We thank Stevens and McManus for pointing out that using unlinked primary care databases may underestimate the absolute risk of myocardial infarction because these databases fail to capture all events.\(^1\)\(^2\) Non-differential under-recording may not cause bias in studies solely focusing on relative effects, but for predicting absolute risks or rates, using a single data source—whether from primary or secondary care settings—may lead to biased estimates. Ideally, all events should be recorded definitively in a single electronic health record. Because this currently doesn’t occur in the NHS or other health systems, we recommend the use of linked data sources, such as those available through the Clinical Practice Research Datalink, to overcome under-recording in individual sources.

In the Cardiovascular disease research using Linked Bespoke studies and Electronic health Records (CALIBER) programme,\(^3\) we are developing prognostic models for patients with coronary disease in a linked dataset, using multiple data sources for outcome ascertainment. We are investigating the use of free text entered by doctors as an additional source of diagnostic information.\(^4\) We are part of the new network of four UK e-health informatics research centres,\(^5\) which will make use of linked datasets available through the Clinical Practice Research Datalink, facilitate further linkages, and enable greater use of electronic health records for research.

Competing interests: None declared.


Cite this as: BMJ 2013;346:f3741
© BMJ Publishing Group Ltd 2013