Change in health inequalities among British civil servants: the Whitehall II study

J E Ferrie, M J Shipley, G Davey Smith, S A Stansfeld, M G Marmot

Study objective: Despite an overall decline in mortality rates, the social gradient in mortality has increased over the past two decades. However, evidence on trends in morbidity and cardiovascular risk factors indicates that socioeconomic differences are static or narrowing. The objective of this study was to investigate morbidity and cardiovascular risk factors in white collar British civil servants.

Design: Self rated health, longstanding illness, minor psychiatric morbidity (General Health Questionnaire (GHQ) 30 score, GHQ caseness and GHQ depression subscale), cholesterol, diastolic and systolic blood pressure, body mass index, alcohol over the recommended limits, and smoking were collected at baseline screening (1985–88) and twice during follow up (mean length of follow up 5.3 and 11.1 years). Employment grade gradients in these measures at each phase were compared.


Participants: White collar women and men aged 35–55, employed in 20 departments at baseline screening. Analyses included 6770 participants who responded to all three phases.

Results: Employment grade gradients were observed for most measures at second follow up. In general, there was little evidence that employment grade gradients have increased over the 11.1 years of follow up, but marked increases in the gradient were observed for GHQ score (p<0.001) and depression (p=0.05) in both sexes and for cholesterol in men (p=0.01).

Conclusions: There is little evidence of an increase in inequality for most measures of morbidity and cardiovascular risk factors in white collar civil servants over the 11.1 years to 1998. Inequalities have increased significantly for minor psychiatric morbidity in both sexes and for cholesterol in men.
senior executive officer, higher executive officer, executive officer, clerical officer, and clerical assistant. Other professional and technical staff were assigned to these grades on the basis of salary. For analysis, unified grades 1–6 were combined into six categories; category 1 represents the highest grade (Grade 1), and presented these in the tables. Respondents in the lowest (Grade 6) and respondents in the highest grade (Grade 1) were restricted to the 6770 respondents (2013 women and 4757 men) who participated in all three Phases.

At Phases 1 and 3, health over the past year was self rated as very good, good, average, poor or very poor. For the purpose of analysis, reports of health as average, poor or very poor were combined to form the outcome of interest. Additionally at Phase 5 this measure was assessed using the self rated health question from the SF36. In this case, health over the past year self rated as fair or poor was compared with health rated as very good, good, average, poor or very poor. For the purpose of analysis, unified grades 1–6 were combined into six categories; category 1 represents the highest grade (Grade 1), and presented these in the tables.

At the screening examination, blood pressure (mm Hg) was measured twice with the participant seated after a five minute rest, using a Hawksley random-zero sphygmomanometer. Blood was taken and serum cholesterol concentration (mmol/l) measured using the cholesterol oxidase/peroxidase colorimetric method (BCL kit). Weight (wt) in kilograms and height (ht) in metres were recorded. Body mass index (BMI) was calculated from these two measures as wt/ht². Further details of all these measures have been reported previously. 1, 11

Ethical approval
Ethical approval for the Whitehall II study was obtained from the University College London Medical School committee on the ethics of human research.

Study sample and statistical analysis
Of the 10 308 respondents who participated in the baseline screening 8354, 81%, participated at Phase 3 (5.3 years follow up) and 7824, 76%, at Phase 5 (11.1 years follow up). Analyses were restricted to the 6770 respondents (2013 women and 4757 men) who participated in all three Phases.

The aim of the analysis was to compare the employment grade gradient at baseline for each measure of morbidity or cardiovascular risk factor, based on baseline grade, for the same measure at 5.3 and 11.1 years. The analyses produce a regression line, which represents the increase in risk for a unit increase in the employment grade. We have used this employment grade gradient to estimate the odds ratio (OR) or difference (Diff) in risk (and 95% confidence intervals) between respondents in the lowest (Grade 6) and respondents in the highest grade (Grade 1), and presented these in the tables.

---

### Table 1  Age adjusted morbidity prevalence and cardiovascular risk factors (means and percentages) after 11.1 years of follow up by civil service employment grade category—women

<table>
<thead>
<tr>
<th>Health measure</th>
<th>Employment grade category</th>
<th>N</th>
<th>1 (high)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6 (low)</th>
<th>p Value test for trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>self rated health fair or poor (SF36)</td>
<td>2013</td>
<td>10.0%</td>
<td>11.3%</td>
<td>10.8%</td>
<td>14.6%</td>
<td>19.0%</td>
<td>23.6%</td>
<td>p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>longstanding illness</td>
<td>1533</td>
<td>52.3%</td>
<td>55.6%</td>
<td>50.0%</td>
<td>51.6%</td>
<td>58.6%</td>
<td>54.9%</td>
<td>p=0.31</td>
<td></td>
</tr>
<tr>
<td>GHQ score</td>
<td>1950</td>
<td>2.35</td>
<td>3.66</td>
<td>4.16</td>
<td>3.88</td>
<td>4.30</td>
<td>p=0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GHQ caseness</td>
<td>1950</td>
<td>20.0%</td>
<td>26.4%</td>
<td>23.6%</td>
<td>30.9%</td>
<td>25.7%</td>
<td>29.3%</td>
<td>p=0.09</td>
<td></td>
</tr>
<tr>
<td>depression</td>
<td>1959</td>
<td>0.80</td>
<td>1.10</td>
<td>1.30</td>
<td>1.11</td>
<td>1.34</td>
<td>1.48</td>
<td>p=0.001</td>
<td></td>
</tr>
<tr>
<td>cholesterol (mmol/l)</td>
<td>1733</td>
<td>5.85</td>
<td>6.12</td>
<td>5.87</td>
<td>5.96</td>
<td>6.04</td>
<td>5.97</td>
<td>p=0.98</td>
<td></td>
</tr>
<tr>
<td>systolic blood pressure (mm Hg)</td>
<td>1791</td>
<td>116.8</td>
<td>121.1</td>
<td>120.0</td>
<td>122.0</td>
<td>123.0</td>
<td>121.7</td>
<td>p=0.29</td>
<td></td>
</tr>
<tr>
<td>diastolic blood pressure (mm Hg)</td>
<td>1791</td>
<td>72.2</td>
<td>75.1</td>
<td>74.2</td>
<td>74.5</td>
<td>75.6</td>
<td>75.6</td>
<td>p=0.02</td>
<td></td>
</tr>
<tr>
<td>body mass index (kg/m²)</td>
<td>1588</td>
<td>24.7</td>
<td>25.5</td>
<td>26.2</td>
<td>25.9</td>
<td>26.4</td>
<td>27.2</td>
<td>p=0.001</td>
<td></td>
</tr>
</tbody>
</table>

Health related behaviours:
- alcohol (15 units/week)               | 1972                      | 39.1% | 30.5% | 23.3% | 19.5% | 11.1% | 6.8%  | p<0.001 |
- smoking                               | 2041                      | 5.4%  | 8.4%   | 12.3% | 10.4% | 11.8% | 6.4%  | p<0.001 |

% of participants in each grade*:
- 4.5%                                   | 9.8%                      | 7.4%  | 16.3% | 20.5% | 41.4% |

*These percentages vary slightly for each measure.

---

### Table 2  Age adjusted morbidity prevalence and cardiovascular risk factors (means and percentages) after 11.1 years of follow up by civil service employment grade category—men

<table>
<thead>
<tr>
<th>Health measure</th>
<th>Employment grade category</th>
<th>N</th>
<th>1 (high)</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6 (low)</th>
<th>p Value test for trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>self rated health fair or poor (SF36)</td>
<td>4757</td>
<td>8.9%</td>
<td>9.5%</td>
<td>9.0%</td>
<td>11.2%</td>
<td>14.6%</td>
<td>16.6%</td>
<td>21.9%</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>longstanding illness</td>
<td>3683</td>
<td>44.7%</td>
<td>45.8%</td>
<td>46.9%</td>
<td>45.2%</td>
<td>53.8%</td>
<td>54.1%</td>
<td>p=0.007</td>
<td></td>
</tr>
<tr>
<td>GHQ score</td>
<td>4696</td>
<td>2.46</td>
<td>2.71</td>
<td>2.48</td>
<td>2.74</td>
<td>3.59</td>
<td>3.35</td>
<td>p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>GHQ caseness</td>
<td>4696</td>
<td>17.5%</td>
<td>20.2%</td>
<td>18.3%</td>
<td>19.8%</td>
<td>24.3%</td>
<td>21.2%</td>
<td>p=0.03</td>
<td></td>
</tr>
<tr>
<td>depression</td>
<td>4697</td>
<td>0.79</td>
<td>0.84</td>
<td>0.80</td>
<td>0.93</td>
<td>1.36</td>
<td>1.44</td>
<td>p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>cholesterol (mmol/l)</td>
<td>4344</td>
<td>5.87</td>
<td>5.95</td>
<td>5.90</td>
<td>5.91</td>
<td>5.79</td>
<td>5.79</td>
<td>p=0.04</td>
<td></td>
</tr>
<tr>
<td>systolic blood pressure (mm Hg)</td>
<td>4383</td>
<td>122.1</td>
<td>124.0</td>
<td>123.2</td>
<td>123.3</td>
<td>124.8</td>
<td>124.8</td>
<td>p=0.02</td>
<td></td>
</tr>
<tr>
<td>diastolic blood pressure (mm Hg)</td>
<td>4383</td>
<td>77.7</td>
<td>78.4</td>
<td>78.6</td>
<td>78.7</td>
<td>78.9</td>
<td>79.5</td>
<td>p=0.008</td>
<td></td>
</tr>
<tr>
<td>body mass index (kg/m²)</td>
<td>3771</td>
<td>25.8</td>
<td>26.0</td>
<td>26.2</td>
<td>25.9</td>
<td>26.1</td>
<td>26.6</td>
<td>p=0.06</td>
<td></td>
</tr>
</tbody>
</table>

Health related behaviours:
- alcohol (15 units/week)               | 4731                      | 33.4% | 27.9% | 27.4% | 26.4% | 24.8% | 17.0% | p<0.001 |
- smoking                               | 4835                      | 3.8%  | 6.0%   | 7.2%  | 11.2% | 12.9% | 24.4% | p<0.001 |

% of participants in each grade*:
- 15.9%                                  | 25.4%                      | 18.6% | 21.9% | 11.8% | 6.4%  |

*These percentages vary slightly for each measure.
Over the 11.1 years are presented in tables 3 and 4.

Employment grade gradients at baseline, 5.3, and 11.1 years, and a test for change in the gradient over 11.1 years. Employment grade gradients at baseline, 5.3, and 11.1 years, and a test for change in the gradient over 11.1 years—women

<table>
<thead>
<tr>
<th>Health measure</th>
<th>OR or Diff (95% CI)</th>
<th>p Value for change in gradient over 11.1 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>self rated health</td>
<td>2.30 (1.6 to 3.2)***</td>
<td>2.00 (1.4 to 2.8)*** 0.56 (Baseline–5.3 years)</td>
</tr>
<tr>
<td>(from SF36 question)</td>
<td></td>
<td>(5.3 years–11.1 years)</td>
</tr>
<tr>
<td>longstanding illness</td>
<td>1.89 (1.1 to 2.6)***</td>
<td>1.89 (1.1 to 2.6)*** 0.64</td>
</tr>
<tr>
<td>GHQ caseness</td>
<td>0.57 (0.4 to 0.8)**</td>
<td>0.57 (0.4 to 0.8)** 0.01 (Baseline–5.3 years)</td>
</tr>
<tr>
<td>depression</td>
<td>0.18 (0.1 to 0.5)***</td>
<td>0.18 (0.1 to 0.5)*** 0.001 (Baseline–5.3 years)</td>
</tr>
<tr>
<td>cholesterol (mmol/l)</td>
<td>0.09 (0.1 to 0.3)***</td>
<td>0.09 (0.1 to 0.3)*** 0.001 (Baseline–5.3 years)</td>
</tr>
<tr>
<td>diastolic blood pressure (mm Hg)</td>
<td>1.38 (1.2 to 4.0)***</td>
<td>1.38 (1.2 to 4.0)*** 0.001 (Baseline–5.3 years)</td>
</tr>
<tr>
<td>body mass index (kg/m²)</td>
<td>1.82 (0.3 to 3.4)***</td>
<td>1.82 (0.3 to 3.4)*** 0.001 (Baseline–5.3 years)</td>
</tr>
</tbody>
</table>

Test of significance of employment grade gradient: *p<0.05; **p<0.01; ***p<0.001.

Odors ratios above 1.0 and differences above 0.0 indicate an inverse grade gradient, that is, greater risk among the lower grades, while odds ratios below 1.0 and differences below 0.0 indicate a positive grade gradient, that is, greater risk among the higher grades. The p value for the change in the employment grade gradient over the 11.1 years from baseline was determined by calculating the difference between the regression coefficients for the gradient at each data collection point and testing this against a null hypothesis of no difference. Different grade gradients for women and men had already been demonstrated at baseline in the Whitehall II cohort, so analyses were conducted separately by sex. Data were analysed using SAS version 6.12 for Windows.

RESULTS

Tables 1 and 2 show age adjusted morbidity and cardiovascular risk factor figures for each employment grade category after 11.1 years of follow up. Employment grade gradients at baseline, 5.3 and 11.1 years, and a test for change in the gradient over the 11.1 years are presented in tables 3 and 4.

At all three time periods there is consistent evidence of an inverse employment grade gradient for self rated health in both sexes, for BMI in women and for depression and diastolic blood pressure in men (p<0.01). BMI in men is also inversely related to grade, but the strength of the association at baseline and 5.3 years (p<0.01) has diminished by 11.1 years (p<0.05). There is little indication of a gradient in longstanding illness at baseline and 5.3 years, although by 11.1 years an inverse gradient has emerged, p<0.05 in men. In both sexes a positive gradient in GHQ score and caseness at baseline has changed to an inverse association after 11.1 years. Among women these positive baseline gradients are statistically significant (p<0.001), but the strength of the association is diminished after 5.3 years, and by 11.1 years there is evidence of a steep inverse gradient in GHQ score (p<0.01). In men the positive gradients at baseline are weak, but after 11.1 years there is evidence of a steep inverse association with GHQ score (p<0.001) and GHQ caseness (p<0.05). There is some indication of an inverse gradient in depression in women at baseline and 5.3 years, and evidence of a strong inverse association after 11.1 years (p<0.001).

In women there is little indication of a gradient in cholesterol concentration. However, among men an inverse association at baseline has changed to a positive gradient after 11.1 years (p<0.05). There is also evidence of an inverse
gradient in diastolic pressure in women at 5.3 and 11.1 years, and in systolic pressure at 11.1 years in men (p<0.05). In both sexes and all time periods there is a positive employment grade gradient in alcohol over the recommended limits (p≤0.01) and a strong inverse gradient for smoking (p<0.001).

Tests of difference between the grade gradient at baseline and 11.1 years for each measure show that inequality in the distribution of morbidity and cardiovascular risk factors between participants in the highest and lowest grades has remained fairly stable, although there is some indication of a tendency to widen. Furthermore, there is evidence of a marked increase in the gradient over 11.1 years for GHQ score (p<0.001), GHQ caseness (p<0.005) see figure 1, and depression (p=0.05) in both sexes, and for cholesterol in men (p=0.01). There is also an indication of an emerging inverse grade gradient in longstanding illness in women (p=0.06).

Examination of the raw data from all three phases shows that the emerging gradient in longstanding illness in women is attributable to a greater increase in morbidity in the lower grades, while change in the cholesterol gradient in men from an inverse to a positive association is attributable largely to a decrease in cholesterol concentration in the lowest grade. In contrast, changes in the gradient in GHQ score, GHQ caseness, and depression are explained by considerable improvements in the higher grades and deterioration in the lower grades, mostly grade 6.

Comparison of employment grade gradients at baseline for those included and those excluded from the analyses showed no statistically significant difference for any measure in women and for most measures in men. Exceptions among men were depression (p=0.05) and cholesterol (p=0.03). The gradient for cholesterol in both groups was non-significant, but was inverse among men included in the analyses and positive among excluded men. There was also a considerably steeper gradient in longstanding illness among the excluded men (p=0.07) (tables available on request).

DISCUSSION

Steep employment grade gradients in measures of morbidity and cardiovascular risk factors have persisted over the 11.1 years follow up of the Whitehall II study. Furthermore, the gap between those in the lowest employment category and those in the highest has widened considerably for measures of minor psychiatric morbidity in both sexes and cholesterol concentration in men.

Methodological considerations

The strengths of this study are that three waves of data are available on the same individuals over a period of 11.1 years, the dataset contains a wide range of measures and, apart from self rated health, the measures used are the same at each data collection point. Of the few recent studies of morbidity and cardiovascular risk factor trends, none has longitudinal cohort data. However, the disadvantage of prospective data from the same people is that effects could merely reflect age differences over time. To examine this we repeated the analyses in a subpopulation of those respondents aged 45–54 at each data collection phase. These analyses showed trends at baseline, 5.3 and 11.1 years were similar to those found in the whole cohort (tables available on request). The other limitation of a cohort aged 35–55 at baseline and almost exclusively white collar is that findings may not apply to wider populations.

Comparison of baseline gradients for participants included and those excluded from the analyses indicate that, in general, gradients are steeper at baseline among excluded participants. In several cases, such as longstanding illness and systolic blood pressure in men, this loss to follow up explains the difference between the non-significant gradient at baseline seen in these data and the statistically significant gradient reported by us previously. However, evidence of differences between those included and those excluded from further analysis, with the exception of depression and cholesterol in men, is weak. In both sexes there was no significant interaction between grade and health in the likelihood of being excluded from the analyses.

Use of employment grade at baseline to determine the gradients means our measure of socioeconomic position is not contemporaneous with our morbidity and cardiovascular risk factor measures at 5.3 and 11.1 years. Baseline grade was chosen as it is available for all participants, and analyses using last known grade produced findings little different from those presented (data not shown). However, it was felt that use of last known grade as the measure of socioeconomic position posed problems in that it reflects mobility for those who remained in the civil service, but cuts short the trajectories of those who have retired early or left the civil service to take up employment elsewhere, 54% of participants by the 11.1 year follow up. Use of baseline grade also minimises the effect of reverse causality where the levels of morbidity at baseline may effect subsequent mobility and hence grade at 5.3 and 11.1 years.

Consistency with other studies

Most recent studies of risk factors for cardiovascular disease report that socioeconomic differences have remained unchanged, with the exception of an increase in the differential in smoking. However, a study among men, which included morbidity in addition to cardiovascular risk factors, showed that the degree of inequality for most measures tended to narrow between 1984 and 1993. Most of the changes were non-significant, but the decrease in inequality in diastolic blood pressure and GHQ caseness was marked (p<0.01). The latter findings are in direct contrast with our own, but this is possibly because men who participated in the study were aged 20–64 and the data come from two unrelated cross sectional studies.

Of particular note in our findings is the shift from a significant positive gradient in GHQ score for both sexes at baseline to a significant inverse gradient 11.1 years later, and the emergence of a significant positive gradient in cholesterol in men. Cholesterol concentration in other studies shows an inverse relation with education which has persisted over time, although there is some evidence of a non-significant positive association with income among men in the Minnesota Heart Survey. We hypothesised that the emergence of a significant, positive grade gradient in cholesterol concentration after 11.1
years could be attributable to low cholesterol associated with ill health among participants in the lower grades. This would have been absent at baseline because of the healthy worker effect. Analyses comparing cholesterol concentrations for men with and without longstanding illness indicated that cholesterol was slightly lower among men with longstanding illness at 11.1 years (p=0.24), but analyses of self rated health showed those in poorer health to have slightly higher cholesterol.

It is difficult to explain why there should have been such improvement in levels of minor psychiatric morbidity in the three highest employment grades. A validation study at baseline showed that the positive gradient in GHQ score in the Whitehall II cohort, which ran counter to most evidence in the literature, was attributable to under-reporting of minor psychiatric disorder on the GHQ amongst civil servants in the lower employment grades. Possibly there was also some over-reporting in the higher grades.

Generally, with increasing age, levels of psychological distress might be expected to diminish in this cohort. This might explain the decline in minor psychiatric morbidity in the higher grades. At the same time, new physical illness might develop with age, particularly affecting the lower employment grades. The increase in minor psychiatric morbidity in the lower grades might thus be subsequent to the development of physical illness. A further contributor to the increase in minor psychiatric morbidity among the lower grades at 11.1 years might have been the growing insecurity and deterioration in conditions of employment and income as participants from these grades were hived off to private sector employers or were shed from the workforce. There are two other potential explanations. It is possible there was some selective early retirement among higher employment grades with minor psychiatric morbidity after 5.3 years that led to an apparent improvement in minor psychiatric morbidity after 11.1 years. Early retirement was more of an option in higher grade employees. The other, less plausible explanation is a deteriorating working conditions and the increased likelihood of physical illness in the lower grades seem the most plausible and demand further analysis in the future.

ACKNOWLEDGEMENTS
The Whitehall II study has been supported by grants from the Medical Research Council; British Heart Foundation; Health and Safety Executive; Department of Health; National Heart Lung and Blood Institute (HL36310); US; NIH; National Institute on Aging (AG13196); US; NIH; Agency for Health Care Policy Research (HS06516); and the John D and Catherine T MacArthur Foundation Research Networks on Successful Midlife Development and Socioeconomic Status and Health. JP was supported by the Economic and Social Research Council (L128253046) during the preparation of this paper. MM is supported by an MRC Research Professorship. MS is supported by a grant from the British Heart Foundation.

We thank all participating Civil Service departments and their welfare, personnel, and establishment officers; the Occupational Health and Safety Agency; the Council of Civil Service Unions; all participating civil servants in the Whitehall II study; and all members of the Whitehall II study team. The authors would like to thank Jenny Head for comments on an early draft of the paper.

Contributors
Jane Ferrie wrote the original and successive drafts of the paper and was involved in the data collection for Phase 3. Martin Shipley performed the analyses and advised on drafts of the paper. George Davey Smith was involved with the data collection at Phase 1 and commented on all drafts of the paper. Stephen Stansfeld was involved with the data collection and commented on drafts of the paper. Michael Marmot designed and directs the Whitehall II study and commented on drafts of the paper.

Authors’ affiliations
J E Ferrie, M J Shipley, M G Marmot, International Centre for Health and Society, Department of Epidemiology and Public Health, University College London Medical School, London, UK
G Davey Smith, Department of Social Medicine, University of Bristol, Bristol, UK
S A Stansfeld, Department of Psychiatry, Queen Mary University of London, London, UK

Conflicts of interest: none.

REFERENCES