Optimal strategies for identifying patients with myocardial infarction in general practice

Peter T Donnan, Hamish T Dougall and Frank M Sullivan


Background. In order to provide evidence-based secondary prevention of coronary heart disease (CHD) in general practice, eligible patients need to be identified. The optimal strategy is one in which all appropriate patients are identified with the least effort.

Objective. The purpose of the study was to determine the optimal strategy to identify subjects with a myocardial infarction (MI) from general practice records using different search criteria.

Methods. The study was a cross-sectional survey of 10 general practices in Tayside, Scotland. A random sample of all subjects aged over 35 (n = 5061) and registered with the general practices was obtained. The main outcome measures were sensitivity, specificity, positive predictive value (PPV) and yield (the number of records that need to be examined to detect a ‘true case’).

Results. Of the sample of 5061, 207 (4.1%) were defined to have had a ‘gold standard’ MI. A Read code for ischaemic heart disease (IHD) had the highest sensitivity (95%) but with a poor PPV (52%). All searches had high specificities. The addition of a record of hospitalization for MI to the Read code for MI gave 100% sensitivity and high yield (1 in 1.11). In situations where the Read coding is of poor quality, the alternative search strategy of a hospital record of MI or receiving aspirin or nitrates was optimum.

Conclusions. Patients who had experienced an MI can be easily identified from a combination of a Read code for MI and a record of hospitalization for an MI giving 100% sensitivity and specificity with a yield of 1 in 1.11.

Keywords. Coronary heart disease, myocardial infarction, Read coding, secondary prevention.

Introduction

Various well-publicized studies have pointed repeatedly to clear deficiencies in the delivery of best care in the secondary prevention of coronary heart disease (CHD).1–3 National guidelines such as the National Service Framework are driving practices to improve care.4 Before a practice can provide evidence-based secondary prevention of CHD, they have to be able to identify eligible patients. The HEARTS initiative (www.hearts.org.uk) was established at the beginning of 2001 to support practices in implementing best evidence-based care for their post-myocardial infarction (MI) patients through the separate strands of audit, education and research.5,6 HEARTS mirrors the successful DARTS (Diabetes Audit and Research in Tayside) project, which has been running in Tayside for the past 7 years.7 This novel initiative is a clinical audit and research tool that provides real time electronically linked data from a multitude of sources on diabetes patients via the NHS-net (www.diabetes-healthnet.ac.uk).8

An earlier study to identify patients with ischaemic heart disease (IHD) rather than MI in general practice demonstrated a sensitivity of 73% using a strategy of a Read code for IHD (G3) and a prescription for a nitrate.9 Sensitivity was increased to 96% but with a concomitantly greater effort by searching through lists of patients receiving aspirin, atenolol, digoxin or a statin. Clearly, the optimal strategy is one that identifies as many patients as possible, i.e. a sensitivity close to 100%, but with the least effort in terms of number of individual searches, i.e. a high yield. We define yield as the number of records that need to be examined to detect a ‘true case’ (see Table 1).9

This strategy for identification of IHD, although developed in England and Wales, could also be used to

Received 17 February 2003; Revised 19 June 2003; Accepted 14 July 2003.
Tayside Centre for General Practice, Kirsty Semple Way, University of Dundee, Dundee DD2 4BF, UK. Correspondence to Dr PT Donnan; E-mail: p.t.donnan@dundee.ac.uk

706
identify patients with an MI in Scotland since Read codes are also used in general practice in Scotland. However, Scotland has the additional advantage that a record of hospital admission for an MI, the Scottish Morbidity Record 1 (SMR1), is easily obtained from the Information and Statistics Division (ISD) of the Scottish Office. The SMR1 contains all data on individual hospital episodes for the whole of Scotland. Since there is little private hospital provision in Tayside, this is considered to be almost 100% complete.

This study aimed to identify the optimal search strategy with the highest sensitivity along with high yield, to enable the identification of MI patients using a random sample of 5000 individuals in the population of Tayside aged 35 and over, and registered with a GP.

**Methods**

**Study population**

The study population consisted of subjects resident in Tayside aged 35 years or over and registered with 10 Tayside general practices from 1 December 2000. All 75 practices in Tayside were asked to participate; 63 agreed initially and 10 were chosen at random by a secretary unconnected with the study team. The total number of patients eligible for inclusion was 50,355. The 12 practices which initially refused to take part had 100% computer usage and all subsequently later were admitted to the study.

**Sample**

A unique 10-digit number (Community Health Number, CHNo.) is used to identify all subjects registered with a GP in Scotland. In Tayside, this number is used for all health care encounters and so facilitates record linkage of data from multiple sources. Using a database with all the CHNos for Tayside, a 1:10 sample of 5061 was randomly selected from the 10 general practices stratified by age (35–44, 45–54, 55–64, 65–74, 75–84 and 85+) and gender according to the age and sex distribution in Tayside, Scotland.

**Searches**

Three trained data facilitators visited the 10 practices. Each practice list was searched independently for an indication of MI by practice registration for an MI; a Read code for an MI; or a Read code for IHD. In addition, each subject was checked separately for appearance on lists of patients receiving aspirin, statins, nitrates or β-blockers. Finally, events of hospitalization for an MI were obtained from the Tayside part of the SMR1 held in the Medicines Monitoring Unit. A second facilitator validated the information from the first facilitator and disagreements were resolved with two clinicians (HTD and FMS). Data were anonymized before analysis to comply with the data protection act. The study was approved by the Tayside Caldicott Guardians and the local Medical Research Ethics Committee.

**Case definition**

A ‘gold-standard’ MI was determined to have occurred where discharge or other clinical documentation recorded at the time of, or immediately following, a cardiovascular event confirmed that an MI had occurred according to adapted WHO MONICA criteria (Fig. 1). Where post-discharge investigations show that the event was not an MI, the latter verdict was considered dominant. Where, at the time of discharge, an MI was suspected but not proven, and where no post-discharge investigations confirm the infarct, a diagnosis of MI was not recorded.

**Statistical methods**

Each individual subject in the random sample was classified (Yes/No) as having had an MI (ICD9 410) according to the ‘gold standard’ criteria described above.
All subjects were also classified by each of the searches for identifying an MI (Yes/No). Sensitivity, specificity, positive predictive value (PPV) and yield and their 95% confidence intervals (CIs) were calculated for each search and combination of searches. The yield is the number of records that need to be examined to detect a ‘true case’ of MI. CIs were calculated adjusting for clustering within practice by calculating the variance inflation factor. All analyses were implemented using the statistical package SPSS (Version 10).

Results

A total of 5061 patients were sampled from the population of the 10 randomly selected practices. Of this total, 207 (4.1%; range 2.4–5.5% across the practices) were found to have had an MI by the ‘gold standard’ definition. Table 2 shows the results of comparing each individual search strategy with the gold standard identification of MI from detailed case notes. Since all the specificities were high and ≥94%, these are not presented in the table. A Read code for IHD had the highest sensitivity of all the single search strategies, although its PPV was not particularly high. This simply reflects that a large proportion of patients with IHD have not had an MI (48.4%). A Read code for MI was, not surprisingly, highly sensitive but also highly specific (>94%) and, in addition, had a high PPV. This search strategy consequently had a good detection rate with a very high yield. An SMR1 record of an MI was also highly specific and had a high yield, but sensitivity was relatively low, reflecting that 41% of MI patients did not have a record of hospitalization for MI. Of the drug searches, aspirin had the highest sensitivity but also the lowest yield. Hence aspirin searches were likely to be good for not missing cases but at the expense of a higher workload for those searching because it produces many false positives.

Table 3 shows combinations of strategies incorporating a Read code for MI with other searches. Virtually all strategies apart from the use of β-blockers identified all cases of MI. The main differences were in PPV, with the combination of Read code for MI and

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Sensitivity (% and 95% CI)</th>
<th>PPV (% and 95% CI)</th>
<th>Yield (100/PPV)</th>
<th>No. of false positives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read code IHD</td>
<td>95 (91–99)</td>
<td>52 (32–71)</td>
<td>1 in 1.94</td>
<td>185</td>
</tr>
<tr>
<td>Read Code MI</td>
<td>83 (71–96)</td>
<td>91 (83–99)</td>
<td>1 in 1.10</td>
<td>17</td>
</tr>
<tr>
<td>Aspirin</td>
<td>70 (55–86)</td>
<td>33 (11–54)</td>
<td>1 in 3.08</td>
<td>303</td>
</tr>
<tr>
<td>SMR1 – MI</td>
<td>59 (52–66)</td>
<td>95 (91–100)</td>
<td>1 in 1.05</td>
<td>6</td>
</tr>
<tr>
<td>Practice Reg.</td>
<td>58 (33–88)</td>
<td>81 (59–103)</td>
<td>1 in 1.24</td>
<td>29</td>
</tr>
<tr>
<td>Nitrates</td>
<td>55 (41–69)</td>
<td>50 (28–73)</td>
<td>1 in 1.98</td>
<td>112</td>
</tr>
<tr>
<td>Statins</td>
<td>52 (42–62)</td>
<td>43 (23–63)</td>
<td>1 in 2.33</td>
<td>143</td>
</tr>
<tr>
<td>β-Blockers</td>
<td>50 (43–57)</td>
<td>29 (4–54)</td>
<td>1 in 3.45</td>
<td>255</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Sensitivity (% and 95% CI)</th>
<th>PPV (% and 95% CI)</th>
<th>Yield (100/PPV)</th>
<th>No. of false positives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read code MI+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMR1 – MI</td>
<td>100</td>
<td>90 (83–97)</td>
<td>1 in 1.11</td>
<td>23</td>
</tr>
<tr>
<td>Nitrates</td>
<td>100</td>
<td>62 (44–80)</td>
<td>1 in 1.60</td>
<td>125</td>
</tr>
<tr>
<td>Statins</td>
<td>100</td>
<td>57 (37–76)</td>
<td>1 in 1.76</td>
<td>157</td>
</tr>
<tr>
<td>Aspirin</td>
<td>100</td>
<td>40 (17–63)</td>
<td>1 in 2.50</td>
<td>310</td>
</tr>
<tr>
<td>β-Blockers</td>
<td>90 (81.99)</td>
<td>41 (17–65)</td>
<td>1 in 2.43</td>
<td>266</td>
</tr>
</tbody>
</table>
SMR1 record of MI having the highest yield and hence lowest workload. In the situation where hospital codes (SMR1 in Scotland) were unavailable, then a Read code for MI in combination with nitrates would be the next optimal strategy (Table 3).

From the experience of the data facilitators, it was found that the quality of the Read code and practice registration records was variable and so it was necessary to explore search strategies that did not use Read codes. For example, the sensitivity of a Read code for MI varied from 33 to 97%, although most were between 80 and 90%. The wide CI for Read code PPV for IHD in Table 2 illustrates the high variability of Read coding. Table 4 compares strategies incorporating an SMR1 record of MI with various drug combinations. The best sensitivity was obtained from the combination of an SMR1 record of MI or aspirin, although at the expense of a poor PPV. Sensitivity was lower for combinations with nitrates or statins, but this was offset by better yields. Finally, a combination of an SMR1 record, or aspirin, or nitrates improved sensitivity to 91.8% with similar PPV. The addition of statins increased the sensitivity slightly but with a consequent worsening of yield, and so was not considered optimum.

Discussion

This study has demonstrated that the optimum search strategy for identifying cases of MI in general practice, with the highest yield and sensitivity, was a combination of a Read code for MI or an SMR1 record of hospitalization for an MI. This strategy gave 100% sensitivity with a yield of 1 in 1.11 records. Hence, this provides a pragmatic way for GPs to identify patients eligible for secondary prevention. If the Read code database was found to be of poor quality, an alternative would be a combination of an SMR1 record of hospitalization for MI or aspirin or nitrates. The latter strategy had a sensitivity of 91.8% and a yield of 1 in 2.83 records. Poor quality of Read coding could be formally defined by carrying out a search based on Read code and finding that the actual number of cases identified was >2 SD from the expected number of cases based on the age and sex distribution of the practice. Failure to Read code is a frequent finding in GP research. It should also be noted the 10 practices in this study were a random sample of volunteers and therefore may represent better quality of coding compared with other practices in Tayside, which did not take part. In situations where a practice does not have access to hospital records, a strategy of a Read code for MI and nitrates would be optimal in terms of sensitivity and yield.

Although the use of unique patient identifiers is fairly unusual in the UK, the methods identified here are, in fact, available in other parts of the world, notably Denmark, Columbia and Australia. Where a unique identifier is not available, probabilistic record linkage methods can achieve similar results, although requiring more resources. In any case, the best approach of unique identifiers should be encouraged, and this study demonstrates the power of such an approach. Of course, the methods described in this study are equally applicable to data stored on paper records, where computers are not available, although such searches would be more prone to human error.

Clinicians and policymakers trying to improve the management of cardiovascular disease in populations need to know who the affected patients are. The NHS IT strategy has proposed several components designed to ensure that practice computing resources assist in the development of useful disease registers. These are:

(i) NHSNet, a communications infrastructure;
(ii) unique patient identifiers (NHS numbers);
(iii) registers of demographic data for all health care sectors;
(iv) READ coding of clinical terms;
(v) national standards for communication and data entry templates in primary care;
(vi) data guardians to ensure security and confidentiality.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Sensitivity (% and 95% CI)</th>
<th>PPV (% and 95% CI)</th>
<th>Yield (100/PPV)</th>
<th>No. of false positives</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMR1 MI+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspirin</td>
<td>87 (82–92)</td>
<td>37 (15–59)</td>
<td>1 in 2.70</td>
<td>307</td>
</tr>
<tr>
<td>Nitrates</td>
<td>77 (71–84)</td>
<td>58 (39–77)</td>
<td>1 in 1.72</td>
<td>116</td>
</tr>
<tr>
<td>Statins</td>
<td>74 (69–80)</td>
<td>51 (32–69)</td>
<td>1 in 1.97</td>
<td>149</td>
</tr>
<tr>
<td>β-Blockers</td>
<td>74 (67–82)</td>
<td>37 (13–61)</td>
<td>1 in 2.69</td>
<td>260</td>
</tr>
<tr>
<td>Aspirin + nitrates</td>
<td>92 (88–96)</td>
<td>35 (13–57)</td>
<td>1 in 2.83</td>
<td>348</td>
</tr>
<tr>
<td>Aspirin + nitrates + statins</td>
<td>93 (89–97)</td>
<td>31 (9–54)</td>
<td>1 in 3.18</td>
<td>422</td>
</tr>
</tbody>
</table>
Several components of this strategy are already operational in Tayside, but many practices still need additional resources to improve the quality of their clinical databases. With adequate computing resources to produce and maintain disease registers, the strategies developed in this study could be used to identify suitable patients for secondary prevention easily and efficiently.

Acknowledgements
We thank Neil McEwan, who set up and maintains the HEARTS database, and the team of facilitators: Julie Beattie, Michelle Hughes, Shona Black, Mandy Andrew, Karen Fletcher, Sarah Trueman, Helen Anderson and Sue Sage. The HEARTS collaboration is supported by NHS Tayside and an unrestricted educational grant from Boehringer Ingelheim.

References
16 Breen KJ. Consent for the linkage of data for public health research: is it (or should it be) an absolute pre-requisite? Aust NZ J Public Health 2001; 25: 423–425.