5–2.5 years.

Studies of possible differential misclassification have shown inconsistent results: one study showed no differences in accuracy of recall between cases and controls (elderly non-pregnant women), some showed slight differences that did not, however, substantially affect associations between smoking and pregnancy outcome away from the null hypotheses; and some showed a clear difference, possibly because of differential misclassification. We found no differential misclassification, yet the results from this study may not apply to women with pregnancy outcomes perceived as more severe than those studied here, such as malformations and spontaneous abortions.

The context in which the questions were asked may be of significance. In both questionnaires questions on smoking behaviour were mixed with a number of other questions. Smoking behaviour has thus not appeared to be the target issue. This could have reduced the tendency towards “deliberate” misreporting, which is one mechanism that may underlie recall bias.

Women who were willing to participate in our case-control study may have been more able and motivated to accurately recall their smoking habits than the non-respondents, making our results slightly optimistic. However, this is unlikely to put a serious limitation to the generalisability of our findings, as the proportion of smokers among respondents does not differ substantially from the proportion of smokers in the cohort.

Recall might further depend on the general acceptance of smoking in pregnancy and the emphasis placed on smoking cessation in antenatal care programmes. Smoking cessation programmes may thus increase focus in antenatal care on the adverse effects of smoking in pregnancy and hence give rise to self reproach in smokers, which subsequently may influence their recall. Smoking cessation trials were not introduced in Aarhus until 1994, and our results may thus reflect the best potential recall of women not influenced by such self reproach.

Key Points

- Information on smoking habits during pregnancy may be accurately obtained retrospectively 0.5–3 years in the past.
- Information on smoking habits may be accurately obtained retrospectively independent of adverse pregnancy outcomes including preterm delivery and intrauterine growth retardation.
- Accuracy of recall seems to diminish with increasing alcohol intake, particularly among women smoking >10 cigarettes/day.
It is interesting to notice that women who drank 5–14 drinks/week and smoked ≥10 cigarettes/day substantially underreported their smoking habits, on average four cigarettes/day, when asked after delivery. It is well known that women tend to decrease their alcohol consumption during pregnancy, and these women may thus represent moderate-heavy drinkers when not pregnant. Still, it is difficult to say more precisely what are the mechanisms underlying this finding. Nevertheless, the finding implies that data from studies using retrospective assessment of these exposures should be interpreted with caution.

These data add to earlier evidence that information on smoking habits may be obtained retrospectively in an accurate manner 0.5–3 years in the past independent of the pregnancy outcomes studied here. Accuracy, however, seems to diminish with increasing alcohol intake, particularly among women smoking ≥10 cigarettes/day. This has potential implications for case-control studies assessing the simultaneous effects of smoking and alcohol exposure in pregnancy.

We would like to thank the staff of the Perinatal Epidemiological Research Unit for their help. Particularly, we are grateful to Tine Brink Henriksen PhD, Morten Hedegård PhD, Professor Niels Jørgen Secher, Jannie Dalby Salvig PhD as well as Data Manager Mr Jakob Hjort for valuable discussions and for providing data from the cohort data base.

Funding: this study was supported by The Danish Medical Research Council (J.No. 12-1082-1) and the Danish Research Foundation.

Conflicts of interest: none.

Smoking habits among pregnant Danish women: reliability of information recorded after delivery.
U Kesmodel and S F Olsen

J Epidemiol Community Health 1999 53: 239-242
doi: 10.1136/jech.53.4.239

Updated information and services can be found at: http://jech.bmj.com/content/53/4/239

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to: http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to: http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to: http://group.bmj.com/subscribe/