Intercensal denominators – feasibility of using the Labour Force Survey to estimate mortality rates by NS-SEC

Brian Johnson and Ann Langford *Office for National Statistics*

Abstract

Background

The purpose of this analysis is to investigate the feasibility of using the Labour Force Survey (LFS) to provide population denominators for the estimation of mortality rates by the National Statistics Socio-economic Classification (NS-SEC). This is in order to contribute to the monitoring of health inequalities over time. Currently it is possible to produce mortality rates only at the time of a census when populations are enumerated by occupation and NS-SEC.

Methods

The Labour Force Survey (LFS) was used to provide population denominators by age, sex and NS-SEC for various time periods. Numerators for mortality rates were derived by obtaining aggregate deaths by age group and NS-SEC analytic class for men aged 25–64. Standardised mortality rates were then derived. Results for 2001–03 were compared with published estimates which used 2001 Census-based denominators.

Results

Significant differences were detected for most NS-SEC classes between mortality rate estimates, based on the 2001 Census, and those based on the contemporary LFS dataset. However the results for different years, based on LFS denominators, suggest that a series of mortality rates using LFS-based denominators appeared to have internal consistency. The LFS-based estimates showed a statistically significant decrease in mortality rates between the periods 2001–03 and 2005–07, across all but one NS-SEC classes. The precision of LFS-based estimates at different levels of geography was also assessed in order to determine the feasibility of using these estimates to monitor inequalities at sub-national level.

Conclusions

The LFS can be used to produce regular population denominators for the estimation of mortality rates, to assess health inequalities by NS-SEC for men at the level of England & Wales, both on an annual basis and over three-year time periods. It is recommended that, in order for the LFS-based estimates to be effective over time, they would have to be related to each other rather than to the census-based ones. At the level of Government Office Region (GOR), it would also be possible to do this, but on a three-year period basis only. Below this level of geography, estimates would be insufficiently precise. Information exists on the LFS to produce a similar analysis for women.

Contents

Abstract.....	3
Introduction	6
Background.....	7
Methods	8
Results	11
Discussion.....	19
Limitations of the analysis	21
Conclusions and Recommendations	21
References.....	22
Annex.....	24

List of Tables

Table 1	LFS based estimates of percentage distribution of population by NS-SEC men aged 25–64.....	11
Table 2	Mortality rate estimates by NS-SEC and confidence intervals using published census-based population denominators and LFS populations for 2001.....	13
Table 3	Typical confidence intervals for mortality rates of routine occupations class at different levels of geography, using the LFS for population denominators.....	18
Table A1	Example showing method for the normalisation of the LFS population counts by NS-SEC to the ONS mid-year population 2001 and the 'health selection' transfer from the residual groups to the NS-SEC analytic classes, for methodological consistency with the published results using census-based denominators.....	25
Table A2	Example showing resultant estimates of populations by NS-SEC and age for 2001, based on LFS counts 2001 mid-year population and the 'health selection' transfer from the unallocated groups to the NS-SEC analytic classes	27

List of Figures

Figure 1	Mortality rates by NS–SEC, denominators based on Census and LFS(2001) estimates, men aged 25–64, 2001–03	12
Figure 2	Mortality rates by NS–SEC, denominators based on Census and LFS(2002) estimates, men aged 25–64 2001–03	14
Figure 3	Mortality rates by NS-SEC, denominators based on Census and LFS(2003) estimates, men aged 25–64, 2001–03	15
Figure 4	Mortality rates by NS-SEC, denominators based on Census and LFS(2001–03) estimates, men aged 25–64 2001–03	16
Figure 5	Mortality rates by NS-SEC, LFS(2001–03) and LFS(2005–07) based estimates, men aged 25–64.....	17
Figure 6	Approximate Confidence Intervals for Mortality Rates by NS-SEC based on a three-year period for different levels of geography	18

Introduction

For several decades, the measurement of inequalities in mortality by socio-economic class in England and Wales focused on estimated mortality rates around the time of the decennial census. The reason for this is that mortality rates require counts of deaths as numerators, and of relevant population counts as denominators. Deaths are routinely coded by occupation of the deceased, thus allowing counts of deaths by socio-economic class. The Office for National Statistics (ONS) compiles annual (mid-year) population estimates for England and Wales based on the latest census, updated to reflect subsequent births, deaths, migration and ageing. However, counts of the total population broken down by occupation and socio-economic class are available only at the census.

Thus through time, it is possible to compare decennial 'snapshots' of mortality by class. However, it has not been possible to produce a continuous annual analysis of mortality rates by occupation-based class, nor has it been possible to produce mortality rates by class at regional or local level, except at the time of the census.

With the increasing demand for health inequality measurement, there has been interest in the development of a continuous intercensal measurement system. The Labour Force Survey (LFS) is one possible source of data for intercensal population denominators classified by socio-economic class. Unlike the census, the LFS covers less than 1 per cent of the population, but is produced quarterly and has all the detail necessary to estimate mortality rates by class. It also has geographic identifiers allowing sub-national analysis to be undertaken.

As a result, it was decided that a study should be undertaken to determine how practical it would be to use the LFS to produce population denominators by the National Statistics Socio-economic Classification (NS-SEC) and to use them, in combination with death registrations, to produce mortality rates by NS-SEC for people of working age. The resultant mortality rates were then compared with published estimates which were based on denominators derived using the census, and are the only currently available estimates of mortality by NS-SEC. Investigations were also undertaken to examine how accurate these estimates were likely to be at sub-national levels of geography.

Therefore the aims of this article are:

1. To compare the mortality rates for men aged 25–64 in England and Wales using the 2001 Census, and the Spring Quarter 2001 LFS sample (where the data for both were collected at a similar time), and identify any substantial differences between the estimates using the two sources
2. To compare mortality rates produced using populations derived respectively from the 2001, 2002 and 2003 annual LFS datasets with rates derived using the 2001 Census, in order to see whether differences between rates derived from the two sources appeared to be random or patterned by class

3. To extend the estimates using LFS-based denominators to the middle of the decade to see how measured mortality rates by class changed over time
4. To assess the accuracy of estimates at lower levels of geography

At this stage, the detailed analysis was restricted to men, although tests were carried out to determine whether such an analysis would be possible for women.

Background

The Public Service Agreement (PSA) life expectancy target¹ in England is based on the difference in life expectancy at birth between the areas with the worst health and deprivation indicators (the Spearhead group) and the England average. This is a target based on socio-economic inequalities, but measured in terms of geographic areas. This is because it is relatively straightforward to compute annual age-specific mortality rates for administrative areas, compared to annual mortality rates by occupation-based class.

The work of the post-2010 strategic review of health inequalities (the Marmot Review)², commissioned by the Department of Health, has drawn attention to the need for new measures of health inequalities, building on past experience including the use of current PSA targets. Consequently, there is likely to be interest in new measures of inequality which are continuously measurable and based on individual characteristics as well as geography, such as mortality rates by socio-economic classification.

The only estimates of mortality rates by NS-SEC for men aged 25–64 currently available were published by ONS in Health Statistics Quarterly 36⁶. These rates apply to the period 2001–03 and were estimated using the 2001 Census to derive the denominators. These estimates cannot be updated by this method until the next decennial census.

This article explores the possibility of using the LFS as an alternative estimate of population denominators by NS-SEC. The LFS, which is designed to provide information on the UK labour market, collects detailed data on occupation for a sample of approximately 60,000 households in the UK each quarter. Respondents are interviewed five times, at 13-week intervals. Interviews are carried out mostly by telephone, although the first interview is face-to-face. LFS data are available at national, regional and local authority levels. Results from the survey are then weighted to ONS estimates of private household population. There is an annual LFS dataset which includes the quarterly data plus a boost of 39,000 households for England, 14,000 for Wales and 17,000 for Scotland³.

The LFS and the census differ in several ways. The most important of these for the current purpose are:

- The LFS is a sample, while the census attempts to enumerate the whole population.
- The LFS is not compulsory – respondents are able to opt out. The weighting system is designed to compensate for the difference in response rate from different population groups.
- The occupation coding process differs. The LFS has an interviewer prompting for information, whereas the census usually does not.

- Communal establishments are not included in the LFS. However, since deaths at communal establishments were only about 2 per cent of all deaths for men under age 65 and there was no identifiable pattern among analytic NS-SEC classes in the proportion of deaths which were at communal establishments, it can be assumed that this will not materially affect the results. Given that these differences exist, this article compares mortality rate estimates computed using these two sources.

Methods

This section describes the NS-SEC in more detail, specifies the data sources used in the analysis and describes the analytical approach adopted in this article.

The National Statistics Socio-economic Classification

The conceptual basis for the NS-SEC is the structure of employment relations operating in modern developed economies⁷. Occupations are differentiated in terms of reward mechanisms, promotion prospects, autonomy and job security. The most advantaged NS-SEC groups (higher managerial and professional occupations), typically exhibit personalised reward structures, have good opportunities for advancement, relatively high levels of autonomy within the job, and are relatively secure. These attributes tend to be reversed for the most disadvantaged group (routine occupations). Box 1 shows the NS-SEC analytic class breakdowns used in this analysis, and provides examples of the occupations included in each class.

An NS-SEC category can be derived from an individual's occupation and employment status and the size of their organisation. Since size of organisation is not collected on the death register, a version of NS-SEC is used which is derived from occupation and employment status alone. This is known as 'reduced NS-SEC' and differs in terms of its typical distribution among NS-SEC classes by less than three per cent⁸. Reduced NS-SEC is used throughout the following analysis.

All people who can be assigned an occupation code and have sufficient information on employment status can be assigned to an NS-SEC category. These are known as 'functional' categories⁷. Aggregated versions of these such as that shown in Box 1 are known as analytic classes. For those who cannot be assigned an occupation based 'functional' NS-SEC category, there are certain other categories, such as 'full-time student', 'never worked' 'long-term unemployed' and 'inadequately described' These are known as 'residual' categories.

Data Sources for analysis

The following data sources were used in the analysis;

- Counts of deaths by NS-SEC, extracted from death registration data for the relevant periods
- The LFS-based populations derived from the weighted LFS datasets
- ONS mid-year population estimates by age group, used in the analysis as control totals
- The census-based mortality rate estimates, taken from the Health Statistics Quarterly article reporting male mortality rates by NS-SEC for the period 2001-03⁶. These estimates were used for comparison purposes

Box 1 National Statistics Socio-Economic Classification Analytic classes

Analytic class	Examples of occupations included
1.1 Large employers and higher managerial	Senior officials in national and local government; directors and chief executives of major organisations; officers in the armed forces
1.2 Higher professional	Civil engineers, medical practitioners, physicists, geologists, IT strategy and planning professionals, legal professionals, architects
2 Lower managerial and professional	Teachers in primary and secondary schools, quantity surveyors, public service administrative professionals, social workers, nurses, IT technicians
3 Intermediate	NCOs and other ranks in the Armed Forces, graphic designers, medical and dental technicians, local government clerical officers, counter clerks
4 Small employers and own account workers	Hairdressing and beauty salon proprietors, shopkeepers, dispensing opticians in private practice, farmers, self-employed taxi drivers
5 Lower supervisory and technical	Bakers and flour confectioners, screen-printers, plumbers, electricians and motor mechanics employed by others, gardeners, rail transport operatives
6 Semi-routine	Pest control officers, clothing cutters, traffic wardens, scaffolders, assemblers of vehicles, farm workers, veterinary nurses and assistants, shelf fillers
7 Routine	Hairdressing employees, floral arrangers, sewing machinists, van, bus and coach drivers, labourers, hotel porters, bar staff, cleaners and domestics, road sweepers, car park attendants

NS-SEC User Manual, Office for National Statistics⁸

Analytic approach

The first two aims of the study are concerned with the comparison of mortality rate estimates, based on LFS population denominators, with estimates of census-based mortality rates published in Health Statistics Quarterly 36⁶ (HSQ36). The latter were produced using 'optimised' population estimates by NS-SEC, which were derived from the 2001 Census and ONS mid-year population estimates and adjusted for the health related under-classification of men to occupations at the census. These adjustments⁶, referred to as health selection adjustments, were calculated using the ONS Longitudinal Study (LS).

In this study, estimates of mortality rates by NS-SEC were produced using the same figures for deaths by age (numerators) that were extracted from the death registers and used in HSQ36. The population figures (denominators) were obtained from the LFS, standardised by age to ONS 2001–03 mid-year population estimates, and adjusted for potential health selection using the same proportional adjustments as were used for the census-based denominators in HSQ36. The age-specific rates for each NS-SEC class were then standardised to the European standard population to produce age-standardised mortality rates for each class.

The method for obtaining the LFS population counts by age and NS-SEC is described in more detail in the Annex.

Outcome measures

Sets of mortality rate estimates were produced based on LFS datasets for the following periods:

- Spring Quarter 2001
- Annual 2001
- Annual 2002
- Annual 2003

From these it was possible to determine;

- The proximity of the Spring Quarter 2001 LFS and Census estimates of mortality rates by NS-SEC for 2001.
- The amount of variation between the three sets of mortality rate estimates based on three individual years of LFS denominator populations. For example, is the mortality rate for NS-SEC class X consistently over or under estimated, or do the errors appear to be random?

Mortality rates for 2005–07 were calculated using deaths registered in 2005–07 and the LFS population estimates for 2005–07. This enabled a comparison over time with figures for 2001–03.

Approximate estimates of the variances of the mortality rates were calculated taking into account the variance of the death counts, the sampling variance of the LFS population estimates, and the additional variation associated with the health selection adjustment. Confidence intervals and standard statistical significance tests were conducted assuming normality of the estimated rates.

In order to assess the level of precision of estimates at local levels of geography, both the number of deaths in a period and the corresponding LFS population estimates were assumed to be proportionately smaller than those for England and Wales, and approximate variances and hence confidence intervals were recalculated on this basis. For instance England and Wales is made up

of the nine Government Office Regions of England, as well as Wales. For this purpose, a typical region was therefore assumed to be one-tenth of the size of England and Wales, having one-tenth of the deaths and one-tenth of the population of England and Wales. These calculations were made both for the annual 2001 estimates, and the aggregate 2001–03 estimates.

Results

NS-SEC population distributions using the census and the LFS

The population percentages, by NS-SEC for males aged 25–64, using the Spring Quarter 2001 and annual 2001 LFS data are shown in Table 1, and are compared with the raw population percentages by NS-SEC from the 2001 Census and the optimised population percentages published in HSQ36⁶. The data for the census and the LFS Spring 2001 survey would have been collected at a similar time, and could therefore be assumed to relate to the same period.

Table 1 **LFS based estimates of percentage distribution of population by NS-SEC¹ men aged 25–64**

England and Wales				Percentages
NS-SEC analytic class	LFS Spring Quarter 2001	LFS Annual 2001	Census ²	Optimised ³
1.1 Higher managerial	8.7	8.8	7.5	7.7
1.2 Higher professional	10.5	10.2	9.1	9.4
2 Lower managerial and professional	21.0	21.2	21.1	22.2
3 Intermediate	5.3	5.2	5.2	5.9
4 Self-employed and own-account workers	12.2	12.6	13.4	13.9
5 Lower supervisory	13.7	13.7	12.2	13.6
6 Semi-routine	9.7	9.6	9.2	10.7
7 Routine	12.4	11.9	12.0	14.4
Full Time Students	1.0	0.9	1.1	0.2
Others	5.6	5.8	9.3	2.1
Total	100	100	100	100

1 Reduced derivation

2 Census based estimates before adjustment from White C, Glickman G, Johnson B and Corbin T (2007) 'Social Inequalities in adult male mortality by NS-SEC, England and Wales, 2001–03', Health Statistics Quarterly 36⁶

3 Optimised population estimates from White C, Glickman G, Johnson B and Corbin T (2007) 'Social Inequalities in adult male mortality by NS-SEC, England and Wales, 2001–03', Health Statistics Quarterly, 36⁶

The LFS classified approximately 6.7 per cent of men to residual NS-SEC classes, compared to 10.4 per cent for the census and only 2.3 per cent for optimised person-years at risk. The difference relative to the census is partly a result of the coverage of the LFS excluding communal establishments and partly of the higher propensity of the LFS to assign people an occupation. The difference relative to the optimised person-years at risk is partly a result of the method used in the HSQ36 analysis which increased the proportion of the population assigned to analytical NS-SEC

classes, by using a 'health selection adjustment'. This adjustment was calculated using ONS Longitudinal Study information on 1991 occupation for those with no occupation recorded at the 2001 Census.

Comparison of the percentages in each NS-SEC analytic class for the LFS Spring Quarter of 2001 and for the whole year suggests very little difference between the two. Only the routine class demonstrated differences that were statistically significant at the 5 per cent level. As a result, LFS annual rather than quarterly data were used for the remainder of this study.

Mortality rates using LFS population denominators

The mortality rates for 2001 using the LFS, adjusted as described above, are compared with the published mortality rates in Figure 1.

Figure 1 **Mortality rates by NS-SEC¹, denominators based on Census² and LFS(2001)³ estimates, men aged 25–64, 2001–03**

England and Wales

Rate per 100,000 person years

1 Reduced derivation

2 From White C, Glickman G, Johnson B and Corbin T (2007), Social Inequalities in adult male mortality by NS-SEC, England and Wales, 2001–03, Health Statistics Quarterly, 36⁶

3 Calculated from LFS estimated populations 2001, mid year census estimates 2001–03, death registrations 2001–03

It can be seen from Figure 1 that the pattern for the estimates based on annual LFS denominators follows the published (census-based) pattern fairly closely. The LFS-based estimates were lower for NS-SEC class 1.1, 1.2 and 6, and higher for NS-SEC class 4. Although the LFS-based

estimates were for most classes within a 6 per cent range of the census-based estimates, for all but one class the two estimates were statistically significantly different (Table 2).

Table 2 **Mortality rate estimates by NS-SEC¹ and confidence intervals using published census-based population denominators² and LFS³ populations for 2001**

England and Wales							Rate per 100,000
NS-SEC analytic class	Published census based estimates			LFS 2001 estimates			Percentage difference ⁴
	estimate	LCL	UCL	estimate	LCL	UCL	
1.1 Higher managerial	182	177	187	168	162	175	-7 **
1.2 Higher professional	206	202	211	187	181	193	-9 **
2 Lower managerial and professional	259	256	262	269	263	274	4 **
3 Intermediate	286	279	294	285	274	296	0
4 Self-employed and own-account	307	303	312	342	335	349	11 **
5 Lower supervisory	374	369	379	363	355	371	-3 *
6 Semi-routine	473	466	479	446	436	456	-6 **
7 Routine	513	508	519	533	523	543	4 **

1 Reduced derivation

2 From White C, Glickman G, Johnson B and Corbin T (2007), Social Inequalities in adult male mortality by NS-SEC, England and Wales, 2001–03, Health Statistics Quarterly, 36⁶

3 Calculated from LFS estimated populations 2001, mid year census estimates 2001–03, death registrations 2001–03.

4 Difference between the two estimates expressed as a percentage of the census based estimate.

* indicates a statistically significant difference at the 5% level, ** indicates a statistically significant difference at the 1% level. (Statistical significance of the difference was tested using the standard errors of both estimates and assuming independence and normality)

The next stage was to compare results for 2002 and 2003 to see whether the 2001 pattern was repeated for particular classes.

Mortality rates for 2002 were calculated with identical numbers for deaths (2001–03) but with the population denominators estimated using the annual LFS dataset for 2002, together with the ONS mid-year population estimates for each age group. The same adjustments were made to the LFS populations as for 2001, to reflect the approach used for the census-based estimates.

Once again, mortality rates using census-based denominators were underestimated by the rates with LFS denominators for classes 1.1, 1.2 and 6, and overestimated for class 4 (Figure 2).

Figure 2 **Mortality rates by NS-SEC¹, denominators based on Census² and LFS(2002)³ estimates, men aged 25–64, 2001–03**

England and Wales

Rate per 100,000 person years

1 Reduced derivation

2 From White C, Glickman G, Johnson B and Corbin T (2007), Social Inequalities in adult male mortality by NS-SEC, England and Wales, 2001–03, Health Statistics Quarterly, 36⁶

3 Calculated from LFS estimated populations 2002, mid year census estimates 2001–03, death registrations 2001–03

A similar pattern was shown for 2003 (Figure 3), although the overestimate of mortality rate for the self-employed class and the underestimate of the rate for the semi-routine class were both smaller than for 2001 and 2002.

It seems from this evidence that there is a systematic pattern for the LFS-based estimates to underestimate mortality for the higher managerial and professional classes and, to a lesser extent, the semi-routine class. There also appears to be a systematic overestimate for the self-employed relative to the census-based estimates.

Estimates using aggregate LFS population estimates for the whole period 2001–03 are shown in Figure 4, where the pattern observed above is repeated.

Figure 3 Mortality rates by NS-SEC¹, denominators based on Census² and LFS(2003)³ estimates, men aged 25–64, 2001–03

England and Wales

Rate per 100,000 person years

1 Reduced derivation

2 From White C, Glickman G, Johnson B and Corbin T (2007), Social Inequalities in adult male mortality by NS-SEC, England and Wales, 2001–03, Health Statistics Quarterly, 36⁶

3 Calculated from LFS estimated populations 2003, mid year census estimates 2001–03, death registrations 2001–03

Figure 4 **Mortality rates by NS-SEC¹, denominators based on Census² and LFS(2001–03)³ estimates, men aged 25–64 2001–03**

England and Wales

Rate per 100,000 person years

1 Reduced derivation

2 From White C, Glickman G, Johnson B and Corbin T (2007), Social Inequalities in adult male mortality by NS-SEC, England and Wales, 2001–03, Health Statistics Quarterly, 36⁶

3 Calculated from LFS estimated populations 2001–03, mid year census estimates 2001–03, death registrations 2001–03

Estimates for 2005–07 are shown in Figure 5 and compared with those for 2001–03.

It appears from Figure 5 that there was a general reduction in the mortality rate between the period 2001–03 and 2005–07. There appears to have been a statistically significant fall in mortality rates in most classes, at the 5 per cent level, but no discernible change for the 'intermediate' class. (This is a relatively small group and has wide confidence intervals).

Figure 5 Mortality rates by NS-SEC¹, LFS(2001–03)² and LFS(2005–07)³ based estimates, men aged 25–64

England and Wales

Rate per 100,000 person years

1 Reduced derivation

2 Calculated from LFS estimated populations 2001–03, mid year census estimates 2001–03, death registrations 2001–03

3 Calculated from LFS estimated populations 2005–07, mid year census estimates 2005–07, death registrations 2005–07

Accuracies for sub-national geographies

Approximate estimates of the variances were calculated for lower levels of geography, and confidence intervals are displayed in Table 3 and Figure 6.

Table 3 Typical confidence intervals for mortality rates of routine occupations class at different levels of geography, using the LFS for population denominators

Geographical area	Estimated mortality rates	Rate per 100,000			
		Three Year Period confidence interval		One Year Period confidence interval	
		+/- number	+/- per cent	+/- number	+/- per cent
England & Wales	542	8	1	10	2
Region	542	24	4	32	6
County	542	59	11	77	14
Unitary authority	542	138	26	183	34

Figure 6 Approximate confidence intervals for mortality rates by NSSEC based on a three-year period for different levels of geography

From Table 3 it can be seen that a typical confidence interval for the mortality rate of those in routine occupations, using a three-year data period, is approximately 1 per cent of the estimate, that is +/-8 deaths per 100,000. At a regional level the corresponding intervals are larger, and larger still at lower levels of geography. Figure 6 demonstrates how the confidence limits increase for all the analytic classes as area size decreases.

Discussion

The purpose of this study was to assess whether it is feasible to use LFS to produce intercensal denominators which are sufficiently stable and consistent over time to be used to monitor inequalities in mortality rates. In order to achieve this, the first test was to compare mortality rate estimates using LFS-based denominators with published ones using the 2001 Census.

The census and the LFS are known to differ in their population estimates for different groups. An article comparing population estimates from the April 2001 Census with the Spring Quarter 2001 LFS sample⁴ found the following differences at the level of the UK:

- The LFS estimates suggested 2 per cent more economically active people than the census
- The LFS estimates suggested 7 per cent fewer economically inactive people than the census
- The LFS and the Census differed by 14 per cent (203,000) in the estimate of the number of people unemployed

The same article states that as ‘The LFS covers all people in private households, but only covers certain types of communal establishments such as student halls of residence and NHS accommodation but not nursing homes, boarding houses and prisons... Given that most groups of people missing from the population estimates used to weight the LFS are likely to have higher proportions of economically inactive, this difference is more likely to affect the estimates of economic inactivity more than those for economic activity’⁴. The authors also found that estimates of the population by major occupation group for the census and the LFS were similar, but differed by as much as 11 per cent for certain groups.

While the results reported in this article are not directly comparable with those above, a similar pattern was found, with the LFS having a lower total population count within the age range studied but having a higher proportion capable of being assigned to a socioeconomic class. Thus there are known differences between census and LFS population estimates that will affect mortality rates calculated from the two sources.

The test of proximity of census and LFS-based mortality rate estimates was not passed. The estimates were statistically significantly different to the published census-based estimates for all but one NS-SEC class and the differences appeared to be systematic. The LFS estimates for 2001–03 consistently underestimated mortality for higher managers and professionals and overestimated mortality for the self-employed relative to those based on the census.

One possible explanation for this is the process of obtaining the LFS responses. Census respondents do not have an interviewer to explain the questions whereas LFS respondents do. This may lead to a number of census responses indicating self-employment when, had they been mediated by an LFS interviewer, they would have been classified as an employee. If true, this would produce a systematically higher estimate of the self-employed population in the census relative to the LFS, and thus a lower estimate of mortality rate for the same number of deaths.

While the LFS-based estimates of mortality rates differ from the census-based ones, they do appear to have internal consistency from year to year. The pattern of mortality rates over the classes in Figures 2–4 are very similar, and Figure 5 also demonstrates an almost parallel pattern

between 2001–03 and 2005–07. Thus the LFS-based estimates could feasibly be used to compare mortality from one three-year period to another, and possibly from year to year.

It would be technically possible to use changes over time in the LFS-based estimates for each class to update the estimates based on the 2001 Census. This would have the advantage that all estimates of mortality by NS-SEC published by ONS would be comparable, and would avoid possible confusion among users. However, the LFS-based and census based estimates are different, for reasons discussed above. Consequently this process would require the further assumption that movement in the LFS-based rates would also apply to the census-based estimates. The need to make this extra assumption, and the danger of introducing methodological artefacts in the adjustment process, are the reasons why this article recommends the use of the LFS-based estimates directly for the study of intercensal changes in mortality within socio-economic classes.

The usefulness of the LFS-based estimates to monitor changes due to policy might be assessed using the results shown in Table 3 together with some very broad assumptions. For example, if a policy was aimed at reducing the mortality rate of the least advantaged class, then by comparing three-year period estimates for England and Wales, differences of above 1.4 per cent between the two periods would be detectable. This is because confidence interval widths on the difference of two independent rates are approximately 1.4 times the width of the confidence limit of a single rate. This level of accuracy would seem to be more than sufficient for most applications, therefore the authors recommend the production, on an experimental basis, of a national series based on a three-year period. On the same basis when using a one-year period, only differences of the order of 2.8 per cent per annum could be detected. This may also be useful, particularly in comparisons of trends over a number of years, and therefore the authors also recommend the production, on an experimental basis, of an annual national series.

The usefulness of the regional estimates to monitor changes in the annual mortality rate is less clear. Since the confidence intervals for the three-year estimate are around plus or minus 4 per cent, the approximate confidence intervals for the difference in two estimates would be about 5.6 per cent. This may well be sufficient over a three-year period. On a similar basis, the confidence intervals for the difference in regional estimates for two single years would be 8.4 per cent. This is greater than the probable change in a given year and so could be deemed to be of marginal value. Therefore, the authors also recommend the production, on an experimental basis, of a regional series based on three-year periods only.

Extension of this approach to county districts suggests that changes in inequalities of less than 15.4 per cent (for a three-year period) would not be statistically detectable and this would appear to be of little use for any monitoring measure. For unitary authorities, the estimates would be even less precise.

Analysis of mortality rates for women

All the analysis presented in this article has been for men. This is because the individual datasets have information on occupation, employment status and (full) NS-SEC for individuals. These datasets also include women who were at the time or had recently been employed. However, the census-based published mortality rates for women by NS-SEC are based on a 'combined' classification⁹, whereby the 'most advantaged' NS-SEC of the woman and her husband, (if she has

one), is used to represent the socio-economic position of the household. To produce estimates analogous to this, it would be necessary to link data for a woman to her husband's data – if such a person existed. From this, a 'combined' NS-SEC could be estimated for females. The technical feasibility of this process has been checked during the course of this work, but no analysis is presented here.

Should this work be developed, repeating the process for women will be a priority.

Limitations of the analysis

Certain assumptions were made in the course of this analysis. The principal ones were introduced as a result of the health selection adjustment, used to correct for the known bias in the assignment of populations to NS-SEC classes, and to place the estimates on the same basis as that used for the published census-based results. These were:

- The percentage of those unallocated to an NS-SEC analytic class to be reallocated to an analytic class varies over time in proportion to the numbers in these categories computed using the LFS for the corresponding period.
- The distribution across analytic classes of the reallocated population should be in the same proportions over time as those derived from tables published in the article in HSQ volume 36⁶. These redistributions will be in the same proportions across all age groups.

There is a case for not making these health selection adjustments. It could be argued that the year to year consistency of the LFS would best be maintained by not making such adjustments. The process would be simpler and require fewer assumptions. The case for proceeding as reported in this article is that health selection effects are real¹⁰, even in cross-sectional analyses and to ignore them would invite a known bias. The authors favour the approach set out in this article because we feel it produces a more accurate picture of socio-economic inequalities.

Another assumption made was that those living in communal establishments and not counted by the LFS did not make a substantive difference in the distribution across NS-SEC analytic classes for men aged 25–64.

Other assumptions were made in the calculation of the confidence limits presented. Although adequate for the purposes of a feasibility study, the presented confidence intervals should be considered as indicative only.

Conclusions and Recommendations

The LFS can be used at the level of England & Wales to measure health inequalities by NS-SEC analytic class for men. National level estimates based on one or three year's deaths and population should be capable of providing useful estimates for the study of trends in health inequalities. This would also be feasible for the three-year period at the Government Office Region level. Below this level of geography, estimates would not be sufficiently reliable.

The mortality rates obtained for 2001–03 using the LFS are sufficiently different to the published census-based estimates for the results from the two systems not to be interchangeable. The LFS estimates would have to be related to each other rather than to the census-based ones.

Using a three-year average comparison, the LFS-based estimates show a statistically significant decrease in mortality rates over time between 2001–03 and 2005–07 across all but one NS-SEC analytic classes.

The technical feasibility of using the same approach to produce mortality rates by NS-SEC for women was tested, and found to be possible in terms of availability and linkage of data. No comparative analysis with the census was undertaken for women, but a similar analysis to that carried out for men would be a priority if this work were to be developed.

It is recommended that, following review of this paper and consultation with interested parties, a measure of mortality by NS-SEC using LFS-based denominators be created for years from 2001 to date at a national level for single years and for three-year periods. At a regional level, it is recommended that a series be produced based on three-year periods.

References

- 1 Department of Health. *Health inequalities guidance and publications: Life expectancy*. Available on the Department of Health website at:
www.dh.gov.uk/en/Publichealth/Healthinequalities/Healthinequalitiesguidancepublications/DH_064240
- 2 The Marmot Review (2001) *Fair Society, Healthy Lives. Strategic Review of Health Inequalities in England post-2010*. Available at:
www.ucl.ac.uk/gheg/marmotreview
- 3 Office for National Statistics (2003) *Labour Force Survey user Guide Volume 1 Background and Methodology*. Available on the Office for National Statistics website at:
www.statistics.gov.uk/statbase/Product.asp?vlnk=1537
- 4 Heap D (2005) 'Comparison of 2001 Census and Labour Force Survey labour market indicators', *Labour Market Trends*. Available on the Office for National Statistics website at:
www.statistics.gov.uk/ccf/article.asp?id=1037
- 5 Office for National Statistics (2002) *Joined up labour market data*. Available on the Office for National Statistics website at:
www.statistics.gov.uk/STATBASE/Product.asp?vlnk=9846
- 6 White C, Glickman M, Johnson B and Corbin T (2007) 'Social inequalities in adult male mortality by the National Statistics Socio-Economic Classification, England and Wales, 2001–03', *Health Statistics Quarterly* **36**, 6-23. Available on the Office for National Statistics website at:
www.statistics.gov.uk/ccf/article.asp?ID=1901
- 7 Rose D and Pevalin D (eds) (2003) *A Researcher's Guide to the National Statistics Socio-economic Classification*. SAGE Publications Ltd.
- 8 Office for National Statistics, *The National Statistics Socio-economic Classification on-line edition*. Available on the Office for National Statistics website at:
www.statistics.gov.uk/statbase/Product.asp?vlnk=13561

9 Langford A and Johnson B (2009) 'Social inequalities in adult female mortality by the National Statistics Socio-Economic Classification, England and Wales, 2001–03', *Health Statistics Quarterly* **42**, 6–21. Available on the Office for National Statistics website at:

www.statistics.gov.uk/cci/article.asp?ID=2191

10 OPCS (1978) *Occupational Mortality 1970-72*, Series DS No 1, Chapter 3, HMSO: London

Annex

Detailed description of methodology for estimating populations by age and NS-SEC using the LFS

1 Use of LFS dataset to produce estimates of private household population by age group and NS-SEC

The LFS has a variable relating to full NS-SEC but it was not possible to use it, since it was necessary to compute *reduced* NS-SEC, using only occupation and employment status, to be compatible with NS-SEC in death registrations. In addition, since the census classifies people by former occupation if they are not currently working, it was decided to derive an NS-SEC classification which took note of an LFS sample member's previous job if they were not currently employed.

Those identified as full-time students by LFS were classified as such in the analysis regardless of any occupation data associated with them.

Thus there were ten NS-SEC categories – eight occupied analytic classes, full-time students and 'other'.

The computations were performed in Stata and weighted to the total private household population estimates for the relevant year by applying the LFS system weighting variable.

2 Health selection adjustments for LFS data

The concept of 'occupation' used to derive socio-economic class for the purpose of estimating mortality rates in successive Registrar General's decennial supplements¹⁰, is that in which the person is currently engaged or most recently was engaged, if not currently employed or economically active. This is partly because using only current occupation tends to understate those who are usually employed and makes classification more subject to macroeconomic fluctuations. In addition, death registration of necessity uses a definition which extends back in time. Thus occupation in this context is not restricted to the employed or even to the economically active. Most importantly, those in the most disadvantaged classes are less likely than the more advantaged to report a former occupation if sick, and this may lead to biased estimates if not adjusted.¹⁰

In the published estimates using census-based denominators, an adjustment was made to counteract a potential selection effect brought about because some of those in poor health may have been selected out of the labour market⁶. This adjustment was made using the ONS Longitudinal Study to estimate the proportion of those with no recorded occupation and therefore no analytic NS-SEC class in 2001, but who had a recorded occupation in 1991. The 1991 NS-SEC distribution of those in each age group unclassified at 2001 was used to redistribute population to the analytic NS-SEC classes. In order to compare results from the current study with the published ones, it was necessary to imitate the adjustments undertaken on the census-based denominators using the published adjustment for those with no occupation recorded in 2001. The adjustment process first estimates the number of people which should be redistributed from the residual to the

analytic classes, and then allocates to classes using the same distribution pattern as the census-based estimates.

3 Illustrative example of derivation of populations by age and NS-SEC using the LFS counts and ONS mid-year populations

The process of translating the raw LFS counts into population estimates comparable with the published census-based figures is shown for the Spring Quarter 2001 in Table A1.

Table A1 Example showing method for the normalisation of the LFS population counts by NS-SEC to the ONS mid-year population 2001 and the 'health selection' transfer from the residual groups to the NS-SEC analytic classes, for methodological consistency with the published results using census-based denominators.

England and Wales										Thousands					
										Total of analytic groups sum (1.1 – 7)	ONS mid year population ¹	Un-allocated population	Proportion residual in optimised table ²	Final residual	Others for reallocating across analytic groups
LFS population counts by NS-SEC analytic class															
Age band	1.1	1.2	2	3	4	5	6	7	a	b	c = b - a	d	e = b x d	f = c - e	
25-29	95	209	364	147	110	239	177	211	1,554	1,742	188	0.045	78	110	
30-34	160	219	432	114	205	285	207	230	1,852	2,016	165	0.030	60	105	
35-39	208	207	437	107	246	284	195	234	1,918	2,050	132	0.024	48	84	
40-44	195	190	386	86	221	255	172	217	1,722	1,831	109	0.020	36	73	
45-49	168	182	341	73	235	206	138	192	1,535	1,644	109	0.018	29	80	
50-54	161	180	386	74	236	243	156	215	1,653	1,771	118	0.015	26	92	
55-59	116	132	291	64	209	187	138	199	1,337	1,492	155	0.011	17	138	
60-64	72	93	191	45	175	149	128	173	1,026	1,252	226	0.012	16	210	
Total	1,176	1,412	2,828	710	1,639	1,847	1,311	1,672	12,595	13,797	1,202		310	892	

1 ONS mid-year population estimates for 2001

2 White C, Glickman M, Johnson B and Corbin T (2007) 'Social inequalities in adult male mortality by the National Statistics Socio-Economic Classification, England and Wales, 2001-03, Health Statistics Quarterly 36, 6-23

The LFS based estimates for each NS-SEC analytic class are shown on the left-hand side of the table and the totals for each age group are shown in column a. The ONS mid-year population totals for 2001 are in column b and the difference (the 'unallocated' population), in column c.

In order to approximate the optimised estimates, it is necessary to simulate the transfer from the unallocated populations to the NS-SEC analytic classes, which was undertaken on the published census data using the health selection adjustment.

Firstly the proportion of the population in each age group which was to be assigned to the residual NS-SEC classes was set equal to that in the census-based 'optimised populations'. (column d in

Table A1). These proportions were multiplied by the ONS mid-year populations to obtain the numbers in the residual categories (column e). The remainder of the unallocated population is obtained by subtracting column e from column c.

Thus for example, the number unclassified aged 25-29 in the optimised population estimates was 4.5% of the total population⁶. Thus the population (in thousands) in the residual is $1742 \times 4.5\% = 78$ (to the nearest thousand). The number available for reallocation is therefore equal to the population for the age group (1742) less the NS-SEC analytic group population estimated from the LFS estimates (1554) less 78 (equals 110).

The resultant population is then reallocated across NS-SEC classes according to the proportions estimated from the gain in person-years by each NS-SEC analytic group from the health selection adjustment in the published study using census-based denominators, (White C Glickman G, Johnson B and Corbin T (2007), Social Inequalities in adult male mortality by NS-SEC, England and Wales, 2001-03, Health Statistics Quarterly 36, Table 2⁶) These are shown in Box A1.

Following the obtaining of matrices of deaths and populations by age group and NS-SEC, the standardised mortality rates were calculated by reference to the European Standard Population.

Box A1 Percentage distribution of unclassified to analytical classes for 'health selection adjustment'

1.1	Large employers and higher managerial	2.0%
1.2	Higher professional	6.9%
2	Lower managerial and professional	15.5%
3	Intermediate	8.1%
4	Small employers and own account workers	10.1%
5	Lower supervisory and technical	10.3%
6	Semi-routine	20.6%
7	Routine	26.5%

The resultant population denominators in this example would be as in Table A2.

Table A2 Example showing resultant estimates of populations by NS-SEC and age for 2001, based on LFS counts 2001 mid-year population and the 'health selection' transfer from the unallocated groups to the NS-SEC analytic classes

England and Wales									Thousands	
	LFS population counts by NS-SEC analytic class									Total
	1.1	1.2	2	3	4	5	6	7	Residual ¹	Population ²
25-29	98	217	381	156	121	250	200	240	78	1,742
30-34	163	226	448	122	215	295	228	258	60	2,016
35-39	210	212	450	114	255	292	212	256	48	2,050
40-44	196	195	397	91	229	263	187	236	36	1,831
45-49	169	187	354	79	243	214	154	214	29	1,644
50-54	163	187	400	81	246	252	175	240	26	1,771
55-59	119	142	312	75	223	201	167	236	17	1,492
60-64	76	107	224	62	196	171	171	229	16	1,252
Total	1,193	1,473	2,966	782	1,729	1,939	1,495	1,909	310	13,797

1 As per Table A1

2 ONS mid-year population estimates for 2001, as per Table A1