Smoking in pregnancy: some notes on the statistical controversy

HARVEY GOLDSCHMIDT*

From the National Children's Bureau, 8 Wakley Street, London

SUMMARY The statistical evidence for a relationship between pregnancy smoking, birthweight, and perinatal mortality is critically examined. Some apparent discrepancies between and within different studies are resolved.

The debate over the relationship between maternal smoking during pregnancy and subsequent damage to the fetus has centred around two issues. The first issue is similar to one often encountered in other debates on smoking, namely, whether a statistical association can be interpreted causally. The second issue is concerned with the reliability of the statistical results themselves.

Unlike most other links between smoking and disease, the case of pregnancy smoking can be submitted to designed experimentation to give a rigorous test of the causal hypothesis. Such an experiment would consist of a group of pregnant women subjected to intensive anti-smoking advice and compared, in terms of the status of their offspring, with a control group not so subjected. The first such experiment which has been adequately designed, although inconclusive in its results, has now been reported (Donovan, 1977) and the purpose of the current paper is to provide a background and summary of previous evidence and to comment on it scientifically.

It has often appeared that the important debate is about the interpretation of causation and not about the statistical findings themselves (Nature, 1973). In fact, however, rival causal interpretations of the statistical associations do not account for most of the controversy. They have certainly been important motivations inspiring different research workers, but the disputes themselves have centred largely on the comparison and validity of the statistical analyses. Moreover, the demonstration of a statistical association is only one stage in the whole scientific process, and if one accepts the views of Popper (1973), the idea of a direct 'causal' link between smoking in pregnancy and fetal outcome is a rather better theory than, say, the 'constitutional' or 'genetic' theories, because it is more open to falsification. I shall describe some unsuccessful attempts to falsify the first theory, whereas I know of no published attempts to falsify the other two types of theory. It is precisely the all-embracing statement of the other two theories that makes them scientifically unattractive, and also, incidentally, drives their proponents into attempting statistical falsifications of the first theory.

There are now many papers that confirm the finding of Simpson (1957) that babies born to mothers who smoke during pregnancy are, on average, about 200 g lighter than those born to non-smokers. There are fewer papers on the relationship between smoking and perinatal mortality and they can be divided into those which detect a sizeable association and those which do not. There are few articles which analyse the relationship between smoking in pregnancy and subsequent child development.

Birthweight

Several studies have demonstrated a dose-effect relationship between mean birthweight and the numbers of cigarettes smoked during pregnancy (Butler et al., 1972; Niswander and Gordon, 1972). They show a large birthweight reduction in moving from non-smokers to those mothers who smoke up to about 15 cigarettes a day but there is a smaller decrease beyond this amount. Since most studies, however, only classify the mothers into smokers and non-smokers this dichotomy will be used in the following discussions.

The proportion of the population who smoke differs quite considerably from one study to another (Table 1) and hence we would expect the average number of cigarettes smoked to differ also, and consequently that the differences between

*Present address: Institute of Education, University of London, 20 Bedford Way, London WC1H OAL.
average birthweights of children of smokers and non-smokers will change. In fact the range of
difference in birthweight is quite small for most
studies, about 160-200 g. There also seems to be
general agreement (Comstock et al., 1971) that the
birthweight distribution for children of mothers who
smoke is almost identical with that for non-smokers
except for a shift in the mean value. Figure 1
illustrates this shift and also the likely consequence
that this will have for the low birthweight babies
when this is defined as those of 2500 g and under;
paradoxically although mothers who smoke will
have a greater proportion of low birthweight babies, the mean birthweight of the smokers' babies in this group will probably be higher than
that of the mean weight of the non-smokers' babies.
Although this will not necessarily be true for all populations, the difference in the mean
birthweight of low birthweight infants of smoking
and non-smoking mothers will be considerably
narrower than the difference in mean birthweights
of all infants. In the British Perinatal Mortality
Survey, for example, low birthweight babies of
mothers smoking 10 or more cigarettes a day have
a mean birthweight of 2072 g, those smoking 1-9 a
day have a mean birthweight of 2011 g, and non-
smokers have a mean birthweight of 2062 g. In fact,
if the birthweight distributions are exactly gaussian
with equal variances, then Fig. 1 represents what
in fact happens for these distributions. Thus
in terms of mean birthweight, the low birthweight
babies of smokers do not suffer a disadvantage, a
fact which is of some technical interest but little
practical relevance. It will be referred to again when
discussing the controversy over perinatal mortality
rates.

Those who claim that the statistical association
is spurious because it is merely a reflection of a
deeper underlying cause hold that there are certain
kinds of mothers who both smoke and happen to
have lower than average birthweight babies, and
that the real task is to discover what kinds of
mothers these are. Although it is not possible
to dispose of this claim by citing evidence only from
epidemiological surveys, it is possible to narrow
the range of plausible explanations using such
evidence. For example, several studies have made
allowance for possible 'mediatory' factors such as
social class, maternal age, and parity. Within the
categories of such factors, however, the difference
between smokers and non-smokers persists virtually
unchanged (Butler et al., 1972). If there is a
mediatory factor it seems that it is likely to be of a
different kind, possibly a genetic one.

Another test of the causal explanation is to study
those women who gave up smoking in the early

Table 1 Proportion of mothers who were smokers, proportion of births <2500 g in smokers and non-smokers,
ratio of mortality in smokers to that in non-smokers in six large samples from white populations

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Sample size</th>
<th>Smokers %</th>
<th>Low birthweight</th>
<th>Mortality ratio smokers/non-smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rantakallio</td>
<td>1969</td>
<td>11 700</td>
<td>21</td>
<td>6.1</td>
<td>1.01 (PNM)</td>
</tr>
<tr>
<td>Yerushalmi</td>
<td>1971</td>
<td>9 800</td>
<td>38</td>
<td>6.4</td>
<td>1.03 (NNM)</td>
</tr>
<tr>
<td>Nisander and Gordon</td>
<td>1972</td>
<td>19 000</td>
<td>54</td>
<td>9.5</td>
<td>1.12 (PNM)</td>
</tr>
<tr>
<td>Ontario Canada*</td>
<td>1967</td>
<td>51 500</td>
<td>43</td>
<td>8.9</td>
<td>1.27 (PNM)</td>
</tr>
<tr>
<td>Butler et al.</td>
<td>1972</td>
<td>223 300</td>
<td>32</td>
<td>9.3</td>
<td>1.28 (PNM)</td>
</tr>
<tr>
<td>Comstock et al.</td>
<td>1971</td>
<td>12 500</td>
<td>37</td>
<td>11.1</td>
<td>1.40 (NNM)</td>
</tr>
</tbody>
</table>

*The Ontario Study included babies born dead before 28 weeks' gestation, these rates have been adjusted to conform to the ICD 1967 definitions of perinatal mortality. This has little effect on the mortality ratio.

PNM: The mortality ratio is based upon perinatal mortality rates.

NNM: The mortality ratio is based upon neonatal mortality rates.
Smoking in pregnancy: some notes on the statistical controversy

part of pregnancy. If smoking acts as a causal agent during pregnancy, then these women ought
to have babies with an average birthweight between
that of the babies of smokers and non-smokers. In fact, those who stated that they had given up
smoking by the fourth month of pregnancy had
babies whose birthweight distribution was virtually
indistinguishable from that of non-smokers; (in
the British Perinatal Mortality Survey the respective
average birthweights are 3367 g and 3350 g). It is,
of course, possible to argue that the woman who
gives up smoking during pregnancy is really a
‘non-smoking’ type. The point, however, is that
this represents another failure to falsify the causal
hypothesis. In addition it suggests that the second
half of pregnancy may be the vulnerable period
but this should be interpreted cautiously because
such information about changes in smoking habits
is usually obtained retrospectively and may not
be reliable concerning timing.

We could test further the causal hypothesis by
comparing babies born to those women who gave
up smoking after delivery with babies of mothers
who were non-smokers. This was attempted by
Yerushalmy (1972) who found a small but statistically
non-significant increase in the proportion
of low birthweight babies among the former
group. He also found that the proportion of low
birthweight babies born to mothers under 25
years old who had their first baby before they started
to smoke was significantly higher than among
the corresponding non-smokers. On the face of it this
seems to provide evidence against the causal
hypothesis, but on closer examination the validity
of these findings turns out to be questionable.

Firstly, the data on smoking habits were obtained
retrospectively going back over a number of years,
with the possibility that many of the mothers who
were classified as becoming smokers only after their
babies’ birth would in fact have been smoking
during the pregnancy. Secondly, the mothers who
smoked not only had their first baby before they
were 25 years old, but actually started to smoke
before that age, so that they would tend to have had
their first babies at younger ages than those mothers
who had never smoked under 25 years. In fact, they
were on average about two years younger in
Yerushalmy’s study (personal communication).
Moreover, several studies, including that by
Yerushalmy, have shown that the proportion of low
birthweight babies born to mothers under 20 years
is about 20% to 25% higher than for mothers aged
between 20 and 24 years. Hence a proper age
standardisation would be necessary before drawing
conclusions. Until a large-scale prospective study
following young women through several pregnancies
can be carried out, this test of the causal hypothesis
cannot be made. It appears, therefore, that all
tries to falsify the causal hypothesis concerning
the effect of smoking on birthweight have so far
proved failures, thus serving to strengthen our
belief in it.

Perinatal mortality

Table 2 sets out the chances, for different sample
sizes, of detecting a statistically significant difference
at the 5% level between the perinatal mortality
rates of infants born to smokers and non-smokers.
This table demonstrates that it is only the very
largest samples which are likely to give a clear
indication of any difference.

Table 2 Sample sizes necessary to achieve a given
chance of obtaining a statistically significant difference
(at the 5% level) in perinatal mortality rates of smokers
and non-smokers

<table>
<thead>
<tr>
<th>Probability of a significant result</th>
<th>Sample size (nearest 1000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>9 000</td>
</tr>
<tr>
<td>0.6</td>
<td>11 000</td>
</tr>
<tr>
<td>0.7</td>
<td>14 000</td>
</tr>
<tr>
<td>0.8</td>
<td>17 000</td>
</tr>
<tr>
<td>0.9</td>
<td>23 000</td>
</tr>
<tr>
<td>0.95</td>
<td>29 000</td>
</tr>
</tbody>
</table>

It is assumed that the mortality rates and proportions of smokers and non-smokers are those found in the British Perinatal Mortality Survey.

The mortality ratio is defined as the ratio of the mortality rate found in infants of mothers who
smoke to the rate found in those who are non-smokers. This mortality ratio may be calculated
either for the perinatal mortality or for the neonatal mortality according to the information available
in a particular study. Butler et al. (1972) and
Comstock et al. (1971) found it was possible to
calculate both ratios and the difference was very
small. The results for six studies are shown in Table I.
The first three studies have low mortality ratios which
do not give a statistically significant difference.
But these three studies also have the lowest
proportion of low-birthweight babies among non-smokers, with a very close association between
this low-birthweight rate and the mortality ratio.
This association provides one possible explanation
for the differing mortality ratios. Figure 2 shows the relationship between perinatal mortality
and birthweight in the British Perinatal Mortality
Survey. It is clear that a reduction of about 200 g
for a baby weighing 3500 g leads to less than a 10% increase in mortality risk, whereas the same
reduction for a 2500 g baby results in about a 30% increase. Clearly therefore, the more low-birthweight
babies there are in a population, the greater the increase in the mortality rate associated with the average decrease of about 200 g due to smoking. One important consequence of this explanation, if we assume a causal relationship, is that the greatest risk from smoking is to the mother who is already at 'high risk', who can expect a lower than average birthweight baby and who is poor, short, high parity, etc.

Perinatal mortality rate by birthweight.
(1958 British Perinatal Mortality Survey, mean = 100.)

Professor Yerushalmy, who was one of the main opponents of the causal hypothesis, raised a further paradoxical finding which he claimed cast doubt on a causal link. He points out that in his own and other studies, the mortality among low birthweight babies of smokers is lower than among low-birthweight babies of non-smokers. Yerushalmy (1972) failed to account for this in terms of downward shift in birthweights of smokers. Unfortunately, his claim that a 'shift' explanation is untenable is not justified, for reasons given in detail elsewhere (Goldstein, 1972). This gist of the argument, however, can be gleaned from Fig. 1, where as pointed out earlier, the low-birthweight babies of smokers tend to have a higher average birthweight than those of non-smokers. There is no reason to be surprised, therefore, if they manifest a lower mortality. Far from being 'paradoxical', such a finding is to be expected, and is not particularly relevant to the main issue.

As in the case of birthweight, the relationship between smoking and mortality persists after allowing for similar mediatory factors, and there is a similar dose-response relationship between perinatal mortality and number of cigarettes smoked. Also, those who had given up by the fourth month had a very similar rate (31·6 per 1000 births in the non-smokers and 32·2 per 1000 births in the smokers). As with birthweight, therefore, these failures to falsify the causal hypothesis lend it further strength.

**Subsequent child development**

Very few studies have followed-up large enough samples of babies to compare adequately the development of those born to smokers with those born to non-smokers. The surviving babies of the British Perinatal Mortality Survey have, however, been followed to the age of 11 years in the National Child Development Study, and small differences still persist at that age (Butler and Goldstein, 1973).

There is, for example, a 1·6 cm difference in height between children of heavy smokers (10 or more cigarettes a day) and non-smokers. Unlike birthweight, however, this difference is reduced by nearly 40% when account is taken of 'mediatory' factors such as social class, family size, and maternal height. This suggests that the association may not be wholly causal, and that further mediatory factors are taken into account, they may explain the major part of this association. In any case, the difference at this age is small when compared with the average differences associated with social class and family size which are of the order of several centimetres. The story is similar for educational attainment. With reading, for example, the difference between children of heavy smokers and non-smokers is about nine months of reading age, which reduces to about four months after allowing for the same mediatory factors.

**Conclusions**

The main objections to the statistical case for a causal relationship between smoking in pregnancy and birthweight and mortality do not bear much scrutiny. Apparent inconsistencies in comparisons of mortality rates have, in fact, straightforward interpretations in terms of the downward shift in the birthweight distribution among smokers. Various attempts to falsify the causal hypotheses have failed, leaving us with good reason for acting as if smoking really did cause a decrease in birthweight and an increased risk of perinatal mortality. On this assumption, it is estimated that in the perinatal period about 1500 babies die in Britain every year because their mothers smoke, out of a total of about 18 000 deaths. In the United States of America about 4600 die out of a total of about 87 000 deaths.
Smoking in pregnancy: some notes on the statistical controversy

a year (United States Public Health Service, 1973.)
Whatever is felt about the usefulness of health education propaganda, there are seen to be good reasons for attempting to dissuade pregnant women from smoking. If properly planned, with adequate controls, health education experiments can be mounted to determine not only whether women respond by changing their smoking habits, but also whether there is any appreciable effect on the pregnancy outcome. The apparent success or failure of such experiments may not only throw light on the causal hypothesis, but may also help to design efficient health education techniques.

My thanks are due to Michael Healy, Ken Fogelman, Maurice Backett, and Neville Butler for their comments and to the National Birthday Trust Fund for the use of some of the material. This work was supported by grants from the Department of Health and Social Security, the Department of Education and Science and the Social Science Research Council (UK).

Reprints from Harvey Goldstein, Department of Statistics and Computing, Institute of Education, 20 Bedford Way, London WC1H 0AL.

References