THEORY AND METHODS

# Is an internal comparison better than using national data when estimating mortality in longitudinal studies?

**T R Card, M Solaymani-Dodaran, R Hubbard, R F A Logan, J West**

*J Epidemiol Community Health* 2006;**60**:819–821. doi: 10.1136/jech.2005.041202

See end of article for authors' affiliations

Correspondence to: Dr T R Card, Division of Epidemiology and Public Health, University of Nottingham Medical School, Queen's Medical Centre, Nottingham NG7 2UH, UK; timcard@doctors.org.uk

Accepted for publication 20 March 2006

**Background:** Discrepancies between the results of different studies looking at mortality in similar disease cohorts led us to consider the impact of methodology upon outcome.
**Methods:** Cohort studies were carried out using age, sex, practice, and calendar time matched control groups in the general practice research database. Data were used on all subjects with inflammatory bowel disease, coeliac disease, or Barrett's oesophagus. Mortality data for the population of England and Wales were obtained from the UK Office for National Statistics. The study compared hazard ratios (HR) for mortality using the matched controls to those found when an indirect standardisation to the mortality experience of England and Wales was carried out.
**Results:** For all three conditions the mortality risk was slightly lower when the national population data were used compared with the internal comparison group (coeliac disease HR 1.33 *v* standardised mortality ratios (SMR) 1.25, Barrett's oesophagus HR 1.32 *v* SMR 1.32, inflammatory bowel disease HR 1.50 *v* SMR 1.34).
**Conclusions:** A bias was found towards underestimating mortality risk when cohort studies use national population death rates as a comparator. Estimates obtained when an internal comparison group has been used are probably more appropriate.

The calculation of standardised mortality ratios (SMR) is an epidemiological technique that aims to permit comparisons of mortality between populations corrected for the confounding effects of age and sex.[1] This method is particularly useful when comparing a subset of a large (for example, national) population with the whole of the population as comparisons can be made using routinely available summary data (mortality figures for age and sex groups) and is widely used.[2 3] In this paper we question whether this technique may now be inappropriate in some settings where large clinical databases provide access to alternative methodologies.

We recently published a study of mortality among patients with inflammatory bowel disease based upon a cohort nested within the general practice research database (GPRD) and using a Cox regression analysis to compare the inflammatory bowel disease and control cohorts.[4] Our study unsurprisingly showed that these people suffering from a chronic disease had a greater risk of death than did the control group. However, in the same issue of the journal publishing our study there appeared another study of mortality in inflammatory bowel disease patients that showed a radically different result based upon an SMR analysis comparing a patient group in Denmark with local population figures.[5] To try to understand the differences between these results we re-examined our own data and carried out a similar SMR analysis. To examine whether any discrepancies between the two methodologies were specific to our inflammatory bowel disease subjects we carried out similar comparisons in two other chronic gastrointestinal disease cohorts in which we have recently examined their mortality.[6]

## METHODS

The GPRD is the world's largest longitudinal, primary care database and contains about 50 million patient years of data collected from computerised UK general practices since 1987. To ensure data quality, contributing practices received data quality training, and are audited to ensure that at least 95% of prescribing and morbidity events are included.[7] The construction of each of the cohorts within GPRD used in this study has been previously described in some detail.[4 6] To summarise we used cohorts of all people with inflammatory bowel disease, coeliac disease, and Barrett's oesophagus in the GPRD. For each of these we selected age, sex, and practice matched controls from the GPRD. From the available data on these subjects we extracted information on age, sex, and death (whether and when it occurred). These data were used initially to construct Cox regression models for each pair of cohorts. To permit the most appropriate comparison with SMR analysis we restricted our regression analyses to correcting for confounding by age (in five year age bands), sex, and calendar period; we also limited our analysis to data gathered before 2000 as population data for the SMR analysis were not available after this (see below). With the exception of this limitation the methodology of this part of the analysis was the same as that described in our previous papers.[4 6]

After performing the regression analyses we went on to perform SMR analyses for each of our cohorts. For the purposes of this study the standard population chosen was

**Table 1** Cox analysis of mortality adjusted for age in five year bands, sex, and calendar period

| | Deaths | Number | HR | 95% CI |
|---|---|---|---|---|
| Barrett's | 145 | 1417 | 1.32 | 1.10, 1.57 |
| Control* | 911 | 11322 | 1 | |
| All IBD | 981 | 16068 | 1.50 | 1.40, 1.61 |
| Control* | 3509 | 80491 | 1 | |
| Coeliac disease | 223 | 4585 | 1.33 | 1.15, 1.54 |
| Control* | 848 | 22894 | 1 | |

*Baseline category. IBD, inflammatory bowel disease.

**Abbreviations:** CI, confidence interval; GPRD, general practice research database; HR, hazard ratio; SMR, standardised mortality ratio

**Table 2** SMR analysis of cases

| | Observed | Expected | SMR | 95% CI |
|---|---|---|---|---|
| Barrett's | 145 | 110 | 1.32 | 1.11, 1.55 |
| All IBD | 981 | 730 | 1.34 | 1.26, 1.43 |
| Coeliac disease | 223 | 178 | 1.25 | 1.09, 1.43 |

IBD, inflammatory bowel disease.

**Table 3** SMR analysis of controls

| | Observed | Expected | SMR | 95% CI |
|---|---|---|---|---|
| Barrett's control group | 911 | 944 | 0.97 | 0.90, 1.03 |
| All IBD control group | 3509 | 3977 | 0.88 | 0.85, 0.91 |
| Coeliac disease control group | 848 | 925 | 0.92 | 0.86, 0.98 |

IBD, inflammatory bowel disease.

the population of England and Wales. Annual age, sex specific mortality rates were obtained for this population from the Office for National Statistics in five year age bands.[8] The analyses were carried out using Stata 7/SE software.

## RESULTS

Our study cohorts comprised of 16 068 inflammatory bowel disease cases matched to 80 491 controls with a median age of 44.1 years and an interquartile range of 31.5 to 60.3, 4588 coeliac cases matched to 22 894 controls with a median age of 44.2 and an interquartile range of 28.4 to 58.7. and 1417 Barrett's cases matched to 11 322 controls with a median age of 65.0 and an interquartile range of 53.7 to 74.2. Among the cases there were 981 deaths in inflammatory bowel disease cases, 145 deaths in Barrett's cases, and 223 deaths in coeliac cases. The figures for the corresponding control groups were 3509, 911, and 848 respectively.

The Cox regression analysis gave hazard ratios for death of 1.50 (95% CI: 1.40, 1.61) for the inflammatory bowel disease cohort when compared with their matched controls, 1.33 (95% CI: 1.15, 1.54) for the coeliac cohort, and 1.32 (95% CI: 1.10, 1.57) for the Barrett's cohort (table 1). The SMR analysis showed a similar level of risk for the Barrett's cohort (SMR 1.32 (95%: CI 1.11, 1.55)) but the SMR for the inflammatory bowel disease and coeliac cohorts both suggested lower levels of risk (SMR 1.34 (95% CI:1.26, 1.43) and 1.25 (95% CI: 1.09, 1.43)) respectively. Table 2 summarises these data.

Examination of the corresponding control cohorts for each of our sets of cases showed that these groups chosen from within GPRD had consistently lower risk of death than that expected from the national figures. The SMR found were 0.88 (95% CI: 0.85, 0.91) for the controls for the inflammatory bowel disease cohort, 0.92 (95% CI: 0.86, 0.98) for the coeliac controls, and 0.97 (95% CI: 0.90, 1.03) for the Barrett's controls (table 3).

## DISCUSSION

In this study of three paired disease cohorts and controls from the GPRD we have shown that indirect standardisation against the national population tends to suggest lower levels of relative risk of death for cases than those found by comparison against internal controls. We have further shown that this is probably attributable to a consistently lower risk of death among our controls than is seen in the population data.

Judging the importance of these results depends on the assessment of two important questions. Firstly, we must examine whether it is probable that the apparent systematic differences between the results gained from the two methodologies are a misleading representation of the data within GPRD because of chance, bias, or confounding? That chance might be an explanation was a possibility we seriously considered when seeing the finding in one cohort. In fact as in each case the 95% confidence intervals for the two methodologies overlap it is clearly a possibility that we cannot entirely exclude. We think it is unlikely however that the same finding would be replicated albeit to a lesser degree purely by chance in a second cohort. Further reassurance that this is not a chance finding comes from the significant differences in control group mortality compared with national figures that explain the differences we found. That our analyses have not dealt with potential confounding factors is undeniable, but as we are concerned here with the difference between the two methodologies, and we ensured that each was similarly limited, this is unlikely to be an explanation. Finally, therefore we must consider the possibility of bias. As selection of cohorts within the GPRD was by a transparently unbiased method the only possible bias would be that of differential recording of the outcome (that is, death) between cases and controls. This is a possibility that we have previously considered and shown not to be an important problem in this dataset for this outcome.[4]

The second important question is whether it is possible that our findings are specific to GPRD or whether they could be more widely generalisable to clinical data from other sources. It has been shown that morbidity in the GPRD is equivalent to other measures of morbidity in general practice in the UK,[9] and hence we think that any finding from GPRD is probably generalisable at least to other general practice data in the UK. Other clinical databases (from tertiary or secondary care) can therefore be expected to be free from the problem highlighted in this paper only if they are truly more reflective of the general population than is general practice data, this seems unlikely.

### What this paper adds

The use of indirect standardisation against national data does not produce results equivalent to the use of an internal control group. There is a bias towards underestimating mortality risk when cohort studies use national population death rates as a comparator. Estimates obtained when an internal comparison group has been used are probably more appropriate.

### Policy implications

When using population based datasets an internal control group drawn from the same dataset is preferable to the use of external rates for comparison.

### What is already know on this topic

The calculation of standardised mortality ratios is an epidemiological technique that aims to permit comparisons of mortality between populations. It is widely used but the advent of large clinical databases provides alternative methodologies.

One further source of potential error in our analysis also requires highlighting. This is the possibility that the differences we have seen are a reflection only of the different statistical methodologies used. It was for this reason that we limited our Cox analysis to omit the examination of confounding (ensuring the greatest similarity between the methodologies), and examined the SMR of our control cohorts. That this alternative analysis supports our findings suggests that any effect of methodology is small. An alternative would have been to derive age and sex specific mortality rates for the whole of the GPRD to use as another standard population. The data required for this were however not available to us and likewise we were unable merely to calculate an SMR for the whole of the GPRD dataset standardised against the national data.

How then could we explain our findings? Why should such a bias arise? We would suggest that there are two main possibilities. These are firstly that GPs under-record deaths in their practice notes in a manner that does not affect national statistics, and secondly that general practice lists do not include some section or sections of the population that have a higher than average risk of death. As the difference found was greater in our younger cohorts it is probable that these missing deaths are in particular among the young. One group of such unregistered people with a high risk of death are the homeless.

It is not revolutionary to suggest that the use of an internal group of controls is preferable to the use of summary population data and the technique of standardisation. There can be no doubt as to the benefits available of using internal controls to account for a wider range of confounders. What has not been previously recognised, but is clearly suggested by our results, is that the use of population data to construct an SMR is liable to produce a bias when the case population under study is from a clinical database. This bias is analogous to the healthy worker effect[10] in occupational studies and, just as in that case, it is easily avoided by selection of controls from an appropriate source rather than from the general population.


## Authors' affiliations

**T R Card, M Solaymani-Dodaran, R Hubbard, R F A Logan, J West,** Division of Epidemiology and Public Health, University of Nottingham Medical School, Queen's Medical Centre, Nottingham, UK



Funding: JW and TRC were supported by The Wellcome Trust (grant numbers 063800 and 060529). The work on Barrett's oesophagus was partly funded by a grant from the Special Trustees for Nottingham University Hospitals. The work on inflammatory bowel disease was partly funded by a grant from the National Association for Colitis and Crohn's Disease.

Conflict of interest: none.

This study was granted ethical approval by the Scientific and Ethical Advisory Group for the General Practice Research Database.


## REFERENCES


1 **Armitage P**, Berry G. *Statistical methods in medical research*. Oxford: Blackwell Scientific, 1994:660–7.
2 **Pui CH**, Cheng C, Leung W, *et al*. Extended follow-up of long-term survivors of childhood acute lymphoblastic leukemia. *N Engl J Med* 2003;**349**:640–9.
3 **Potterat JJ**, Brewer DD, Muth SQ, *et al*. Mortality in a long-term open cohort of prostitute women. *Am J Epidemiol* 2004;**159**:778–85.
4 **Card T**, Hubbard R, Logan RF. Mortality in inflammatory bowel disease: a population-based cohort study. *Gastroenterology* 2003;**125**:1583–90.
5 **Winther KV**, Jess T, Langholz E, *et al*. Survival and cause-specific mortality in ulcerative colitis: follow-up of a population-based cohort in Copenhagen County. *Gastroenterology* 2003;**125**:1576–82.
6 **West J**, Logan RF, Smith CJ, *et al*. Malignancy and mortality in people with coeliac disease: population based cohort study. *BMJ* 2004;**329**:716–19.
7 **Walley T**, Mantgani A. The UK general practice research database. *Lancet* 1997;**350**:1097–9.
8 **Statistics OfN**. 20th century mortality (England and Wales 1901–2000) (CD ROM). London: UK Government Statistical Service).
9 **Hollowell J**. The general practice research database: quality of morbidity data. *Popul Trends* 1997;**87**:36–40.
10 **McMichael AJ**. Standardized mortality ratios and the ''healthy worker effect'': Scratching beneath the surface. *J Occup Med* 1976;**18**:165–8.