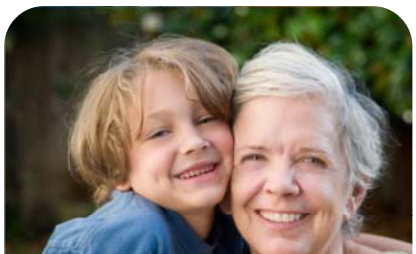
Office for National Statistics

Health Statistics Quarterly

Spring 2011

Edition No.: **49**

Office for National Statistics



49

ISSN 2040-1574

A National Statistics publication

National Statistics are produced to high professional standards set out in the Code of Practice for Official Statistics. They are produced free from political influence.

About us

The Office for National Statistics

The Office for National Statistics (ONS) is the executive office of the UK Statistics Authority, a non-ministerial department which reports directly to Parliament. ONS is the UK government's single largest statistical producer. It compiles information about the UK's society and economy, and provides the evidence-base for policy and decision-making, the allocation of resources, and public accountability. The Director-General of ONS reports directly to the National Statistician who is the Authority's Chief Executive and the Head of the Government Statistical Service.

The Government Statistical Service

The Government Statistical Service (GSS) is a network of professional statisticians and their staff operating both within the Office for National Statistics and across more than 30 other government departments and agencies.

Contacts

This publication

For information about the content of this publication, contact
Steve Brown
Tel: 01633 455447
Email: hsq@ons.gsi.gov.uk

Other customer enquiries

ONS Customer Contact Centre
Tel: 0845 601 3034
International: +44 (0)845 601 3034
Minicom: 01633 812399
Email: info@statistics.gsi.gov.uk
Fax: 01633 652747
Post: Room 1.101, Government Buildings,
Cardiff Road, Newport, South Wales NP10 8XG
www.ons.gov.uk

Media enquiries

Tel: 0845 604 1858
Email: press.office@ons.gsi.gov.uk

Copyright and reproduction

Contents

In brief

Young Statisticians' Meeting 2011

On 12–14 April 2011 the Young Statisticians' Meeting will be held at the University of Southampton. This is a two and a half day conference that brings together young statisticians – those in the first ten years of their statistical career – from across the civil service, industry and academia.

Registration for this event is now open and delegates are invited to submit abstracts for oral and poster presentations for a number of parallel sessions and a poster presentation session. The YSM 2011 also provides delegates with an excellent opportunity to experience a variety of statistical topics currently being researched and to explore the debates around wider issues in statistics.

This year, the conference will focus on the debate around the extent to which personal data should be held by the Government. There will be a special "Dinner and Debate" session on the first day of the conference where guest speakers - including Simon Briscoe, Statistics Editor of The Financial Times and David Bond, a freelance journalist known for his reporting on the use of personal data - will sit on a panel to discuss their opinions on this topic and answer delegates' questions. David recently made a Channel 4 film that was aired in cinemas nationwide and was broadcast on TV in late spring 2010. This documentary focused on the confidentiality issues around personal data and looked at how much of a 'surveillance society' the UK has become. The documentary will be screened prior to dinner on day one to give delegates some background information, and to spark debate. In this session, delegates will sit down to dinner and wine followed by the panellists putting forward their views on "Personal Data Held by Government". There will then be an opportunity for delegates to ask the panel questions and debate the issues raised. In addition, there will be workshops on the Census and Data Visualisation.

Full details of the conference are available at: YSM 2011

Consultation - Conception statistics

ONS is conducting a review of conception statistics to reduce the cost of producing conception outputs while ensuring that the published outputs continue to meet user requirements. Changes are proposed to both the timing and content of conception outputs.

Conception statistics and the consultation document are available on the ONS website at: Conception Statistics, England and Wales

Email: vsob@ons.gsi.gov.uk

Users' views received as a result of the consultation will be considered before implementing any changes to conception outputs. Further information will be published later in the year once responses from the consultation have been considered.

Consultation - Avoidable Mortality

On 16th February 2011 ONS launched an eight week consultation on avoidable mortality. The consultation outlines some key issues in the definition of avoidable mortality, explains new definitions proposed by ONS, and seeks the views of statistics users and experts.

Further information, including the consultation paper and how to respond can be found on the ONS website at: Open consultations: Office for National Statistics

Responses to the consultation are invited up to the 12th April 2011. A summary of the comments made will be published approximately 2 months after the consultation closes.

One release of birth and death summary tables for 2011

Provisional annual birth and death summary tables for 2009 were released on 25 May 2010 with final data published on 21 July 2010. To ensure the effective use of resources and avoid confusion among users, the requirement for provisional data in May was explored. No strong user requirement was identified so **provisional** annual data will no longer be published.

Final annual birth and death summary tables for 2010 will be published in July 2011 with more detailed data published over subsequent months. Any user comments are welcomed.

Email: vsob@ons.gov.uk

Injury and poisoning mortality, England and Wales

Statistics on the main causes of deaths from injury and poisoning mortality in England and Wales 2005–09 will be published on 1 March 2011. The tables present data using a matrix of mechanism by intent developed by the International Collaborative Effort (ICE) on injury statistics.

Mortality Statistics: Injury and Poisoning (Series DH4)

Adult Dental Health Survey preliminary findings

The 2009 Adult Dental Health Survey (ADHS) is the fifth in a series of national dental surveys that have been carried out every ten years since 1968. The main purpose of the survey is to establish the condition of the natural teeth and supporting tissues; to investigate dental experience, knowledge about and attitudes towards dental care and hygiene; to examine changes over time in dental health, attitudes and behaviour and to monitor the extent to which dental hygiene targets set by the government are being met. ONS have conducted the survey on behalf of the NHS Information Centre, working in consortium with the National Centre for Social Research and academics at the Universities of Birmingham, Cardiff, Dundee, Newcastle and University College London.

A total of 11,380 adults were interviewed for the survey, and 6,469 adults were examined. The survey covers England, Wales and Northern Ireland.

A First Release with preliminary findings was published in December 2010 by the NHS Information Centre. A series of more detailed reports are due to be published in March 2011 alongside a Foundation Report with technical details of the survey.

Key findings include:

- Overall, 94 per cent of adults in England, Wales and Northern Ireland were dentate[1], only 6 per cent had not maintained this basic threshold of oral health.
- The proportion of people who were edentate[2] varied by country. In England 6 per cent of adults were edentate compared with 10 per cent in Wales. Seven per cent of adults in Northern Ireland were edentate.
- Over the last 30 years the proportion of adults in England who were edentate has fallen by 22 percentage points from 28 per cent in 1978. In Wales, the proportion of adults who were edentate has fallen 27 percentage points from 37 per cent in 1978. In Northern Ireland, the proportion of adults who were edentate has fallen 26 percentage points from 33 per cent in 1979[3].
- 86 per cent of dentate adults had 21 or more natural teeth.
- The proportion of dentate adults with 21 or more natural teeth varied by country. In England 86 per cent of dentate adults had 21 or more natural teeth compared with 80 per cent of dentate adults in Wales. Eighty-four per cent of dentate adults in Northern Ireland had 21 or more natural teeth.

1 Dentate, that is having at least one natural tooth
2 Edentate, that is having no natural teeth
3 The Northern Ireland ADHS was conducted in 1979
4 This question was asked for the first time in Northern Ireland in 1988

- Over three-fifths of dentate adults said they attended the dentist for regular check-ups which had increased in all three countries between 1978 and 2009 (1988 to 2009 in Northern Ireland[4]). The greatest increase was observed in Wales, up 30 percentage points from 39 per cent in 1978 to 69 per cent in 2009.

- 12 per cent of all adults (who had ever been to the dentist) were classified as having extreme dental anxiety.

- Extreme dental anxiety was more prevalent among women than men, 17 per cent compared with 8 per cent respectively.

- There was a clear pattern of higher levels of dental anxiety among younger adults. Levels of extreme dental anxiety ranged from 15 per cent of adults aged 16 to 24 to 9 per cent of adults over 85.

Adult Dental Health Survey - 2009

Life Opportunities Survey Interim Report

The Life Opportunities Survey (LOS) is a new large-scale longitudinal survey of disability in Great Britain. LOS is the first major social survey in Great Britain to explore disability in terms of social barriers to participation, rather than only measuring disability in terms of impairments or health conditions. LOS compares the experiences of people with and without impairments across a range of areas, including education and training, employment, transport, leisure, social and cultural activities, and social contact. The interim findings were published in December 2010.

On LOS, an adult has a participation restriction if they experience at least one social barrier to taking part in at least one life area. The key findings of LOS between June 2009 and March 2010 are presented below.

- 17 per cent of adults with impairments experienced participation restrictions in their learning opportunities compared with 9 per cent of adults without impairments

- 56 per cent of adults with impairments experienced restrictions in the type or amount of paid work they did, compared with 26 per cent of adults without impairments

- 74 per cent of adults with impairments experienced restrictions in using transport compared with 58 per cent of adults without impairments

- 83 per cent of adults with impairments experienced a participation restriction in leisure, social and cultural activities compared with 78 per cent of adults without impairments

- 24 per cent of adults with impairments experienced a participation restriction to social contact (that is, being able to meet with close contacts as much as they would like) compared with 22 per cent of adults without impairments

- 12 per cent of adults with impairments experienced difficulty accessing rooms within their home or difficulty getting in or out of their home compared with 1 per cent of adults without impairments
- 29 per cent of adults with impairments experienced a participation restriction to accessing buildings outside their home compared with 7 per cent of adults without impairments
- 45 per cent of households where at least one person had an impairment were unable to afford expenses or make loan repayments. This compares with 29 per cent of households without any people with impairments

The LOS also reports on a variety of social barriers, such as lack of help or assistance, attitudes of other people and caring responsibility. Barriers to education and training are presented here as an example of the social barriers reported by respondents. Of those adults experiencing participation restrictions to learning opportunities:

- The main barrier to learning for adults with impairments was financial reasons (46 per cent). This was also the main barrier among adults without impairments (52 per cent)
- "Too busy/not enough time" was the second most common barrier to learning for adults with impairments (23 per cent). This was also the second most common barrier for adults without impairments (40 per cent)

Life Opportunities Survey

LOS User Guide (162Kb - Pdf)

Introduction of ICD–10 v2010

In January 2011 the Office for National Statistics (ONS) updated the software used to code cause of death from the International Classification of Diseases, Tenth Revision (ICD–10) v2001.2 to v2010.

The main changes in ICD–10 v2010 are amendments to the modification tables and selection rules. Modification tables and selection rules are used to ascertain a causal sequence and consistently assign underlying cause of death from the conditions recorded on the death certificate.

Mortality statistics for England and Wales for the 2011 data year will be coded to ICD–10 v2010. A bridge coding study has been carried out to help users understand the impact of the change. Overall, the impact is small although some cause groups are affected more than others.

Results from the ICD–10 v2010 bridge coding study

Results from the ICD–10 v2010 bridge coding study: stillbirths and neonatal deaths

ONS Longitudinal Study updated to 2009

The annual refresh of vital events linked to the ONS Longitudinal Study (ONS LS) will be complete in March 2011.

Researchers using the ONS LS will have access to an extra year of vital event information to use in their research. Information on births and deaths will be available for 2009, while cancer registrations will be available up to 2008.

The ONS LS contains linked census and vital event data for a 1 per cent sample of the population of England and Wales. Information about more than 500,000 study members has been captured at each of the censuses since 1971, and plans to incorporate data from the 2011 Census are well advanced. More than one million people have been recorded in the ONS LS sample at some point.

ONS LS data processing involves close collaboration between ONS staff based in Titchfield and their former colleagues in Southport, who are now part of the Medical Research Information Service within the NHS Information Centre for Health and Social Care. This has been the fourth full year of processing following ONS independence. Thanks to hard work and diligence of staff in both organisations, the transition has been seamless.

The ONS actively promotes wide use of the ONS LS while maintaining the confidentiality of the individuals in the sample. ONS LS records available for analysis are anonymised but the database contains individual-level data that have not been aggregated or disguised. To ensure confidentiality, these microdata can only be accessed at ONS sites through the Virtual Microdata Laboratory (VML). Support officers are available to help extract and use the data. ONS LS users can be sent aggregated data in the form of tabulations that have been checked to ensure there is no risk of identifying an individual.

For further information or for a discussion about using the LS please contact:

Academic users:
Centre for Longitudinal Study Information & User Support (CeLSIUS)
Tel: +44 (0) 20 7299 4634
Email: celsius@lshtm.ac.uk
CeLSIUS - Centre for Longitudinal Study Information and User Support

Other users:
Microdata Analysis and User Support Team
Tel: +44 (0) 1633 45 5844
Email: maus@ons.gsi.gov.uk
The ONS Longitudinal Study

New and improved ONS website to launch 30 April 2011

The Office for National Statistics (ONS) is developing a new website, with a focus on putting users' needs first. When it goes live on 30 April 2011 the new ONS website will deliver the following improvements:

- quicker and easier to find information, including better search and navigation
- easier to use the information, by downloading data, charts and graphs
- improved accessibility to ONS content for users with sight or other impairments
- prompt release of outputs at 9.30am sharp

Content will be transferred from the current website to the new one, but some existing bookmarks and links will no longer work.

After the initial launch, there will be further developments to the new website. These include:

- an online data explorer tool, allowing users to customise, interact with and download datasets
- an Application Programming Interface, enabling re-use of ONS data by others

Web Development

Deriving trends in life expectancy by the National Statistics Socio-economic Classification using the ONS Longitudinal Study

Brian Johnson ***Office for National Statistics***

Abstract

Background

Health inequalities among socio-economic groups are well documented. One of the measures used to track inequalities over time is the series 'Trends in life expectancy by social class, 1972–2005', on the Office for National Statistics website. In 2001 the National Statistics Socio-economic Classification (NS-SEC), replaced Registrar General's social class (RGSC) for the purposes of official statistics. This paper describes the challenges involved in producing an analogous series of trends in life expectancy by NS-SEC to that by RGSC, the approach adopted, and publishes the first results of the new series.

Methods

NS-SEC was devised in the 1990s and introduced in 2001. Like RGSC, it is an occupation-based measure. In order to produce a series of trends over more than 20 years based on NS-SEC, it is necessary to classify people according to NS-SEC based on their occupation at the 1981 and 1991 Censuses and then to measure subsequent mortality rates for different classes. The 1981 Census preceded the construction of the NS-SEC classification system by nearly 20 years, and there was no recognised way of classifying 1981 Census respondents by NS-SEC. This paper describes how an approximation to allow such a classification was derived. The ONS Longitudinal Study was used to provide the data from which mortality and survival rates by NS-SEC class could then be estimated.

Results

The results are presented in terms of life expectancy at birth and at age 65 by five-year calendar periods, from 1982–86 to 2002–06. A social gradient was found using NS-SEC, similar to the one found using RGSC. For most classes for all periods studied, life expectancy improved for both males and females but inequalities persisted between classes. There was a difference of around six years for males between the most and least advantaged classes in expectation of life at birth and about four years for females in the

period 2002–06. The estimates suggested a widening of inequalities over the study period for men, which appeared to end after 2001. For women, no overall trend could be detected, but there were no signs of any narrowing of the gap in the most recent period.

Conclusions

NS-SEC can be used to provide medium-term trends in life expectancy by occupation based class, which will be capable of extension over time, although certain approximations are necessary. It is important that work should continue on investigating other means of classification, particularly for women, for example based on educational attainment and on household rather than individual-based measures.

Contents

List of Figures

List of Tables

Introduction

Health inequalities among socio-economic groups are well documented. Since 1921 when the Registrar General's social class schema was introduced for adults, in the form that was to survive almost to the end of the century (Rose, 1995), differences in mortality rates among classes have been observed. These differences have persisted over time and have tended to prove resistant to the attempts of policymakers to reduce them. Influential reports on the subject have been produced by (among others), Black (1980), Acheson (1998) and most recently Marmot (2010) in '*Fair Society, Healthy Lives (The Marmot Review).*' Together with policy recommendations, the latter provides a comprehensive overview of the situation in England and a wealth of references.

The conventional method for estimating mortality rates by class has been to use the decennial census of population to provide age and socio-economically stratified denominators, and death registrations coded by occupation for years close to the census to provide numerators. (Up to and including 1991 only a sample of census households had their occupants coded by occupation).

This method, while producing consistent findings over a long period, had a number of drawbacks:

Firstly, there are differences in the method of recording of occupation at death and at the census, and secondly, mortality rates can only be estimated at working ages, since occupation is not coded at the census for people over 74 years of age and sparsely coded between the ages of 65 and 74. (OPCS, 1978). Given that a decreasing minority of deaths occur below age 65, it became increasingly important to have available a means of classifying older people to NS-SEC using occupation recorded earlier in their lives. In this way, life expectancy can take account of life course exposures of those dying at older ages.

These shortcomings constituted a major part of the motivation for creating the ONS (formerly OPCS) Longitudinal Study (LS). This has eliminated numerator-denominator bias and allowed the classification of sample members according to their census characteristics often at working age, and the subsequent follow up to measure death and survival rates for each class over time. This allows the computation of age-specific mortality rates at older ages. This in turn has allowed the computation of life tables for different socio-economic classes to be produced.

'Expectation of life by social class' was first published by Hattersley (1997) and developed and updated by Donkin *et al*. (2002) and makes use of the LS, together with the continuity provided by the occupation-based RG social class classification. The series has been extended at intervals since then, the most recent update being in 2007, when the last period for which estimates were published was 2002–05.

In 2001 RG Social Class was replaced by the National Statistics Socio-economic Classification (NS-SEC) for the purposes of official statistics. It was possible to extend the series on trends in life expectancy by RG social class by a few years following the census, but it was clear that, since neither deaths nor census returns were to be coded by RGSC any longer, it was imperative to develop a new series using NS-SEC.

Analyses of mortality rates by NS-SEC have been undertaken using death registrations as numerators and figures based on the 2001 Census population as denominators in 'cross-sectional' studies (White *et al*., 2007, Langford and Johnson, 2009), but these are only possible for ages up to 64 and therefore are not suitable for estimating life expectancy.

The key component of this study was to provide an NS-SEC code for occupations recorded at the 1981 Census using the occupational coding undertaken at the time. If such a classification could be produced, it would be possible to undertake a prospective study whereby those present at the 1981 Census could be classified at, or shortly after working age, and then followed up through to the 2000s. This step was particularly important for those in retirement age in the 1990s and 2000s, since the 1991 or 2001 Census may well not have captured their occupation.

The main difficulties with this approach were:

- conceptual – the change in the nature of work and contracts of employment over time make the application of NS-SEC in 1981 problematic
- practical – the lack of appropriate granularity in the 1981 Census classification system to allow mapping to modern occupation codes

It would not have been cost effective to go back to the original census returns and attempt to recode them, so some kind of 'bridge-coding' was necessary

This article describes how ONS went about constructing this bridge and the methodology adopted to overcome the problems of discontinuity. It addresses some questions related to the attribution of any socio-economic classification system, particularly for women, and presents the first results for 'life expectancy by NS-SEC'.

Methods

This section describes the NS-SEC in more detail, the reasons why it replaced the Registrar General's social class classification in official statistics, and the analytical approach adopted in this article. The key component of the analysis is the creation of an NS-SEC classification using data from the ONS Longitudinal Study (LS) relating to the 1981 Census. This process is described in detail following a brief overview of the NS-SEC and the data sources used.

The National Statistics Socio-economic Classification

Registrar General's social class based on occupation remained almost unchanged in structure for most of the 20th Century. It was based on a hierarchical view of occupational skills and general social standing, and above all, the divide between non-manual and manual occupations. While proving effective in discriminating between classes in empirical studies, changes in economic structure, such as the decline in manufacturing and the growth of the service sector, have provided practical reasons for dispensing with the classification in official statistics. Just as important, it was criticised for having no universally agreed underlying concept, (Rose & Pevalin, 2003), and therefore it was claimed that it is not clear how to interpret the meaning of differences between classes revealed by measuring mortality or other characteristics.

The conceptual basis for the NS-SEC is the structure of employment relations operating in modern developed economies (Rose & Pevalin, 2003), based on a schema developed by Goldthorpe (1997). Occupations are differentiated in terms of reward mechanisms, promotion prospects, notice periods and job security. Those occupations exhibiting a high level of such characteristics are said to be operating on a 'service contract'. Those with the least of these attributes are said to be operating under a 'labour contract'. While not designed as a hierarchy, there are differences in social advantage across the classes. 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.

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, Civil Service administrative officers and local government clerical officers, counter clerks, school and company secretaries
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 Statistics (2002)

An individual can be assigned to an NS-SEC class based on their occupation, employment status and the size of their organisation, (employment status relating to whether they are an employer, a manager, supervisor or ordinary employee). A version of NS-SEC can be 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 about two per cent (ONS NS-SEC User Manual 2002). Reduced NS-SEC was used throughout the following analysis since size of establishment is not available for all sources of information used in this study, (for example father's or mother's occupation based classification at the birth of a child). Reduced NS-SEC was also used in previously published cross-sectional analyses of mortality rates by NS-SEC (White *et al.,* 2007), since size of establishment is not included in the data collected at death registration.

All people for whom there is sufficient information at a census on their occupation and employment status can be assigned an NS-SEC 'analytic' category such as those which appear in Box 1. The lookup table, which allows the transformation of combinations of occupation and employment status, is known as a 'derivation matrix'. A section of a derivation matrix for NS-SEC in 2001 is shown in Box 2. Every decade for which NS-SEC is to be assigned requires a derivation matrix which converts the contemporary occupation definitions and employment statuses into NS-SEC classes.

Box 2 Extract of derivation matrix for NS-SEC using occupation (SOC2000) and employment status to determine analytic class

SOC2000	Standard Occupational Classification 2000	Employers	self-employed	Manager	Supervisor	Employee
Occupa	tional Unit Group					
1111	Senior officials in national government	1.1	1.1	1.1	1.1	1.1
1112	Directors and chief executives of major organisations	1.1	1.1	1.1	1.1	1.1
1113	Senior officials in local government	4	4	1.1	1.1	1.1
1114	Senior officials of special interest organisations	4	4	2	2	2
1121	Production, works and maintenance managers	4	4	1.1	1.1	1.1
1122	Managers in construction	4	4	2	2	2
1123	Managers in mining and energy	4	4	1.1	1.1	1.1
1131	Financial managers and chartered secretaries	1.2	1.2	1.1	1.1	1.1
1132	Marketing and sales managers	4	4	1.1	1.1	1.1
1133 Purcha	sing managers	4	4	1.1	1.1	1.1
1134	Advertising and public relations managers	4	4	1.1	1.1	1.1
1135	Personnel, training and industrial relations managers	4	4	1.1	1.1	1.1
1136	Information and communication technology managers	4	4	1.1	1.1	1.1
1137	Research and development managers	1.2	1.2	1.1	1.1	1.1
1141 Quality	assurance managers	4	4	2	2	2

Source: NS-SEC User Manual, Office for National Statistics (2002)

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’ ‘inadequately described’. These are known as ‘residual’ categories. Residual categories are reported in this analysis in aggregate as ‘Unclassified’, both for completeness and because a substantial proportion of them are likely to be among the most deprived, having no occupation assigned. The NS-SEC sub-divisions, such as ‘Never worked’ and ‘Long-term unemployed’ were not available at censuses before 2001. As a result, ‘Unclassified’ constitutes a heterogeneous category, the composition of which may change over time. Consequently, it was decided to restrict the measurement of trends in inequality to the occupation-based classes.

Further aggregations of the classes are possible, and a useful summary measure is ‘condensed’ NS-SEC which consists of:

1. Managerial and professional (consisting of analytic classes 1.1, 1.2 and 2, the 'service contract' groups.
2. Intermediate (consisting of analytic classes 3 and 4)
3. Routine and manual (consisting of analytic classes 5, 6 and 7, the 'labour contract' groups)

These groupings are larger and less volatile from one time period to the next than the more refined analytic classes, and sometimes give a clearer picture of trends and turning points in inequalities. For this reason, the analysis was carried out both for the seven analytical classes and the condensed NS-SEC.

Data Sources for analysis

The ONS Longitudinal Study (LS):

The LS contains linked census and vital event data for one per cent of the population of England and Wales. Information from the 1971, 1981, 1991 and 2001 Censuses has been linked together, along with information on events such as births, deaths and cancer registrations. The unique advantage of this data source for the type of analysis described in this article is that sample members can be classified according to their census characteristics, often at working age, and then followed up to measure death and survival rates for each class over time.

This allows the computation of age-specific mortality rates at older ages based on occupational data recorded at or shortly after working age, which is not possible using a cross-sectional approach, since occupation is not coded at the census for people over 74 years of age. Father's and mother's occupation is recorded for children present at a census and at birth for new babies. The LS has occupation and employment status coded for those present at censuses from 1971 to 2001 and NS-SEC for those present at the 2001 Census.

Derivation matrices for NS-SEC.

The matrix developed for the 2001 Census producing reduced NS-SEC is available from the NS-SEC user manual (ONS 2002). An approximation to NS-SEC, derived using the SOC90 occupation classification also exists and can be found in the same publication. This allows the NS-SEC classification of LS members at the 1991 Census, or for the parents of children born in the decade between the 1991 and 2001 Censuses.

As yet there is no derivation matrix for NS-SEC using the 1981 occupational classification. Therefore, the main component of the current study was to produce and validate such a matrix for use with LS data so that members may be classified at the 1981 Census or at birth between 1981 and 1991.

Analytical approach

As a result of changes in economic structure and the nature of occupations over time, the occupational classification system is reviewed every 10 years. Box 3 shows the nomenclature for the occupational and socio-economic classifications used at each census since 1981.

Box 3 The nomenclature for the occupational and socio-economic classifications used at each census since 1981

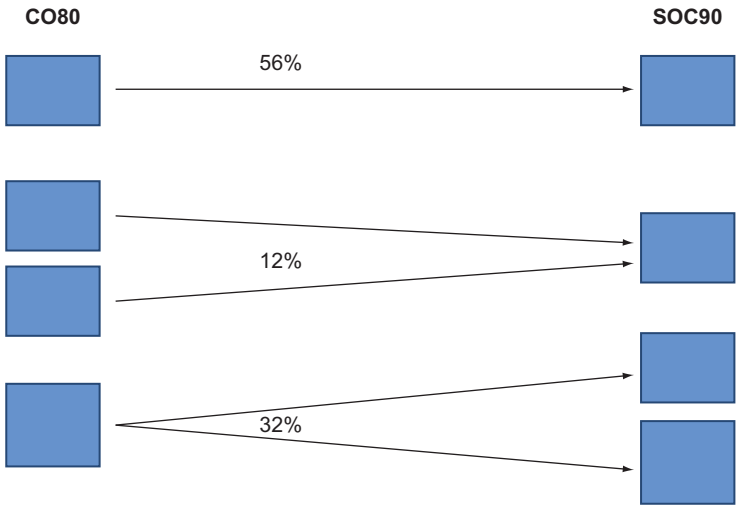
Census	Occupational Classification	Socio-economic Classification
1981	CO80	Registrar General's Social Class (RGSC)
1991	SOC90	RGSC NS-SEC90
2001	SOC2000	NS-SEC

The 1991 and 2001 derivation matrices were produced in part via detailed analysis of specially commissioned questions included in the 1996/97 Labour Force Survey (Rose and Pevalin (2003).

It would be neither practical nor economically viable to repeat this detailed analysis for the census of 1981. However, the existing degree of linkage between the 1981 occupational classification system (known as CO80) and the SOC90 classification used for the 1991 Census, from which NS-SEC90 can be derived, is sufficient to allow an approximation for NS-SEC using 1981 occupations.

Hence the NS-SEC90 derivation matrix of occupation and employment status could be used as a template from which an 'NS-SEC80' matrix can be derived. If there was a unique mapping between CO80 (1981) occupations and the SOC90 ones, this would be straightforward. However changes in the occupation classification scheme over the 10-year period result in some degree of discontinuity. The Standard Occupational Classification, Volume 3 (OPCS 1991) estimated that 56 per cent of a 0.5 per cent sample of occupations drawn from the 1981 Census population had a directly comparable SOC90 code. A further 12 per cent was capable of mapping to a SOC90 code if CO80 codes were aggregated. This left roughly 30 per cent of the 1981 population not capable of mapping to a unique SOC90 code (see Box 4).

Box 4 Correspondence between 1981 (CO80) occupational classification and SOC90



In the current study it is the 1981 Census population which needs to be mapped to a SOC90 code and via that, to an NS-SEC class – it does not matter if it is not possible to map SOC90 categories to unique CO80 ones, (Box 4). That is to say, it is occupations in 1991 for which there is no 1981 equivalent, which provided the problem (for example: Software engineers who were grouped together with Computer systems and data processing managers and with Computer analyst/programmers in 1981).

There were three categories of CO80 occupation:

1. Those which mapped directly to a single SOC90 code
2. Those which mapped to more than one SOC90 codes, but all relevant SOC90 codes produced a unique NS-SEC analytic class when combined with employment status
3. Those which mapped to more than one SOC90 code, at least two of which were associated with different NS-SEC classes (even when employment status was the same)

Only the third category presents problems, since the other two produce unique analytic classes.

The scale of the problem could be tested empirically, since the 1991 Census occupations were dual coded on the LS both to the 1981 and 1991 classifications. For each 1981 occupation code, where there was no unique mapping to SOC90, the population assigned to each analytic NS-SEC class could be measured. This allowed the measurement of the size of the population for which the combination of 1981 occupation and employment status did not uniquely determine NS-SEC class.

For example, CO80 occupation number 43, (nurse administrators and nurses), maps to SOC90 occupations 340 (nurses), 640 (assistant nurses, nursing auxiliaries) and 643 (dental nurses). The number in the sample having this occupation and an employment status of employee as follows:

NS-SEC	Class 2 (lower professional)	4,461
	Class 3 (intermediate)	1,857
	Class 6 (semi-routine)	240

(The heterogeneity of this CO80 occupation group and the problems it posed for social classification was referred to by Elias (1997)).The simplest approach is to choose the NS-SEC class with the greatest population (in the above case class 2), for that combination of occupation and employment status and apply that class in the appropriate cell of the derivation matrix. In cells where the ratios between the most and least represented NS-SEC classes were very high (10 to 1 or more) and the minority group had less than 10 members in the sample, this approach was used.

However, to use this simple majority method in the above instance would result in the incorrect assignment to NS-SEC of a substantial number of sample members. Additional information, which was available in 1981 and could be empirically tested using the dual coded 1991 Census data, was sought for these problematic cells of the matrix. Only ‘industry code’ and ‘industry size’ were found to discriminate in cells where there were multiple potential NS-SEC classes. In the example given above, nearly all the NS-SEC class 6 population related to dental nurses. These were not separated from other nurses by CO80, but where the industry code was ‘dental practices’, almost all the population was assigned to class 6 in 1991. Where industry code was found to be useful in this context, it was used to adjust the cells of the derivation matrix. The impact of industry code on the precision of the classification is reported in the Results section below.

A further test of validity of the method was performed by applying it to the dual coded deaths registrations for 1993. In each case it was possible to determine the proportion of erroneous classifications made using the 1981 derivation matrix constructed as described above.

Assignment of NS-SEC class

Once the 1981 derivation matrix was produced, it was possible to classify LS members present at 1981 by their own or by their spouse’s, father’s or mother’s class, if the occupation of any of the latter was recorded at the census.

The standard approach to assigning class for the ‘life expectancy by social class’ series was to classify sample members on entry to the study, after Hattersley (1997). Where possible, an individual’s own NS-SEC class was used. For those with no assigned occupation and who were married, spouse’s class was used. For those who were children at a census, father’s class was used, and failing that mother’s. For new births where the child was not recorded at the following census, father’s class at birth was used, and failing that, mother’s. Immigrants were assigned an NS-SEC class at the first census at which they were present. If they were not present at any census they were excluded from the study since there was no way of assigning a class to them and they may have gone abroad without this fact being recorded, and therefore no longer ‘at risk’ (Johnson and Blackwell, 2007). If none of this information resulted in the sample member being assigned a class, they were coded as ‘Unclassified’

The main rationale for assigning class at the earliest possible point after entry to the study is that, in a prospective study, it is important that the assignment of class is not influenced by the fact of

death or survival. An additional benefit is that the longer the period of follow up, the less prevalent are health selection effects. These occur when some sample members cannot be classified by occupation owing to poor health when assignment of class takes place. These members have a relatively high probability of dying soon after assignment of class. This health selection out of the labour market disproportionately affects Routine and manual occupations, so a bias may be introduced where those in the least advantaged classes appear to have a lower mortality rate relative to others, in the period immediately following classification, than they otherwise would have done (Fox *et al.,* 1985).

The earlier that assignment takes place, and therefore the longer the period of follow up, the less effect selection will have upon the estimates, since those sample members who were sick at the point of assignment will tend either to die or to recover over time.

There is also some evidence that socio-economic position in early years is more influential on health outcomes than status in later life (The Black Report 1980, White *et al.,* 2005).

Taking into account the above and the history of these estimates, the following approach was decided upon: NS-SEC was derived for sample members and their spouses for each census between 1981 and 2001 at which they were present, or at birth for children born after census day in 1981, based on their father's or mother's class, if available. Therefore, for a given sample member, there was frequently more than one class that they could potentially be assigned to. The next stage was to apply a set of rules assigning class to each sample member.

The method of assignment followed the standard approach described above.

A second method classified people by the most advantaged of their own and their spouse's class, and children by the most advantaged of their father's and mother's class. This 'combined' approach is based on a concept of Erikson (1984) and was used in the cross-sectional study of inequalities in mortality rates of women of working age by NS-SEC (Langford and Johnson, 2009). It perhaps provides a more relevant assessment of a person's socio-economic position than one based on primacy of one's own class. While this is not a true 'household' measure of socio-economic position, it is useful to determine the sensitivity of the life expectancy estimates to a more family based assignment of class, particularly for women, for some of whom, access to social and economic resources are less defined by the labour market than they are for men.

Calculating life expectancy

Period life expectancy for a particular NS-SEC class is the average number of years a person would live, if he or she experienced the age-specific mortality rates for that time period, (for example 2002–06), for that class throughout his or her life. It makes no allowance for any future changes in age-specific mortality rates. Period life expectancies are a useful measure of mortality rates actually experienced in a given period and provide an objective means of comparison of the trends in mortality over time, between areas of a country and with other countries. Official life tables in the UK and in other countries which relate to past years are generally period life tables for these reasons. Cohort life expectancies, which attempt to project the actual number of years someone living now can expect to live, usually require projected mortality rates for their calculation and hence, in such cases, involve an element of judgement about future trends.

Life tables are driven by age-specific mortality rates which were derived (for five-year age bands) using deaths and person years survived at risk for each NS-SEC class occurring in each five year period from 1982, calculated using Stata software. The tables, including variances and confidence intervals were derived using ONS standard methodology based on Chiang's method for deriving q_x, (the probability of death in each interval) and variances (Chiang, 1968).

While the LS does not suffer attrition in the way that many surveys do, it is well documented that there are 'losses to follow up' where existing sample members are not present at a census and yet have no record of death or embarkation. Methods for dealing with this problem were explored in depth in Johnson and Blackwell (2007), and analogous methods were used in this paper. Sensitivity tests were carried out to measure the impact on the results of different approaches to excluding sample members who were possibly no longer at risk.

Results

Table 1 shows the prediction of reduced NS-SEC using the 1981 derivation matrix against those based on the 1991 matrix and SOC90 occupation codes for males and females under aged 75 in 1993.

Table 1 Cross-tabulation of death registrations in 1993 classified according to NS-SEC[1] based on 1981 occupation (NS-SEC80) and 1991 occupation (NS-SEC90)

England and Wales

	1981 occupation (NS-SEC80) and 1991 occupation (NS-SEC90)								
NS-SEC based on 1981 occupation	**NS-SEC90**								
	1.1	**1.2**	**2**	**3**	**4**	**5**	**6**	**7**	
1.1 Large employers & higher managers	4,823	20	50	0	0	0	0	0	4,893
1.2 Higher professional	4	6143	38	33	0	0	0	0	6,218
2 Lower managerial & professional	216	89	24,049	828	6	40	33	0	25,261
3 Intermediate	0	11	40	20,771	0	0	50	36	20,908
4 Small employers & own a/c workers	0	18	0	0	15,026	0	0	0	15,044
5 Lower supervisory & technical	0	0	0	0	0	22,701	113	242	23,056
6 Semi-routine	0	0	0	191	0	76	33,056	636	33,959
7 Routine	0	0	0	0	0	16	187	46,615	46,818
Total	**5,043**	**6281**	**24,177**	**21,823**	**15,032**	**22,833**	**33,439**	**47,529**	**176,157**

1 Reduced derivation, children classified according to mother's occupation.

The proportion of the population in off-diagonal cells represents the 'error' in assignment when 1981 occupational information is used to derive NS-SEC, using the simple majority method for assigning NS-SEC to ambiguous cells. The percentage error measured in this way was 1.7 per cent.

A slightly greater error of 2.5 per cent was found when applying a similar method to the dual-coded LS 1991 Census occupation data – the constructed 1981 derivation matrix was used to predict NS-SEC90 codes for sample members based on their occupation code . This scale of error is fairly small relative to the generally accepted levels of routine error in form completion and coding. However, since these errors are systematic rather than random, it is important that they are as small as possible.

The error was reduced to just under 1 per cent when industry codes were used to discriminate in cases where the NS-SEC class was ambiguous.

Tables 2a and 2b show the main results of the study for males at birth and at age 65 respectively, using the conventional method for assignment of socio-economic class, (as used for the 'Trends in life expectancy by Social Class' series). The results are shown both for the seven analytic NS-SEC classes and for the condensed three-class schema. Class 1 (Higher managerial and professional) was split into its two components: 1.1 Large employers and higher managerial and 1.2 Higher professional occupations), since previous studies have found significant differences between the two subgroups. However, the confidence intervals for both classes 1.1 and 1.2 were relatively wide

and so in estimating the range in life expectancy, from most to least advantaged, the whole of NS-SEC class 1 was used in order to reduce the volatility from one period to the next resulting from sampling variance. The life expectancy of those unclassified by occupation is also shown. Owing to the potential change in composition of this group over time and selection effects, it was not used in estimating the range from most to least advantaged.

Figure 1 shows graphically life expectancy at birth by analytic class NS-SEC for men.

Table 3 shows the change in life expectancy for males for each analytical class over the whole study period and between the periods 1997–2001 to 2002–06 to aid comparison of classes over time.

Table 4 shows the corresponding estimates by social class for comparison with those by NS-SEC. These were previously published for the period 1972–2005 and are now updated to include 2006 in the latest reported period.

Tables 5a and 5b, 6 and 7 and Figure 2 show the corresponding information for women.

Table 2a Life expectancy by NS-SEC class, males at birth

England and Wales — Years

NS-SEC	1982–86 LE	1982–86 95% CI (+/-)	1987–91 LE	1987–91 95% CI (+/-)	1992–96 LE	1992–96 95% CI (+/-)	1997–2001 LE	1997–2001 95% CI (+/-)	2002–06 LE	2002–06 95% CI (+/-)
	At birth									
Analytic classes										
1. Higher managerial & professional	75.6	*0.7*	76.6	*0.6*	77.5	*0.6*	78.8	*0.7*	80.4	*0.6*
1.1 Large employers & higher managers	75.1	*1.1*	75.6	*1.0*	77.0	*1.0*	79.0	*1.0*	79.9	*0.9*
1.2 Higher professional	76.5	*1.1*	77.7	*0.9*	78.0	*0.9*	78.5	*0.9*	80.6	*0.8*
2 Lower managerial & professional	74.3	*0.7*	75.4	*0.6*	76.5	*0.6*	78.2	*0.5*	79.6	*0.5*
3 Intermediate	73.3	*0.8*	74.5	*0.8*	75.3	*0.8*	76.8	*0.8*	78.5	*0.8*
4 Small employers & own a/c workers	73.6	*0.8*	74.4	*0.8*	75.6	*0.7*	76.6	*0.7*	77.8	*0.7*
5 Lower supervisory & technical	72.3	*0.6*	73.2	*0.6*	73.8	*0.6*	75.3	*0.6*	76.8	*0.6*
6 Semi-routine	71.3	*0.6*	71.7	*0.6*	72.4	*0.6*	74.0	*0.6*	75.1	*0.6*
7 Routine	70.7	*0.5*	71.5	*0.5*	71.6	*0.5*	72.6	*0.5*	74.6	*0.5*
Range highest- lowest	4.9		5.1		5.9		6.2		5.8	
Condensed NS-SEC										
Managerial & professional	74.8	*0.5*	75.9	*0.4*	77.0	*0.4*	78.4	*0.4*	80.0	*0.4*
Intermediate	73.5	*0.5*	74.5	*0.5*	75.5	*0.5*	76.7	*0.5*	78.1	*0.5*
Routine & manual	71.4	*0.3*	72.0	*0.3*	72.5	*0.3*	73.8	*0.3*	75.4	*0.3*
Range highest- lowest	3.4		3.9		4.5		4.6		4.6	
Unclassified	60.2	*1.4*	60.3	*1.2*	65.8	*1.3*	67.2	*1.2*	71.5	*1.1*
All Men	71.7	*0.2*	72.6	*0.2*	73.8	*0.2*	75.2	*0.2*	77.0	*0.2*

Source: ONS Longitudinal Study

Table 2b Life expectancy by NS-SEC class, males at age 65

England and Wales Years

	1982–86		1987–91		1992–96		1997–2001		2002–06	
NS-SEC	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)
	At age 65									
Analytic classes										
1. Higher managerial & professional	15.2	*0.6*	15.9	*0.5*	16.6	*0.5*	18.1	*0.5*	18.8	*0.4*
1.1 Large employers & higher managers	14.5	*0.8*	15.1	*0.7*	16.2	*0.7*	18.0	*0.7*	18.7	*0.6*
1.2 Higher professional	16.2	*0.8*	16.7	*0.7*	17.0	*0.7*	18.3	*0.6*	18.9	*0.6*
2 Lower managerial & professional	15.1	*0.5*	15.4	*0.4*	16.0	*0.4*	17.2	*0.4*	18.2	*0.4*
3 Intermediate	13.9	*0.5*	15.0	*0.5*	15.7	*0.5*	16.4	*0.5*	17.5	*0.5*
4 Small employers & own a/c workers	14.0	*0.5*	14.7	*0.5*	15.5	*0.5*	16.1	*0.5*	17.5	*0.5*
5 Lower supervisory & technical	13.4	*0.4*	13.7	*0.3*	14.5	*0.3*	15.3	*0.3*	16.4	*0.4*
6 Semi-routine	12.9	*0.4*	13.3	*0.3*	13.8	*0.3*	14.7	*0.4*	15.6	*0.4*
7 Routine	12.9	*0.3*	13.1	*0.3*	13.4	*0.3*	14.0	*0.3*	15.3	*0.3*
Range highest- lowest	2.3		2.8		3.2		4.1		3.5	
Condensed NS-SEC										
Managerial & Professional	15.1	*0.4*	15.6	*0.3*	16.2	*0.3*	17.6	*0.3*	18.4	*0.3*
Intermediate	13.9	*0.4*	14.8	*0.4*	15.6	*0.4*	16.2	*0.3*	17.5	*0.3*
Routine & Manual	13.0	*0.2*	13.3	*0.2*	13.9	*0.2*	14.6	*0.2*	15.8	*0.2*
Range highest- lowest	2.1		2.3		2.4		3.0		2.7	
Unclassified	11.1	*0.3*	10.2	*0.4*	11.2	*0.6*	12.0	*0.7*	14.2	*0.7*
All Men	13.1	*0.1*	13.7	*0.1*	14.5	*0.1*	15.5	*0.1*	16.7	*0.1*

Source: ONS Longitudinal Study

Table 2a shows that for men, there was a clear gradient in life expectancy at birth from the most to the least advantaged, at both birth and age 65, as was seen for RG Social Class.

However, higher professionals, (class 1.2) appeared to have a higher life expectancy than large employers and higher managers (class 1.1) in most periods. Differences between the two classes were not statistically significant for periods from 1992 onwards.

Only the differences between NS-SEC class 4, the self-employed and own account occupations, and class 5 (lower supervisory and technical occupations), and between class 5 and class 6 (semi-routine) were statistically significant in every period, but the pattern of the gradient through time suggests that it is a persistent phenomenon. The largest gap between any two adjacent analytical classes in 2002–06 was between lower supervisory and technical occupations and semi-routine occupations (1.7 years). The range in years between the classes with the highest and the lowest life expectancies at birth widened over time from 4.9 years in 1982–86 to 6.2 years in 1997–2001. The gap narrowed to 5.8 years in 2002–06 but this was not a statistically significant change.

The pattern of the range over time was similar for the 3-class condensed version of NS-SEC, but this did not exhibit narrowing in the latest period. The reason for the difference was that the least advantaged of the analytic classes (Routine occupations) was responsible for a high proportion of the increase in the life expectancy of the 'labour contract' group. Table 3 shows that the Routine occupations class had the second highest increase in life expectancy of any class between the periods 1997–2001 and 2002–06, while the Lower supervisory and technical class and semi-routine occupations class had only average increases. Table 3 also shows that, over the whole study period, the Routine class had the second lowest increase over the whole period and the Routine and manual grouping had the lowest aggregate increase of the condensed NS-SEC classes.

Table 3 Change in life expectancy at birth and at age 65 by NS-SEC, males

England and Wales — Years

NS-SEC	At birth	Change between 1982–86 and 2002–06	Change between 1997–01 and 2002–06
Analytic classes			
1. Higher managerial & professional		4.8	1.6
1.1 Large employers & higher managers		4.8	0.9
1.2 Higher professional		4.1	2.1
2 Lower managerial & professional		5.3	1.4
3 Intermediate		5.2	1.7
4 Small employers & own a/c workers		4.2	1.2
5 Lower supervisory & technical		4.5	1.5
6 Semi-routine		3.8	1.1
7 Routine		3.9	2.0
Condensed NS-SEC			
Managerial & professional		5.2	1.6
Intermediate		4.6	1.4
Routine & manual		4.0	1.7
Unclassified		11.3	4.3
All Men	5.3		1.8
	At age 65		
Analytic classes			
1. Higher managerial & professional		3.6	0.7
1.1 Large employers & higher managers		4.2	0.7
1.2 Higher professional		2.7	0.6
2 Lower managerial & professional		3.1	1.0
3 Intermediate		3.6	1.1
4 Small employers & own a/c workers		3.5	1.4
5 Lower supervisory & technical		3.0	1.1
6 Semi-routine		2.7	0.9
7 Routine		2.5	1.3
Condensed NS-SEC			
Managerial & professional		3.3	0.9
Intermediate		3.6	1.3
Routine & manual		2.8	1.2
Unclassified		3.1	2.2
All Men	3.6		1.3

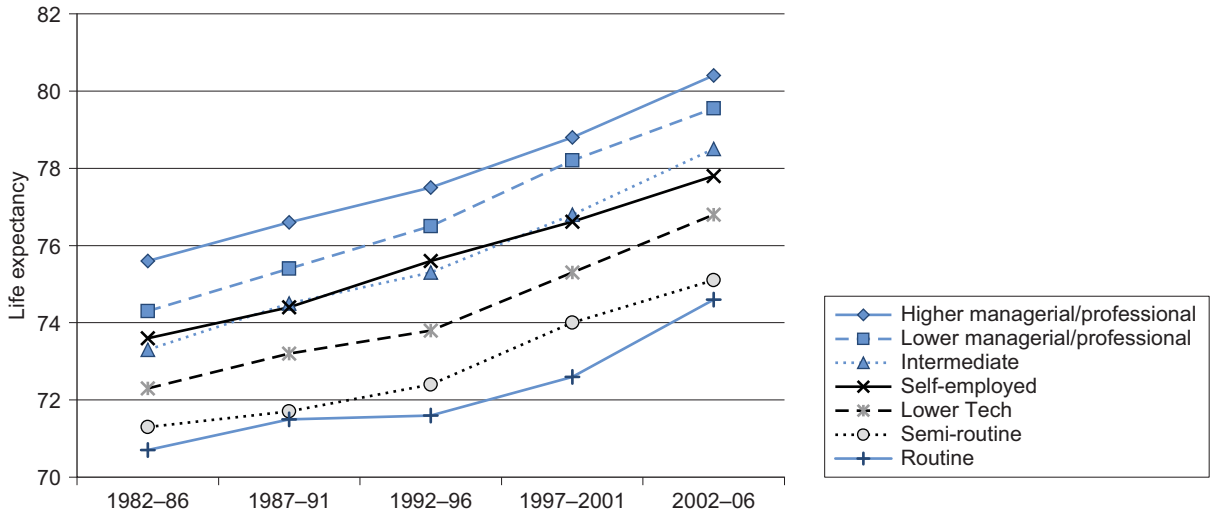
Figure 1 illustrates this pattern. The Intermediate and the self-employed, and own account occupations had the closest path over time. There appear to be clear gaps between occupations with the three major types of employment contract. There was also a persistent gap of between 1.5 and 2 years between Class 5 (Lower supervisory and technical) and Class 6 (Semi-routine).

Table 2b suggests that a similar pattern across the classes was observed age 65 as at birth. The gap widened from 2.3 years in 1982–86 to 4.1 years in 1997–2001 and then narrowed to 3.5 years in 2002–06.

Those unclassified by occupation had consistently lower life expectancy than the analytic classes.

Figure 1 Life expectancy by NS-SEC class, males at birth

England and Wales **Years**



Source: ONS Longitudinal Study

Table 4 Life Expectancy at birth and at age 65 by social class, males

England and Wales — Years

	1972–76		1977–81		1982–86		1987–91		1992–96		1997–2001		2002–06	
Social Class	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95 % CI (+/-)
	At birth													
I Professionals	71.9	*1.2*	74.4	*1.2*	75.1	*1.1*	76.2	*1.0*	77.9	*0.9*	79.5	*0.9*	80.4	*1.0*
II Managerial & Tech	71.9	*0.6*	72.6	*0.6*	74.1	*0.5*	75.0	*0.5*	76.0	*0.5*	77.9	*0.5*	79.4	*0.5*
IIIN Skilled non-manual	69.5	*0.8*	71.1	*0.7*	72.5	*0.7*	74.4	*0.7*	75.2	*0.7*	76.9	*0.7*	78.6	*0.7*
IIIM Skilled manual	70.0	*0.4*	70.0	*0.4*	71.7	*0.4*	72.7	*0.4*	73.7	*0.4*	74.7	*0.4*	76.5	*0.4*
IV Semi-skilled manual	68.3	*0.7*	69.0	*0.6*	71.0	*0.6*	70.8	*0.6*	72.8	*0.6*	73.4	*0.6*	75.7	*0.6*
V Unskilled manual	66.5	*1.1*	67.4	*1.1*	67.7	*1.0*	68.6	*1.0*	68.5	*1.0*	71.3	*0.9*	73.0	*1.1*
Range	5.4		7.0		7.3		7.6		9.4		8.2		7.4	
All men	**69.3**	***0.3***	**70.1**	***0.3***	**71.7**	***0.2***	**72.6**	***0.2***	**73.7**	***0.2***	**75.2**	***0.2***	**77.0**	***0.2***
Non-manual	71.2	*0.4*	72.3	*0.4*	73.7	*0.4*	75.0	*0.4*	76.1	*0.4*	77.8	*0.4*	79.3	*0.4*
Manual	69.1	*0.3*	69.4	*0.3*	71.0	*0.3*	71.7	*0.3*	72.9	*0.3*	74.0	*0.3*	75.9	*0.3*
Difference	**2.1**	***0.5***	**2.9**	***0.5***	**2.7**	***0.5***	**3.3**	***0.5***	**3.2**	***0.5***	**3.8**	***0.5***	**3.4**	***0.5***
	At age 65													
I Professionals	14.0	*0.9*	15.5	*0.9*	15.5	*0.8*	15.8	*0.7*	17.0	*0.7*	18.2	*0.6*	18.6	*0.6*
II Managerial & Tech	13.3	*0.4*	14.2	*0.3*	14.5	*0.3*	15.0	*0.3*	15.6	*0.3*	17.0	*0.3*	18.1	*0.3*
IIIN Skilled non-manual	12.6	*0.4*	13.3	*0.4*	13.6	*0.4*	14.3	*0.4*	15.3	*0.4*	16.6	*0.4*	17.5	*0.5*
IIIM Skilled manual	12.2	*0.2*	12.5	*0.2*	13.1	*0.2*	13.6	*0.2*	14.3	*0.2*	15.2	*0.2*	16.3	*0.3*
IV Semi-skilled manual	12.2	*0.3*	12.1	*0.3*	12.7	*0.3*	12.8	*0.3*	13.9	*0.3*	14.0	*0.3*	15.8	*0.4*
V Unskilled manual	11.6	*0.4*	11.8	*0.4*	11.6	*0.4*	12.0	*0.5*	12.5	*0.5*	13.1	*0.5*	14.5	*0.7*
Range	*2.5*		3.7		3.9		3.7		4.5		5.1		4.1	
All men	**12.3**	***0.1***	**12.6**	***0.1***	**13.1**	***0.1***	**13.7**	***0.1***	**14.5**	***0.1***	**15.5**	***0.1***	**16.7**	***0.2***
Non-manual	13.1	*0.3*	14.0	*0.3*	14.3	*0.2*	14.9	*0.2*	15.7	*0.2*	17.1	*0.2*	18.0	*0.3*
Manual	12.1	*0.2*	12.3	*0.2*	12.7	*0.2*	13.2	*0.2*	14.0	*0.2*	14.6	*0.2*	16.0	*0.2*
Difference	**1.0**	***0.3***	**1.7**	***0.3***	**1.6**	***0.3***	**1.7**	***0.3***	**1.7**	***0.3***	**2.5**	***0.3***	**2.0**	***0.3***

Source: ONS Longitudinal Study

Comparison with Table 4 showing life expectancy by social class for males, shows that the ranges are similar for both NS-SEC and RG social class. The range between the most and least advantaged classes is not as great for NS-SEC as for social class, (5.8 years in 2002–06 for NS-SEC compared with 7.4 years for social class). The main reason is that Social Class V, 'unskilled manual occupations', is a smaller class than 'Routine occupations' in the NS-SEC schema. In both cases there is a narrowing of the range in the most recent period. This narrowing was statistically significant for social class at age 65. At birth the widest gap between social classes I and V was in 1992–96, rather than 1997–2001. This was almost certainly a result of the small size and consequently wide confidence interval (in excess of one year) for social class V.

Table 5a shows that there was also a persistent socio-economic gradient for females. As for men, higher professionals tended to have a higher life expectancy than higher managers, although the difference between these classes was only statistically significant for the most recent period. The range between the highest and lowest life expectancy was smaller than for men (4.2 years at birth and 3.2 years at age 65, compared with 5.8 years and 3.5 years for men in the period 2002–06).

Table 5a Life expectancy by NS-SEC class, females at birth

England and Wales — Years

	1982–86		1987–91		1992–96		1997–2001		2002–06	
NS-SEC	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)
	At birth									
Analytic classes										
1. Higher managerial & professional	80.9	*1.1*	81.7	*1.0*	82.3	*0.9*	82.6	*0.8*	83.9	*0.7*
1.1 Large employers & higher managers	79.9	*1.7*	82.3	*1.5*	82.7	*1.3*	81.9	*1.2*	83.0	*1.1*
1.2 Higher professional	82.1	*1.5*	81.5	*1.3*	82.0	*1.3*	83.0	*1.1*	84.8	*1.0*
2 Lower managerial & professional	79.7	*0.7*	81.0	*0.6*	81.2	*0.5*	82.2	*0.5*	83.4	*0.5*
3 Intermediate	79.6	*0.7*	81.1	*0.7*	81.4	*0.6*	81.5	*0.6*	82.7	*0.6*
4 Small employers & own a/c workers	79.1	*1.0*	79.9	*0.9*	80.7	*0.9*	80.8	*0.8*	82.6	*0.8*
5 Lower supervisory & technical	78.5	*0.9*	78.1	*0.8*	79.4	*0.7*	79.5	*0.7*	80.4	*0.7*
6 Semi-routine	78.1	*0.6*	78.5	*0.6*	79.2	*0.6*	79.6	*0.5*	80.6	*0.6*
7 Routine	77.1	*0.6*	77.5	*0.6*	78.3	*0.5*	78.6	*0.5*	79.7	*0.5*
Range highest- lowest	3.8		4.2		4.0		4.0		4.2	
Condensed NS-SEC										
Managerial & professional	80.1	*0.6*	81.2	*0.5*	81.5	*0.5*	82.3	*0.4*	83.5	*0.4*
Intermediate	79.6	*0.6*	80.7	*0.5*	81.1	*0.5*	81.4	*0.4*	82.7	*0.4*
Routine & manual	77.7	*0.4*	78.0	*0.4*	78.9	*0.3*	79.2	*0.3*	80.2	*0.3*
Range highest- lowest	2.4		3.2		2.6		3.1		3.3	
Unclassified	71.5	*1.1*	73.1	*1.0*	74.2	*1.2*	75.8	*0.9*	76.9	*0.9*
All Women	77.4	*0.2*	78.3	*0.2*	79.2	*0.2*	79.9	*0.2*	81.1	*0.2*

Source: ONS Longitudinal Study

Table 5b Life expectancy by NS-SEC class, females at age 65

England and Wales — Years

	1982–86		1987–91		1992–96		1997–2001		2002–06	
NS-SEC	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)
	At age 65									
Analytic classes										
1. Higher managerial & professional	19.7	*1.0*	20.1	*0.8*	20.8	*0.7*	20.9	*0.6*	21.7	*0.6*
1.1 Large employers & higher managers	19.4	*1.5*	20.6	*1.4*	20.9	*1.1*	20.4	*0.9*	21.1	*0.8*
1.2 Higher professional	20.0	*1.3*	19.8	*1.0*	20.6	*1.0*	21.4	*0.9*	22.2	*0.8*
2 Lower managerial & professional	18.9	*0.6*	19.4	*0.5*	19.3	*0.4*	20.3	*0.4*	21.1	*0.4*
3 Intermediate	18.3	*0.6*	19.5	*0.5*	19.6	*0.4*	19.8	*0.4*	20.5	*0.4*
4 Small employers & own a/c workers	18.6	*0.8*	18.9	*0.7*	19.6	*0.6*	19.1	*0.6*	20.5	*0.6*
5 Lower supervisory & technical	18.7	*0.7*	17.6	*0.5*	18.3	*0.5*	18.3	*0.5*	18.8	*0.5*
6 Semi-routine	17.4	*0.5*	17.6	*0.4*	18.3	*0.4*	18.4	*0.3*	19.4	*0.3*
7 Routine	16.7	*0.4*	17.4	*0.4*	17.7	*0.3*	17.8	*0.3*	18.5	*0.3*
Range highest- lowest	3.0		2.7		3.1		3.1		3.2	
Condensed NS-SEC										
Managerial & professional	19.1	*0.5*	19.6	*0.4*	19.7	*0.4*	20.5	*0.3*	21.3	*0.3*
Intermediate	18.4	*0.5*	19.3	*0.4*	19.6	*0.4*	19.6	*0.3*	20.5	*0.3*
Routine & manual	17.3	*0.3*	17.5	*0.2*	18.0	*0.2*	18.1	*0.2*	18.9	*0.2*
Range highest- lowest	1.7		2.1		1.6		2.3		2.4	
Unclassified	16.2	*0.2*	16.2	*0.3*	16.3	*0.3*	16.6	*0.4*	17.4	*0.5*
All Women	17.0	*0.1*	17.5	*0.1*	18.0	*0.1*	18.5	*0.1*	19.5	*0.1*

Source: ONS Longitudinal Study

The range did not show the same pattern as for men, with no identifiable trend and no sign of a narrowing between the most and least advantaged. The greatest relative change among classes appeared to be the growth of the life expectancy of Classes 2 (Lower professional) and 4 (Self-employed and own-account occupations). The latter class closed the gap between it and the intermediate class from 1.2 years in 1987–91 to a negligible 0.1 years in 2002–06. As for men, there was a clear difference between the 'labour contract' group, (the least advantaged), and the intermediate group, but in the case of women, this difference was greater as a proportion of the overall range across the classes. Table 6 suggests that the three labour contract analytic classes were those with the least increase in life expectancy over the study period.

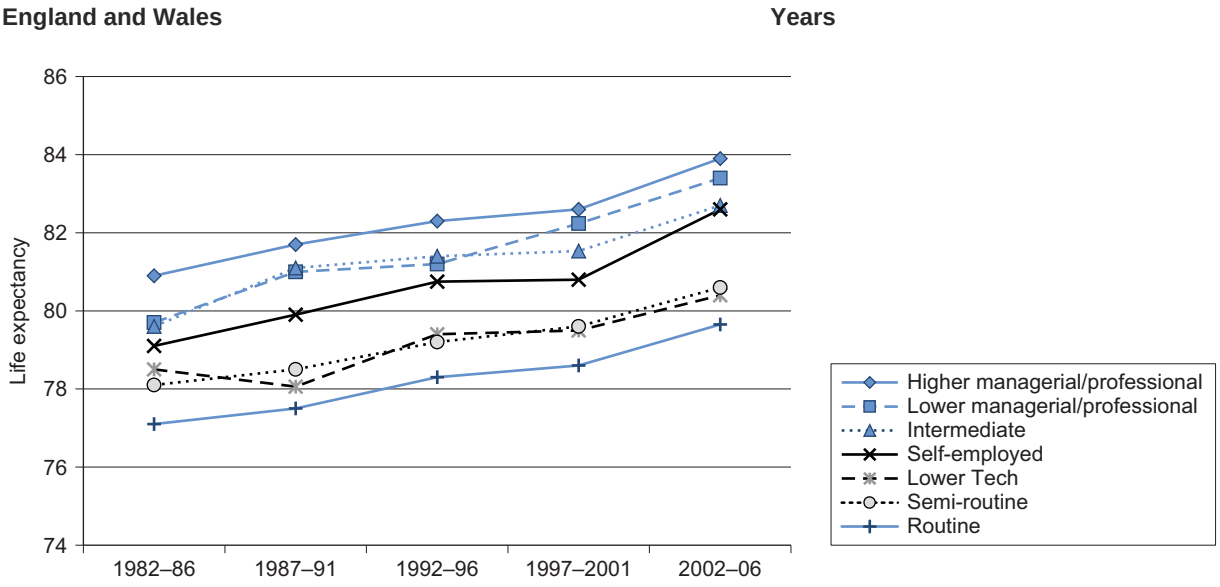
Table 6 Change in life expectancy at birth and at age 65 by NS-SEC, females

England and Wales — Years

NS-SEC	At birth	Change between 1982–86 and 2002–06	Change between 1997–2001 and 2002–06
Analytic classes			
1. Higher managerial & professional		3.0	1.3
1.1 Large employers & higher managers		3.1	1.1
1.2 Higher professional		2.7	1.8
2 Lower managerial & professional		3.7	1.2
3 Intermediate		3.1	1.2
4 Small employers & own a/c workers		3.5	1.8
5 Lower supervisory & technical		1.9	0.9
6 Semi-routine		2.5	1.0
7 Routine		2.6	1.1
Condensed NS-SEC			
Managerial & professional		3.4	1.2
Intermediate 3.1			1.3
Routine & manual		2.5	1.0
Unclassified 5.4			1.1
All Women		3.7	1.2
	At age 65		
8 analytic classes			
1. Higher managerial & professional		2.0	0.8
1.1 Large employers & higher managers		1.7	0.7
1.2 Higher professional		2.2	0.8
2 Lower managerial & professional		2.2	0.8
3 Intermediate		2.2	0.7
4 Small employers & own a/c workers		1.9	1.4
5 Lower supervisory & technical		0.1	0.4
6 Semi-routine		2.0	1.0
7 Routine		1.8	0.7
Condensed NS-SEC			
Managerial & professional		2.2	0.8
Intermediate 2.1			0.8
Routine & manual		1.6	0.8
Unclassified 1.2			0.8
All Women	2.5		0.9

Figure 2 illustrates the relative increase in the life expectancy of the self-employed and own account workers and the persistent gap between the labour contract occupations and the rest.

Figure 2 Life expectancy by NS-SEC class, females at birth



Source: ONS Longitudinal Study

Table 7 Life Expectancy at birth and at age 65 by social class, females

England and Wales — Years

	1972–76		1977–81		1982–86		1987–91		1992–96		1997–2001		2002–06	
Social Class	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)	LE	95% CI (+/-)
	At birth													
I Professionals	79.0	*1.9*	80.1	*1.7*	80.7	*1.3*	81.1	*1.1*	83.8	*1.1*	82.5	*1.0*	85.2	*1.1*
II Managerial & tech	77.1	*0.6*	78.3	*0.6*	78.8	*0.5*	80.7	*0.5*	81.3	*0.5*	81.7	*0.5*	83.1	*0.5*
IIIN Skilled non-manual	78.3	*0.7*	78.2	*0.7*	79.0	*0.6*	80.0	*0.5*	80.7	*0.5*	81.6	*0.5*	82.5	*0.5*
IIIM Skilled manual	75.2	*0.6*	76.3	*0.5*	77.3	*0.5*	77.9	*0.5*	79.1	*0.4*	79.4	*0.4*	80.6	*0.5*
IV Semi-skilled manual	75.3	*0.7*	75.9	*0.6*	77.5	*0.6*	77.4	*0.5*	78.1	*0.5*	78.7	*0.5*	79.9	*0.6*
V Unskilled manual	74.2	*1.2*	75.6	*1.0*	75.9	*0.9*	76.6	*0.9*	77.4	*0.9*	77.8	*0.9*	78.3	*1.2*
Range	4.8		4.4		4.8		4.5		6.4		4.7		6.9	
All women	75.3	*0.3*	76.5	*0.2*	77.4	*0.2*	78.3	*0.2*	79.1	*0.2*	79.9	*0.2*	81.1	*0.2*
Non-manual	77.7	*0.4*	78.4	*0.4*	79.0	*0.4*	80.4	*0.3*	81.2	*0.3*	81.7	*0.3*	82.9	*0.3*
Manual	75.2	*0.4*	76.0	*0.4*	77.2	*0.3*	77.5	*0.3*	78.4	*0.3*	78.9	*0.3*	80.0	*0.3*
Difference	2.5	*0.6*	2.4	*0.5*	1.9	*0.5*	2.9	*0.5*	2.8	*0.4*	2.8	*0.4*	2.9	*0.5*
	At age 65													
I Professionals	19.1	*1.7*	20.0	*1.5*	18.9	*1.1*	19.0	*0.8*	21.1	*0.9*	20.7	*0.8*	22.0	*0.9*
II Managerial & tech	17.2	*0.4*	17.8	*0.4*	18.2	*0.4*	19.2	*0.3*	19.7	*0.3*	20.0	*0.3*	21.0	*0.3*
IIIN Skilled non-manual	17.9	*0.6*	17.7	*0.4*	18.2	*0.4*	18.7	*0.3*	19.2	*0.3*	19.8	*0.3*	20.1	*0.3*
IIIM Skilled manual	16.4	*0.5*	16.9	*0.4*	16.9	*0.3*	17.2	*0.3*	18.1	*0.3*	18.2	*0.3*	18.8	*0.4*
IV Semi-skilled manual	16.9	*0.4*	16.8	*0.3*	17.5	*0.3*	17.3	*0.3*	17.4	*0.3*	17.9	*0.3*	19.0	*0.3*
V Unskilled manual	16.6	*0.7*	16.4	*0.6*	16.2	*0.5*	16.3	*0.5*	16.6	*0.5*	16.8	*0.5*	17.7	*0.6*
Range	2.5		3.6		2.7		2.7		4.5		3.9		4.2	
All women	16.3	*0.1*	16.7	*0.1*	17.0	*0.1*	17.5	*0.1*	18.0	*0.1*	18.5	*0.1*	19.5	*0.2*
Non-manual	17.5	*0.3*	17.9	*0.3*	18.2	*0.2*	18.9	*0.2*	19.5	*0.2*	19.9	*0.2*	20.6	*0.2*
Manual	16.6	*0.3*	16.8	*0.2*	17.0	*0.2*	17.1	*0.2*	17.5	*0.2*	17.8	*0.2*	18.7	*0.2*
Difference	0.9	*0.4*	1.1	*0.4*	1.2	*0.3*	1.8	*0.3*	2.0	*0.3*	2.1	*0.3*	1.9	*0.3*

Source: ONS Longitudinal Study

In comparison with the results by social class, shown in Table 7, the range between the most and least advantaged was smaller and more stable for NS-SEC than for social class, again possibly because of the small size and greater variation in social class V. Condensed NS-SEC showed a similar range and pattern to the non-manual/manual divide by social class at birth, although the pattern at age 65 was slightly different.

The results of using a 'combined' measure of classification of spouses or fathers and mothers of children are shown in Tables A1 and A2 of the Appendix. The main effect of using the combined measure of NS-SEC was to widen the gap slightly between the most and the least advantaged, both for men and for women. In 2002–06 this widening was 0.2 years for males at birth and 1.1 years for females. This was a result of the decrease in the life expectancy estimates of the Routine class under the 'combined' rule in both cases. The narrowing of the range between highest and lowest in the most recent period was more pronounced for men under the combined classification, reducing from 6.8 years to 6.0. For men at age 65, the reduction in the gap in the most recent period was from 4.4 years to 3.8 years.

For women, the effect was to accentuate the apparent increase in range in the most recent period Otherwise, the effect on the estimates of the combined classification was small.

Discussion

These results continue the pattern shown in previously published estimates of trends in life expectancy by RG social class. There was a clear social gradient in the life expectancy estimates for both males and females, which persisted through the period 1982–86 to 2002–06. The overall range between the most advantaged (higher managers or higher professionals), and the least advantaged (Routine occupations), was slightly less than the range across the occupied RG social classes for 2002–06, (around six years compared with over seven years for RG social class for males at birth). The range appeared to increase in the 1990s and, for males, to narrow slightly in the early part of the 21st Century. The range in life expectancy at birth between NS-SEC class 1 (Higher managerial and professional occupations) and class 7 (Routine occupations) fell from 6.2 years in 1997–2001 to 5.8 years in 2002–06, and from 4.1 years to 3.5 years at age 65. While this only represents a return to the inequality of the 1990s, it does appear to be a change from the earlier trend between 1982-86 and 1997–2001. It is dangerous to attach too much significance to the figures of a single period, but these results are consistent with the hypothesis that inequalities were no longer rising for males in the early 2000s, as they had done up to 1997–2001. It would be necessary to observe another period with a substantial relative improvement in life expectancy for the Routine class before there could be confidence that a turning point had been reached for inequalities in male mortality. The relative improvement in the Routine class for men was consistent with the findings of Langford and Johnson (2010), using a cross-sectional method with the Labour Force Survey forming the basis for population denominators. This suggested that mortality rates at working age for Routine class males fell by the most of any of the NS-SEC analytic classes between 2001 and 2008.

No corresponding pattern was visible for females, for whom those assigned to Routine and manual occupations seemed to continue to decline in relative life expectancy.

Lower managerial and professional occupations was the class with the greatest increase in life expectancy at birth over the study period for both males (5.3 years) and females (3.7 years). The Semi-routine class had the lowest increase for men (3.8 years) and the Lower supervisory and technical class had the lowest increase for women (1.9 years).

Those not classified by NS-SEC had a lower expectation of life than those who could be assigned a class based occupation. This group includes full-time students; the long-term unemployed; those who never worked; and those whose occupation was inadequately described or were unclassified for other reasons. Since the method used should be capable of classifying those who were unemployed for a short time or recently retired, and since spouse's class was used if the sample member could not be classified, and father or mother's class was used for children, a high proportion of those who remain unclassified could be thought to be composed of the least advantaged in society, since there is no link to the labour market apparent from available sources.

It is not surprising, therefore, that the unclassified group had low life expectancy estimates. However there is almost certainly a selection effect operating. Unclassified men had a life expectancy of 60.2 years in 1982–86. There was a negligible increase to 60.3 in 1987–91, but then the life expectancy jumped to 65.8 in 1992–96. Fox *et al.* (1985) showed that in the period from 5 to 10 years post-classification, the manual social classes had substantially higher mortality and the unoccupied lower mortality compared with the period immediately post-classification. In the current study, selection seemed to have a longer lasting effect, whereby it appeared to take at least 10 years for the effect of those with incipient ill health to wear off via their death or recovery. Changes in the relative life expectancy of the unclassified may also be partly explained by changes in the composition of the group over time. However, the trend over time in the life expectancy of the unclassified is not adequately explained and further analysis is necessary.

Differences within NS-SEC class 1

A slightly unexpected result was the higher life expectancy for higher professionals (NS-SEC class 1.2) than for higher managers (class 1.1). Cross-sectional studies using the 2001 Census and death registrations have usually found lower mortality (at working age) for class 1.1. A test of the impact of each occupation which was mainly classified to classes 1.1 and 1.2 did not suggest that any one occupation had a disproportionate effect on the relative life expectancies of the two groups. Fitzpatrick (2003) found that use of reduced NS-SEC had a substantial effect on the size and composition of class 1.1. As a result, size of establishment was included as a variable where available, as a sensitivity test for the current study to approximate full NS-SEC and substituted for the reduced NS-SEC used to produce the life expectancy estimates. However the greater life expectancy of higher professionals relative to managers persisted irrespective of the inclusion of establishment size and industry data.

The confidence intervals around the estimates for Classes 1.1 and 1.2 were relatively wide at (or greater than) a year, and this can produce volatility in the ranking of the estimates. The persistence of the phenomenon makes it more likely that it is an artefact of the classification system used in 1991 and hence 1981 rather than a real reversal of the survival probabilities of managers and professionals between 1981 and the 21st Century. It should also be remembered that with the Registrar General's social class schema, class I was 'Professionals' and class II was 'managerial', (and subsequent to 1991 'managerial and technical' (OPCS, 1991)), and historically, class I tended to have higher life expectancy than class II.

Assignment of class

The combined approach to classification was based on the concept of 'dominance' in the labour market, as proposed by Erikson (1984). The main effect of using the 'combined' measure of NS-SEC was to widen the gap slightly between the most and least advantaged both for men and women. This increase in gradient was observed in the analysis of mortality rates of women of working age using a 'combined' measure in relation to using women's own NS-SEC (Langford and Johnson, 2009). All measures which attempt to include members of the household other than the sample member tend to have the effect of selecting from the more disadvantaged classes, those who are at the most socio-economic disadvantage, and often those who live alone. This tends to increase the average mortality rates of the least advantaged group. Nevertheless, estimates for most classes were not significantly different using the combined method for either males or females.

This contrasts with the cross-sectional analysis of female mortality at working ages for 2001–03 (Langford and Johnson, 2009), where applying such a combined approach tended to steepen the gradient substantially for all causes of death. This is not surprising since, in the latter case, the method allowed women to be classified by occupation at death where otherwise they would not have been. In the current study, the combined method did not bring people into analytic NS-SEC categories who would have otherwise remained unclassified, but simply reordered the hierarchy of rules by which a class was assigned.

Other methods of assignment to an NS-SEC class were considered. It has been suggested that classifying someone by their occupation-based class at, for example, age 22 might not adequately express a person's access to resources. In addition to a natural trajectory that occurs in some careers, there is a degree of social mobility which is not accounted for when classifying people at entry to the study. Donkin *et al*. (2002) attempted to classify sample members at later censuses. They found that for those classified in 1991 (the most recent census at the time) a high proportion could not be assigned a class since the census coded occupational information only for those under 75 years old. This, when combined with the clear evidence of health selection – 'life expectancy immediately following classification appears to be enhanced for those in social class V' (the least advantaged), tends to militate against constantly updating classifications in longitudinal studies of this type. The Black report (1980, p 358) suggested that …childhood illness 'cast long shadows forward'. Power *et al*. (2005) indicated the link between socio-economic position in childhood and early life and the risk of mortality. Blane *et al.* (1999) observed that the upwardly mobile tend to have higher mortality and the downward mobile tend to have lower mortality than the average for their destination classes. White *et al*. (2005) found that social class in 1971 continued to have a strong influence on mortality risk in 1995–2001, after controlling for interim social mobility and unemployment, level of neighbourhood deprivation, and other factors. However, these authors acknowledge that it is an accumulation of disadvantage through the lifecourse which perpetuates these inequalities and that social mobility plays some role in health outcomes.

One compromise might be to update the classification for all those present at a census and capable of being assigned an occupation based socio-economic class, while retaining the existing class for those who are unclassified at the latest census. This would not completely eradicate any selection bias in that someone might have changed to a less demanding occupation owing to poor

health, but it would not exacerbate the main problem - that arising from the effect of those selected out of the workforce altogether for health reasons.

Use of occupation-based measures

All methods of assigning socio-economic position based on occupation are subject to the weakness of occupation in describing access to material and social resources. NS-SEC has provided a conceptual basis for an occupation-based system whereby an explanation can be made of differences in mortality and survival, based on job stress, type of employment contract and other job-related factors, but is not designed to evaluate the role of social capital and other non-occupation based factors, especially for women.

There is the additional problem that occupational structures are always changing. The current study required a construct of NS-SEC for 1981 occupations for which it was not designed. The empirical validity of this has been discussed in this article. While there is also a conceptual problem, it must be remembered that social class – while giving the appearance of continuity – had to change over time, owing to the existence of new occupations, the deskilling of others and various other structural changes. It was found that approximately 14 per cent of the 1971 Census cohort of men between ages 45 and 64 would have been assigned a different social class by the 1981 Census without any change of occupation (Goldblatt and Whitehead, 2000). Thus, any occupation-based measure is mutable in some way.

Educational attainment might be a better all round measure of 'social capital', and studies in Europe in particular have made use of these measures (Mackenbach *et al.,* 2008, Corsini, 2010). However, these too have shortcomings, with types of qualification and the prevalence of qualifications changing over time. There is also a practical problem in England and Wales, with very limited educational attainment data being available from the 1981 and 1991 Censuses.

Future developments

The occupational classification system has recently been revised in time for the 2011 Census. It is not feasible to set out the impact of these changes here, but the main effects will be to reduce the number of occupations which qualify as 'managerial', and to increase the number of lower professionals. It is reasonable to assume that some further degree of discontinuity will result. However, this was always the case with Registrar General's social class and as NS-SEC becomes more recognised and understood, comparisons over time will become less problematic.

Limitations of the analysis

All occupation-based classifications change over time with changes in technology and economic structures. NS-SEC is no exception to this. Donkin et al. (2002) found only 84 per cent agreement at the analytic class level used in this study between NS-SEC90 and NS-SEC2001, when applied to dual coded 2001 first quarter deaths. Of the 16 per cent discrepancy, approximately 9 per cent was a result of occupational class changes and 7 per cent to new employment status coding. The latter was more likely to record the deceased as a supervisor than earlier systems. This had a particular impact on Classes 2 and 5 (lower managerial and professional and lower supervisory and technical), where the supervisor classification often raises the subject from the Intermediate or

Semi-routine classes respectively. While the effect of such changes on a longitudinal study such as the current one are less than the numerator-denominator problems occurring for cross-sectional studies, the changes in classification rules at each census detract to some extent from the consistency of the results over time.

Any errors arising from the non-unique mapping between 1981 and 1991 occupation codes are not random. These are small however, as has been shown.

There was substantial social mobility over time, as denoted by sample members' own NS-SEC as measured in 1981 and 2001, for those who were at both censuses. While there is much evidence referenced in this paper, to suggest that early years have the greatest effect on health outcomes, subsequent employment and other experiences, including social mobility play a role in the accumulation of relative advantage and disadvantage (Marmot Review, 2010).

Immigrants and people who move frequently in and out of the country are under-represented in the analysis. Immigrants could only be assigned an NS-SEC class if they were present at a census. Therefore, many people, in particular those who entered the country since 2001, could not be assigned an occupation-based class.

The selection effect (referred to in the Discussion) may have depressed the range across the analytic classes in the first and possibly the second period analysed.

Conclusions

The analysis suggests that an approximation to NS-SEC for the 1981 Census can be constructed to produce a valid series of life expectancy by socio-economic class over time.

A clear social gradient persists over time, with remarkable consistency. The only minor exception is the relative ranking of Higher managers and Higher professionals for whom the ranking changed in one period, although Higher professionals were estimated to have the highest life expectancy in most periods. For men, the range from most to least advantaged class was approximately six years at birth and for women four years, both slightly less than the range when using the former occupation based measure of Registrar General's social class. For men, inequality appeared to increase from the mid-1980s and then to decrease slightly in the period 2002–06. However, it is too soon to say whether this represents a turning point for men. It is worth noting however, that the Routine class, the most disadvantaged of the classes whose members could be assigned by a current or former occupation, showed the second greatest improvement in life expectancy of all the classes in the most recent period.

For women there was no sign of a decline in inequality, either at birth or at age 65. Women classified as lower professional or as self-employed or own account workers had the greatest increase in life expectancy since the mid-1980s.

Using the summary three-class NS-SEC the managerial and professional class had the highest life expectancy, and had experienced the greatest growth in life expectancy since the start of the study period. The reverse was the case for the Routine and manual class. The intermediate group tended to have life expectancy estimates closer to those for managerial and professional occupations than to those for routine and manual occupations.

Comparison with the social class measure of life expectancy suggested that the range from highest to lowest exhibited less volatility from one period to the next using NS-SEC. Results obtained using the three-class condensed NS-SEC were broadly similar in range and pattern to those obtained using the non-manual and manual groupings for social class.

The degree of social mobility between 1981 and 2001 suggests that, while there were good reasons to adhere to the conventional early assignment of class as described in this article, further investigation should be made into a conditional revision of class of a sample member at the beginning of each decade, but structured to avoid the worst effects of health selection.

Owing to the difficulties involved in classifying people by occupation – especially for women, some of whom may have a weaker attachment to the labour market – other methods of classification are being sought. This study attempted a 'combined' measure for husbands and wives which attributed the most advantageous class for a couple, but this did not make a substantive difference to the results. It also does not reflect the range of modern family structures. Work on household-based measures and those which better reflect a person's more recent occupational status is continuing.

An alternative approach may be to investigate educational attainment rather than occupation as a proxy for social capital, but this too has associated problems.

Subject to consultation, the series – based on social class – which is being published simultaneously with this article will be discontinued and the NS-SEC series updated periodically.

Acknowledgements

This work is an extension of that first published by Lin Hattersley and developed by Peter Goldblatt, Angela Donkin, Kevin Lynch, Alec Ross, Bola Akinwale and Louisa Blackwell among others. The author has built upon the work of these people, but is obviously responsible for any errors in the current analysis. The author is also grateful for the support of various colleagues during the course of this work.

References

Acheson D (1998) Independent inquiry into inequalities in health, *The Stationery Office*: London

Black D, Morris J, Smith C and Townsend P (1980) Inequalities in health: report of a research working group. *Department of Health and Social Security*: London. (P.358)

Blane D, Harding S and Rosato M (1999) 'Does social mobility affect the size of the socio-economic mortality differential?: evidence from the Office for National Statistics Longitudinal Study.

Chiang C L (1968) 'Introduction to Stochastic Processes in Biostatistics', New York: *John Wiley & Sons, Inc.,* Chapter 9, pp 189–214

Corsini V (2010) 'Highly educated men and women live longer' Eurostat, Luxembourg

Department of Health (2010) Post-2010 strategic review of health inequalities (*The Marmot Review*)

Donkin A, Goldblatt P and Lynch K (2002) 'Inequalities in life expectancy by social class, 1972–99' *Health Statistics Quarterly* 15, 5–15. nswebcopy/StatBase/Product.asp?vlnk=15342

Donkin A, Lee YH and Toson B (2002) 'Implications of changes in the UK social and occupational classifications in 2001 for vital statistics', *Population Trends* 107, 23–34

Elias P (1997) 'Social Class and the Standard Occupational Classification' in Rose D and O'Reilly K (eds) Constructing Classes – towards a new social classification for the UK, ESRC/ONS, Swindon

Erikson R (1984) 'Social class of men, women and families', *Sociology*, 18, 500–14

Fitzpatrick J (2003) 'Examining mortality rates by the NS-SEC using death registration data and the 1991 Census', in Rose D and Pevalin D (eds) A Researcher's Guide to the National Statistics Socio-economic Classification*. SAGE Publications Ltd.*

Fox A, Goldblatt P and Jones D (1985) 'Social class mortality differentials: artefact, selection or life circumstances?' *Journal of Epidemiology and Community Health,* 39, 1–8

Goldblatt P and Whitehead M (2000) 'Inequalities in health – development and change', *Population Trends* 100, 13–19

Goldthorpe J (1997) 'The Goldthorpe' Class Schema: some observations on conceptual and operational issues in relation to the ESRC review of government social classifications', in Rose D and O'Reilly K (eds) *Constructing Classes: Towards a new social classification for the UK.* Swindon: ESRC/ONS.

Hattersley L (1997), Expectation of life by social class in Drever F and Whitehead M (eds) Health Inequalities: Decennial Supplement Series DS No. 15, *The Stationery Office*: London

Johnson B and Blackwell L (2007) 'Review of methods for estimating life expectancy by social class using the ONS Longitudinal Study' *Health Statistics Quarterly* 35, 28–36.

Langford A and Johnson B (2009) 'Social inequalities in female mortality by the National Statistics Socio-economic Classification, England and Wales, 2001–03', Health Statistics Quarterly 42, 6–21

Langford A and Johnson B (2010) 'Trends in social inequalities in male mortality, 2001–08. Intercensal estimates for England and Wales'. Health Statistics Quarterly 47, 5–32.

Mackenbach J, Stirbu I, Roskam A-J, Schaap M, Menvielle G, Leinsalu M and Kunst A (2008) 'Socio-economic inequalities in health across 22 European Countries', *New England Journal of Medicine,* 358 2468–81

Office for National Statistics (2002) The National Statistics Socio-economic Classification. User Manual, on-line edition. Available on the Office for National Statistics website: statistics.gov.uk/statbase/Product.asp?vlnk=13561&Pos=1&ColRank=2&Rank=816

Office for National Statistics report (2007) 'Trends in Life Expectancy by social class 1972–2005'.Available on the Office for National Statistics website at: statistics.gov.uk/statbase/Product.asp?vlnk=8460

Office of Population Censuses and Surveys (1978) *'Occupational Mortality 1970-72. Decennial Supplement'* Series DS no 1, HMSO

Office of Population Censuses and Surveys (1991) '*Standard Occupational Classification'* Volume 3, London: HMSO

Power C, Hypponen E and Davey Smith G (2005) 'Socio-economic position in childhood and early adult life and risk of mortality: A prospective study of the mothers of the 1958 British Birth Cohort', *American journal of Public Health* 95, No 8 1396–1402

Rose D (1995) *Official social classifications in the UK,* Social Research Update, Issue 9, University of Surrey

Rose D and Pevalin D (Eds) (2003) A Researcher's Guide to the National Statistics Socio-economic Classification. *SAGE Publications Ltd.*

Stata statistical software (Release 9) – Statacorp, College Station, TX, USA

White C, Edgar G and Siegler V (2008) 'Social inequalities in male mortality for selected causes of death by the National Statistics Socio-economic Classification, England and Wales, 2001–03', *Health Statistics Quarterly* 38, 19–31. Available on the Office for National Statistics website at: statistics.gov.uk/statbase/Product.asp?vlnk=6725

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: statistics.gov.uk/STATBASE/Product.asp?vlnk=6725

White C, Wiggins R, Blane D, Whitworth A and Glickman M (2005) 'Person, place or time? The effect of individual circumstances, area and changes over time on mortality in men, 19952001',

Health Statistics Quarterly 28, 18–26. Available on the Office for National Statistics website: statistics.gov.uk/STATBASE/Product.asp?vlnk=6725

Appendix

Table A1a Life expectancy by NS-SEC class, males at birth, using a 'combined' classification

England and Wales — Years

	1982–86		1987–91		1992–96		1997–2001		2002–06	
	LE	95% CI +/-	LE	95% CI +/-	LE	95% CI +/-	LE	95% CI +/-	LE	95% CI +/-
	At birth									
Analytic classes										
1. Higher managerial & professional	75.7	*0.7*	76.6	*0.7*	77.6	*0.6*	78.7	*0.6*	80.4	*0.6*
1.1 Large employers & higher managers	75.1	*1.0*	75.5	*1.0*	76.9	*0.9*	78.7	*0.9*	79.9	*0.9*
1.2 Higher professional	76.5	*1.1*	77.8	*0.9*	78.2	*0.9*	78.6	*0.9*	80.7	*0.8*
2 Lower managerial & professional	74.4	*0.6*	75.2	*0.6*	76.3	*0.5*	77.8	*0.5*	79.2	*0.5*
3 Intermediate	73.7	*0.6*	74.8	*0.7*	75.5	*0.6*	76.8	*0.6*	77.9	*0.7*
4 Small employers & own a/c workers	73.2	*0.8*	74.2	*0.8*	75.3	*0.8*	76.4	*0.7*	77.7	*0.8*
5 Lower supervisory & technical	71.9	*0.7*	72.9	*0.7*	73.2	*0.7*	74.9	*0.6*	76.4	*0.7*
6 Semi-routine	71.4	*0.6*	72.0	*0.6*	72.1	*0.6*	73.8	*0.6*	74.9	*0.6*
7 Routine	69.9	*0.6*	70.4	*0.6*	71.1	*0.6*	71.9	*0.7*	74.4	*0.6*
Range highest- lowest	5.8		6.2		6.5		6.8		6.0	
Condensed NS-SEC										
Managerial & professional	74.9	*0.5*	75.8	*0.4*	76.8	*0.4*	78.1	*0.4*	79.7	*0.4*
Intermediate	73.5	*0.5*	74.5	*0.5*	75.4	*0.5*	76.7	*0.5*	77.7	*0.5*
Routine & manual	71.0	*0.4*	71.7	*0.4*	72.1	*0.4*	73.4	*0.4*	75.2	*0.4*
Range highest- lowest	3.9		4.1		4.7		4.6		4.5	
Unclassified	60.2	*1.4*	60.3	*1.2*	65.8	*1.3*	67.2	*1.2*	71.5	*1.1*
All Men	71.7	*0.2*	72.6	*0.2*	73.8	*0.2*	75.2	*0.2*	77.0	*0.2*

Source: ONS Longitudinal Study

Table A1b Life expectancy by NS-SEC class, males at age 65, using a 'combined' classification

England and Wales — Years

	1982–86		1987–91		1992–96		1997–2001		2002–06	
	LE	95% CI +/-	LE	95% CI +/-	LE	95% CI +/-	LE	95% CI +/-	LE	95% CI +/-
At age 65										
Analytic classes										
1. Higher managerial & professional	15.2	*0.6*	15.9	*0.5*	16.6	*0.5*	18.1	*0.5*	18.8	*0.4*
1.1 Large employers & higher managers	14.7	*0.8*	15.1	*0.7*	16.2	*0.7*	17.9	*0.6*	18.6	*0.6*
1.2 Higher professional	16.1	*0.8*	16.7	*0.7*	17.1	*0.7*	18.2	*0.6*	19.0	*0.6*
2 Lower managerial & professional	14.9	*0.4*	15.4	*0.4*	15.9	*0.3*	17.0	*0.3*	17.9	*0.3*
3 Intermediate	13.9	*0.4*	15.0	*0.5*	15.5	*0.4*	16.2	*0.4*	17.4	*0.4*
4 Small employers & own a/c workers	14.0	*0.5*	14.6	*0.5*	15.4	*0.5*	15.9	*0.5*	17.3	*0.5*
5 Lower supervisory & technical	13.3	*0.4*	13.3	*0.4*	14.3	*0.4*	15.2	*0.4*	16.2	*0.4*
6 Semi-routine	12.8	*0.3*	13.3	*0.3*	13.7	*0.3*	14.5	*0.3*	15.6	*0.4*
7 Routine	12.8	*0.3*	12.9	*0.3*	13.2	*0.3*	13.7	*0.3*	15.0	*0.4*
Range highest- lowest	2.4		3.0		3.4		4.4		3.8	
Condensed NS-SEC										
Managerial & professional	15.1	*0.3*	15.5	*0.3*	16.2	*0.3*	17.4	*0.3*	18.3	*0.3*
Intermediate	13.9	*0.3*	14.8	*0.3*	15.5	*0.3*	16.1	*0.3*	17.3	*0.3*
Routine & manual	12.9	*0.2*	13.2	*0.2*	13.7	*0.2*	14.4	*0.2*	15.6	*0.2*
Range highest- lowest	2.2		2.3		2.5		3.0		2.7	
Unclassified	11.1	*0.3*	10.2	*0.4*	11.2	*0.6*	12.0	*0.7*	14.2	*0.7*
All Men	13.1	*0.1*	13.7	*0.1*	14.5	*0.1*	15.5	*0.1*	16.7	*0.1*

Source: ONS Longitudinal Study

Table A2a Life expectancy by NS–SEC class, females at birth, using a 'combined' classification

England and Wales — Years

	1982–86		1987–91		1992–96		1997–2001		2002–06	
	LE	95% CI +/-	LE	95% CI +/-	LE	95% CI +/-	LE	95% CI +/-	LE	95% CI +/-
	At birth									
Analytic classes										
1. Higher managerial & professional	81.2	*1.0*	82.4	*0.8*	83.0	*0.8*	82.8	*0.6*	84.3	*0.6*
1.1 Large employers & higher managers	80.5	*1.6*	83.0	*1.3*	83.2	*1.1*	82.4	*1.0*	83.7	*0.9*
1.2 Higher professional	82.0	*1.4*	82.2	*1.1*	82.7	*1.1*	83.1	*0.9*	84.9	*0.8*
2 Lower managerial & professional	79.9	*0.7*	81.1	*0.6*	81.2	*0.5*	82.1	*0.5*	83.3	*0.4*
3 Intermediate	79.7	*0.7*	80.8	*0.7*	81.1	*0.6*	81.1	*0.6*	82.0	*0.6*
4 Small employers & own a/c workers	79.3	*0.9*	80.1	*0.9*	80.7	*0.8*	80.9	*0.7*	82.2	*0.8*
5 Lower supervisory & technical	78.6	*0.9*	78.4	*0.7*	79.7	*0.7*	79.7	*0.6*	80.5	*0.7*
6 Semi-routine	77.7	*0.7*	77.7	*0.6*	78.6	*0.6*	79.0	*0.6*	80.0	*0.6*
7 Routine	76.4	*0.7*	76.9	*0.7*	77.4	*0.6*	77.9	*0.7*	79.0	*0.6*
Range highest- lowest	4.8		5.5		5.6		4.9		5.3	
Condensed NS-SEC										
Managerial & professional	80.3	*0.5*	81.5	*0.5*	81.7	*0.4*	82.4	*0.4*	83.6	*0.4*
Intermediate	79.6	*0.6*	80.5	*0.5*	80.9	*0.5*	81.1	*0.5*	82.1	*0.4*
Routine & manual	77.4	*0.4*	77.6	*0.4*	78.5	*0.4*	78.8	*0.4*	79.8	*0.4*
Range highest- lowest	2.9		3.9		3.2		3.5		3.9	
Unclassified	71.5	*1.1*	73.1	*1.0*	74.2	*1.2*	75.8	*0.9*	76.9	*0.9*
All Women	77.4	*0.2*	78.3	*0.2*	79.2	*0.2*	79.9	*0.2*	81.1	*0.2*

Source: ONS Longitudinal Study

Table A2b Life expectancy by NS–SEC class, females at age 65, using a 'combined' classification

England and Wales — Years

	1982–86		1987–91		1992–96		1997–2001		2002–06	
	LE	95% CI +/-	LE	95% CI +/-	LE	95% CI +/-	LE	95% CI +/-	LE	95% CI +/-
	At age 65									
Analytic classes										
1. Higher managerial & professional	19.6	*1.0*	20.3	*0.7*	21.1	*0.7*	20.9	*0.5*	21.9	*0.4*
1.1 Large employers & higher managers	19.5	*1.5*	20.8	*1.3*	21.2	*0.9*	20.7	*0.7*	21.6	*0.6*
1.2 Higher professional	19.7	*1.3*	20.1	*0.9*	20.9	*0.9*	21.0	*0.7*	22.2	*0.6*
2 Lower managerial & professional	18.9	*0.6*	19.6	*0.5*	19.3	*0.4*	20.3	*0.4*	21.0	*0.3*
3 Intermediate	18.3	*0.6*	19.3	*0.5*	19.4	*0.5*	19.7	*0.4*	20.2	*0.4*
4 Small employers & own a/c workers	18.6	*0.8*	18.8	*0.7*	19.5	*0.6*	19.1	*0.5*	20.3	*0.5*
5 Lower supervisory & technical	18.5	*0.6*	17.7	*0.5*	18.3	*0.5*	18.4	*0.4*	19.2	*0.4*
6 Semi-routine	17.3	*0.5*	17.2	*0.4*	18.1	*0.4*	18.0	*0.4*	18.8	*0.4*
7 Routine	16.6	*0.4*	17.3	*0.4*	17.4	*0.4*	17.2	*0.4*	17.9	*0.4*
Range highest- lowest	3.0		3.0		3.7		3.7 4.0			
Condensed NS-SEC										
Managerial & professional	19.0	*0.5*	19.7	*0.4*	19.8	*0.3*	20.5	*0.3*	21.3	*0.3*
Intermediate	18.4	*0.5*	19.1	*0.4*	19.4	*0.4*	19.5	*0.3*	20.2	*0.3*
Routine & manual	17.3	*0.3*	17.4	*0.2*	17.9	*0.2*	17.9	*0.2*	18.6	*0.2*
Range highest- lowest	1.8		2.4		1.9		2.6 2.8			
Unclassified	16.2	*0.2*	16.2	*0.3*	16.3	*0.3*	16.6	*0.4*	17.4	*0.5*
All Women	17.0	*0.1*	17.5	*0.1*	18.0	*0.1*	18.5	*0.1*	19.5	*0.1*

Source: ONS Longitudinal Study

Linking maternity data for England, 2005–06: methods and data quality

Nirupa Dattani, Preeti Datta-Nemdharry and Alison Macfarlane ***City University London, Department of Midwifery and Child Health.***

Abstract

Introduction

Maternity Hospital Episode Statistics (HES) data were linked to birth registration and NHS Numbers for Babies (NN4B) data to bring together some key demographic and clinical data items not otherwise available at a national level. This project added to earlier work involving linkage of birth registration records to NN4B records.

Methods

Birth registration and NN4B records were linked to Maternity HES delivery records and also Maternity HES baby records using the NHS Number or other indirect identifiers if NHS Number was missing.

Data quality and completeness of Maternity HES were assessed in relation to birth registration data wherever possible. For information not collected at registration, NN4B data were used to validate quality of Maternity HES.

Results

Overall, 91 per cent of Maternity HES delivery records could be linked to the birth registration/NHS Numbers for Babies records and 84 per cent of Maternity HES baby records were linked.

In 2005 only 3 per cent of Maternity HES records had mother's NHS number missing, compared with 30 per cent in the NN4B dataset. This did not reflect the extent to which Maternity HES data items were missing or discordant.

Over a quarter of all linked Maternity HES records for singleton babies had one or more of the following data items missing: birthweight, gestational age, birth status, sex, and date of birth of the baby. On the other hand, for data items where information was stated such as birthweight, birth status, and sex for singleton babies, there was good agreement between Maternity HES and linked birth registration and NN4B data.

Although NN4B records the ethnic category of the baby as defined by the mother, and Maternity HES records mother's ethnic category, 87 per cent of the linked records had the same ethnic group.

Conclusion

Even though a good linkage rate was obtained, the method used will be simplified before data for 2007 are linked. To gain the maximum benefit from this linkage in future years, improvements are urgently needed in the quality and completeness of the data contained in Maternity HES.

Contents

List of Tables

Introduction

The data recorded at birth registration are mainly socio-demographic such as names, address of residence, place of birth, occupation of the parents and date of birth (shown in Office for National Statistics publication series DH3). As a result some important items needed for demographic and clinical purposes are not available at a national level. The opportunity to obtain gestational age and ethnicity nationally was provided by the introduction of the NHS Numbers for Babies (NN4B) Service in 2002. This service collects a small dataset which contains key items which are not recorded at birth registration. Information on gestational age at birth is of key importance. Babies born preterm, before 37 completed weeks of gestation, are at particular risk of morbidity and mortality in early years of life (Brocklehurst P, 1999; ISD Scotland report 2004; Confidential Enquiry into Maternal and Child Health, 2004)

Clinical information on maternity care at delivery could only be obtained from the Maternity Hospital Episode Statistics (HES) dataset for births that occurred in England and from the Community Child Health database (CHD) and Patient Episode Database for Wales (PEDW) for births that occurred in Wales.

Therefore a collaborative project between City University London, Office for National Statistics (ONS) and the Welsh Assembly Government was set up in 2004 to link these datasets for all births that occurred in England and Wales from 2005 to 2007. Stage 1 of the project involved linkage of birth registration data with NN4B dataset and assessment of data quality and completeness of the NN4B data. This is reported elsewhere (Hilder *et al*., 2007 and Moser K *et al*., 2008).

Stage 2 of the project, involving linkage of the dataset created in stage 1 to Maternity HES and the corresponding Welsh records for 2005 and 2006, and then the assessment of data quality and completeness by comparison with birth registration or NN4B, where possible. At the time, 2007 birth registration – NN4B linked data were not available. Therefore these data will be linked to Maternity HES and corresponding Welsh records at a later date using the experience gained in linking the first two years' data. This article describes linkage to Maternity HES records and reports on its data quality and completeness. The Welsh linkage for the first two years will be reported later.

Several data items are common to all three sources (Maternity HES, birth registration and NHS Numbers for Babies) as shown in Box 1. In addition, some data items are unique to each data source and linkage will enable new analyses across these data sources. For example, it will be possible to analyse caesarean section rates by father's socio-economic classification, compare time of birth with birth outcomes, and report on time and method of delivery by day of the week. Once the linkage has been completed and checked, the next stage of the project will be to undertake some of these analyses.

Box 1 Availability of selected data items from birth registration, NN4B and Maternity HES

Data items	Data sources		
	Birth registration	NN4B	Maternity HES
Baby's NHS number	+	+	+
Mother's NHS number		+	+
Birth date of baby	+	+	+
Delivery time		+	
Birthweight	+	+	+
Gestational age (stillbirth)	+	+	+
Gestational age (live birth)		+	+
Sex of baby	+	+	+
Number of babies born	+	+	+
Live or stillbirth	+	+	+
Parity (all births)			+
Baby/mother's postcode of usual residence	+	+	+
Ethnic category of baby		+	
Ethnic category of mother			+
Country of birth of mother	+		
Country of birth of father	+		
Father's socio-economic status	+		
Type of delivery place	+	+	+
Mother's date of birth	+	+	+
Marital status of mother	+		
Method of delivery			+
Complications in pregnancy			+

Methods

Source data

Birth registration

It is a legal requirement to register all live births within 42 days of birth. The definition of a live birth, the legal basis, the process and a complete list of data items collected, are described in detail elsewhere (Office for National Statistics publication series DH3). The information is obtained, usually from one or both parents, by the local registrar of births, marriages and deaths. The local child health department passes some information from the birth notification it receives from the midwives to the registrar to verify the birth. Since 1975 this has included the baby's birthweight and since 2002 the NHS number (National Health Service Act 1977).

The process for registration of stillbirths is similar to that for live births, except that registrars do not retain the NHS number for a stillbirth and the informant will also give the registrar a medical certificate of stillbirth completed by the attending midwife or doctor. This certifies the cause of death and includes an assessment of gestational age at birth and birthweight.

NHS Numbers for Babies (NN4B) service

The NHS Numbers for Babies Service was set up to issue an NHS number to all babies soon after birth: it went live in 2002. Under this system a standardised set of information about each birth is notified electronically, usually by the midwife attending the birth, to the Central Issuing System which checks for duplicates. If there is no existing record with the same details, the system returns a newly generated NHS number to the agency notifying the birth. If a number has already been issued, the system will return a message indicating that a record with matching details already exists. The notifier is required to check this record against the details of the baby they have just entered and use the existing number if a match is found. If no match is found then the notifier will proceed to obtain the NHS number using the NN4B Child Health Browser, or remotely by telephone from the NHS Number Issue Helpdesk if the browser is unavailable.

Records are held for six months on the Central Issuing System and then deleted. When the NHS number is issued, a copy of the information is sent to the local child health system where the birth occurred and a limited dataset, including the NHS number, to the NHS Central Register.

In the first part of the project, arrangements were made for ONS to receive a subset of variables from the NN4B dataset from 2005 onwards for linkage to birth registration record.

Maternity Hospital Episode Statistics (HES)

Maternity Hospital Episode Statistics contains data for all births occurring in England, including those at home and in non-NHS hospitals. There are however, some data items not recorded for births outside NHS hospitals such as admission/discharge dates or patient classification. HES includes a wide range of information on maternity such as details on how the baby was born (method of delivery), method of onset of labour, complications, gestation and ethnicity as well as information about the baby, such as date of birth, gender, birthweight and geographical information on where the baby was born.

Most records of admission to hospital will be classed as general inpatient episodes, even when heavily pregnant women are admitted to maternity wards with the clear intention of giving birth. However, as soon as the mother has given birth, the record becomes a maternity record and is updated and extended before being submitted to HES. If the pregnant woman does not give birth, for whatever reason, before the episode ends, the record remains as a general inpatient record (known as a general episode in HES).

There are two types of maternity records in HES, the delivery record and the birth record. Both types of records consist of an admitted patient care record with an additional 19 fields, in an appended baby 'tail'.

- A delivery record is a HES record for the mother containing the same data as a general record, but has a baby tail for information about the delivery
- A birth record is a HES record for the baby. Again it has the same format as a general record but it also has a baby tail containing exactly the same information recorded in the corresponding tail of the delivery record

The baby tail data coverage is not as complete as the rest of the HES data (HES website). There are a number of reasons for the incompleteness and data quality issues, such as:

- Trusts submitting a significantly higher number of delivery episodes compared to birth episodes
- Trusts failing to submit data on the number of birth episodes where they record a high number of delivery episodes
- Trusts failing to submit delivery records – the reason for this is that approximately 20 trusts have a stand-alone maternity system which is not linked to the Patient Administration System
- Trusts reporting a high number of maternity beds available, but not recording any information about deliveries or births
- Trusts reporting that they have no maternity beds available, but recording a high number of birth and delivery episodes

Over the years from 2001–02 to 2005–06, coverage of deliveries was 86.2 per cent on average. The problems with coverage are significantly compounded by data quality issues (HES website)

HES Patient ID (HESID) is an identifier used to track patients through the HES database, or for linkage to other datasets such as mortality, rather than using patient identifiers such as NHS Number. It is a pseudonymised number which uniquely identifies each patient without the necessity of viewing or using patient identifiable information such as the NHS Number. HESID is derived using a matching algorithm which looks at various combinations of the following patient identifiable fields:

- NHS Number (fieldname NEWNHSNO)
- Date of Birth (fieldname DOB)
- Sex (fieldname SEX)
- Postcode (fieldname HOMEADD)
- Provider code (fieldname PROCODET)
- Local patient ID (fieldname LOPATID)

HESIDs are stored in the HES index. This is updated monthly and older versions are not kept. Hence the most recent HES index (22 June 2010) was used to link birth registration/NHS Numbers for Babies data for 2005 and 2006.

For each episode of care with a particular consultant or midwife, a HES record is created, but each time this record is updated with new information, a new version of the record is created. As a result, several versions of the record of the same episode are created.

Record linkage

Record linkage was carried out by Northgate Solutions, which processes HES records under contract with the NHS Information Centre. For details on linkage of registration and NN4B linked data to the mother and baby records in HES, see Appendix.

The linked data for each year (2005 and 2006) consisted of previously linked registration and NN4B data linked to the mother record in HES, and a second file based on linkage of previously linked registration and NN4B data to the baby record in HES. These were provided to ONS and were accessed by researchers from City University London in the secure environment of the Virtual Microdata Laboratory (VML) facilities at ONS. Outputs of analyses undertaken in the VML were released by ONS in the form of disclosure controlled tables.

Data Quality

Review of quality of Maternity HES was based on completeness and consistency of the HES data in relation to birth registration data wherever possible. Since all babies born in England and Wales have to be registered, information collected at registration is subject to quality checks (Office for National Statistics publication series DH3). However, where information was not available from registration, NN4B data were used to validate the quality of Maternity HES. The quality of the NN4B data in comparison to birth registration data is reported elsewhere (Moser *et al*., 2008). Completeness of the main data items in all three sources was measured by identifying missing data.

The linked data for the mothers file was split into singleton and multiple births (using multiple birth status information from registration) to ease assessment of data quality. In some instances the results are reported separately.

Results for 2005 are reported in this paper and 2006 results are available on the ONS website in a spreadsheet

Data analyses were carried out using SAS version 9 and SPSS version 16.

Results

Mother file

The Maternity HES record is a mother-based record containing the mother's details with a maternity tail and a baby tail which can accommodate up to nine babies born in one maternity. In contrast, the registration and NN4B linked data consists of one record per baby. Therefore, the linkage was based on baby to mother record.

Northgate solutions returned 582,680 records that had linked to 2005 registration and NN4B linked data. For 2006, 601,623 HES records were linked. These included multiple records for the same mother for each episode. Records with the most information were selected to ensure one-to-one linkage to registration and NN4B linked dataset. This gave a file of 566,240 records for 2005 and 584,795 records for 2006.

Due to the way the linkage was done, the same Maternity HES record was linked to multiple registration and NN4B linked records. Maternity HES records where the postcode, mother's or baby's date of birth and birthweight were all inconsistent, when compared to registration and NN4B linked dataset were deleted. A total of 264 and 364 Maternity HES links were deleted for 2005 and 2006 data respectively.

In the 2005 registration and NN4B linked data there were 617,613 babies who were either born in England or resident in England. The resident in England category was used for births that occurred at home in the registration and NN4B linked data.

Around 64 per cent of the linked registration and NN4B records were linked to Maternity HES records using mother's NHS number and partial date of birth.

A further 27 per cent of the linked registration and NN4B records were matched to Maternity HES using mother's postcode and date of birth. Registration and NN4B linked records that were not linked to HES accounted for 8 per cent of all records. Of the Maternity HES records, 3 per cent did not link to registration and NN4B linked records.

Overall, there were 50,380 singletons and 1,265 multiples in the registration and NN4B linked data that were not linked to HES records. This gave a total 565,968 Maternity HES records that were linked to the registration and NN4B linked records giving a linkage rate of 91.6 per cent. The linkage rate for 2006 was similar at 91.3 per cent.

Baby file

The baby file was much more straightforward than the mother file as it involved a one to one link between baby records in registration and NN4B linked data, and Maternity HES.

For 2005, a total of 686,087 HES baby records were linked to registration and NN4B linked data by Northgate solutions. These included 128,482 records that were general episodes. After omitting these, there were 557,605 HES baby records that had been linked to registration and NN4B linked data. There were multiple HES birth records for the same baby linked to registration and NN4B linked record. Again only the records with the most information were kept and others were deleted.

After deletion, 524,536 records remained. On further examination of the data, 4,009 records had to be deleted where the same HES record was matched to another registration and NN4B linked record. A similar process was applied for the 2006 data with 674,534 records sent by Northgate and 540,283 linked HES records were left after deletion of multiple HES baby record. A further 1,958 records were deleted due to same Maternity HES records being linked to more than one registration and NN4B linked record.

In the 2005 registration and NN4B linked data, there were 617,613 babies who were either born in England or resident in England; 515,434 registration and NN4B linked records were linked to HES baby records using NHS number, partial date of birth, and sex; and 4,597 were linked using date of birth, postcode, and sex. Over 15 per cent of registration and NN4B linked records were not linked to HES baby records. Around 1 per cent of HES baby records did not link to registration and NN4B linked data.

Overall 520,527 out of 617,613 records were linked giving a linkage rate of 84.3 per cent. For 2006 the linkage rate was 84.1 per cent.

Data Quality

For HES, missing and discordant data were assessed only in the mother record – as this included information on the baby, and also because the linkage rate was far better than the baby record. For multiple births information was recorded only on the first baby; data on other babies was either missing or the same as the first baby suggesting there were problems in the linkage process in HES. Hence singleton and multiple births were analysed separately and only results for singletons are reported here.

Missing data

The mother's NHS number is recorded only on the NN4B record and not recorded at birth registration. For singleton births, 30 per cent of registration and NN4B linked records did not have the mother's NHS number compared with three per cent in the Maternity HES records. In Maternity HES, birthweight and gestational age information was missing for 25 per cent and 48 per cent of singletons respectively. Status of baby, date of birth of baby and sex of baby was missing in over 25 per cent of the records (Table 1). Similar results were observed in 2006 (Table 2).

Table 1 Number of linked records[1] for singleton births with missing data items in common data fields, 2005

England

	NHS Numbers for Babies		Registration		Maternity HES	
	Number	Percent	Number	Percent	Number	Percent
NHS number of mother	164,458	*30.0*	NA	*NA*	16,685	*3.0*
Date of birth of mother	960	*0.2*	0	*0.0*	0	*0.0*
Ethnicity	59,865	*10.9*	NA	*NA*	77,771	*14.2*
Postcode	10,101	*1.8*	10	*<0.1*	693	*0.1*
Birthweight	2,721	*0.5*	874	*0.2*	135,144	*24.6*
Gestational age	3,829	*0.7*	NA	*NA*	264,877	*48.3*
Status	615	*0.1*	0	*0.0*	176,455	*32.2*
Date of birth of baby	615	*0.1*	0	*0.0*	139,414	*25.4*
Sex of baby[2]	1,098	*0.2*	0	*0.0*	144,115	*26.3*

1 Number of linked records 548,857
2 Includes 483 cases with indeterminate sex in NN4B and 15 cases in maternity HES

Source: HES, registration and NHS Numbers for Babies

Table 2 Number of linked records[1] for singleton births with missing data items in common data fields, 2006

England

	NHS Numbers for Babies		Registration		Maternity HES	
	Number	Percent	Number	Percent	Number	Percent
NHS number of mother	147,782	*26.1*	NA	*NA*	14,440	*2.6*
Date of birth of mother	465	*0.1*	0	*0.0*	0	*0.0*
Ethnicity	56,521	*10.0*	NA	*NA*	90,839	*16.0*
Postcode	8,209	*1.5*	14	*<0.1*	343	*0.1*
Birthweight	3,123	*0.6*	2,839	*0.5*	142,696	*25.2*
Gestational age	4,104	*0.7*	NA	*NA*	255,437	*45.1*
Status	189	*0.0*	0	*0.0*	155,754	*27.5*
Date of birth of baby	189	*0.0*	0	*0.0*	146,183	*25.8*
Sex of baby[2]	704	*0.1*	0	*0.0*	153,015	*27.0*

1 Number of linked records 566,313
2 Includes 515 cases of indeterminate sex in NN4B and 1,529 cases in Maternity HES

Source: HES, registration and NHS Numbers for Babies

Discordant data

Discordance in common individual data items

Discordance in each of the common data fields in the linked records was assessed using information from birth registration rather than NN4B. However, for data items that were not recorded at birth registration, NN4B data were used.

For singletons, using birth registration and Maternity HES data, baby's date of birth was discordant in 0.2 per cent of records in 2005 and 2006. Postcode did not agree in 7 per cent of records in 2005 and 2006.

Discordance in multiple birth status

There were 13,850 records that were identified as multiple births in birth registration and Maternity HES. Multiple birth status was discordant between the two data sources in 1,806 records in 2005 and 1,913 records in 2006 (Table 3). In 2005, 91,030 Maternity HES records had unknown multiple birth status.

Table 3 Comparison of plurality[1] between registration and maternity HES, 2005 and 2006

England

Year	Maternity HES plurality	Registration plurality		
		Singleton	Multiple	Total
2005	Singleton	459,282	178	459,460
	Multiple	1,628	13,850	15,478
	Total	460,910	14,028	474,938
2006	Singleton	461,912	267	462,179
	Multiple	1,646	14,543	16,189
	Total	463,558	14,810	478,368

1 Plurality was missing in maternity HES for 91,030 records in 2005 and 106,051 records in 2006.

Source: HES and registration

Discordance in live or stillbirth status

For the records which had a stated live or still birth status in both data sources, 0.05 per cent and 0.2 per cent did not agree in 2005 and 2006 (Table 4) respectively. Out of all the linked records for 2005, around 32 per cent of Maternity HES records did not have information on birth status. This fell to 27.5 per cent in 2006.

Table 4 Comparison of birth status for singleton births between registration and maternity HES, 2005 and 2006

England

Year	Maternity HES birth status	Registration birth status			
		Live birth	Still birth	Total	
		Number	Number	Number	Percent
2005	Live	370,939	99	371,038	*67.6*
	Still birth: ante-partum	56	1,021	1,077	*0.2*
	Still birth: intra-partum	9	137	146	*0.0*
	Stillbirth: indeterminate	6	135	141	*0.0*
	Not Known	175,578	877	176,455	*32.2*
	Total	546,588	2,269	548,857	
2006	Live	408,178	114	408,292	*72.1*
	Still birth: ante-partum	63	1,212	1,275	*0.2*
	Still birth: intra-partum	14	178	192	*0.0*
	Stillbirth: indeterminate	661	139	800	*0.1*
	Not Known	154,883	871	155,754	*27.5*
	Total	563,799	2,514	566,313	

Source: HES and registration

Discordance in baby's sex

The sex of the baby recorded on birth registration for singleton births was used to compare with Maternity HES. In 2005 where baby's sex was stated in both data sources, an agreement of 99 per cent was observed (Table 5). A similar percentage was noted in the 2006 data. In the 2006 Maternity HES data, sex was coded incorrectly in 301 records to codes 4, 5, 6, and 7, and there were 1,529 records where sex was coded as indeterminate, compared with 15 records with indeterminate sex in 2005.

Table 5 Comparison of sex for singleton births between registration and maternity HES, 2005 and 2006

England

			Registration			
Year			Male	Female	Total	Percentage
2005	Maternity HES[1]	Male	204,613	791	205,404	*50.8*
		Female	2,814	196,524	199,338	*49.3*
		Total	207,427	197,315	404,742	
2006	Maternity HES[2]	Male	208,162	586	208,748	*50.7*
		Female	2,978	200,043	203,021	*49.3*
		Total	211,140	200,629	411,769	

1 There were 34 records classified as not known, 78 as not specified, 15 as indeterminate, 143988 records had sex missing.
2 There were 2323 records classified as not known, 1450 as not specified, 1529 as indeterminate, 148941 records had sex missing and 301 records had sex coded incorrectly.

Source: HES and registration

Discordance in birthweight

Birthweights in birth registration data were grouped into 500g groups and compared with grouped birthweights from Maternity HES. In terms of concordance between the two data sources, 99 per cent of records with stated birthweight were in the same 500g birthweight group. This amounts to only 75 per cent of all linked records because birthweight was not stated on a large number of records. In Maternity HES, birthweight was missing in 25 per cent of the records compared with only 0.2 per cent in birth registration data (Table 6). Similar findings were also observed with the 2006 data (99 per cent concordance between the two data sources where birthweight was stated) and these can be found in Table 1 on the web.

Table 6 Comparison of birthweight distribution for singleton births between registration and maternity HES, 2005

England

Registration Birthweight (grams)	Maternity HES birthweight (grams)														
	<500	500–999	1000–1499	1500–1999	2000–2499	2500–2999	3000–3499	3500–3999	4000–4499	4500–4999	5000–5499	5500+	Not stated	Total	Percentage
<500	**233**	4	4	5	4	3	10	5	10	5	1	.	140	424	0.1
500–999	5	**1,533**	.	7	7	9	4	4	4	2	2	17	693	2,287	0.4
1000–1499	7	6	**2,322**	11	2	7	17	7	5	1	.	.	954	3,339	0.6
1500–1999	11	3	12	**4,638**	6	21	12	7	2	.	.	.	1,716	6,428	1.2
2000–2499	18	3	7	17	**16,259**	250	308	174	65	10	1	3	5,408	22,523	4.1
2500–2999	30	7	6	15	112	**68,622**	162	236	22	6	6	14	21,888	91,126	16.6
3000–3499	93	14	31	16	121	160	**149,372**	558	100	5	17	24	48,764	199,275	36.3
3500–3999	62	10	13	11	30	205	508	**119,583**	97	13	8	21	39,781	160,342	29.2
4000–4499	30	1	11	3	10	16	69	64	**39,334**	39	6	7	13,051	52,641	9.6
4500–4999	8	.	.	2	1	2	13	29	42	**6,387**	.	.	2,155	8,639	1.6
5000–5499	.	1	.	1	.	1	6	5	1	.	**636**	.	216	867	0.2
5500 +	1	7	1	2	2	3	1	2	.	.	.	**47**	26	92	0.0
Not stated	6	14	5	7	32	93	203	117	39	5	1	.	352	874	0.2
Total	504	1,603	2,412	4,735	16,586	69,392	150,685	120,791	39,721	6,473	678	133	135,144	548,857	
Percentage	0.1	0.3	0.4	0.9	3.0	12.6	27.5	22.0	7.2	1.2	0.1	0.0	24.6	100.0	

Source: HES and registration

Discordance in gestational age

Information about gestational age for all births was available from the NN4B and Maternity HES. In 2005 approximately 89 per cent of records with a recorded gestational age were the same in both data sources (Table 7). For Maternity HES data 48 per cent of records had gestational age information missing. Gestational age differed by one week in around 7 per cent of the records, and two weeks or more in about 4 per cent of the records. There was a large variation in gestational age between the two data sources in the 'tails' for those under 30 weeks and over 42 weeks, but only 3 per cent of births occurred in these gestational age groups. The difference ranged from 20 to 85 per cent for those under 30 weeks. At 43 weeks, gestational age differed in about half of the records and it decreased further to a third of all records at 44 weeks and over.

For 2006 there was agreement between the two sources in 90 per cent of records which contained gestational age from Maternity HES and NN4B (Table 2 on the web).

Table 7 Comparison of gestational age for all births between NN4B and maternity HES, 2005

England

NN4B	Maternity HES gestational age (weeks)																									
Gestational age (weeks)	**<22**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44 +**	**Not stated**	**Total**
<22	**74**	.	.	.	2	1	1	.	2	.	.	.	.	1	1	2	.	1	2	7	3	.	.	.	121	218
22	2	**44**	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1	.	.	.	.	74	123
23	7	2	**89**	7	2	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	1	.	.	.	114	223
24	6	.	5	**235**	8	.	.	.	.	.	.	.	.	3	.	1	.	2	.	2	1	.	.	.	323	586
25	12	.	.	6	**253**	9	3	.	3	.	.	.	.	.	2	.	.	.	.	2	2	1	.	.	269	562
26	9	1	1	1	6	**312**	11	1	1	.	1	.	1	1	.	4	.	1	1	.	1	.	.	.	373	726
27	11	.	.	3	.	7	**356**	10	.	.	1	.	.	.	.	.	2	.	3	3	2	2	.	.	374	774
28	10	1	.	.	1	5	9	**465**	9	.	1	.	.	2	1	1	2	11	2	2	4	1	.	.	521	1,048
29	5	.	.	.	2	.	.	11	**470**	21	1	1	4	1	.	.	3	.	5	5	.	.	.	.	626	1,155
30	12	.	.	.	.	.	1	1	15	**611**	16	6	6	1	1	.	1	3	7	14	.	.	.	.	788	1,483
31	14	.	.	1	.	.	2	.	3	18	**770**	25	.	3	3	.	.	3	6	17	3	.	.	.	910	1,778
32	1	17	.	1	.	.	.	2	.	4	20	**1,168**	21	9	1	6	2	5	1	24	3	.	.	1	1,241	2,527
33	1	.	36	.	.	.	.	1	1	1	9	33	**1,529**	54	10	8	9	8	7	16	2	.	.	1	1,824	3,550
34	.	.	45	3	2	2	.	1	1	2	6	3	57	**2,474**	85	19	19	18	14	37	3	.	.	.	2,814	5,605
35	3	.	.	58	1	.	.	.	.	2	5	2	14	72	**3,826**	125	27	17	16	63	16	2	.	1	4,119	8,369
36	4	.	1	.	102	.	1	3	1	1	2	5	6	24	132	**7,125**	227	62	33	141	20	11	1	9	7,616	15,527
37	5	.	.	.	229	1	6	2	2	.	3	7	5	7	26	269	**15,049**	611	162	312	67	28	7	13	15,697	32,508
38	7	3	.	1	2	473	2	4	3	11	6	11	24	23	28	80	751	**36,057**	1,405	881	114	50	26	36	37,091	77,089
39	6	.	2	.	.	2	881	1	3	10	7	9	9	18	32	68	169	1,610	**55,910**	2,915	406	123	74	83	59,341	121,679
40	13	.	3	1	4	1	3	983	1	5	3	9	5	21	25	70	98	331	2,444	**71,064**	2,593	368	147	180	72,955	151,327
41	7	.	.	2	2	2	1	865	1	4	9	4	7	11	14	44	63	121	307	4,411	**49,742**	1,275	192	228	52,850	110,162
42	2	1	.	.	.	1	.	.	252	.	.	2	2	5	2	6	24	54	81	548	1,290	**8,452**	62	49	11,205	22,038
43	.	.	.	.	.	.	1	1	.	12	2	.	.	.	1	2	7	22	51	102	74	42	**543**	7	582	1,449
44+	.	.	.	.	.	.	1	2	.	6	.	.	.	3	2	2	6	9	26	73	43	26	.	**278**	397	874
Not stated	.	1	.	3	3	5	5	8	8	11	10	14	15	36	53	115	239	458	679	780	616	88	1	4	1,436	4,588
Total	211	70	183	322	619	821	1,284	2,361	776	719	872	1,299	1,705	2,769	4,245	7,947	16,698	39,404	61,164	81,420	55,006	10,469	1,053	890	273,661	565,968

Source: HES and registration

Discordance in ethnicity

Baby's ethnicity from the NN4B dataset and mother's ethnicity from Maternity HES were compared (Table 8). There was agreement in 87 per cent of the records which had a stated ethnic category. Out of all the linked records, 19 per cent of records had no ethnicity recorded in Maternity HES, and in 11 per cent of records ethnic group was not stated in the NN4B dataset.

For 2006 data the ethnic category was the same in 87 per cent of the linked records. In Maternity HES 16 per cent of records had no ethnicity recorded and 10 per cent of records had ethnic group not stated in the NN4B dataset (Table 3 on the web).

Table 8 **Comparison of baby's ethnicity from NN4B with mother's ethnicity from maternity HES for all births, 2005**

England

	Mother's ethnicity																	
Babies ethniciy	**White British**	**White Irish**	**Any other White**	**White and Black Caribbean**	**White and Black African**	**White and Asian**	**Any other mixed**	**Indian**	**Pakistani**	**Bangladeshi**	**Any other Asian**	**Caribbean**	**African**	**Any other Black**	**Chinese**	**Any other**	**Not stated**	**Total**
White British	**281,674**	1,036	7,385	179	75	186	441	254	255	90	183	183	209	99	118	1,440	68,309	362,116
White Irish	683	**832**	76	2	.	1	6	.	1	.	3	3	3	.	3	16	306	1,935
Any other White	5,746	170	**14,963**	20	46	42	403	26	23	6	130	17	75	36	31	1,607	4,262	27,603
White and Black Caribbean	2,577	24	134	**558**	34	5	110	1	2	1	4	503	79	151	.	146	853	5,182
White and Black African	1,003	13	144	21	**278**	1	42	6	3	2	9	49	700	170	1	84	596	3,122
White and Asian	1,723	34	220	9	2	**274**	65	342	120	32	392	4	10	7	117	260	921	4,532
Any other mixed	2,259	32	645	175	67	81	**615**	180	87	31	249	168	134	150	204	729	1,351	7,157
Indian	290	.	36	23	3	39	36	**9,553**	855	220	872	18	37	15	2	275	2,370	14,644
Pakistani	300	1	34	.	3	44	22	439	**17,515**	224	732	7	27	38	2	202	2,716	22,306
Bangladeshi	92	1	12	2	1	3	7	120	427	**5,974**	175	4	8	6	.	68	526	7,426
Any other Asian	132	1	76	9	6	57	65	647	361	104	**2,721**	18	52	32	125	951	1,298	6,655
Caribbean	283	6	38	218	25	1	53	5	6	2	15	**4,195**	274	592	1	147	788	6,649
African	234	1	107	28	541	3	67	34	28	18	97	292	**11,585**	1,111	1	500	2,376	17,023
Any other Black	182	3	71	74	81	4	47	24	11	5	51	390	591	**2,278**	2	181	666	4,661
Chinese	62	2	16	.	1	5	24	2	.	1	68	.	2	.	**1,323**	135	380	2,021
Any other	1,198	26	1,249	80	62	77	255	163	130	35	837	92	378	163	149	**4,032**	2,233	11,159
Not stated	35,701	197	2,058	185	121	118	267	2,076	915	698	498	326	1,097	316	263	1,192	15,749	61,777
Total	334,139	2,379	27,264	1,583	1,346	941	2,525	13,872	20,739	7,443	7,036	6,269	15,261	5,164	2,342	11,965	105,700	565,968

Source: HES and registration

Discussion

Northgate Solutions designed part of the linkage algorithm based on its previous experience of linkage and this was enhanced by the authors to improve the linkage rate. Northgate also linked Maternity HES records (mother and baby records) to registration and NN4B linked data. There were however, some issues with the linked file as in some cases more than one registration and NN4B linked records were linked to the same Maternity HES record using different combinations of patient identifiable information. Also in the mother file, delivery records were provided (as requested) but in the baby file both birth and general episodes were provided when only the former were requested. Furthermore, HES data included multiple records for each episode. These had to be reviewed to identify the record with the most information that should be kept for analysis and the rest were deleted. Hence, a considerable amount of time was spent in cleaning the files. The final files consisted of one maternity record linked to a registration and NN4B linked record for the mother, and one HES baby record linked to a registration and NN4B linked record for the baby. The linkage method used by Northgate Solutions will be simplified before data for 2007 are linked, so that a maternity HES record is linked to a single linked registration and NN4B record.

Two-thirds of the registration and NN4B records were linked to the HES mother record using NHS number and partial date of birth. This was not surprising as mother's NHS number was missing in nearly a third of the registration and NN4B linked records and also a very small proportion of Maternity HES records. A further quarter of the registration and NN4B linked records were linked using date of birth or month and year or birth, and postcode. There were concerns about using postcode in the linkage algorithm, as the HES index used for linkage is derived using current postcode of residence of the mother and the postcode on registration and NN4B linked data was recorded at the time of registration. It is possible the mother could have moved since having the baby and this variable is also subject to recording and reporting errors. Despite this, an overall linkage rate of over 90 per cent was achieved for both 2005 and 2006. This could have improved further if there was a shorter delay before linkage was carried out, as HESID would be less likely to have changed, or if HESID at birth was retained as a separate field for linkage.

The linkage rate for registration and NN4B linked data to HES baby record was slightly lower than the linkage rate for the mother record. This was not surprising, as large proportion of 'baby tails' are known to be missing in Maternity HES (see HES website)

HES mother records include information on the baby, and as the linkage rate for registration and NN4B linked data to HES mother records was better than the baby records quality of information on HES was assessed using the mother records. There were however, issues with multiple births in the HES mother record. Information was often recorded for the first baby only and for the remaining multiples it was either missing or the same as the first baby. But in very few cases it was different. Multiple birth status was also unknown in a fifth of the records. Further work to assess quality of data on multiple births is necessary before it could be used for any analyses.

Discrepancy in the recording of live/still birth status for singleton babies was found in 5 in 10,000 linked records in 2005, and 2 in 1,000 records in 2006. Classification at birth registration is judged to be more reliable as a medical certificate of stillbirth is required to register a stillbirth. Also a third of the HES records did not have any information on birth status; this is a much higher proportion

than found in pilot study involving linkage of Maternity HES data for one month to birth registration (Abraham C *et al.*, 2002).

Birthweight was missing in a quarter of all linked Maternity HES records for singletons babies, compared with only 0.2 per cent at birth registration. There was however, good concordance between the two data sources where birthweight was stated, as 99 per cent of records were in the same 500g birthweight group. Missing birthweights are investigated by ONS by going back to registrars and also to child health departments. Therefore the quality of birthweight information on registration is expected to be better and more reliable than in Maternity HES.

Gestational age is not recorded at registration for live births but is available from the NN4B data. This records gestational age in weeks 'calculated from relevant menstrual data held within the maternity system' whereas Maternity HES specifies 'time from the first day of the last menstrual period (LMP)'. Where this is not available an estimate is supposed to be recorded. However, it is likely the gestational age assessed by ultrasound is now used because second trimester scans are a routine part of antenatal assessment in Britain. A study of births at 27/28 weeks of gestational age in England, Wales, and Northern Ireland between 1998 and 2000 showed that 79 per cent of the mothers had had an ultrasound before 20 weeks gestation, and 85 per cent had had menstrual history recorded (Confidential Enquiry into Stillbirths and Deaths in Infancy, 2001). Gestational age distributions have been shown to differ depending on the method used to assess gestational age. Studies have shown that if second trimester ultrasound is used rather than LMP, then the mean gestational age is one week less. However, gestational age differed by one week in only 7 per cent of the linked records. Of all the linked records, nearly half of the HES records had no information on gestational age, and where gestational age was stated it was in good agreement with NN4B in majority of the records.

The NN4B system requests information on the ethnic category of the baby, as defined by the mother using the 2001 Census categories (Moser K *et al*., 2008). On the Maternity HES record it is the mother's ethnicity which is self-selected using the 2001 Census categories. It is however, unclear in both of these data sources, whether the mother was involved in defining the ethnic category or the health professional decided what to record rather than asking the mother. In practice it is likely to be a mixture of both, and although the ethnic category of the baby is requested in NN4B it is not possible to know whose ethnicity was actually recorded, the mother's or the baby's. A further consideration is that people's identification with an ethnic group is not always straightforward and individual responses, whether self-reported or not, may vary according to circumstances and over time. Despite these limitations, in over 80 per cent of the linked records mother's ethnicity was the same as that of the baby. In 4.2 per cent of records mother's ethnicity has been categorised as 'White British' and baby's ethnicity has been categorised as 'White other' or vice versa. This suggests that father's ethnicity may have been taken into consideration in recording the baby's ethnic category on the NN4B data and this is more likely to have been defined by the mother.

Conclusion

This study shows that it is possible to link the majority (90 per cent) of Maternity HES records to registration and NN4B linked records, but the method used for linkage by Northgate Solutions needs to be amended before it is used to link data for future years. Linkage would be beneficial, and should be carried out routinely, if data quality and completeness improves in Maternity HES. However, at national level, information such as parity, method of delivery, complications in pregnancy is only available from Maternity HES, so linkage would be necessary to access this information, together with the data obtained from birth registration and NN4B.

Birth registration and NN4B are more reliable sources of data than Maternity HES. But where data have been recorded they are in good concordance with birth registration or NN4B, but there were a large proportion of linked records where information was not recorded on Maternity HES.

Acknowledgement

This work forms part of the linkage, analysis and dissemination of national birth and maternity data for England and Wales project, funded by the Medical Research Council as part of the Joint Wellcome Research Councils Electronic patient data linkage initiative. We would like to thank Northgate Solutions, in particular Jonathan Low, for linking the datasets; Julie Messer at ONS for providing the birth registration – NHS Numbers for Babies linked data to Northgate Solutions to link to Maternity HES records, making the linked data accessible in the VML system and for releasing outputs; Chris Roebuck and Tony Childs at the NHS Information Centre, for their advice and support. Collaborators in the original National Gestational Age project included, in addition to the authors, Lesz Lancucki, formerly Maternity Hospital Episode Statistics, Community Health Statistics and Surveys, NHS Information Centre and Tony Couch, formerly Head of Information Products, Health Solution Wales who we would like to thank for their help in the earlier stage of the project. We are grateful to Gwyneth Thomas, Health Statistics and Analysis Unit, Welsh Assembly Government and Martin Ward Platt, Clinical Director, Regional and Maternity Surveys Office, North East Region for their help and support in this project.

References

Abrahams C and Davy K (2002). Linking HES maternity records with ONS birth records. *Health Statistics Quarterly* 13, 22–30

Brocklehurst P (1999). Infection and preterm delivery. *British Medical Journal*, 318, 548–549.

Confidential Enquiry into Maternal and Child Health (2004) Stillbirth, neonatal and postneonatal mortality 2000–02, England, Wales and Northern Ireland.

Confidential Enquiry into Stillbirths and Deaths in Infancy (2001) *8th Annual Report*, Maternal and Child Health Research Consortium: London.

Hilder L, Moser K, Dattani N and Macfarlane A (2007). Pilot linkage of NHS Numbers for Babies data with Birth registrations. *Health Statistics Quarterly* 33,25–33.

ISD Scotland and Scottish Programme for Clinical Effectiveness in Reproductive Health (2004) *Scottish Perinatal and Infant Mortality and Morbidity Report 2003*, SPERCH Publication No 21, NHS Scotland: Edinburgh

'Maternity data in HES' available on the Information Centre website at: www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=925

Moser K and Hilder L (2008). Assessing quality of NHS Numbers for Babies data and providing gestational age statistics. *Health Statistics Quarterly* 37,15–23

Moser K, Stanfield K M and Leon D A (2008). Birthweight and gestational age by ethnic group, England and Wales 2005: introducing new data on births. *Health Statistics Quarterly* 39,22–31.

National Health Service Act 1977, S124 (4–7).

Office for National Statistics, *Mortality Statistics: Childhood, infant and perinatal, England and Wales, 2007.* Series DH3 No. 40. Available on the ONS website www.statistics.gov.uk/StatBase/Product.asp?vlnk=15362

Appendix

For linkage of registration and NHS Numbers for Babies (NN4B) linked data to the mothers records (also known as delivery record) in HES, a file consisting of a small subset of data items including the mother's NHS number, date of birth and age, the baby's date of birth, their postcode and a unique ID compiled by ONS was sent to Northgate Solutions. A second file consisting of baby's NHS number, date of birth, postcode of residence and the unique id was sent to link to the baby records, known as 'baby tails'. Northgate extracted the mother and baby records from HES using the filters shown in Box A1. These records were then linked to the registration and NN4B linked records.

Box A1 The filters used by Northgate Solutions to extract HES records for mother and baby

Mother records	Baby records
Episode type = 2 or 5 (delivery episodes)	Episode type = 3 or 6 (birth episodes)
AND	AND
Main procedure/Intervention Code between R14 and R27	Well baby index = YES
OR	OR
Main diagnosis code between ICD codes O00 and O99 (O = pregnancy and childbirth and Puerperium conditions).	Neonatal level of care in (0,1,2,3,4,9)
	OR
	Main diagnosis code between ICD codes P00 and P99 (Conditions originating in perinatal period)
	OR
	Main diagnosis code between ICD codes Q00 and Q99 (Congenital malformations, deformities, chromosomal abnormalities).

For the registration and NN4B linked data, indirect identifiers were used for linkage of records where the mother's NHS number was missing. These included different combinations of mother's date of birth, postcode, and baby's date of birth. A pilot study, using registration and NN4B linked data for all babies that were born in January 2005, was carried out to test the algorithm compiled by Northgate Solutions, involving stages 1–3 as shown in Table A1. This was based on their previous experience of linking mortality registration records to HES.

Examination of the unlinked registration and NN4B linked data and maternity HES records showed that the linkage rate could be improved by using partial information, such as month and year of birth or first four characters of postcode. Therefore further stages were added to the algorithms

used for linking the mother and baby records before the annual registration and NN4B linked data files for 2005 and 2006 were sent to Northgate Solutions. The final algorithm used to link mothers records to the registration and NN4B linked data involved 9 stages, of which linked records were obtained only from stages 1, 2 and 5 (Table A1). A variable indicating stage of algorithm used for linking HES records to the registration and NN4B linked dataset was provided by Northgate Solutions.

Table A1 Number and percentage of registration/NN4B records that were linked to HES mothers records by algorithm, 2005[1] and 2006[2]

England

Stage	Variables used	Records linked			
		2005		2006	
		Number	Percentage	Number	Percentage
1	Exact NHSNO + partial mother's DOB	395,855	*64.1*	431,296	*67.4*
2	Exact mother's DOB + exact postcode	168,704	*27.3*	152,040	*23.8*
3	Exact NHSNO + exact mother's DOB	-	-	-	
4	Exact NHSNO + month and year of mother's DOB	-	-	-	
5	Exact postcode + month and year of mother's DOB	1,409	*0.2*	1,083	*0.2*
6	Exact baby's DOB + baby sex + exact POSTCODE	-	-		-
7	Exact baby's DOB + exact POSTCODE	-	-		-
8	Exact baby's DOB + first four characters of POSTCODE	-	-		-
9	Month and year of Baby's DOB + exact POSTCODE	-	-		-

1 2005 - there were 51,645 (~8.36%) registration/NN4B records that did not match.
2 2006 - there were 55,852 (~8.72%) registration/NN4B records that did not match.

Source: HES

Linkage for babies' records consisted of five stages of an algorithm which included a mix of baby's NHS number, date of birth and postcode. The records were linked using all the stages (Table A2).

Table A2 Number and percentage of registration/NN4B records that were linked to HES baby records by algorithm, 2005[1] and 2006[2]

England

Stage	Variables used	Records linked			
		2005		2006	
		Number	Percentage	Number	Percentage
1	Exact SEX + Partial DOB[3] + Exact NHSNO	515,434	*83.5*	525,897	*82.1*
2	Exact SEX + Exact DOB + Exact POSTCODE	4,597	*0.7*	10,213	*1.6*
3	Exact DOB + Exact POSTCODE	176	*<0.1*	377	*0.1*
4	Exact DOB + first 4 characters of POSTCODE	319	*0.1*	1,838	*0.3*
5	Exact POSTCODE + month and year of DOB	1	*<0.1*	-	-

1 2005 - there were 97086 (15.7%) registration/NN4B records that did not match.
2 2006 - there were 101946 (15.9%) registration/NN4B records that did not match.
3 DOB = Baby's Date of Birth.

Source: HES

In some instances, after the first appropriate stage of the algorithm was applied leading to a HES record being linked to a registration and NN4B linked record, using further stages of the algorithm, the same HES record was linked again to a different registration and NN4B linked record. These were identified using the HESID. Records that were uniquely linked using either stages 1 or 2 of the algorithm were kept for analysis and the rest were examined manually to see whether data items such as date of birth of mother, date of birth of baby and postcode were consistent on HES and registration and NN4B linked file. Linkage to HES was deleted for records where the data items were inconsistent.

In addition, the data sent by Northgate Solutions contained multiple records for each HESID. These were identified using the variable Epikey. Multiple epikeys were present for both the mother and baby records. The record with the highest epikey was kept as this contained information for most of the variables required for analysis and all others were deleted.

Northgate Solutions had used the HES index to link the Maternity HES records to registration and NN4B linked data. Registration and NN4B linked records for 2005 and 2006 were linked using the most recent HES index and this might have contributed to some discrepancies in the linkage rate.

Narrative verdicts and their impact on mortality statistics in England and Wales

Chris Hill and Lois Cook *Office for National Statistics*

Abstract

Background

Annually, there are around 30,000 coroner's inquests held in England and Wales that conclude with a verdict. 'Short form' verdicts such as accident or misadventure; natural causes; suicide; and homicide make up the majority of all verdict conclusions. 'Narrative' verdicts can be used by a coroner or jury, instead of a short form verdict, to express their conclusions as to the cause of death following an inquest. Since 2001 narrative verdicts have been more widely used, with over 3,000 narrative verdicts returned in 2009. In some cases, it can be difficult to code the underlying cause of death from the information provided in the narrative. For some time, the Office for National Statistics (ONS) and other organisations have been concerned about the impact of narrative verdicts on the quality of the statistics on cause of death. Our research investigated the impact of narrative verdicts on trends for deaths attributed to injury and poisoning in England and Wales.

Methods

The research considered narrative verdicts received by ONS between 2001 and 2009. All available information provided by the coroner from the narrative verdict, together with the underlying cause of death, was used in the analysis. All causes of death where a narrative verdict was returned were investigated. More in-depth analysis of accidental deaths was undertaken, as classification of these deaths by intent is more difficult when the information from the coroner is imprecise. A sensitivity analysis of suicide rates (intentional self-harm and event of undetermined intent) was carried out. This involved using two different scenarios of reclassifying selected proportions of accidental hanging and poisoning deaths, where a narrative verdict was returned, as intentional self-harm. An exercise to measure the consistency of coding cause of death from narrative verdicts was also undertaken.

Results

The increasing proportion of narrative verdicts involving injury and poisoning has not significantly affected published mortality rates for suicide (intentional self-harm and injury or poisoning of undetermined intent). However, if the rise in narrative verdicts continues at the same rate, the accurate reporting of injury and poisoning deaths, including suicides, is likely to be affected. The exercise to establish the consistency of coding the cause of death by ONS cause coders showed that the current coding rules were being applied uniformly.

Conclusions

The increase in the use of narrative verdicts by coroners has not had a statistically significant impact on published suicide rates in England and Wales and so no revision to these rates is needed. A review of current coding practices and the handling of narrative verdicts will be undertaken by ONS with particular reference to deaths from intentional self-harm. A recommendation has been made to coroners to consider ways of recording narrative verdicts to allow more accurate coding of cause of death. This will ensure that mortality statistics are maintained to the highest standards.

Contents

List of Figures

List of Tables

Introduction

Mortality statistics provide a rich source of information for population health surveillance. They are used to assess the effectiveness of health and social care services and to evaluate the success of policies to reduce mortality rates. The reliability and accuracy of mortality statistics is crucial in the assessment of their impact. It is essential that the ONS is able to classify the cause of death accurately from information received from the certifier of the death.

Deaths can be divided into two main cause groups – those that are disease related, or have other internal conditions as their underlying cause of death; and those that are the result of an injury or poisoning and are attributed to an external cause of death. This distinction applies to all deaths, including those that have been the subject of an inquest. Where a narrative verdict is returned, the potential for misclassification of the injury and poisoning is greater if there is insufficient information from the certifier. If the narrative verdict does not contain sufficient detail about who caused the initial event leading to the death, or their intention at the time of the action, then the coder has to assign the death as an accident in accordance with international coding rules. This may affect the quality of mortality statistics.

Background

Coroners are independent judicial officers appointed and paid for by the relevant local authorities. They are responsible for investigating violent, unnatural deaths or sudden deaths of unknown cause that are reported to them, including all deaths occurring in custody. Coroners are also responsible for registering the cause of death for those deaths they investigate. Unnatural deaths exclude those deaths that have disease and other internal conditions as their underlying cause.

A coroner's inquest is an inquiry to establish the facts surrounding these types of death. The verdict following the inquest is usually returned in short-form format that fits into one of a series of established categories, for example, suicide, accident or misadventure. Alternatively, a narrative verdict can be returned by the coroner or jury. Narrative verdicts record the facts surrounding the death in more detail, but do not use a standard verdict to express their conclusions as to the cause of death. In some cases, it can be difficult to code the underlying cause of death from the information provided by the coroner in the narrative verdict. The aim of the research was to assess the impact of narrative verdicts on National Statistics relating to cause of death.

Since 2001 there has been a downward trend in the number of deaths registered in England and Wales. However, there has been an increase in the number of coroner's inquests over the same period. According to Ministry of Justice statistics (Ministry of Justice, 2010), the number of inquests opened in England and Wales has increased each year since 2001. In 2009 inquests were opened on around 31,000 deaths reported to coroners, compared with about 25,800 in 2001. 'Non-specific verdicts' (which include narrative verdicts) formed less than 1 per cent of all verdicts in 2001, but accounted for more than 13 per cent of verdicts in 2009. Recent case law may be responsible for the increased number of narrative verdicts in recent years, including the House of Lords 'Middleton' judgement (Ministry of Justice, 2010). The use of narrative verdicts by individual coroners varies throughout England and Wales.

Selecting the underlying cause of death

The selection of the underlying cause of death is based on the Tenth Revision of the *International Statistical Classification of Diseases and Related Health Problems,* (ICD-10) (WHO, 1992–1994) and is derived from the condition or conditions reported by the certifier, as recorded on the medical certificate of cause of death. The underlying cause of death is defined by the World Health Organisation (WHO) as:

- the disease or injury that **initiated** the train of events directly leading to death, or
- the circumstances of the accident or violence that produced the fatal injury

In order to preserve the international comparability of statistics, and to take account of different medical and legal practices, internationally agreed rules have been devised to code the underlying cause of death from the information recorded on the death certificate and other supporting documentation.

While the majority of deaths in England and Wales have a cause of death assigned automatically by coding software, around 20 per cent (100,000) of deaths are manually coded, including all deaths which have been the subject of an inquest.

When coding injury and poisoning deaths, both the intent and the mechanism determines how the death is assigned to an ICD code. The 'manner of death' (which equates to verdict in England and Wales) is also recorded. This includes accident, homicide, suicide, or event of undetermined intent if the intent could not be determined following investigation. The mechanism includes, for example, suffocation, drowning, poisoning or transport-related. As the majority of injury and poisoning deaths are investigated by a coroner, it is the verdict following the inquest, which should provide this information. Where a narrative verdict has been returned, details about the initial intent behind the act leading to the death, together with the mechanism used, may not always be clear from the facts provided by the coroner.

Intentional self-harm

To return a verdict of suicide, the coroner must be satisfied beyond reasonable doubt that the deceased intended to kill him/herself. However, in recognition of the different international medical and legal practices, the ICD code 'intentional self-harm' (which includes suicide), requires only that the certifier states that, to the best of their knowledge and belief, the initial event which led to the death was that the individual intended to harm himself/herself. As a high proportion of people who kill themselves are reported to be suffering from a mental health problem, (Harris and Barraclough, 1997; Pirkis and Burgess, 1998) it may be difficult to establish at an inquest whether the deceased was capable of understanding the consequences of their actions. In England and Wales it has been customary to assume that most injury and poisoning deaths, where the intent could not be determined following investigation, are those where the harm was self-inflicted, but there was insufficient evidence to prove that the deceased deliberately intended to kill themselves (Brock and Griffiths, 2003). Therefore, the definition of suicide used in ONS publications is any death given an underlying cause of intentional self-harm, or injury/poisoning of undetermined intent (Box 1).

Box 1 National Statistics definition of suicide

ICD–10 code	Description
X60–X84	Intentional self-harm
Y10–Y34[1]	Injury/poisoning of undetermined intent

1 From 2007 onwards, ONS has used ICD–10 code U50.9 for deaths with adjourned inquests. In the past, these deaths were coded to Y33.9. The principal reason for the change was to exclude these deaths from the Y10–Y34 range, making the tabulation of events of undetermined intent easier.

Determining intent from narrative verdicts

In some cases, the intent and mechanism are clearly stated in the narrative verdict. However, in a proportion of injury and poisoning deaths, where a narrative verdict has been returned, ONS has no indication from the information provided by the coroner of whether the fatal injury, or toxic substance, was self-administered, or if there was deliberate intent to self-harm. The rules of coding cause of death dictate that, where no indication of intent has been given by the certifier, deaths from injury or poisoning must be coded as accidents. Consequently, the net effect of the increase in narrative verdicts may have been to inflate the number of deaths classified as accidents and decrease the number classified as intentional self-harm. This is important, as trends in these types of death are closely monitored.

Methods

The analyses in this report focus on deaths with a narrative verdict received by ONS between 2001 and 2009 in England and Wales. Narrative verdicts were identified using the death registrations database. Variables selected for the analyses included age, sex, underlying cause of death and all coroner's text received by ONS. Descriptive statistics were produced and further analysis of the narrative text carried out to identify examples of text where the intent behind the initiating act causing the death was unclear.

Sensitivity analysis

To assess the potential impact of coroners providing insufficient detail in narrative verdicts to establish the intent, simulated age-standardised mortality rates (ASMRs) were calculated for each of the years 2001 to 2009 using two different assumptions. Firstly, mortality rates were calculated assuming all deaths, where a narrative verdict had caused the death to be coded as an accidental poisoning or hanging, were intentional self-harm (scenario 1). Secondly, rates were calculated assuming half of these deaths were intentional self-harm (scenario 2). Hangings and poisonings were used as these are the two most common methods for intentional self-harm in England and Wales. The ASMRs were calculated for all ages and were directly standardised using the European Standard Population.

Consistency of coding cause of death

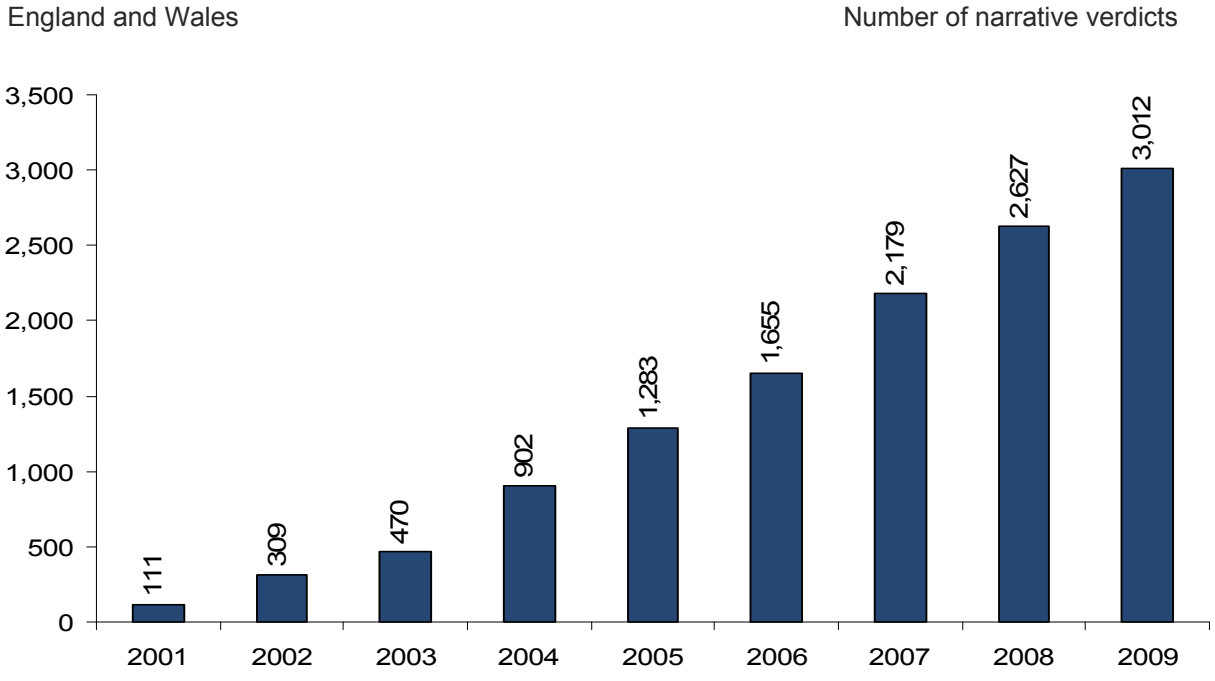
In order to assess the consistency of cause of death coding within ONS, an exercise was set up to evaluate any variation in the coding of narrative verdicts over time. All narrative verdicts, with the exception of neonatal deaths and transport accidents (V01–V99), were selected for the years 2004 to 2008. Neonatal deaths were excluded because they do not have an underlying cause of death. The reliability of coding transport accidents was assumed to be sufficiently high to allow the removal of these records. A total of 7,914 records were coded by eight experienced ONS cause coders who were able to see the coroner's text and cause of death text for each record, but not the original underlying cause of death code. Each case was coded and the results entered on a test database. A comparison could then be made between the original cause and the new cause, and any differences identified. A manner of death code, or verdict outcome, was also recorded by the coders in the exercise.

The coding frame used by ONS to classify the manner of death was also reviewed to ensure that it included common phrases used by coroners in their narrative verdicts.

Results

Since 2001 there has been a downward trend in the number of deaths registered in England and Wales, with the total number of registrations falling by around 8 per cent between 2001 and 2009. Despite this trend, the number of coroner's inquests has been rising since 2001, together with the proportion of these with a narrative verdict.

Figure 1 **Number of narrative verdicts: by year of death registration, 2001–09**



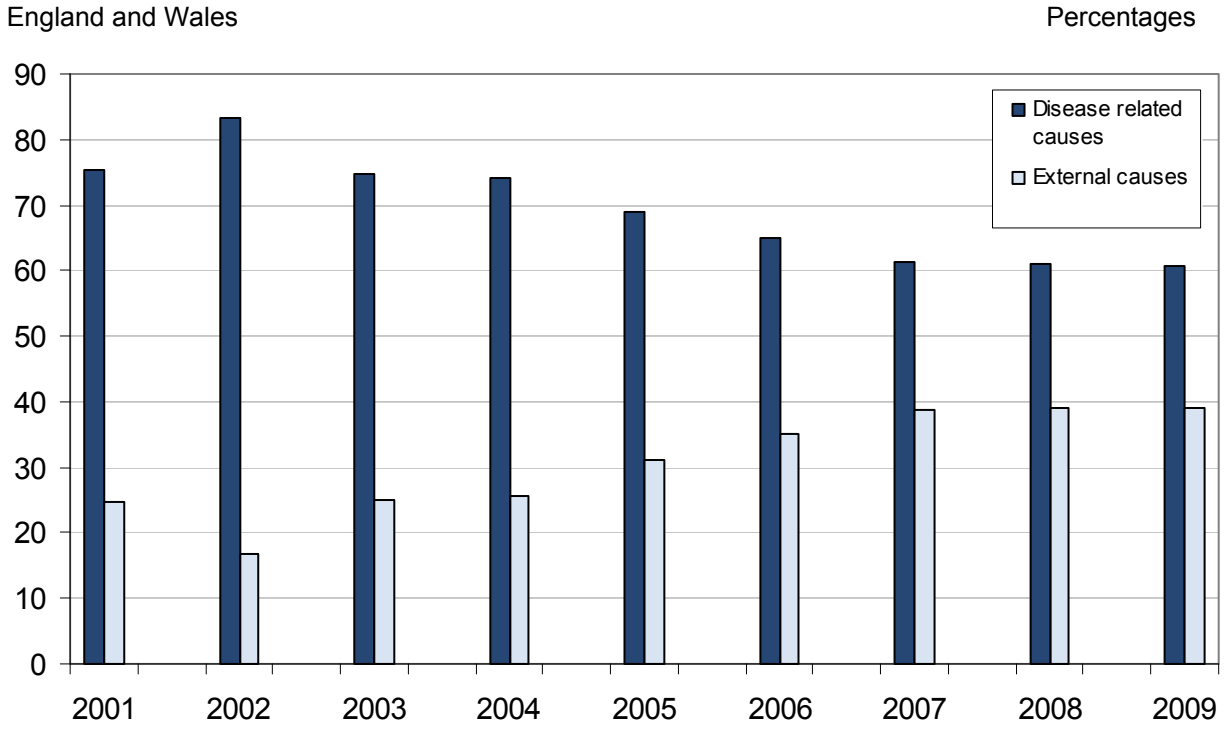
Source: ONS

Between 2001 and 2009 there has been a 27-fold increase in narrative verdicts coded by ONS (Figure 1).

Narrative verdicts and cause of death

The increase in narrative verdicts has been associated with an increase in the proportion of these which are deaths from external causes, especially from 2005 onwards (Figure 2).

Figure 2 Narrative verdicts by cause of death, 2001–09



Source: ONS

Currently around 40 per cent of deaths with a narrative verdict are classified as having an external cause of death, compared with around 25 per cent in 2001. While the proportion of all deaths which are from external causes has remained stable over the period (around 3.5 per cent), there has been a slight increase in the proportion of these deaths classified as being accidental, from 2.1 per cent in 2001 to 2.4 per cent in 2009 (ONS, 2010).

Table 1 Narrative verdicts classified as accidental deaths, 2001–09

England and Wales — Number of verdicts, percentages

	2001	2002	2003	2004	2005	2006	2007	2008	2009
All narrative verdicts	111	309	470	902	1,283	1,655	2,179	2,627	3,012
Narrative verdicts coded to accident	21	25	58	155	295	457	728	908	1,002
Percentage of narrative verdicts where the cause of death was accident	*18.9*	*8.1*	*12.3*	*17.2*	*23*	*27.6*	*33.4*	*34.6*	*33.3*

Source: ONS

Table 1 shows the increase in the proportion of narrative verdicts classified as having an accidental cause of death between 2001 and 2009. The proportion of narrative verdicts classified as being accidental deaths has risen from around a fifth (19 per cent) in 2001 to a third (33 per cent) in 2009. As mentioned above, unless there is sufficient information from the coroner to state otherwise, the coder has to classify the death as an accident in accordance with international coding rules.

Types of accident

Table 2 gives the aggregate numbers of accidental deaths for narrative verdicts by selected mechanism. Between 2001 and 2009 the two most common mechanisms for accidental deaths were transport accidents and falls, accounting for 50 per cent of the total. As these deaths are more likely to be accidental, the analyses concentrated on deaths from poisoning or suffocation (predominately hangings), these being the next two most common mechanisms. For these types of death, the intent behind the initial act leading to the death can be more difficult to determine from the information contained in a narrative verdict. It should also be noted that the majority of accidental deaths with an unspecified mechanism are likely to have been falls (Griffiths and Rooney, 2003).

In England and Wales the most common methods of suicide are hangings (including strangulation and suffocation) and poisonings. In 2009 these two mechanisms accounted for 60 per cent and 19 per cent of intentional self-harm deaths respectively (ONS, 2010). Where a narrative verdict had been returned, it is possible that a proportion of accidental deaths involving these two mechanisms could be classified as intentional self-harm, but there is insufficient detail in the narrative to confirm the intent.

Table 2 Narrative verdicts by selected mechanisms for all accidental deaths, 2001–09

England and Wales

Mechanism	Number	Percentages
Falls	1,085	*29.7*
Transport accidents	745	*20.4*
Poisoning	679	*18.6*
Suffocation	408	*11.2*
Drowning	109	*3.0*
Unspecified factor	330	*9.0*
Other accidents	293	*8.0*
All	**3,649**	***100.0***

Source: ONS

Analysis of coroner's text

Narrative verdicts vary in both length and detail ranging from a single line of text to detailed factual records, especially if the death occurred in custody. Some narrative verdicts contain all the information required to code accurately the underlying cause of death. In other cases, crucial detail is not provided.

The following examples illustrate deaths, where a narrative verdict has been returned, which are difficult to code. An example of the simplest narrative verdict record is where the cause of death is given as 'hanging', or 'suspension by ligature around neck' and the accompanying coroner's narrative verdicts reads 'On date at x, the deceased was found with a ligature around her neck.' The text does not provide any information as to whether there was a third party involved, nor if there was an intent to self-harm. With this vital detail missing, the cause of death code has to be accidental hanging as laid out in the ICD rules.

An example of a narrative verdict following a death in custody is 'Mr x, after being found hanging in his cell at x youth offenders institution on date**,** died on date at x infirmary. It was a serious omission by x young offenders institute not to have informed x's parents on each occasion that x had self-harmed. The jury's verdict is that x died from hanging which caused his death.' With this example, there is some evidence that there was intent to self-harm because previous instances of this are mentioned in the narrative. However, as there is no mention of the intent surrounding the initiating act that caused the death, the death has to be classified as accidental.

There may also be references to use of drugs or alcohol, together with mentions of 'mental illness' which may cast doubt about whether the deceased was in control of their own actions. Again, without a clear mention of intent behind the hanging, these records must classified as accidental deaths. It should be noted that a proportion of hangings will be accidental, for example, a person

playing a game, or an auto-erotic act that went too far. However, further analysis of over 400 records in this study involving hangings revealed that these events were rare.

For poisoning deaths, the information provided in the narrative verdict about the events surrounding the death can be less clear, especially where there is an overdose of drugs. A verdict of suicide will only be returned by the coroner if this is proved beyond reasonable doubt. For drug poisoning, there may not be any clear indication of intent to self-harm as the drugs may have been taken for recreational purposes. Similarly, there may not be sufficient detail in the narrative verdict regarding the involvement of any third party.

An example of a narrative verdict using the word ‘overdose’ is ‘Died as a result of taking an overdose of heroin with alcohol’. The cause of death text on the coroner’s form was ‘Overdose of heroin with alcohol’. Again, there is no indication within the narrative verdict to verify if the overdose was self-administered and no indication of intent to self-harm. A similar example is ‘Died following self-administration of an excessive amount of non-prescribed medication and alcohol’. Although this narrative text clearly states that the drug was self-administered there is doubt regarding the intent surrounding the taking of the non-prescribed medication with the alcohol. The intent may have been to ‘get high’ or to relieve pain, depending on the drug involved. A self-administered overdose of heroin, codeine or co-proxamol, may tacitly point to accidental death, but this could still have been a deliberate act. As nothing can be assumed, the only option is to classify such deaths as accidental according to the ICD rules. Use of the word ‘overdose’ can be subjective when the quantities taken following the investigation are not mentioned in the narrative verdict.

Poisonings by other noxious substances were also reviewed. Accidental deaths from carbon monoxide poisoning were generally well documented with the narrative detailing a faulty gas boiler or blocked flue. However, where a person was found dead in a car from carbon monoxide poisoning, additional information to classify the cause of death could be absent with no indication of whether the deceased intended to kill her/himself.

Sensitivity analysis – Hangings and poisonings

The research investigated the impact of potential misclassification on trends in deaths attributed to intentional self-harm and events of undetermined intent. This is the National Statistics definition used by ONS to monitor trends in suicide deaths in England and Wales.

Figure 3 (scenario 1) shows the results of combining all accidental hangings and poisonings from narrative verdict records with existing suicide death rates (intentional self-harm deaths and event undetermined intent). This was carried out for hangings and poisonings both separately and jointly. Between 2001 and 2004, the inclusion of accidental hangings and poisonings had no effect on the suicide rate for England and Wales. In 2005 adding hangings increased the mortality rate from 84 deaths per million population, to 85 deaths per million population. A similar increase in the rate was seen in 2007. In 2008 the effect was greater, increasing the rate from 81 deaths per million population, to 83 deaths per million population, and likewise in 2009 the effect was to increase the rate from 82 deaths per million population to 84 deaths per million population.

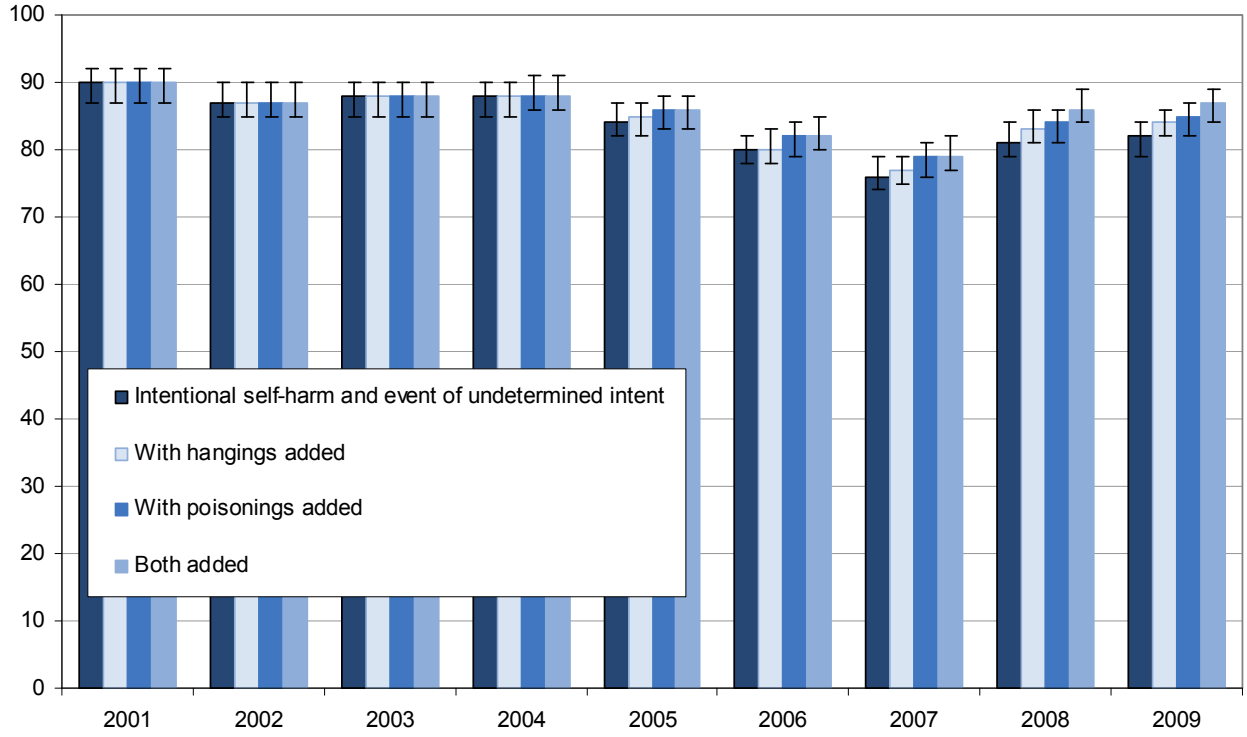
Between 2007 and 2009 a similar effect was seen when poisonings were included. In 2007 the rate increased from 76 deaths per million population to 79 deaths per million population. In 2008 the rate increased from 81 deaths per million population to 84 deaths per million population, and in 2009, adding poisoning, increased the rate from 82 deaths per million population to 85 deaths per million population.

The effect of adding both accidental hangings and poisonings had greatest overall effect on the suicide mortality rates for 2008 and 2009.

Figure 3 Simulations[1] of age-standardised mortality rates for suicide deaths[2], 2001–09 (scenario 1)

England and Wales

Rate per million population



1 Addition of all deaths where the verdict was a narrative and the underlying cause of death was accidental hanging or poisoning.

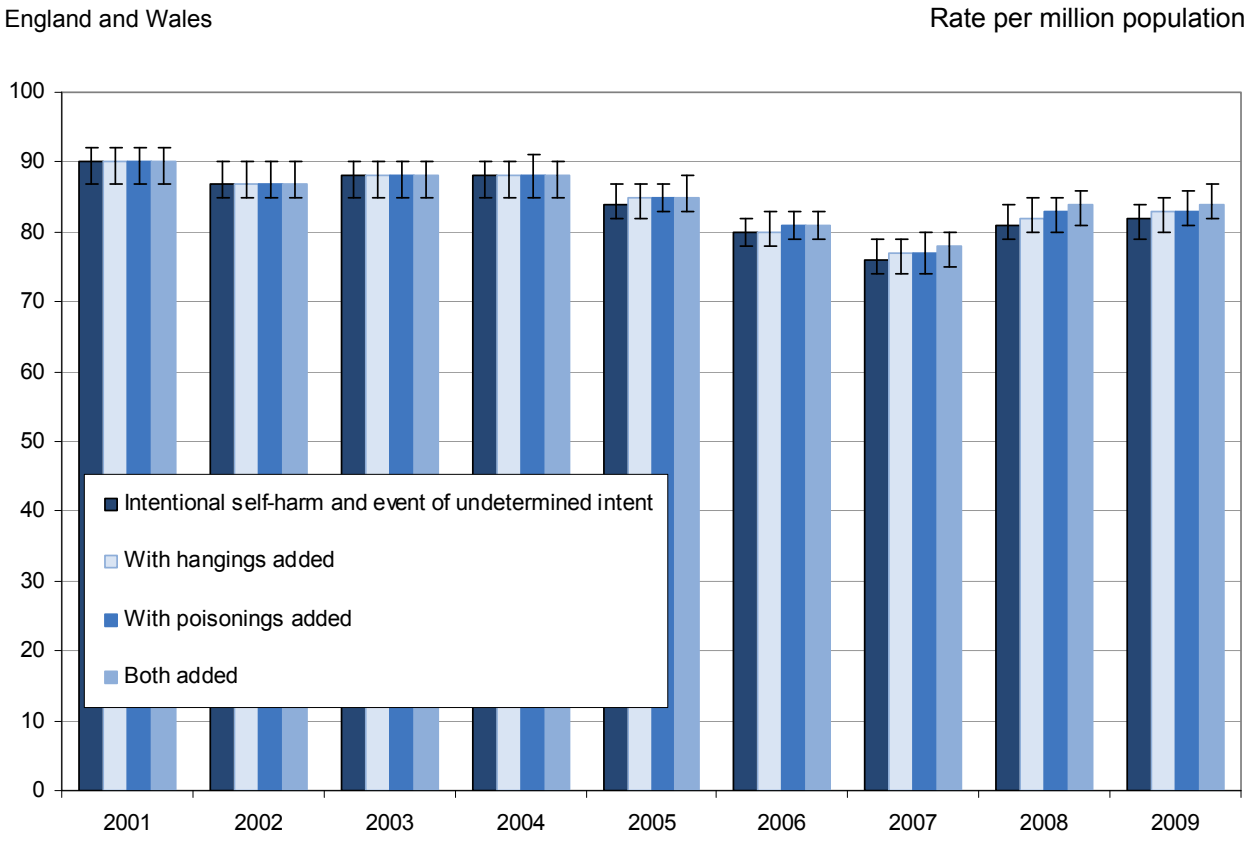
2 Includes deaths with an underlying cause of death intentional self-harm and event of undetermined intent.

Source: ONS

In 2009 the effect of adding both accidental hangings and poisonings would increase the mortality rate from 82 deaths per million population, to 87 deaths per million population, the highest rate since 2004 when the suicide rate was 88 deaths per million population. The differences in the mortality rates are not statistically significant. This simulation could be considered the worst-case scenario as it assumes that all narrative verdicts coded as accidental hangings and poisonings should be classified as intentional self-harm.

Figure 4 shows how the suicide rate would be change if half the accidental hanging and poisoning deaths with a narrative verdicts were added to those coded as intentional self-harm and events of undetermined intent (scenario 2). As in scenario 1, there is no effect on the rates between 2001 and 2004. For the remaining years the effect is less noticeable. In 2008 the effect of adding half of accidental hangings would increase the mortality rate from 81 deaths per million population, to 82 deaths per million population. In 2009 the effect was to increase the rate from 82 to 83 deaths per million population.

Figure 4 Simulations[1] of age-standardised mortality rates for suicide deaths[2], 2001–09 (scenario 2)



1 Addition of 50 per cent of deaths where the verdict was a narrative and the underlying cause of death was accidental hanging or poisoning.

2 Includes deaths with an underlying cause of death intentional self-harm and event of undetermined intent.

Source: ONS

Similar rises could be seen when half of accidental poisonings were included for years 2008 and 2009: with the mortality rate increasing to 83 deaths per million population, from 81 and 82 deaths per million population respectively for these years. Combining both hangings and poisonings with the intentional self-harm deaths would increase the mortality rate from 82 deaths per million population, to 84 deaths per million population in 2009. Again, the differences in the mortality rates are not statistically significant. While it is likely that a higher proportion of hangings than poisonings are intentional self-harm deaths, without sufficient information from the coroner, these deaths have to be classified as accidental.

Gender specific rates were calculated for all years and the results were similar to the all persons figures shown in the simulations above. Tables 4 and 5 show the figures for the two simulations for males and females for 2009. For both scenarios, the effect for males is greater than for females. Table 3 shows the reported ASMRs together with the revised ASMRs once all accidental hangings and poisonings are added. For males the effect is most visible when both hangings and poisonings are included, increasing the rate from 127 deaths per million population to 134 deaths per million population. For females, the effect of combining these two mechanisms increased the rate from 38 deaths per million population to 41 deaths per million population.

Table 3 Simulations of age-standardised mortality rates for suicide deaths: by sex, 2009 (scenario 1)

England and Wales — Rate per million population

2009	Suicide (intentional self-harm and event of undetermined intent)			Addition of deaths where the verdict was a narrative and the underlying cause of death was accidental hanging			Addition of deaths where the verdict was a narrative and the underlying cause of death was accidental poisoning			Addition of deaths where the verdict was a narrative and the underlying cause of death was accidental hanging or poisoning		
	ASMR	Lower 95 per cent CI	Upper 95 per cent CI	New ASMR	Lower 95 per cent CI	Upper 95 per cent CI	New ASMR	Lower 95 per cent CI	Upper 95 per cent CI	New ASMR	Lower 95 per cent CI	Upper 95 per cent CI
Males	127	(123)	(131)	131	(126)	(135)	131	(126)	(135)	134	(130)	(139)
Females	38	(35)	(40)	38	(36)	(41)	40	(37)	(42)	41	(38)	(43)

Source: ONS

Table 4 shows the effect of adding 50 per cent of accidental hangings and poisonings with a narrative verdict. Although the combination of both hangings and poisonings increases the rates by four deaths per million population for males from 127 to 131 deaths per million population, the differences in the rates are not statistically significant. The effect on the female rates was lower, and it is only the inclusion of the accidental poisonings that increases the rate from 38 deaths per million population, to 39 deaths per million population.

Table 4 Simulations of age-standardised mortality rates for suicide deaths: by sex, 2009 (scenario 2)

England and Wales — Rate per million population

2009	Suicide (intentional self-harm and event of undetermined intent)			Addition of 50 per cent deaths where the verdict was a narrative and the underlying cause of death was accidental hanging			Addition of 50 per cent of deaths where the verdict was a narrative and the underlying cause of death was accidental poisoning			Addition of 50 per cent of deaths where the verdict was a narrative and the underlying cause of death was accidental hanging or poisoning		
	ASMR	Lower 95 per cent CI	Upper 95 per cent CI	New ASMR	Lower 95 per cent CI	Upper 95 per cent CI	New ASMR	Lower 95 per cent CI	Upper 95 per cent CI	New ASMR	Lower 95 per cent CI	Upper 95 per cent CI
Males	127	(123)	(131)	129	(124)	(133)	129	(124)	(133)	131	(126)	(135)
Females	38	(35)	(40)	38	(36)	(40)	39	(36)	(41)	39	(37)	(41)

Source: ONS

Consistency of cause coding for narrative verdicts

The results of the cause coding exercise are shown in Table 5. Analysis of the new cause codes against the existing codes showed that 22 per cent of records (1,731) were not an exact match at the fourth digit level of the ICD–10 cause code and 10 per cent of records (797) had moved to another cause. Of the 797 records that had moved cause grouping, nearly a third (31 per cent) were now classified as external causes of mortality. A further 18 per cent had moved to diseases of the circulatory system and 9 per cent to both diseases of the digestive system and diseases of the respiratory system. The remainder were spread across other causes.

Although 247 records were reclassified as external causes of death, 211 records originally coded as external causes moved to other cause groups with the majority of these (46 per cent) classified as diseases of the circulatory system or the respiratory system. There were 10 per cent now classified as mental and behavioural disorders and a similar proportion (9 per cent) as diseases of the digestive system. Although the movement in and out of the external cause group experienced the greatest change, overall there was only a net gain of 36 records for external causes.

There were only five new records with an underlying cause of intentional self-harm. These records had originally been classified as accidental, demonstrating a high consistency of coding these potentially difficult to code records. For accidental hangings and poisonings, there were match rates of 93 and 92 per cent respectively.

The research showed that although in some the cases the coders assigned a different underlying cause of death, overall they applied the ICD classification rules consistently.

The examination of the manner of death coding frame showed that additional terms commonly used by coroners to describe intentional-self harm could be included to improve the accuracy of coding some narrative verdicts. Improvements could also be made to identify more easily that the manner of death had been obtained from a narrative verdict.

Table 5 **Changes in the ICD–10 broad cause of death group**

Broad cause of death group (ICD-10 capters) / Existing cause of death code	New cause of death code: Certain infections and parasitic diseases	Neoplasms	Diseases of blood forming organs/immune mechanism	Endocrine, nutritional and met diseases	Mental and behavioural disorders	Diseases of nervous system	Diseases of the ear and mastoid processes	Diseases of the circulatory system	Diseases of respiratory system	Diseases of digestive system	Diseases of the skin	Diseases of muscuskeletal system and connective tissue	Diseases of the genitourinary system	Pregnancy and childbirth	Conditions originating in perinatal period	Congenital malformations	Symptoms signs and abnormal clinical and lab findings	External causes of mortality	Total
Certain infections and parasitic diseases	99	3	1	2	0	1	0	2	3	1	0	0	2	0	1	0	0	8	**123**
Neoplasms	1	1260	1	0	0	1	0	8	2	12	0	0	1	0	0	0	0	12	**1298**
Diseases of blood forming organs/immune mechanism	0	2	20	0	0	0	0	0	0	1	1	0	0	0	0	0	0	3	**27**
Endocrine, nutritional and met diseases	1	2	0	112	3	0	0	3	0	2	0	3	1	0	1	0	0	5	**133**
Mental and behavioural disorders	1	0	1	1	102	1	0	7	3	2	1	0	1	1	0	0	2	14	**137**
Diseases of nervous system	0	0	0	1	4	105	0	3	0	0	0	2	0	1	0	0	0	7	**123**
Diseases of the ear and mastoid processes	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	**2**
Diseases of the circulatory system	3	3	1	6	4	0	0	1316	7	17	1	8	6	0	0	4	1	77	**1454**
Diseases of respiratory system	4	5	1	0	2	2	0	12	238	2	1	1	0	0	1	2	1	12	**284**
Diseases of digestive system	9	7	0	3	2	1	0	26	9	760	0	6	2	0	2	1	2	65	**895**
Diseases of the skin	2	0	0	1	0	0	0	3	0	0	23	0	1	0	0	0	0	3	**33**
Diseases of musculoskeletal system and connective tissue	2	0	0	0	0	4	0	8	6	10	1	345	1	0	0	1	0	18	**396**
Diseases of the genitourinary system	1	2	1	3	1	1	0	5	1	5	0	1	146	0	0	1	0	16	**184**
Pregnancy and childbirth	0	0	0	0	0	0	0	0	0	0	0	0	1	25	0	0	0	0	**26**
Conditions originating in perinatal period	0	0	0	0	0	0	0	1	0	0	0	0	0	0	11	0	0	2	**14**
Congenital malformations	0	0	0	0	0	2	0	4	2	2	0	0	3	0	0	111	1	2	**127**
Symptoms signs and abnormal clinical and lab findings	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	192	3	**197**
External causes of mortality	7	9	3	6	22	11	1	63	35	20	2	10	10	0	0	2	10	2250	**2461**
Total	**130**	**1293**	**29**	**135**	**140**	**130**	**3**	**1462**	**306**	**834**	**30**	**376**	**175**	**27**	**16**	**122**	**209**	**2497**	**7914**

Source: ONS

Discussion

There has been a sharp increase in the number of narrative verdicts returned by coroners in England and Wales since 2001. If there is insufficient information received from the coroner, the death has to be coded as accidental. The results of the exercise on coding consistency compare well with other studies of cause of death coding reliability (Harteloh *et al*., 2010). However, it is apparent that some difficult to code injury and poisoning records are being coded as accidents when they may be intentional self-harm deaths.

The simulation of higher suicide mortality rates by adding half or even all accidental deaths involving hangings and poisonings to intentional self-harm deaths did not significantly affect the reported mortality rates for 2001 to 2009. However, if the current trend in the use of narrative verdicts continues, and the proportion of these that are difficult to code also increases, this is likely to affect significantly reported mortality rates in the future.

Cause of death coders are reliant solely on the description of the events surrounding the death contained in the narrative text, together with the cause of death, to classify the death according to ICD rules. The ICD rules do not allow the coder to infer anything, and without clear reporting by the coroner of the intent, or involvement of a third party, the death has to be coded as an accident. However as the use of narrative verdicts is increasing, action is required to ensure that the results of coroners' investigation of deaths are used to provide the best possible information to code the underlying cause of death.

It is clear that trying to establish intent, or the involvement of third party, is difficult if there is limited information from the coroner. Form Rev 99 B (Part V) is used by coroners to provide information for statistical purposes to allow the cause of death to be coded according to ICD rules. However, at present the coroner is only asked to provide further information if the death is an accident or misadventure, although many coroners complete this section for all deaths that they register. As all the information provided on the Part V is used to code cause of death, it would be logical to add tick boxes to this form to establish the manner of death. The categories for manner of death would be:

- natural disease
- accidental injury and poisoning
- intentional self-harm
- assault
- event undetermined intent
- legal intervention
- operation of war
- complication of medical or surgical care

ONS is currently working with coroners to consider ways of recording information from narrative verdicts to allow more accurate coding of cause of death.

Limitations were also found in the coding frame used to code the manner of death. Additional common phrases used by coroners in their narrative verdicts were identified and additions to the terms allowed for classification of intentional self-harm will be added to the existing coding frame. Improvements will also allow better identification that the death was registered with a narrative verdict. Analysis of the impact of these changes on mortality statistics is planned for 2011.

Limitations of the analysis

The research focused on the two dominant methods for intentional self-harm – hanging and poisoning. Narrative verdicts classified as accidental deaths involving these two mechanisms were investigated and revised mortality rates calculated using simulations. Other mechanisms for intentional self-harm, for example drowning, were not considered in the analysis as the numbers of deaths using these methods were found to be relatively small. However, a combination of all possible mechanisms may affect the sensitivity analysis although this was not investigated.

The cause coding exercise only used selected narrative verdict records for 2004 to 2008. This was due to the time involved in carrying out such an exercise and the fact that processed records for 2009 were not available for use at the time the coding exercise was carried out. Some of the records may have been recoded by the coder who originally coded the death. This could have had an effect on the results, as a proportion of records may have been familiar to the coder involved. However, it was not possible to measure the level of overlap from the data supplied.

Although all available text received from coroners was reviewed, sometimes the narrative verdict was missing from the record provided at death registration. For a small proportion of deaths, this information cannot be obtained even after requests from the relevant coroner. If the information simply mentions that there was a narrative verdict, then the coder has to classify the death as accidental according to ICD rules.

In order to establish any regional variation in the use of narrative verdicts, further research by coroner district will be needed. If a high percentage of narrative verdicts are used by a small number of coroners this could have an effect on mortality rates at local area level, especially if the level of detail from these coroners does not allow for accurate coding of the intent. The effectiveness of measures to reduce intentional self-harm at local area level could therefore be misleading, if trends are based on statistics that over report accidental deaths. Further research by geography is planned for the future.

Conclusions

This article reviews the impact of narrative verdicts on mortality statistics in England and Wales from 2001 to 2009. The findings show that, while the coding of narrative verdicts has been consistent over the period, the lack of clear information from some coroners, relating to the intent, or involvement of a third party, can affect the coding of the underlying cause of death. This means that some deaths may have been classified as accidental, when they are more likely to be intentional, but the effect on reported suicide mortality rates is not statistically significant. Consequently, there will not be a revision of previously published mortality rates for these years. However, if current trends in the use of narrative verdicts continue, the effect on mortality rates may become large enough to affect the reliability of National Statistics. To help overcome this, some improvements have been made to the coding frame used by ONS to record the manner of death following an inquest. ONS is also working with coroners to consider ways of recording information from narrative verdicts to allow more accurate coding of cause of death.

Acknowledgements

The authors would like to thank Dr. Cleo Rooney and all of the ONS cause coding team for their help with this research.

References

Brock A and Griffiths C (2003) 'Trends in suicide by method in England and Wales, 1979–2001', *Health Statistics Quarterly* 20, pp 7–18, available at: www.statistics.gov.uk/downloads/theme_health/HSQ20.pdf

Griffiths C and Rooney C (2003) 'The effect of the introduction of ICD–10 on trends in mortality from injury and poisoning in England and Wales, 1999', *Health Statistics Quarterly* 14, pp 75–83, available at: www.statistics.gov.uk/downloads/theme_health/HSQ14_v4.pdf

Harris E C, Barraclough, B (1997), 'Suicide as an outcome for mental disorders', *British Journal of Psychiatry* 170, pp 205–228.

Harteloh P, de Bruin K, Kardman J (2010) 'The reliability of cause-of-death coding in The Netherlands', *European Journal of Epidemiology* 25 (8), pp 531–538.

Ministry of Justice (2010), *Statistics on deaths reported to coroners England and Wales, 2009*, available at: www.justice.gov.uk/publications/coronersannual.htm

Office for National Statistics (2010) Mortality statistics: Deaths registered in 2009, available at: www.statistics.gov.uk/statbase/product.asp?vlnk=15096

Pirkis J, Burgess P. (1998) 'Suicide and recency of health care contacts. A systematic review', *British Journal of Psychiatry* 173, pp 462–474.

World Health Organisation (WHO) (1992–94) *International Statistical Classification of Diseases and Related Health Problems,* volumes 1, 2 and 3 (Tenth Revision). WHO: Geneva.