Correction for Biases in a Population-based Study of Family History and Coronary Heart Disease

The Newcastle Family History Study I

Jonathan S. Silberberg,1,2 John Wlodarczyk,1 Jayne Fryer,1 Cheryl D. Ray,1 and Michael J. Hensley1

In this paper, the authors report on the design of a population-based case-control study of family history as a risk factor for coronary heart disease (CHD). They studied the characteristics of subjects who completed a detailed family history questionnaire in 1992–1994 as well as the accuracy of recall of family history in order to quantify both selection and recall biases. Coronary disease cases were enrolled through the Newcastle MONICA Project (Monitoring Trends and Determinants in Cardiovascular Disease), which registered all suspected heart attacks and sudden cardiac deaths in the Lower Hunter region of New South Wales, Australia, between August 1984 and March 1994. Controls were selected at random from the New South Wales electoral roll. The response rate was 76% in cases and 62% in controls; the major factor associated with participation in the study was perceived family history of CHD, more so in the control series than in the case series. Accuracy was determined by comparing information obtained from the proband with that recorded on death certificates. In first-degree relatives, sensitivity of CHD recall was 85% (95% confidence interval (CI) 74–92%) in cases and 95% (95% CI 84–99%) in controls, while specificity was 59% (95% CI 49–69%) and 74% (95% CI 65–82%), respectively. The net bias in both selection and recall is toward the null and hence the comparisons provide a conservative estimate of risk of CHD associated with a positive family history. Am J Epidemiol 1998;147:1123–32.

bias (epidemiology); case-control studies; coronary disease; family characteristics

The goals of the Newcastle Family History Study conducted in New South Wales, Australia, are to define a high-risk strategy (1) for the prevention of coronary heart disease (CHD) in the population based on family history of CHD. We determined the prevalence of a positive family history, defined in various ways, and the risk of a coronary disease event associated with fulfillment of these definitions.

In this paper, we describe the design of the study and consider the likely direction of biases. We have estimated the biases in order to assess their potential impact on the case-control estimates of risk associated with a positive self-reported family history of CHD (2).

Based on the characteristics of respondents, we calculate population selection probabilities (3, p. 194), the relative probability of a person participating in the study, according to their perception of their family history. We also investigate bias in recall of family history of CHD.

MATERIALS AND METHODS

Selection of cases

Coronary disease cases were identified through the Newcastle MONICA Project (4). MONICA (Monitoring Trends and Determinants in Cardiovascular Disease) registered all suspected cases of heart attack and sudden cardiac death in persons aged between 25 and 69 years in the Lower Hunter region of New South Wales between August 1984 and March 1994. This case series is directly comparable with the source population (residents of the five local government areas in the Lower Hunter region).

Recruitment to the Family History Study began on February 1, 1992 and ended on March 31, 1994. We enrolled subjects with a MONICA diagnosis (5) of definite or possible myocardial infarction in six age and sex strata (males under age 65 years, females...
under age 70 years). MONICA cast a wide net and recruited patients admitted to hospital whose presentation was consistent with suspected myocardial infarction, including some with arrhythmias, valvular heart disease, etc. Some of those enrolled subsequently had a final MONICA diagnosis of “not myocardial infarction.”

Because the majority of MONICA registrants are older males, we recruited study subjects sequentially and stopped recruiting older males on August 31, 1993. Recruitment of other participants continued throughout the study period. When the patient was interviewed by the MONICA nurse in hospital, the Family History Study introductory brochure was given to the patient directly. Subjects who were discharged before being seen by the MONICA nurse were sent the brochure by mail. In either case, the introductory questionnaire was completed without the assistance of the MONICA nurse.

MONICA also monitored the death registries to identify out-of-hospital deaths, and sent a questionnaire to their family to ascertain the prior health status of the deceased. We obtained a short family history by proxy from the surviving spouse or next-of-kin.

Selection of controls

In order to represent the source population, we invited a stratified random sample from the nine Hunter districts of the New South Wales electoral roll which correspond to the five local government areas covered by the MONICA survey. Voter registration is required by law in New South Wales and we conducted a single mail-out in 1992, using three age strata as for the case series. The theoretical sampling frame for the study is shown in figure 1, according to the “study base” schema described by Miettinen (6).

Validation of detailed family history

We used a modified form of the family history questionnaire developed in Utah (7), in which subjects first mapped out all their first- and second-degree relatives and then provided detailed information about each. An excerpt from the family history questionnaire is shown in figure 2. All participants who returned the questionnaire by June 1993 were eligible for the validation substudy.

Death certificate searches. We sought to obtain death certificates on all relatives reported by the proband to be dead or for whom vital status was unknown. Searches were conducted in batches with the collaboration of the Registrar of Deaths in each Australian state. Information required for effective location of death certificates was name, date and place of death, and names of parents.

Coding: CHD and other disease status

Death certificates. Death certificates were coded by one experienced person according to the World Health Organization International Classification of Diseases, 9th revision (ICD-9). Up to six causes of death were entered using the four digit ICD-9 code. ICD-9 codes 410–414 inclusive and 429.7 were classified as “definite CHD” while “all heart disease” included the codes for possible CHD, heart failure, rheumatic or non-rheumatic valvular disease and other conditions. We included as “cancer” all neoplasms, whether benign or malignant (ICD-9 codes 140–239 inclusive).

Family history questionnaire. The cause of death as reported by the proband was coded by one experienced physician. For CHD, we defined definite and possible categories; cancer sites were coded individually and the three most common sites (lung, breast, and colon) were combined.

Follow-up of nonrespondents. Two reminder letters were sent in the event that the introductory questionnaire was not returned. For comparison, we report demographic and risk factor profiles of both respondents and nonrespondents of similar age and sex in an earlier Risk Factor Prevalence Study in the region (8).
Number

<table>
<thead>
<tr>
<th>I know nothing at all about this person. (If so, tick box and go to next relative)</th>
</tr>
</thead>
</table>

1. Male □ Female □ 2. Year of birth, or Age at Present.
3. Is this person a blood relative of yours? (Tick one) □ Yes □ Half □ No
4. In what state (or country) was he/she born?
5. Is he/she still alive? (Tick one) □ Yes □ No □ Don't know
   If Don't Know: In what year are you last sure he/she was alive?
   If No: Date of Death: (Give Day, Month & Year, if known)
   Where did he/she die? In which hospital?
   Or... in which suburb/state?
   (If suburb unknown, give town or city.)
6. At present, or at the time of death, is/was this person...
   □ a lifelong nonsmoker? □ an ex-smoker? □ a current smoker?
   If an ex-smoker, at what age did he/she stop?
7. Has he/she ever been told by a doctor or other medical person that he/she suffers from any of the following health problems?
   Stroke.......................................................... YES NO DON'T KNOW AGE FIRST DIAGNOSED
   Asthma............................................................
   Rheumatic Fever (St. Vitus' Dance)..........................
   Heart Attack (Myocardial Infarction)......................
   Angina (Heart Pain)..............................................
   Coronary Bypass Surgery....................................
   High Blood Pressure...........................................
   High Blood Cholesterol....................................... 8
   Diabetes (Sugar)................................................
   Lung Cancer....................................................
   Breast Cancer..................................................
   Colon (Bowel) Cancer......................................... 9
   Skin Cancer.....................................................
   Other............................................................

8. Did you contact anyone else for this information? ........... □ Yes □ No
   If Yes: whom?
   □ Person themselves □ Their spouse/partner □ Other relative

**FIGURE 2.** Excerpt from the detailed family history questionnaire used in the Newcastle Family History Study. A separate page was completed for each relative.

**Statistical methods**

**Selection.** We used logistic regression analysis (SAS Institute, Cary, North Carolina) to study which variables independently predicted participation among people who returned the introductory questionnaire. To study the impact of selection bias on the estimates of coronary risk associated with a positive family history, we estimated selection probabilities according to the method of Kleinbaum et al. (3, p. 194). We expected that rates of both participation and completion would depend on the subjects' family history of CHD. We studied the bias leading to participation (i.e., joining the study at all) and completion (returning the detailed questionnaire) in separate analyses. Using the notation of Kleinbaum et al. (3), we denote $\alpha$, $\beta$, $\gamma$, and $\delta$ as the selection probabilities, $OR^*$ as the estimated odds ratio based on information obtained from the study population (those who participated or completed, respectively) and $OR$ as the estimated odds ratio had information been obtained from the whole source population. Then

$$\frac{\text{odds ratio bias}}{OR} = \frac{OR^* - OR}{OR} = \frac{\alpha\delta}{\beta\gamma} - 1.$$  

There is no net bias if $\alpha/\gamma = \beta/\delta$.

**Recall.** In separate analyses, we studied accuracy of recall of the three categories of disease—CHD, all heart disease, and cancer—using the diagnosis recorded on the death certificate as the gold standard. In the main analysis, we included only those relatives for whom disease status was definitively reported as "yes" or "no." In sensitivity analyses, we reclassified "don't know" as "no."
Crude analysis. We first used the relative as the unit of analysis, ignoring the cluster effect resulting from information about relatives being provided by the proband. We quantified the sensitivity and specificity of disease reporting separately for first- and second-degree relatives and computed exact binomial confidence intervals (9). For the difference between the MONICA and electoral series, 95 percent confidence limits were calculated using the Gaussian approximation (10). To study other determinants of accuracy, we performed multivariate logistic regression analysis with agreement (yes/no) as the outcome variable and two groups of explanatory variables: proband factors (including MONICA or electoral status, age, sex, education level, and ethnic origin) and relative factors (degree of relationship, year of death, and place of death).

Cluster analysis. Because responses about relatives were obtained from the proband rather than the relative directly, these are not independent. We adjusted for the intra-family correlation in accuracy by the method of Zeger and Liang (11, 12) using the SAS macro GEE (SAS Institute, Cary, North Carolina).

Adjustment for recall bias. To evaluate the impact of differential misclassification on the case-control estimates of risk associated with a positive self-reported family history of CHD, we used the matrix correction methods described by Greenland and Kleinbaum (13) and Drews and Greenland (2). In this analysis, a CHD event in any relative could render the family history “positive” and thus agreement could arise even when a different relative was affected. We also studied the extent of bias based on comparisons of the individual relative (i.e., accurate reporting that any particular relative was affected).

RESULTS

Recruitment of MONICA cases

During the study period, 1,444 eligible persons admitted to hospital with suspected acute myocardial infarction in the target age range were reported to us by the MONICA Project. Of these, 1,335 (92 percent) had a final MONICA diagnosis of “definite” or “possible” myocardial infarction. A total of 118 (8 percent) had a short hospital stay and were not seen by the MONICA Project nurse prior to discharge and 13 (1 percent) had died prior to registration.

Demographic characteristics of those who did not reply (nonrespondents) and those who declined to participate (nonparticipants) are given in table 1, along with similar data for those who consented and who completed the detailed family history questionnaire. Overall, 1,092 (76 percent) responded to the introductory brochure, with the lowest response rate being among the subjects not seen in hospital. A total of 831 persons (58 percent) consented to the study, of whom 432 (30 percent) completed the detailed questionnaire.

In regard to the final MONICA diagnosis, nonrespondents had similar frequencies of “definite” or “possible” myocardial infarction. They were more likely to be younger males and to have been admitted to one of the peripheral Hunter hospitals, and less likely to be nonsmokers. Nonparticipants included a higher proportion of older females and those of lower education status. A “positive” family history of CHD was reported by 82 percent (680/831) of participants and 67 percent (175/261) of nonparticipants. The positive family history was attributed to report of CHD in a first-degree relative in 74 percent and 63 percent, respectively. These characteristics were similar in those who provided a short family history only and those who completed the detailed questionnaire.

Recruitment of electoral roll controls

A total of 1,026 persons from the Hunter region population were invited. Of these persons, 637 (62 percent) replied and returned the short questionnaire. Out of the 637, a total of 379 persons (37 percent) consented and 248 (24 percent) completed the detailed family history questionnaire; the breakdown of their demographic characteristics is shown in table 2.

As observed with the case series, older females and persons of lesser education status were less likely to participate. A “positive” family history of CHD was reported by 72 percent (273/379) of participants and 50 percent (130/258) of nonparticipants, and was attributed to a first-degree relative in 58 percent and 44 percent respectively. These characteristics were similar among persons who provided a short family history only and those who completed the detailed questionnaire.

For comparison, table 3 gives age- and sex-weighted frequencies of key variables in both the 1988–1989 (8) and 1994 (14) Hunter Risk Factor Prevalence Studies, the 1989 National Heart Foundation Risk Factor Prevalence Study (15), and the 1987–1988 Newcastle Health Survey (16). The response rates in these surveys ranged from 64 to 80 percent. As shown in table 3, the frequencies of major risk factors in the Family History Study nonparticipants were similar to these surveys.

Multivariate analysis: predictors of participation

Among the people who returned the introductory questionnaire, participation was significantly more likely among the MONICA subjects than the electoral

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Nonrespondents</th>
<th>Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
</tr>
<tr>
<td>No.</td>
<td>352</td>
<td>100</td>
</tr>
<tr>
<td>Age (years), mean (SD*)</td>
<td>54.5 (9.0)</td>
<td>57.5 (9.3)</td>
</tr>
<tr>
<td>Male</td>
<td>212</td>
<td>60</td>
</tr>
<tr>
<td>Admission diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>91</td>
<td>33</td>
</tr>
<tr>
<td>Unstable angina pectoris</td>
<td>71</td>
<td>26</td>
</tr>
<tr>
<td>Chest pain for investigation</td>
<td>55</td>
<td>20</td>
</tr>
<tr>
<td>MONICA diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Definite MI*</td>
<td>137</td>
<td>39</td>
</tr>
<tr>
<td>Possible MI</td>
<td>187</td>
<td>53</td>
</tr>
<tr>
<td>No MI</td>
<td>26</td>
<td>7</td>
</tr>
<tr>
<td>Self-reported past history</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angina pectoris</td>
<td>124</td>
<td>48</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>34</td>
<td>13</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>High blood pressure</td>
<td>145</td>
<td>56</td>
</tr>
<tr>
<td>Current smoker</td>
<td>51</td>
<td>20</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>117</td>
<td>45</td>
</tr>
<tr>
<td>Highest level of education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary school</td>
<td>38</td>
<td>15</td>
</tr>
<tr>
<td>High school</td>
<td>175</td>
<td>67</td>
</tr>
<tr>
<td>&gt;High school</td>
<td>35</td>
<td>13</td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retired</td>
<td>44</td>
<td>14</td>
</tr>
<tr>
<td>Employed full-time</td>
<td>94</td>
<td>29</td>
</tr>
<tr>
<td>Home duties</td>
<td>90</td>
<td>28</td>
</tr>
<tr>
<td>Family history of CHD*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>175</td>
<td>67</td>
</tr>
<tr>
<td>Don’t know</td>
<td>30</td>
<td>11</td>
</tr>
</tbody>
</table>

* MONICA, Monitoring Trends and Determinants in Cardiovascular Disease; SD, standard deviation; MI, myocardial infarction; CHD, coronary heart disease.

roll subjects, females, persons aged 40–49 years, nonsmokers, those who had completed education beyond high school and those with a "positive" family history of CHD (OR = 1.7). The detailed questionnaire was significantly more likely to be completed by electoral roll subjects, nonsmokers, and persons of advanced educational status. Persons who answered "don’t know" to the initial question about family history were less likely to return the detailed questionnaire (OR = 0.5).

Population selection probabilities according to family history of CHD

In both the MONICA series and the electoral roll series, the most important participation bias was that associated with a family history of CHD. The actual sampling frame which operated is shown in figure 3, according to the schema of "target" and "actual" populations described by Kleinbaum et al. (3). The relative frequencies of participating in this study according to family history of CHD and source (MONICA or electoral roll) are given in table 4. For both participation and completion, the observed bias is toward the null (−17 percent and −10 percent, respectively).

Validation substudy

The demographic characteristics of the 124 MONICA participants and 135 electoral roll participants included in the substudy were similar to those of the participants in the overall study, except that they included a higher proportion of older males.

Death certificate search. "Don’t know" responses were more common in the MONICA series. Accordingly, retrieval was somewhat better for the relatives of electoral roll subjects (first-degree, 86 percent; second-degree, 82 percent) than the MONICA subjects (first-degree, 77 percent; second-degree, 70 percent). Death certificates were obtained for 906 MONICA relatives and 1,001 electoral roll relatives.

Cause of death: proband's reports vs. death certificates. The comparisons of sensitivity and specificity when “don’t know” responses were excluded are...

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Nonrespondents</th>
<th>Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
</tr>
<tr>
<td>No.</td>
<td>389</td>
<td></td>
</tr>
<tr>
<td>Age (years), mean (SD*)</td>
<td>50.6 (9.5)</td>
<td>53.0 (10.0)</td>
</tr>
<tr>
<td>Male</td>
<td>217</td>
<td></td>
</tr>
<tr>
<td>Self-reported past history</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angina pectoris</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>High blood pressure</td>
<td>57</td>
<td>22</td>
</tr>
<tr>
<td>Current smoker</td>
<td>36</td>
<td>14</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>56</td>
<td>22</td>
</tr>
<tr>
<td>Highest level of education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary school</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td>High school</td>
<td>145</td>
<td>71</td>
</tr>
<tr>
<td>Advanced</td>
<td>41</td>
<td>20</td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retired</td>
<td>40</td>
<td>16</td>
</tr>
<tr>
<td>Employed full-time</td>
<td>54</td>
<td>25</td>
</tr>
<tr>
<td>Home duties</td>
<td>60</td>
<td>23</td>
</tr>
<tr>
<td>Family history of CHD*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>130</td>
<td>50</td>
</tr>
<tr>
<td>Don’t know</td>
<td>18</td>
<td>7</td>
</tr>
</tbody>
</table>

* SD, standard deviation; CHD, coronary heart disease.

TABLE 3. Percent comparison of Hunter Region, New South Wales, Australia, community risk factor surveys*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current smoker</td>
<td>19.4</td>
<td>19.9</td>
<td>20.8</td>
<td>25.1</td>
<td>18</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>29.7</td>
<td>27.2</td>
<td>27.8</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>Nonsmoker</td>
<td>50.9</td>
<td>52.9</td>
<td>50.7</td>
<td>54</td>
<td>54</td>
</tr>
<tr>
<td>High blood pressure</td>
<td>32.1</td>
<td>34.2</td>
<td>15.1</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>2.8</td>
<td>2.4</td>
<td>2.8</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Short family history of CHD†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any relative</td>
<td>50</td>
<td>73</td>
<td></td>
<td>43</td>
<td>58</td>
</tr>
<tr>
<td>First-degree relative</td>
<td>43</td>
<td>58</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Age- and sex-adjusted estimates were obtained by direct standardization to the age and sex distribution of the Newcastle Family History Study.
† RFPS, risk factor prevalence survey; NHF, National Heart Foundation; CHD, coronary heart disease.

shown in table 5. For first-degree relatives, recall of both CHD and all heart disease was consistently better in the electoral series. The greatest source of misclassification was report of CHD by the proband and report on the death certificate of "other heart disease" or "not heart disease." Of relatives who were wrongly labeled as dying from CHD, 19 percent died from other cardiac diseases and 21 percent from non-cardiac causes. For patients who died from heart diseases other than CHD, the proband reported this as CHD in about 60 percent. In second-degree relatives, both sensitivity and specificity of CHD recall were similar in the two series. Recall of cancer was much better than recall of CHD (overall accuracy, 93 percent in first-degree relatives and 89 percent in second-degree relatives).

Cluster analysis. We examined the impact of the cluster nature of responses about relatives by repeating all analyses using the generalized estimating equations of Liang and Zeger (11, 12). The largest cluster effect (within-proband correlation = 0.12, $z = 3.6, p < 0.001$) was for specificity of recall of heart disease. All...
results were similar to the crude analysis, which assumed independence.

"Don't know" responses. To examine the impact of “don’t know” responses, we reclassified the “don’t know” responses as “no.” When we did this, sensitivity fell in all strata, with a parallel improvement in specificity and with overall accuracy little affected.

Extent of recall bias. For the definition “one or more first-degree relatives with CHD at any age,” the estimated recall bias was −45 percent; for “one or more first-degree relatives with CHD before age 60,” it was −37 percent (a negative bias is toward the null, i.e., conservative). However, we found that the bias was unfavorable when analyzed under a “worst case” scenario. That is, with sensitivity and specificity in cases and controls set at the 25th or 75th centile (whichever was least favorable), the direction of recall bias was reversed.

In figure 4, we describe the analysis based only on the correct relative being identified. We present these as graphs which show the maximum case-control differences in sensitivity and specificity which can be tolerated without biasing the odds ratio by more than 20 percent. This shows that, for the range of sensitivity and specificity observed, differences in recall are unlikely to exert any substantial bias on the estimated risk associated with a positive family history of CHD.

DISCUSSION

In this paper, we have attempted to quantify the selection biases inherent in studies of this type. First, because subjects were invited to join the “Family History Study,” we expected that those who considered themselves to have a positive family history of CHD would be more inclined to participate. This participation bias was found to be differential in that it was greater in the electoral roll series than in the MONICA series. The bias is toward the null and adds a safety margin for the case-control comparison to estimate the risk of CHD associated with a positive family history.

Second, case-control studies are often criticized for the potential for recall bias (17). However, empirical studies (18, 19) have shown that recall bias is often less of a problem than expected. We found that recall of CHD was better among the controls than among the cases. Such an observation may appear to be counter-intuitive, because admission to hospital should provide a memory jog to improve recall of family history of heart disease. This is likely to be due to the selection factors encountered. That is, in the electoral roll series, subjects who were unable to provide accurate information tended not to participate in the first instance. The observed differences in sensitivity and specificity bias the odds ratio in opposite directions, but the overall effect is conservative. On the one hand, false negative recall by cases leads to underreporting of affected relatives. On the other hand, false positive responses increase the frequency of reported family history. The net effect results because death from CHD is a common exposure and the odds ratio is not so sensitive to case-control differences in specificity (2).

The results of the crude and cluster analyses are essentially the same, even though we found a significant cluster effect when we examined specificity of recall of heart disease. This showed that a small num-

**TABLE 4. Selection probabilities: relative likelihood of a respondent participating and completing the study, by source and family history, Newcastle Family History Study, Australia, 1992–1994**

<table>
<thead>
<tr>
<th>Source</th>
<th>Family history of CHD*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Participation</strong></td>
<td></td>
</tr>
<tr>
<td>MONICA* subjects</td>
<td>680/855</td>
</tr>
<tr>
<td>Electoral roll subjects</td>
<td>274/404</td>
</tr>
<tr>
<td>Participation bias (%)†</td>
<td>-17</td>
</tr>
<tr>
<td><strong>Completion</strong></td>
<td></td>
</tr>
<tr>
<td>MONICA* subjects</td>
<td>357/855</td>
</tr>
<tr>
<td>Electoral roll subjects</td>
<td>178/404</td>
</tr>
<tr>
<td>Completion bias (%)†</td>
<td>-10</td>
</tr>
</tbody>
</table>

* CHD, coronary heart disease; MONICA, Monitoring Trends and Determinants in Cardiovascular Disease.
† Calculated by the method of Kleinbaum et al. (3). Minus sign indicates bias toward the null.

Hunter population

![FIGURE 3. Actual sampling frame for the Newcastle Family History Study. The coronary heart disease (CHD) and source populations are stratified by family history of CHD; the central box represents a 2 × 2 table of study participants. The terms α, β, γ, and δ represent the probability of appearing in the four cells. For instance, α is the proportion of cases with a family history of CHD who participated, etc. Family history of CHD: MONICA, α = A/A; electoral roll, γ = C/C. No family history of CHD or “don’t know": MONICA, β = B/B; electoral roll, δ = D/D.](https://academic.oup.com/aje/article-abstract/147/12/1123/93169)
TABLE 5. Sensitivity and specificity (in percent) of reporting of coronary heart disease (CHD), all heart disease, and cancer, with “don’t know” responses excluded, Newcastle Family History Study, Australia, 1992–1994

<table>
<thead>
<tr>
<th>Type of disease and relationship</th>
<th>MONICA* subjects</th>
<th>Electoral roll subjects</th>
<th>Difference (MONICA subjects minus electoral roll subjects)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Value</td>
<td>95% CI*</td>
<td>Value</td>
</tr>
<tr>
<td>CHD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First-degree relatives</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>85</td>
<td>74 to 92</td>
<td>95</td>
</tr>
<tr>
<td>Specificity</td>
<td>59</td>
<td>49 to 69</td>
<td>74</td>
</tr>
<tr>
<td>Accuracy</td>
<td>70</td>
<td>63 to 77</td>
<td>80</td>
</tr>
<tr>
<td>Second-degree relatives</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>80</td>
<td>71 to 87</td>
<td>80</td>
</tr>
<tr>
<td>Specificity</td>
<td>68</td>
<td>61 to 75</td>
<td>65</td>
</tr>
<tr>
<td>Accuracy</td>
<td>72</td>
<td>67 to 77</td>
<td>70</td>
</tr>
<tr>
<td>All heart disease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First-degree relatives</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>83</td>
<td>75 to 90</td>
<td>90</td>
</tr>
<tr>
<td>Specificity</td>
<td>70</td>
<td>58 to 80</td>
<td>83</td>
</tr>
<tr>
<td>Accuracy</td>
<td>78</td>
<td>71 to 83</td>
<td>86</td>
</tr>
<tr>
<td>Second-degree relatives</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>76</td>
<td>69 to 82</td>
<td>76</td>
</tr>
<tr>
<td>Specificity</td>
<td>80</td>
<td>72 to 86</td>
<td>74</td>
</tr>
<tr>
<td>Accuracy</td>
<td>77</td>
<td>72 to 82</td>
<td>75</td>
</tr>
<tr>
<td>Cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First-degree relatives</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>89</td>
<td>75 to 97</td>
<td>100</td>
</tr>
<tr>
<td>Specificity</td>
<td>93</td>
<td>88 to 97</td>
<td>93</td>
</tr>
<tr>
<td>Accuracy</td>
<td>92</td>
<td>87 to 96</td>
<td>95</td>
</tr>
<tr>
<td>Second-degree relatives</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>97</td>
<td>90 to 100</td>
<td>88</td>
</tr>
<tr>
<td>Specificity</td>
<td>85</td>
<td>80 to 90</td>
<td>90</td>
</tr>
<tr>
<td>Accuracy</td>
<td>88</td>
<td>84 to 92</td>
<td>89</td>
</tr>
</tbody>
</table>

* MONICA, Monitoring Trends and Determinants in Cardiovascular Disease; CI, confidence interval.

Because our goal is to derive population-based estimates of CHD risk associated with family history, we would like to know the extent to which our study sample reflects the Hunter region population. We were unable to estimate the response bias associated with family history, because no family history information is available for persistent nonrespondents. Instead, we examined their demographic data and found those data to be similar. We have calculated that $\delta$ would have to be as high as 0.55 to offset the observed safety margin and the direction of all observed biases would have to be reversed.

The final case-control comparison of precise family history definitions, described in the companion paper, Paper II (25), is based only on those who completed the detailed questionnaire. We have shown that although the participation and completion rates are fairly low, there are no important differences between these and other eligible subjects. Furthermore, selection factors led to a study population in which recall bias was generally in a conservative direction. The net biases are toward the null for the case-control comparison.

Am J Epidemiol Vol. 147, No. 12, 1998
Acknowledgments

Supported in part by grant no. G91S3280 from the National Heart Foundation of Australia.

The authors wish to acknowledge the help of Hilary Alexander (Project Manager) and the staff of the Newcastle MONICA team, as well as that of Graham Austin and the staff at the New South Wales Registry for Births, Deaths, and Marriages.

References


