Investigation of under-ascertainment in epidemiological studies based in general practice

Dinesh Sethi, Jerry Wheeler, Laura C Rodrigues, Susan Fox and Paul Roderick

Background
One of the aims of the Study of Infectious Intestinal Disease (IID) in England is to estimate the incidence of IID presenting to general practice. This sub-study aims to estimate and correct the degree of under-ascertainment in the national study.

Methods
Cases of presumed IID which presented to general practice in the national study had been ascertained by their GP. In 26 general practices, cases with computerized diagnoses suggestive of IID were identified retrospectively. Cases which fulfilled the case definition of IID and should have been ascertained to the coordinating centre but were not, represented the under-ascertainment. Logistic regression modelling was used to identify independent factors which influenced under-ascertainment.

Results
The records of 2021 patients were examined, 1514 were eligible and should have been ascertained but only 974 (64%) were. There was variation in ascertainment between the practices (30% to 93%). Patient-related factors independently associated with ascertainment were: i) vomiting only as opposed to diarrhoea with and without vomiting (OR 0.37) and ii) consultation in the surgery as opposed to at home (OR 2.18). Practice-related factors independently associated with ascertainment were: i) participation in the enumeration study component (OR 1.78), ii) a larger number of partners (OR 0.3 for 7–8 partners); iii) rural location (OR 2.27) and iv) previous research experience (OR 1.92). Predicted ascertainment percentages were calculated according to practice characteristics.

Conclusion
Under-ascertainment of IID was substantial (36%) and non-random and had to be corrected. Practice characteristics influencing variation in ascertainment were identified and a multivariate model developed to identify adjustment factors which could be applied to individual practices. Researchers need to be aware of factors which influence ascertainment in acute epidemiological studies based in general practice.

Keywords
Under-ascertainment, epidemiological, general practice, incidence, infectious intestinal disease

Accepted 17 March 1998

The incidence of disease is a central measure of descriptive epidemiology. It forms the basis for comparisons between populations and time periods and hence for defining trends and generating hypotheses about causation. It is also a key measure in disease surveillance and service planning. Under-ascertainment of such studies whether disease registries, disease surveillance systems or ad hoc studies is a key problem but one that has received little attention.1,2 It may lead to lower absolute rates and may bias time, sub-group and geographical comparisons if it varies differentially by these factors.

There are two main strategies for tackling under-ascertainment: prevention, and estimation and correction. The former includes completeness of data collection for example, by using multiple sources or by routine data systems such as general practitioner (GP) records with electronic transmission and random checks.3–5 Another approach is to estimate the level of undercounting using capture-recapture methods.6–8 Levels of ascertainment have rarely been estimated and when they have substantial undercounting has been found.9–11 In
spite of the importance of complete enumeration, only 65% of
cancer registries had some form of evaluation of undercounting;
similar problems are found in myocardial infarction and
connective tissue registries. Under-ascertainment has also
been found in infectious disease surveillance and statutory
notification. 

This paper describes how under-ascertainment was estimated
and then used to correct disease incidence rates in the primary
care-based study of infectious intestinal disease (IID) in England.

Objectives
The objectives of the study were to: estimate the level of under-
ascertainment in the Study of IID in England; examine the
practice and patient characteristics associated with ascertainment;
provide adjustment factors for analysis of presentation
rates; consider implications for epidemiological studies in primary
care.

Methods
Setting
This sub-study is one of nine components of the Study of IID
in England. The general practices participating were selected
from the Medical Research Council’s (MRC) General Practice
Research Framework which are taking part in either the
case-control (n = 34) or enumeration (n = 36) components.
Practices were chosen to be representative of the population of
England by geographical area, Jarman deprivation score, and
urban or rural location. The Jarman score is a composite score
for underprivileged areas which is derived from weighting
several census variables relating to social conditions and is
available for all electoral wards in England.

The case-control component aims to estimate the incidence,
risk factors and aetiological agents of IID presenting to GPs and
the enumeration component aims to estimate the incidence of
cases of IID presenting to general practice and the proportion of
loose stools sent routinely for microbiological testing. In both
components, cases of IID who consulted their GP, at the surgery
or at home, were ascertained over a one-year period. GPs
notified all patients who fulfilled the case definition of IID (see
below) to the research nurse (RN) by using an ascertainment
form. The RN, who was often also the practice nurse, had been
trained in study procedures by the MRC General Practice
Research Framework regional nurse trainer. She requested
further information and specimens from the case and sent data
to the MRC Epidemiology and Medical Care Unit. Cases that
were seen during a domiciliary visit were ascertained by the
GP to the practice RN. The RNs telephoned deputizing agencies
which provided out of hours medical cover to obtain details
of any cases seen. RNs were asked to regularly remind GPs to
ascertain cases. Each practice received regular feedback reports
of all cases ascertained from the MRC coordinating centre in-
cluding comment if this was lower than expected, based on the
pilot study results.

Twenty-six practices with computer diagnosis (CD) were
selected from the 70 participating in the overall study solely
because they routinely entered diagnoses on the practice computer
for all consultations using the Read Code system of
classification. The Read Code is a computerized disease diagnostic
classification system devised for primary care and is widely used
in general practice in England. Almost all general practices in
England have a practice computer which is used for admin-
istrative purposes and increasingly to enter diagnoses using the
Read Code.

Characteristics of practices with and without CD were com-
pared to determine whether the 26 were representative of the
whole study. Five of the 26 practices were spotter practices
belonging to the Royal College of General Practitioners Weekly
Returns Service which provides comprehensive weekly morbidity
surveillance data. Twelve of the practices had underta-
taken previous research with the MRC.

Design
Patients that should have been ascertained to the national study
were those with the following case definition of IID: loose stools
or significant vomiting lasting less than 2 weeks, in the absence of a
known non-infectious cause and preceded by a symptom-free period
of 3 weeks. Vomiting was considered significant if it occurred
more than once in a 24-hour period and if it incapacitated the case or was
accompanied by other symptoms such as cramps or fever. Cases that
fulfilled this definition should have been ascertained; those that
were not represented the undercount.

For the national study, the first practice commenced its one-
year prospective ascertainment of cases in August 1993 and
the last one completed it in January 1996. Practice visits for
the ascertainment study were carried out in the latter half of
1995 so that cases with IID could be identified retrospectively
from each practice. Wherever possible, two time periods were
selected from each one-year period of practice participation, in
order to get a representative sample and assess practice per-
formance over time. At each practice data were gathered on
about 70 cases; this number was feasible for a one-day visit and
represented a reasonable fraction (20–100%) of the total cases
in each practice. Information was recorded on personal and
clinical details, presence of a non-infectious cause, treatment,
place of consultation (home or GP surgery), being seen by a
deputizing service (used by six practices) and hospital admis-
sion. This was done using a proforma, which had been piloted
in live practices. Quality checks on completeness and accuracy
of data were carried out during this pilot phase and were of a
high standard.

At each practice patients with ‘presumed IID’ were identified
from the practice computer using Read Codes compatible with
IID, diarrhoea, vomiting, food poisoning and gastroenteritis.
These codes were chosen because they were over-inclusive. In
order to determine whether patients were eligible for inclusion,
clinical details on practice computer records and case notes
were studied. Personal and clinical details of cases were then
compared with the practice RN’s records to determine if they
had been ascertained to the MRC. Cases that should have been
ascertained but were not represented the undercount.

Data entry and statistical analysis
Information was collected at the GP surgery using a scannable
proforma completed by the RN. Data were entered using a scan-
cer and quality assurance checks of completeness and internal
validity were made. The representativeness of practices with CD
was tested using the $\chi^2$ test. Univariate analysis of potential
factors influencing under-ascertainment was performed. Nine
age groups were used to investigate the effect of age (0–1, 2–4, 5–9, 10–19, 20–39, 40–59, 60–79, 80–99 years). Logistic regression analysis was then used to identify independent factors influencing ascertainment. All factors were reassessed in the logistic regression regardless of significance in the univariate analysis. Factors with a significance of <0.05 by the likelihood ratio test were dropped from the logistic model. Selected interactions between the variables were investigated.

Results

Representativeness of GP practices

The characteristics of the 26 GP practices with CD which took part in the under-ascertainment study were compared with those of all practices in the IID study. There was no significant difference in terms of study component, urban or rural setting, geographical location, Jarman score and number of partners.

Case ascertainment

The total number of records examined was 2021 at 26 practices (median 78, range 37–83). Of these, 1516 (75%) were eligible, based on the case definition of IID. Further analysis was restricted to these eligible cases, all of which should have been ascertained.

Of the 1516 eligible cases only 974 (64.3%) were ascertained to the study. This could be used simplistically, to derive an overall adjustment factor (100/64.3 = 1.56) for the true presentation rate of IID to general practice. Variation in ascertainment between practices ranged from 30% to 93%.

Univariate analysis

Univariate analysis showed that of patient-related factors there was no strong evidence that age and sex were associated with ascertainment (Table 1). Males had a slightly higher ascertainment than females ($P = 0.06$). Cases who complained of loose stools, with or without vomiting, were more likely to be ascertained rather than vomiting alone (70% versus 43%, $P < 0.001$). Consultation in the surgery as opposed to home (67% versus 49%, $P < 0.001$), and by the GP as opposed to deputizing service (65% versus 36%, $P < 0.001$) were also associated with higher ascertainment.

Practice characteristics were examined to see if they were associated with ascertainment. This was higher in enumeration practices versus case-control (70% versus 58%, $P < 0.001$), in rural practices versus urban (75% versus 53%, $P < 0.001$). It decreased as the number of partners per practice increased, from 70% with 1–2 partners to 44% in practices with 7–8 partners ($P < 0.001$ test for trend). Ascertainment varied with practice Jarman score but there was no trend. It was lowest in practices located in the North.

Ascertainment did not vary according to whether the period examined was early or late in the study period. There was lower ascertainment in spotter practices belonging to the Royal College of General Practitioners Weekly Returns Service when compared to those that did not (50% versus 68%, $P < 0.001$).

Logistic regression model

Logistic regression modelling was used to identify variables which were independently associated with ascertainment after taking other variables into account. Of the factors related to patients (age, sex, symptoms, home visit and seen by a deputizing service) those that were independent were: i) vomiting only as opposed to diarrhoea with and without vomiting (odds ratio [OR] 0.37, confidence interval [CI] : 0.28–0.49) and ii) consultation in the surgery as opposed to at home (OR 2.18, CI : 1.63–2.90).

Patient level characteristics were not used to predict ascertainment because: i) the severity of the symptoms may have influenced the GP’s interpretation of whether the case definition of IID was met and ii) it would have been difficult to obtain denominator data for patients visited at home for adjustment of ascertainment rates.

Practice-related factors independently associated with ascertainment were: i) study component (enumeration versus case-control OR 1.78), ii) the number of partners (OR 0.3 for 7–8 partners, test for trend $P < 0.001$), iii) urban or rural location (urban versus rural OR 2.27) and iv) previous research experience (OR 1.92) (Table 2). Factors that were no longer significant included sex ($P = 0.14$), Jarman score ($P = 0.46$) and geographical region ($P = 0.21$). There were no significant interaction terms.

Predicted ascertainment according to practice characteristics

The predicted percentage ascertainment according to practice characteristics were calculated from the final logistic regression model and are shown separately for the case-control and enumeration practices in Table 3. Predicted under-ascertainment was higher in urban practices, with a large number of partners and without previous research experience. These characteristics were then used to correct for the undercount at practice level.

Discussion

Main findings

Our findings show that under-ascertainment was significant and non-random in a primary care-based study of incidence of IID. This was despite working within a research framework with motivated GPs and RNs committed to primary research. Only 64% of potential cases of IID were ascertained; variation between practices ranged from 30% to 93%. Patient and practice characteristics that might explain this variation in ascertainment were identified.

There was no ascertainment bias on the basis of age and sex. Fewer patients with vomiting alone were ascertained when compared to when diarrhoea was also present. This may be explained by the fact that vomiting is a less specific symptom for IID than diarrhoea, and GPs may have considered these patients ineligible as opposed to patients with more severe symptoms.

Our suspicion that a smaller proportion of patients would be ascertained at a home visit was confirmed. This may be because GPs did not always carry ascertainment forms on home visits, and failed to ascertain cases to the RN. However, there did not appear to be any bias due to age, sex or clinical presentation. Similarly not all cases seen by a deputizing service may have been identified to the RN.

The practice characteristics identified in our study which were independently associated with under-ascertainment were: urban location, a larger number of partners, no previous research experience and belonging to the case-control component. The
practice Jarman deprivation score, and its geographical location were not independently associated with ascertainment. Practices with many GPs had a smaller proportion of cases ascertained. In these practices it is possible that not all GPs were equally committed to the study, leading to less comprehensive ascertainment. Moreover, communication may have been better in smaller practices. These potential sources of bias have been identified and adjustments can be made to the estimates of incidence.

Possible explanations for lower ascertainment in urban practices may be greater demands on time, and exclusion of patients thought by practice staff to be unsuitable for inclusion.

Table 1  Univariate analysis of patient and practice related factors influencing ascertainment of cases of infectious intestinal disease in general practice

<table>
<thead>
<tr>
<th>Factor</th>
<th>Categories</th>
<th>n</th>
<th>Percentage ascertained</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td></td>
<td>1516</td>
<td>64.3</td>
<td></td>
</tr>
<tr>
<td><strong>Patient related factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>male</td>
<td>728</td>
<td>66.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>female</td>
<td>781</td>
<td>62.1</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>7</td>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td>Age</td>
<td>age groups males</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>age groups females</td>
<td></td>
<td></td>
<td>0.23</td>
</tr>
<tr>
<td>Symptom</td>
<td>loose stool/stool+vomiting</td>
<td>1185</td>
<td>70.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>vomiting only</td>
<td>162</td>
<td>42.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>169</td>
<td>45.5</td>
<td></td>
</tr>
<tr>
<td>Home visit</td>
<td>yes</td>
<td>269</td>
<td>48.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>no</td>
<td>1231</td>
<td>67.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deputizing service</td>
<td>yes</td>
<td>30</td>
<td>36.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>no</td>
<td>1457</td>
<td>65.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Practice related factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study component</td>
<td>case-control</td>
<td>753</td>
<td>58.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>enumeration</td>
<td>763</td>
<td>70.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Number of partners</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1–2</td>
<td>451</td>
<td>70.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3–4</td>
<td>718</td>
<td>65.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5–6</td>
<td>226</td>
<td>60.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>7–9</td>
<td>121</td>
<td>44.2 (trend)</td>
<td></td>
</tr>
<tr>
<td>Jarman score</td>
<td>&lt;-5</td>
<td>328</td>
<td>56.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-5–10</td>
<td>784</td>
<td>68.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;10</td>
<td>404</td>
<td>62.8</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Geographical region</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>north</td>
<td>344</td>
<td>50.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>midlands/south west</td>
<td>777</td>
<td>69.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>south east</td>
<td>395</td>
<td>66.5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Location</td>
<td>urban</td>
<td>717</td>
<td>82.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rural</td>
<td>799</td>
<td>74.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GP spotter scheme</td>
<td>yes</td>
<td>275</td>
<td>49.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>no</td>
<td>1241</td>
<td>67.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>scanned</td>
<td>1242</td>
<td>63.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>coded</td>
<td>274</td>
<td>67.7</td>
<td>0.21</td>
</tr>
<tr>
<td>Time period</td>
<td>early in practice's year</td>
<td>760</td>
<td>63.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>late in practice's year</td>
<td>756</td>
<td>61.2</td>
<td>0.50</td>
</tr>
<tr>
<td>Previous research experience</td>
<td>yes</td>
<td>744</td>
<td>65.4</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>no</td>
<td>772</td>
<td>63.3</td>
<td></td>
</tr>
</tbody>
</table>

a The age groups were 0–1, 2–4, 5–9, 10–19, 20–39, 40–59, 60–79, 80–99.
RNs in the case-control practices had a larger workload than in the enumeration component—this included requests for questionnaires and stool specimens. GPs would have had to spend longer with patients in explaining the need for questionnaires and stool samples. In practices with previous research experience staff ascertained a higher proportion of cases, suggesting that previously acquired epidemiological skills were transferable.

Possible methodological weaknesses in this study

The 26 practices participating in this study were those which routinely used CD in contrast to the 44 which did not. Although they were representative in terms of characteristics such as study component, urban or rural location, geographical region, Jarman score and the number of partners, it is conceivable that they may have differed in other ways. However, this was the only method available to identify effectively potential cases of IID.

There may have been inconsistent data entry by GPs leading to incomplete computer datasets. However, practices were only selected if they claimed to enter data on diagnoses routinely. Of interest, spotter practices belonging to the Royal College of General Practitioners Weekly Returns Service ascertained a lower proportion of cases compared to those who did not. This failed to reach significance by logistic regression analysis. These practices conform to quality control checks and have previous experience of CD. This makes major inconsistencies in data entry less likely.

Another approach would have been to take a stratified random sample of all consultations from practice registers held in reception. However, the sample size required would have been cumbersome because of the low rate of IID. Our approach was pragmatic and reasonably accurate.

### Table 2

<table>
<thead>
<tr>
<th>Variable</th>
<th>Category</th>
<th>Odds ratio</th>
<th>P-value</th>
<th>Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study component</td>
<td>case-control</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>enumeration</td>
<td>1.78</td>
<td>&lt;0.001</td>
<td>1.39–2.29</td>
</tr>
<tr>
<td>Number of partners</td>
<td>1–2</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3–4</td>
<td>0.63</td>
<td>0.55–0.73</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5–6</td>
<td>0.41</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7–9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>urban</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rural</td>
<td>2.27</td>
<td>&lt;0.001</td>
<td>1.81–2.86</td>
</tr>
<tr>
<td>Previous research</td>
<td>no</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>yes</td>
<td>1.29</td>
<td>&lt;0.001</td>
<td>1.49–2.48</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>No. of partners</th>
<th>Rural location</th>
<th>Urban location</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Previous research experience</td>
<td>No previous research experience</td>
</tr>
<tr>
<td><strong>Case-control component</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–2</td>
<td>88.0</td>
<td>79.2</td>
</tr>
<tr>
<td>3–4</td>
<td>82.3</td>
<td>70.8</td>
</tr>
<tr>
<td>5–6</td>
<td>74.6</td>
<td>60.6</td>
</tr>
<tr>
<td>7–8</td>
<td>65.2</td>
<td>49.2</td>
</tr>
<tr>
<td><strong>Enumeration component</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1–2</td>
<td>92.9</td>
<td>87.1</td>
</tr>
<tr>
<td>3–4</td>
<td>89.2</td>
<td>81.2</td>
</tr>
<tr>
<td>5–6</td>
<td>83.9</td>
<td>73.3</td>
</tr>
<tr>
<td>7–8</td>
<td>76.9</td>
<td>63.4</td>
</tr>
</tbody>
</table>

* Practice level characteristics initially entered in the logistic regression model were study component, number of partners, Jarman score, geographical region, urban or rural location, GP spotter scheme, whether the questionnaire was scanned or coded, time period and previous research experience. Only those attaining 5% significance are presented in the Table.
Implications for the design of our study

The level of ascertainment to the Study of IID in England was 64%. An adjustment factor of 1.56 derived from this could then be applied to the presentation rate of IID. Logistic regression identified practice characteristics independently associated with ascertainment. Predicted percentage ascertainment were calculated according to these variables and adjustment factors could then be applied according to individual practice characteristics. This approach would provide a more accurate adjustment than a global adjustment factor. There are few published studies where multivariate modelling has been used in this manner.

General implications for epidemiological studies based in primary care

Under-counting is a problem in epidemiological studies. Various methods have been employed such as using multiple data sources in order to make data collection as complete as possible. This option was not possible; unlike for chronic disease there are no disease registers for IID, and few are seen by hospital specialists. Capture-recapture methods have been used successfully for a variety of conditions. However, a fundamental assumption of the methodology is that the datasets used are independent. This was not the case in our study. Others have used random checks of completeness as a means of improving ascertainment. Future studies may be able to use computerized general practice systems as a prospective double check. We were not able to because only 26 of the 70 national study practices had CD; there would have been problems with getting timely diagnostic data, particularly with the need to send stool samples in the acute phase. Our study therefore has implications for similar studies based in primary care (e.g. to investigate respiratory tract infection) where routine data sources can be used to estimate the degree of ascertainment but not as a primary data source.

Conclusion

Our study has identified that only 64% of potential cases of IID were ascertainment. Under-ascertainment was common and non-random and had to be corrected for. Practice characteristics that might explain variation in ascertainment were identified. A multivariate model was developed to identify adjustment factors which could be applied to individual practices. Future researchers need to be aware of the factors which influence ascertainment. The methodology used here may be applicable to other epidemiological research carried out in general practice.

Acknowledgments

We wish to thank the MRC EMCU staff, Mr N Graves, Ms M Goldsborough, Ms A Williams, Ms L Hands, Ms E Marshall, Ms P Allen, Ms F Symes, and Ms J Elwood for their invaluable contribution to this study. This work was funded by a Department of Health grant.

We are grateful to the general practices in the MRC’s General Practice Research Framework who made this study possible: Backwell; Street; Coleford; Hearham; Todmorden; Ipswich; Abingdon; Whitehaven; Yaxley; Folkestone; Kidderminster; Camberley; Northampton; Sutton-in-Ashfield; Brightlingsea; Biggleswade; Liskeard; Barrow-in-Furness; Ledbury; Henley; Widnes; Durham; Walsall; Aston Clinton; Torrington Park; Luton; Darlington; Scunthorpe; St Johns; Bloomsbury; Ormskirk; Stockport; Atherton; Ampleforth; Bolton; Weymouth; Tewkesbury; Tunbridge Wells; Guildford; Credington; Tooting; Birmingham; Bournemouth; Stratford upon Avon; Workington; Hull; Fovant; Cradley; Birtstall; Richmond; Chapel en le Frith; Bingley; Pinner; Sheffield; Basingham; Chessington; Temple Fortune; Hampton Wick; Bracknell; East Grinstead; Newcastle; Settle; Fremington; Limehouse; Leeds; Bishop Auckland; Tonbridge; Bromley; Willesden; Beckton.

References

Appendix