A possible artefactual component in specific cause mortality gradients
Social class variations in the clinical accuracy of death certificates

M L SAMPHIER,1 C ROBERTSON,2 AND M J BLOOR3
From the Lothians Area Health Board, 11 Drumsheugh Gardens, Edinburgh1; Department of Mathematics, University of Strathclyde2; and MRC Medical Sociology Unit, University of Glasgow3

SUMMARY This paper investigates one possible avenue of artefactual influence on the production and/or concealment of social class gradients in specific cause mortality rates, namely, the possibility of social class biases in the accuracy of diagnosis of cause of death and the systematic misallocation of certain social groups to particular diagnoses. Information on this topic was obtained by matching occupational data gathered at death registration with data on the accuracy of diagnosis of cause of death (measured by diagnostic agreement between clinician and pathologist) collected in a prospective study of 1152 hospital necropsies. Extrapolation from these data to national mortality rates should be cautious, but it appears that in the majority of the most common causes of death grouped by ICD chapter (neoplasms, cerebrovascular and digestive) social class gradients would be steeper if mortality data were based on pathologists' rather than clinicians' diagnoses. Only in the respiratory chapter would the gradient be reduced, with the gradient in cardiovascular deaths unaffected.

In recent years there has been increasing interest in the role of artefactual influences in the production of social class differences in specific cause mortality rates. Thus, studies of coronary deaths in status-graded occupations such as the civil service1 and the army2 show much steeper mortality gradients than those found in the national coronary heart disease mortality data.3 This has led to the suggestion that social class differences in mortality rates in the national data may be diluted artefactually by the necessary aggregation of heterogeneous occupations into the same social groupings.

However, there are two potential channels of artefactual influence on class gradients in specific cause mortality rates: one concerns the measurement of social class (as in the above studies); the other concerns the diagnosis of cause of death. It is the second channel of artefactual influence that is our topic in this paper. Specifically, our interest lies in establishing whether there may be social class biases in the accuracy of diagnosis of cause of death: if misdiagnosis occurs differentially across social classes then class gradients in specific cause mortality rates may be diluted or inflated artefactually by the systematic misallocation of certain social groups to particular diagnoses. No previous study of this topic has been undertaken.

That there is a prima facie case for such artefactual influence can be established by evidence from diverse sources. Firstly, there is the evidence from a series of studies, all of which have shown high levels of disagreement between clinicians and pathologists in the diagnosis of the major underlying cause of death in hospital patients sent for necropsy4–7; the high levels of disagreement found offer considerable scope for systematic misallocations to occur between social classes. Secondly, there is epidemiological evidence suggesting that in the recent past working class patients and female patients were disproportionately likely to be diagnosed at death as suffering from "other myocardial degeneration", while middle class and male patients were more likely to receive a more specific diagnosis such as angina pectoris.8 9 Finally, it is well known that systematic variations in the diagnostic and certifying process are an important artefactual influence on international and regional differences in specific cause mortality rates.10–13

Accordingly, the present study was undertaken to examine this possible source of artefactual influence on specific-cause mortality rates by matching two data-sets: data comparing clinicians' and pathologists' diagnoses in respect of a consecutive series of 1152 hospital necropsy cases, and data drawn from death
certificates on the occupations of patients dying in hospital over the same period.

Methods

The methods used to compare clinicians' and pathologists' diagnoses have been reported fully elsewhere2 and will be recapitulated only briefly here.

Clinicians requesting necropsies on hospital patients in the South Lothian District were asked to complete 'dummy' death certificates which were then compared with pathologists' findings. Comparative data were thus available on 80% of all hospital necropsies completed in the district between May 1975 and June 1977 (1152 cases).

Unlike in previous studies of this nature, data were gathered prospectively and the investigators had access to casenotes and were able to discuss cases with the clinicians. Consequently, it was possible to eliminate disagreements due to incorrect ordering of the findings on the certificate, to inadvertent omissions, and to merely semantic differences. Clinical diagnoses which would not be associated with diagnostic morphological features at necropsy (e.g., diabetes mellitus) were accepted without question. Moreover, where a difference in the order of importance of diagnoses on the certificate was simply a matter of opinion, then judgement was given in favour of the clinician.

Despite this elimination of artificial and questionable diagnostic disagreements, the level of disagreements on the major underlying cause of death between clinicians and pathologists remained high at 39% (449 cases). However, some of these disagreements were contained within the same ICD chapter (infective, neoplasms, endocrine, etc). Accordingly, a highly restrictive measure of significant diagnostic disagreement was adopted, namely, that significant disagreements were those where the underlying cause of death was assigned to different chapters—these represented 22% of all cases (252 cases in all). Note that ICD chapter VII, diseases of the circulatory system, has been subdivided into cardiovascular (VIIa) and cerebrovascular (VIIb).14

Where we use the terms 'accuracy' or 'disagreement' in the analysis to follow we refer always to this measure of significant diagnostic disagreement across ICD chapters. The term 'accuracy' is only used where the substitution of the term 'disagreement' would lead to infelicities of expression; we do not mean to imply that pathologists possess a monopoly of diagnostic wisdom.

The allocation of cases to the different diagnostic chapters by clinicians and pathologists is shown in Table 1. Those cases where there was diagnostic agreement, according to our above operational measure, are shown within the diagonal lines in the table.

Information on the occupation of the deceased was obtained from the General Register Office for Scotland in respect of all hospital deaths in South Lothian District during the period of the study and coded according to the 1970 Classification of Occupations. These data represent information supplied by the next-of-kin at death registration and are known to be a more accurate source of occupational information than hospital records. However, no information was available on occupation in respect of 48% of female necropsied patients, and so analysis was restricted to male patients where occupational information was available in all but nine cases, leaving 637 cases in all. As is well known, there are in any case considerable problems associated with the social class analysis of female occupational data.15

Social class was assigned on the basis of the Registrar General's classification of occupations. As

Table 1 Allocation of 637 clinicians' and pathologists' diagnoses to ICD chapters

<table>
<thead>
<tr>
<th>Clinicians' diagnoses</th>
<th>Pathologists' diagnoses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
</tr>
<tr>
<td>I Neoplasms</td>
<td>160</td>
</tr>
<tr>
<td>VIIa Cardiovascular</td>
<td>10</td>
</tr>
<tr>
<td>VIIb Cerebrovascular</td>
<td>4</td>
</tr>
<tr>
<td>VIII Respiratory</td>
<td>10</td>
</tr>
<tr>
<td>IX Digestive</td>
<td>4</td>
</tr>
<tr>
<td>Other ICD chapters*</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>191</td>
</tr>
</tbody>
</table>

*Other ICD chapters' were: I Infections, III Endocrinial, IV Blood, V Mental, VI Nervous system, X Genitourinary, XI Complications of pregnancy, XII Skin, XIII Musculoskeletal, XIV Congenital abnormalities, XV Perinatal, XVI III-defined conditions, XVIII Injury and poisoning.
we stated earlier, the measurement of social class is itself a potential avenue for artefactual influence on mortality gradients, but this is not a topic that we could address in this research. In order to provide sufficient numbers for statistical analysis it was necessary to aggregate some of the Registrar General's classes; this left us with a non-manual category (classes I, II, and IIIa), a skilled manual category (class IIIb), and a semi and unskilled category (classes IV and V), containing 183 cases, 238 cases, and 216 cases respectively.

Using a series of logistic models, analysis sought to explore the associations between significant diagnostic agreement and five independent variables. These five variables were social class, clinician's diagnostic chapter (ie, diagnostic chapter assigned on the basis of the clinician's diagnosis), pathologist's diagnostic chapter (ie, diagnostic chapter assigned on the basis of the pathologist's diagnosis), age of deceased, and length of stay in hospital. All but the first of these variables were derived from earlier analyses of these data, where diagnostic chapter, age of deceased, and length of stay had been shown to be associated with diagnostic disagreement.7

Table 2 Relative effects of the study variables on the proportions of agreement

<table>
<thead>
<tr>
<th></th>
<th>Deviance (G²)</th>
<th>df</th>
<th>Change in deviance df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinicians' diagnostic chapter</td>
<td>614.5</td>
<td>631</td>
<td>46.1</td>
</tr>
<tr>
<td>Pathologists' diagnostic chapter</td>
<td>629.0</td>
<td>631</td>
<td>31.6</td>
</tr>
<tr>
<td>Age of deceased</td>
<td>645.4</td>
<td>633</td>
<td>15.3</td>
</tr>
<tr>
<td>Length of stay</td>
<td>650.5</td>
<td>628</td>
<td>10.1</td>
</tr>
<tr>
<td>Social class</td>
<td>656.7</td>
<td>634</td>
<td>3.9</td>
</tr>
</tbody>
</table>

Table 3 Magnitude and direction of effects of variables on diagnostic agreement within diagnostic chapters

<table>
<thead>
<tr>
<th>Mean</th>
<th>Clinicians' diagnostic chapter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Neoplastic (187 cases)</td>
</tr>
<tr>
<td></td>
<td>Cardiovascular (225 cases)</td>
</tr>
<tr>
<td></td>
<td>Cerebrovascular (51 cases)</td>
</tr>
<tr>
<td></td>
<td>Respiratory (68 cases)</td>
</tr>
<tr>
<td></td>
<td>Digestive (57 cases)</td>
</tr>
<tr>
<td></td>
<td>Other (49 cases)</td>
</tr>
<tr>
<td>Age of deceased</td>
<td></td>
</tr>
<tr>
<td>&lt;55 (96 cases)</td>
<td>0</td>
</tr>
<tr>
<td>55-64 (183 cases)</td>
<td>0</td>
</tr>
<tr>
<td>65-74 (228 cases)</td>
<td>0</td>
</tr>
<tr>
<td>75+ (130 cases)</td>
<td>0</td>
</tr>
<tr>
<td>Social class</td>
<td></td>
</tr>
<tr>
<td>non-manual (183 cases)</td>
<td>0</td>
</tr>
<tr>
<td>skilled manual (238 cases)</td>
<td>0</td>
</tr>
<tr>
<td>semi- and unskilled manual (216 cases)</td>
<td>0</td>
</tr>
<tr>
<td>Length of stay</td>
<td></td>
</tr>
<tr>
<td>&lt;1 (58 cases)</td>
<td>0</td>
</tr>
<tr>
<td>1-3 (154 cases)</td>
<td>0</td>
</tr>
<tr>
<td>4-14 (228 cases)</td>
<td>0</td>
</tr>
<tr>
<td>15+ (147 cases)</td>
<td>0</td>
</tr>
</tbody>
</table>

Results

The application of a series of simple logistic models enabled the assessment of the significance of the effects of the previously mentioned variables. The resultant goodness of fit statistics are presented in table 2. By subtraction from the baseline constant model it is possible to derive an indication of the effect of each variable in turn. It can be seen that the most powerful effects are exerted by the type of disorder, whether measured by 'clinicians' diagnostic chapter' or by 'pathologists' diagnostic chapter'. Of the other three variables, age of the deceased is the only one with a significant effect.

Given the strength of the effect of diagnostic category, subsequent analyses were conducted within these categories, at least where sufficient data were available. We have chosen to represent the data here according to the clinicians' rather than the pathologists' chapter. Our grounds for this decision are, firstly, that the former accounted for more of the diagnostic variability than the latter; and, secondly, that national mortality data (to which most epidemiological interest attaches) are, of course, based overwhelmingly on clinicians' rather than pathologists' diagnoses of cause of death.

In table 3, the parameter estimates are shown for the logistic model (constant + age-group + social class + length of stay) applied within each diagnostic chapter. The first category of each variable is adopted as the reference category, and the estimates for the remaining categories indicate the direction and magnitude of the effect relative to the reference category. The parameter estimates show the concurrent effects of the independent variables. Thus there is decreasing diagnostic agreement with...
A possible artefactual component in specific cause mortality gradients

Table 4 Neoplasms chapter: Relative effects of sociodemographic variables

<table>
<thead>
<tr>
<th>Model</th>
<th>Deviance</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant (C)</td>
<td>149.90</td>
<td>183</td>
</tr>
<tr>
<td>Age (AG)</td>
<td>142.16</td>
<td>180</td>
</tr>
<tr>
<td>Social class (SC)</td>
<td>141.58</td>
<td>181</td>
</tr>
<tr>
<td>Length of stay (LS)</td>
<td>147.91</td>
<td>181</td>
</tr>
</tbody>
</table>

Table 5 Neoplasms chapter: Effects of two-variable interactions

<table>
<thead>
<tr>
<th>Model</th>
<th>Deviance</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) C + AG + SC + LS</td>
<td>134.00</td>
<td>176</td>
</tr>
<tr>
<td>(2) C + AG + SC + LS</td>
<td>118.03</td>
<td>160</td>
</tr>
</tbody>
</table>

increasing patient age—most clearly evident in the cardiovascular chapter, and with increasing length of stay in hospital.

An analysis of the pooled data, excluding the composite category 'other' diagnostic chapters (representing 49 cases in all), confirms the overall pattern and reveals a significant social class effect. With the exception of the 'other' column, the estimates in table 3 indicate that the proportion of skilled manual workers with agreed diagnoses is consistently larger than that of non-manual workers. In three out of the five chapters this is also true for the unskilled group. This unexpected inverse social class gradient in diagnostic accuracy is strongest in the neoplasms ICD chapter ($\chi^2 = 8.3$, 2 df).

Table 4 represents the goodness of fit statistics for the neoplasms chapter and the estimates of the significance of each variable. Social class exerts the strongest effect, with age of patient a close second. Length of stay in hospital is not significant in this chapter (note that only three patients died within one day of admission, two with agreed diagnoses; these cases were omitted from this particular analysis). Further analysis demonstrated that there were no significant two-way interactions between class, age, and length of stay. Model 2 in table 5 is the most complex we are able to consider with three interactive variables. Model 2 does not represent a significant improvement in model 1 ($\chi = 16$, 16 df).

As the main interest that attaches to these data lies in their implications for the accuracy of class gradients in specific cause mortality, it is relevant to represent the degree of over- and under-diagnosis by social class for each of the most populous diagnostic chapters (table 6).

Table 6 Percentage over/under-diagnosis by social class

<table>
<thead>
<tr>
<th>Clinician</th>
<th>Pathologist</th>
<th>Non-manual</th>
<th>Skilled manual</th>
<th>Semi- and unskilled manual</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neoplasms</td>
<td>42</td>
<td>78</td>
<td>67</td>
<td>-7.0%</td>
<td>-2.0%</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>87</td>
<td>73</td>
<td>65</td>
<td>+3.2%</td>
<td>-0.9%</td>
</tr>
<tr>
<td>Cerebrovascular</td>
<td>9</td>
<td>24</td>
<td>18</td>
<td>+64.0%</td>
<td>+59.0%</td>
</tr>
<tr>
<td>Respiratory</td>
<td>13</td>
<td>31</td>
<td>24</td>
<td>+4.3%</td>
<td>+1.5%</td>
</tr>
<tr>
<td>Digestive</td>
<td>23</td>
<td>15</td>
<td>19</td>
<td>-32.2%</td>
<td>-23.0%</td>
</tr>
<tr>
<td>Others</td>
<td>9</td>
<td>17</td>
<td>23</td>
<td>+21.0%</td>
<td>+6.5%</td>
</tr>
<tr>
<td>(N)</td>
<td>183</td>
<td>238</td>
<td>216</td>
<td>637</td>
<td></td>
</tr>
</tbody>
</table>
being considerably underdiagnosed overall; it appears that the extent of such under- and over-diagnosis varies between social classes. In the cardiovascular chapter the extent of under-diagnosis overall is very slight, and this is similarly the case within each social class. The most interesting patterns are set by the respiratory and neoplasms chapters: in both cases there is little over- or under-diagnosis overall, but there is much greater over- and under-diagnosis within individual social classes.

Discussion

As we have already seen, there is some indirect evidence that points to systematic social class differences in the accuracy of diagnoses of cause of death. However, this evidence could be read as indicating that misdiagnosis was greatest among manual workers.9 Likewise, there are few studies of social class variations in the accuracy of diagnosis in other areas of medical practice, but what work there is indicates that misdiagnosis may be inversely related to social class. Thus the ability of general practitioners to detect psychiatric symptoms was found to vary directly with the level of patient education.18 It was therefore surprising, in the present study, to find that while there were indeed social class differences in diagnostic accuracy (as measured by pathologists' agreement with clinicians' diagnoses), the differences were not those expected. Instead, skilled manual workers were the class with the highest level of diagnostic agreement in every one of the most populous diagnostic chapters. In order to investigate the replicability of the analysis an alternative measure of diagnostic agreement was developed, namely, a scored rating of the degree of confirmation of all clinical diagnostic entries on the death certificate. This alternative analysis (not reproduced here) produced comparable social class differences.

Two possible explanations might be advanced for the social class associations found—one artefactual and one clinical—but both appear unlikely in the face of further evidence. A possible artefactual explanation would focus on the differential social selection of hospital deaths for necropsy: it might be hypothesised that—for whatever reasons—middle class patients might be examined post mortem less frequently and only when the clinician was most uncertain about his or her diagnosis. But this possibility must be discounted since the class composition of the necropsy population in the study was no different from the class composition of hospital deaths in the same period.17

A clinician explanation might be advanced if one found a differential class incidence of well diagnosed and poorly diagnosed conditions within each of the most populous diagnostic chapters. For example, the class differences in diagnostic agreement within the neoplasms chapter might be the result of the presence of well-diagnosed cancers that were found disproportionately among manual workers, or (the opposite) a result of the presence of poorly diagnosed cancers that were found disproportionately among non-manuals.

Such an explanation seems unlikely because, although numbers were, of course, very small, the same class differences in diagnostic agreement appeared to hold true at the level of individual disease conditions as well as at the level of diagnostic chapters. Thus 40% of all the clinicians' neoplasms diagnoses were lung cancers; this diagnosis was relatively less common among the non-manual group, but pathologists disagreed with proportionately more of the clinicians' diagnoses in the non-manual group (3 out of 13) than in the manual groups (7 out of 61)—a reflection of class differences in diagnostic agreement found in the neoplasms chapter as a whole.

Furthermore, any clinical explanation advanced would face the difficulty of being a specific explanation for a general phenomenon. Table 3 shows that, although the class differences in diagnostic agreement were statistically significant only in the neoplasms chapter, the same class gradient in agreement was evident in all the most populous diagnostic chapters. A whole series of specific explanations of the differential class incidence of readily diagnosed and/or poorly diagnosed conditions would thus be required.

In short, no straightforward explanation of these social class differences in diagnostic agreement presents itself, and it must be left as a problem for further investigation.

There remains the question of the relevance of these findings for our understanding of social class gradients in specific cause mortality rates. Recent analyses of national mortality data show that, despite a general fall in mortality rates, class inequalities are increasing both in all-cause mortality and in respect of the most common causes of death—coronary heart disease, cancers, and cerebrovascular disease.19 No previous study has investigated whether or not there are systematic social class biases in the diagnosis of cause of death which might affect these class mortality gradients.

Considerable caution should be exercised in any attempt to extrapolate from our findings to national mortality data because of the small numbers, the location of the study in a single area of Scotland, the use of an autopsied (and therefore selected) population, and the inclusion of a retired population who are normally excluded from class analyses of national mortality data because of the problem of numerator/denominator biases. Nevertheless, the attempt at extrapolation should be made since it is difficult to envisage any study being mounted
that could overcome all these problems.

Taking the most populous chapter (cardiovascular) first, there is no indication of substantial inter-class differences in diagnostic agreement and therefore no reason to suppose that the known social class gradients are being inflated or diminished by systematic diagnostic biases. In the case of the neoplasms chapter there is some suggestion of such systematic biases: although the levels of agreement between clinicians and pathologists were quite high overall, there was a tendency for clinicians to overdiagnose relative to the pathologists in the non-manual group and to underdiagnose in both manual groups. Our evidence indicates that, if systematic biases in the diagnosis of cancer play any role at all in respect of social class gradients, it is that of partial concealment—social class gradients might be steeper if they were based on pathologists’ rather than clinicians’ diagnoses. In the remaining chapters all numbers are much reduced, but it appears that cerebrovascular diseases are overdiagnosed by clinicians overall and that this overdiagnosis may be greatest among non-manual workers—once again social class gradients might be steeper if they were based on pathologists’ diagnoses. The end-result is similar in respect of the digestive chapter—digestive diseases are underdiagnosed overall but this underdiagnosis is greatest in the two manual groups, thus serving to steepen class gradients. Only in the respiratory chapter in table 6 is there any suggestion of diagnostic disagreement making for a levelling of social class gradients.

The Black Report on Inequalities in Health suggested three possible explanations for social class mortality gradients—artefact explanations, social selection, and material deprivation. In respect of specific cause mortality gradients, our evidence suggests that the role of artefact explanations (in the shape of systematic diagnostic bias) is a limited one. Although diagnostic agreement does indeed vary with the social class of the patient, the variation is small and in all the major diagnostic chapters except respiratory diseases the effect of correcting such diagnostic biases would be either negligible or to steepen existing class gradients.

MLS and MJB acknowledge the support of the Medical Research Council. We are indebted to the General Register Office for Scotland for supplying us with occupational data, and to Hector Cameron and Phemie McGoogan, of the Department of Pathology, University of Edinburgh, for granting us access to the data from their necropsy study. We also thank Russell Ecob, Kate Hunt, Lindsay Prior, Sally Macintyre, and Graham Watt for their comments on an earlier draft of this paper.

Address for reprints and correspondence: M J Bloor, MRC Medical Sociology Unit, 6 Lilybank Gardens, Glasgow, G12 8QQ

References


Accepted for publication January 1988