General practitioner notes as a source of information for case-control studies in young women

C E D Chivers, M C Pike, C N Taylor, C Hermon, B Crossley, S J Smith, and the UK National Case-Control Study Group

Abstract

Study objective – The UK National Case-Control Study was carried out to investigate the relationship between oral contraceptive use and breast cancer risk. This study investigates whether general practitioner notes could be used as the sole data source for epidemiological studies of young women and what the effect would be on non-response and recall bias.

Design – Case-control study with data on gynaecological, obstetric, and contraceptive history collected at interview and from general practitioners’ notes. Information from these two sources was compared.

Setting – This was a population-based study.

Participants – Altogether 755 women with breast cancer aged under 36 years at diagnosis, each with an age-matched control, participated in the study. Response rates at interview were 72% and 89% for cases and controls but GP data were available for 90% of the 1049 case and first-selected control pairs.

Main results – There was generally good agreement between the two data sources with respect to obstetric and gynaecological procedures (hysterectomy, oophorectomy, and tubal ligation). The use of intra-uterine devices, or diaphragm, and partner’s vasectomy were not reliably recorded in the GP’s notes. The overall results of the UK study would have been qualitatively the same with respect to the relationship between oral contraceptive use and breast cancer risk if GP notes only had been used, in spite of the fact that only about half of all oral contraceptive usage was recorded in the notes. Response rates would have been higher, recall bias eliminated, and the cost of the study halved.

Conclusions – When planning case-control studies in young women, the possibility of using GP notes as the primary data source should be considered. Lack of data on potential confounding factors is a possible drawback to such use. The practice of destroying GP’s notes shortly after the death of patients seriously restricts the possibility of using these notes when studying rapidly fatal conditions.

The UK National Case-Control Study Group was set up in 1983 to investigate the relationship between oral contraceptive use and breast cancer risk in young women. Sources of information included personal interviews and general practitioners’ (GP) notes. The objectives of using GP notes were twofold: to supplement information relating to oral contraceptive brand and usage obtained at interview and as a means of checking for non-response bias and recall bias. Here we compare the information derived from interview with that abstracted from GP notes, evaluate certain aspects of the extent and nature of non-response and recall bias, and determine whether GP notes may be used as the sole data source for epidemiological studies in young women.

Methods

The UK National Case-Control Study has been described in detail elsewhere. Briefly, all women who were diagnosed as having breast cancer between 1 January 1982 and 31 December 1985 and who were resident in any of 11 health regions in the UK were included, provided that their breast cancer diagnosis was before their 36th birthday. For every case, one control subject was chosen, effectively at random, from the list of that case’s GP. The control’s date of birth was matched to within six months of the date of birth of the case, and the control had to have been registered with the GP before the date of diagnosis of the case. If a case could not be interviewed, no attempt was made to interview her matched control. If the chosen control could not be interviewed a second (or further) control was selected in the same manner. For both cases and controls, the study was further restricted to white women with no previous malignancy, severe mental handicap, or psychiatric condition. Every control was given a “pseudodiagnosis” date, the date on which she was exactly the same age as her matched case had been at diagnosis.

Personal interviews were carried out in the women’s homes by trained interviewers. The same interviewer interviewed both members of each case-control pair. The interview included questions on basic demographic details, and reproductive and contraceptive history, including brands of oral contraceptives used. Questions were asked up to diagnosis or pseudodiagnosis date. After interview, data on obstetric and oral contraceptive history were abstracted from the GP notes by the same interviewers. Entries in the notes dated after
the diagnosis/pseudodiagnosis date were
ignored, even if they referred to events before
that date. The notes were abstracted onto a
structured form and interviewers were
instructed to use all information (including
correspondence) in the notes. Where an item
in the notes was undated, the abstractor used
her judgment as to whether it happened before
or after diagnosis/pseudodiagnosis date by
using the chronology of the notes. Other
sources of information kept by the GP (for
example the notes of a woman’s husband or
children) were not abstracted.

We wanted complete GP data on all eligible
cases and their first selected controls, regard-
less of whether or not the interviews had been
done. Where we were unable to interview a
case, for whatever reason, permission was
sought from the consultant and GP to abstract
the information from her GP notes. Per-
mission was sought from the GP, and if
required by the GP, from the control also, to
abstract the same information from her GP
notes. If a first-selected control could not be
interviewed for any reason, permission was
also sought to abstract information from her
GP notes even though a second (or subse-
quent) control had been successfully inter-
viewed. Information was also sought from any
family planning clinics that a woman recalled
attending. The data from all sources were used
to contruct the lifelong contraceptive calendar
used for the main analyses of the study data.12

STATISTICAL METHODS
Agreement between the interview and the GP
data was measured by means of the kappa
statistic.3–5 The analyses of breast cancer risk
using interview data alone and GP data alone
were carried out using multivariate logistic
regression methods for individually matched
case-control studies.6 Relative risks were esti-
ated by odds ratios. Significance levels quoted
are two-sided.

| Table 1 Comparison of medical and contraceptive history (excluding oral
contraceptives) at interview and in general practitioner notes: 754 case-control pairs |
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Hysterectomy:</td>
</tr>
<tr>
<td>Cases</td>
</tr>
<tr>
<td>Controls</td>
</tr>
<tr>
<td>All</td>
</tr>
<tr>
<td>Oophorectomy:</td>
</tr>
<tr>
<td>Cases</td>
</tr>
<tr>
<td>Controls</td>
</tr>
<tr>
<td>All</td>
</tr>
<tr>
<td>Tubal ligation:</td>
</tr>
<tr>
<td>Cases</td>
</tr>
<tr>
<td>Controls</td>
</tr>
<tr>
<td>All</td>
</tr>
<tr>
<td>RU4 use, ever:</td>
</tr>
<tr>
<td>Cases</td>
</tr>
<tr>
<td>Controls</td>
</tr>
<tr>
<td>All</td>
</tr>
<tr>
<td>Diaphragm use, ever:</td>
</tr>
<tr>
<td>Cases</td>
</tr>
<tr>
<td>Controls</td>
</tr>
<tr>
<td>All</td>
</tr>
<tr>
<td>Partner’s vasectomy:</td>
</tr>
<tr>
<td>Cases</td>
</tr>
<tr>
<td>Controls</td>
</tr>
<tr>
<td>All</td>
</tr>
</tbody>
</table>

* I+ = recalled at interview; I- = not recalled at interview; N+ = recorded in general prac-
titioner notes; N- = not recorded in general practitioner notes.
† All p < 0.001.
‡ Test for difference in agreement between cases and controls: p < 0.01.

Results
RESPONSE RATES
Response rates and reasons for non-response
have already been reported.1 Of the 1049 el-
igible cases, 755 (72%) were interviewed and
GP notes were abstracted for 754 of these 755.
GP notes were also abstracted for 222 (76%) of
the 294 cases not interviewed. Reasons for not
abstracting the remaining 72 sets of notes were
consultant or GP refusal (35), case refusal (3),
and failure to trace the notes (32).

Of the 755 first-selected controls, 675 (89%) were
interviewed and the GP notes of 674 of these were
abstracted. The 80 first controls not
interviewed were replaced by second or sub-
sequent choices and GP notes were abstracted
for these also. Fifty three (66%) of the 80 first-
selected controls who were not interviewed also
had their GP notes abstracted. Reasons for
not abstracting the remaining 27 sets of notes
were GP refusal (15), control refusal (11), and
impossible to trace (1). We also abstracted 219 sets
(99%) of notes for first-selected controls matched to the 222 non-
interviewed cases whose notes had been suc-
cessfully abstracted. (Subsequent controls
were substituted for the three remaining con-
tral women who refused permission for note
abstraction.)

We compared information obtained from
interview with that obtained from GP notes for
the 754 pairs where both sets were available;
one GP refused permission for abstraction. GP
notes were abstracted for a total of 946 (90%)
of the 1049 case and first-selected control
pairs.

COMPARISON OF INFORMATION OBTAINED FROM
GENERAL PRACTITIONER NOTES AND AT
INTERVIEW
Agreement on whether a woman had ever been
married was 86-0% (1297 of 1508) overall, and
similar for cases and controls. In 204 instances
the woman reported at interview that she had
been married at that time but there was
nothing in the GP notes to indicate that she
had ever been married. Seven women report-
ing themselves never married were recorded as
ever-married in the GP notes. The notes of
these seven women were rechecked; two of the
women reported themselves as cohabiting at
interview and one of these had had two pre-
vious names in the GP notes. For four women
the GP notes implied that they were married,
and one was an error.

Agreement between interview and GP notes
for medical and contraceptive history (exclud-
ing oral contraceptives) was at least 85% for all
items (table 1). Agreement, as measured by the
kappa statistic, varied between 0.24 and 1.00
with a good agreement on surgical procedures,
and much poorer recording of contraceptive
use in the GP notes. There was little difference
between the rates of agreement for cases and
controls, but certain aspects of the results
were clearly warranted further investigation. Some
87-0% (20 of 23) of hysterectomies and 93-8% (15 of 16) oophorectomies reported at inter-
view were recorded in the GP notes. Two
Our assumption that a pregnancy with an unknown outcome in the notes was most likely to be a livebirth was wrong for 15 pregnancies (13 miscarriages and two terminations). One baby not reported at interview had been adopted and one reported as a neonatal death was a miscarriage. For one case a letter reporting a pregnancy filed in the notes referred to another woman. We identified six errors but could not explain 11 discrepancies.

Approximately half the total duration of oral contraceptive use recalled at interview was recorded in GP notes. Where a mention of oral contraceptive use was recorded without a duration or prescription, two alternative assumptions were made: that each mention referred to six or three cycles of use (these being the two most common prescriptions). Making the assumption that a mention was equivalent to a six cycle prescription, the mean durations recalled at interview and reported in the notes were 64.2 months and 35.1 months, respectively, for cases and 53.9 and 29.9 months, respectively, for controls. Making the assumption of three cycle prescriptions, the mean duration reported in the notes were 31.2 months and 26.2 months for cases and controls respectively.

Table 3 cross-tabulates the number of additional brands of oral contraceptive recorded in the GP notes but not recalled at interview, and the number of additional brands recalled at interview but not recorded in the notes. There was agreement between interview and GP notes with respect to known brands for 34.5% (521 of 1508) of women. A total of 138 women had no use of oral contraceptives either reported at interview or recorded in their GP notes. Nine women had a prescription in their GP notes but had never actually taken the pills prescribed. (This was established by re-contacting the women concerned and the total of 147 never-users corresponds with our previous report.12) The data from cases and controls were comparable, the percentages with agreement on brands being 34.1% and 35.0% respectively. Some women reported use of oral contraceptives but could not recall the brand, and use of oral contraceptives of unknown brand was also recorded in the GP notes. These episodes of use have not been included in the above analysis.

Table 3: Number of additional known brands of oral contraceptive recalled at interview by number of additional known brands recorded in general practitioner notes

<table>
<thead>
<tr>
<th>No of additional known brands recorded in notes</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>≥3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>257</td>
<td>107</td>
<td>47</td>
<td>19</td>
<td>403</td>
</tr>
<tr>
<td>1</td>
<td>75</td>
<td>71</td>
<td>25</td>
<td>182</td>
<td>260</td>
</tr>
<tr>
<td>2</td>
<td>34</td>
<td>37</td>
<td>11</td>
<td>2</td>
<td>62</td>
</tr>
<tr>
<td>≥3</td>
<td>32</td>
<td>18</td>
<td>9</td>
<td>1</td>
<td>51</td>
</tr>
<tr>
<td>Total</td>
<td>398</td>
<td>233</td>
<td>92</td>
<td>31</td>
<td>754</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Controls</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>124</td>
<td>43</td>
<td>26</td>
<td>19</td>
<td>208</td>
</tr>
<tr>
<td>1</td>
<td>28</td>
<td>35</td>
<td>15</td>
<td>3</td>
<td>83</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>10</td>
<td>7</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>≥3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>268</td>
<td>113</td>
<td>53</td>
<td>7</td>
<td>433</td>
</tr>
</tbody>
</table>

ComparISON OF RELATIVE RISKS OF BREAST CANCER IN RELATION TO ORAL CONTRACEPTIVE USE USING INTERVIEW AND GP DATA

In the main analyses of these data, combined data from the questionnaire and GP notes were used (see the first and second columns of table 4).12 The combination of data from the two sources (supplemented by use of family planning clinic records) effectively allowed us to identify most of the unknown brands used. We have now also calculated odds ratios using the interview data alone and the GP data alone (see third and fourth columns of table 4); similar results were obtained whichever dataset was used. This result is at first consideration surprising since only about half the information
Table 4 Comparison of relative risks of breast cancer by duration of oral contraceptive (OC) use using data from general practitioner notes alone, interview alone, and combined data

<table>
<thead>
<tr>
<th>OC use (mth)</th>
<th>Interview and GP notes Unadjusted</th>
<th>Adjusted*</th>
<th>Interview Unadjusted</th>
<th>GP notes Unadjusted</th>
<th>754 pairs unadjusted†</th>
<th>946 pairs unadjusted†</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1·00</td>
<td>1·00</td>
<td>1·00</td>
<td>1·00</td>
<td>1·00</td>
<td>1·00</td>
</tr>
<tr>
<td>1·00-0·90</td>
<td>0·89</td>
<td>0·89</td>
<td>1·10</td>
<td>1·10</td>
<td>1·10</td>
<td>1·10</td>
</tr>
<tr>
<td>1·00-0·90</td>
<td>1·18</td>
<td>1·18</td>
<td>1·50</td>
<td>1·50</td>
<td>1·50</td>
<td>1·50</td>
</tr>
<tr>
<td>1·00-0·90</td>
<td>1·69</td>
<td>1·69</td>
<td>1·74</td>
<td>1·74</td>
<td>1·74</td>
<td>1·74</td>
</tr>
</tbody>
</table>

* Adjusted for age at menarche, nulliparity, age at first full-term pregnancy, breast feeding (ever, never), and family history of breast cancer.
† Assuming menarche = 6 months.

recalled at interview is in the GP notes. This would bias the GP results upwards but is effectively compensated for by the random error in using the GP notes. When the GP data are analysed for the 946 cases each matched to their first selected control the odds ratios are slightly higher than for the 754 interviewed pairs.

Discussion
Agreement between information obtained from GP notes and at interview was good for surgical procedures (hysterectomy, oophorectomy, and tubal ligation) but less good for contraceptive use: partner’s vasectomy was relatively rarely reported in the woman’s notes and there was under-recording of IUD and of diaphragm use. The under-recording of contraceptive use is partly due to the use of non-GP family planning clinics. Differences in agreement between cases and controls were generally small, although for oophorectomy and hysterectomy the agreement was better for cases than controls.

There was “almost perfect” (see Landis and Koch) agreement between the obstetric information obtained at interview and recorded in the notes for six of 12 items (that is, kappa statistics greater than 0·80). Termination of pregnancy was as likely to be under-reported as it was to be under-recorded. This was in spite of very careful questioning on pregnancy and its outcome, and our awareness that under-reporting of terminations of pregnancy would be likely to occur. In practice, as the GP notes are unlikely to record all terminations, the rate of under-reporting at interview will be higher than shown in table 2.

The GP notes varied enormously in quality with, at one extreme, impeccable recording of all items including contraception, and at the other, notes that must be regarded as inadequate. It may well be that in some instances notes had been “lost” on registration with a new GP, but it would still be reasonable to expect that basic information on obstetric and medical history might be recorded at a first visit to a new doctor. Changes in the contracts for general practitioners are likely to lead to more consistency and completeness in GP notes.

Studies comparing information recorded in medical records and recalled at interview have generally had the objective of validating the information recalled,8 11 but the premise that medical records are the gold standard for comparison has also been questioned.15 16 The results of studies comparing retrospective recall with information collected prospectively by participants,17 18 with family planning clinic records9 or hospital records19 have been reviewed by Harlow and Linet. More recently comparisons between hospital records, postal questionnaires, and interviews with respect to information on pregnancy and childbirth,20 and of records with interview reports of ovarian surgery21 have been made.

Rates of agreement in our study for oophorectomy and hysterectomy were similar to those reported previously.20 21 A comparison of recall of surgical history using a postal survey with GP records showed that respondents provided accounts which were more reliable than GP records.21

Reproductive history has been compared with antenatal records12 18 and with prospectively collected personal records,18 neither of which are directly comparable to our data. Similar results (93% agreement on number of children compared to our 91·0%) were found in the retirement community study where history was taken at entry rather than being recorded contemporaneously in the GP notes.10

Studies which have compared oral contraceptive use in medical records with recalled information have generally included only self-reported users of oral contraceptives,8 10 11 although one study did include reported non-users.22 Agreement on various parameters of oral contraceptive use was generally good. The questionnaire used in our study (which included a contraceptive calendar and album of photographs of oral contraceptive packaging) has been compared with an earlier questionnaire which included only a list of brand names.11 The former was demonstrably better for total duration of use but results were similar for other parameters. The women chosen had all used oral contraceptives yet when a comparison with GP notes was made, 32% of notes had no mention of oral contraceptive use and only 32% of changes of brand were recorded. It would be expected that concordance in the study reported here would be better because oral contraceptives were prescribed by both GPs and family planning clinics, whereas in the earlier comparison11 contraceptives were obtained only from family planning clinics. Although GPs are notified of clinic prescriptions it is clear that these are not always filed in the notes. The similarity of cases and controls with respect to agreement between interview and GP notes, and numbers of additional brands reported suggests that recall by cases is no better than controls (see also discussion in UK national study).7

The main results from the UK national study were based on a combination of data from the two sources and from family planning clinic records with re-interviews where there were major differences or inconsistencies between the data sources.1 Our finding that only 50% of recalled use of oral contraceptives is recorded in the notes, is in agreement with the conclusion of Coulter et al13 that GP records are insufficient for obtaining a complete
oral contraceptive history. Yet, essentially we obtained the same relative risk result using the GP notes alone (table 4). A previous case-control study of malignant melanoma also reported consistent relative risks for ever-use of oral contraceptive agents from GP records and postal questionnaire, and a case-control study of risk of liver cancer in relation to oral contraceptive use using GP notes alone had results compatible with those obtained in a questionnaire based case-control study. There were, however, certain methodological difficulties in the latter study (different questionnaires administered and different distributions of usual place of residence for cases and controls).

It may be argued that in the same way that the interviewers could not be blind to the case-control status of interviewees, it was not possible for them to be blind to case-control status when carrying out note abstractions. We felt that the only way to deal with any possible bias was firstly by having highly structured questionnaires and secondly by carefully training the interviewers. GP notes are laborious to extract and our interviewers reported that the concentration needed to abstract notes that may be disorganised and voluminous is such that case-control status becomes an irrelevance.

Using notes alone would have had a major impact on response rates. The response rates for cases would have been much greater: 93% (976 of 1049) versus 72% (755 of 1049) and as the major cause of inability to abstract notes was failure to trace them, the rate would be even higher had there been fewer delays in identifying cases. (Untraced notes usually belonged to deceased cases and 16-6% of cases had died before they could be interviewed.) The abstraction rate for first-selected controls was 97% (946 of 977). The main reason for failure to abstract control notes was that the control women who refused interview were less likely to give consent for note abstraction. If an interview had not been requested the abstraction rate would have been higher: the notes of 219 (99%) of the 222 first-selected controls for non-interviewed cases where an interview was not requested were successfully abstracted. Permission to abstract GP notes was always sought at interview from cases and controls. Where cases and controls were not interviewed, permission was only obtained where individual GP notes required it. Disclosure of information from notes for medical research does not require the explicit consent of the patient provided that confidentiality is scrupulously safeguarded. It is clear that at the time that this study was carried out response rates would have been higher using GP notes than for the interview study and potential non-response bias would be reduced and recall bias eliminated. Written permission for note abstraction is, however, now required much more often than during the mid-1980s, and the response rates for controls might now be similar to the 89% for the interview study. Whether the same proportion of women who were successfully interviewed would be willing to give permission to have their notes abstracted only is not known. If the approach to them was made through their GP it is likely, in our opinion, to be the same or even higher than the 89% that we obtained for the interview study. Using GP notes only, the cost of the study would be reduced by at least 50%. The major disadvantage of using notes alone would be the lack of information on potential confounding factors. In our main analyses of risks associated with oral contraceptive use we adjusted the odds ratios for age at menarche, nulliparity, age at first full-term pregnancy, breastfeeding (ever, never), and family history of breast cancer. Of these factors, information on parity and age at first full-term pregnancy would be available from the notes. We did not investigate the frequency with which family history of breast cancer in a first degree relative was recorded in the GP notes. In our study the adjusted odds ratios were only slightly higher than the unadjusted odds ratios (table 4) but it would be unwise to assume that this would always be the case.

In conclusion, we could have carried out a much less expensive study of the association between oral contraceptive use and breast cancer risk and would still have found a strong positive trend in risk with increasing duration of use. The magnitude of the risk estimates would, however, be imprecise because, on average, only about half of the oral contraceptive use recalled was recorded in the notes. A “validation” sub-study could, however, be used to adjust the results obtained from GP notes, and this possibility should be further explored. Information on potential confounding factors would have been insufficient for an adjusted analysis to have been carried out and this must be regarded as an important limitation on studies using GP notes. The value of GP notes for epidemiological research. In Britain the system whereby primary care notes follow an individual, in theory at least, from the cradle to the grave, is unique. At present, the Department of Health recommends that Family Health Services Authorities retain records for seven years after death and then destroy them. In practice, this interval may be reduced because of lack of storage space. A valuable resource is then lost. With the computerisation of general practice notes it is unclear to what extent a paper copy of computerised information will be stored in the notes passed on should a patient re-register elsewhere. If paper copies (of prescriptions, for example) were to be routinely stored in the notes, then it is likely that information would be more reliably recorded than at present. If not, there will be a deterioration in the completeness of data relevant to future research.

Other principal investigators in this study were: Klir McPherson (Department of Public Health and Policy, London School of Hygiene and Tropical Medicine), J Peto (Section of Epidemiology, Institute of Cancer Research, Sutton, Surrey), M P Vessey (Department of Public Health and Primary Care, University of Oxford).

The UK National Case-Control Study was funded by the Imperial Cancer Research Fund, the Cancer Research Campaign and the Medical Research Council. SJS is supported by the Trent Regional Health Authority. This paper was drafted while CEDC was Visiting Professor at the Department of Social and
Preventive Medicine, University of Queensland. We thank Dr Christopher Baird and Dr Christopher Del Mar for helpful suggestions and Vineta O’Malley and Melanie Compston for manuscript preparation.

7 Landis JR, Koch GG. The measurement of observed agreement for categorical data. Biometrics 1977;33:159-74.