The University of Manchester

MANCHESTER 1824

Cathie Marsh Centre for Census and Survey Research

Developing Strategies for Deriving Small Population Mortality Rates

CCSR Working Paper 2006-10
Lee Williamson
Lee.williamson @manchester.ac.uk

This working paper is intended as a resource for users intending to create small area and/or ethnic group mortality rates. It presents methods, data and various strategies for creating mortality rates for use in population projections and introduces the ways of modelling mortality rates, including literature from the actuarial field. The methods review includes mortality laws, or mathematical models, and splines. Examples are presented using data from the Bradford district and outlines data sources that could be used to assist in the creation of mortality rates for Bradford wards and/or ethnic groups (including country of birth (COB)). While no ethnic-specific age-specific mortality rates (ASMRs) are created, two strategies of using relational-type approaches are presented for the Bradford wards.

www.ccsr.ac.uk

Developing Strategies for Deriving Small Population Mortality Rates

Summary

This working paper is intended as a resource for users intending to create small area and/or ethnic group mortality rates. It presents methods, data and various strategies for creating mortality rates for use in population projections. The working paper introduces the ways of modelling mortality rates, including literature from the actuarial field. The methods review includes mortality laws, or mathematical models, and splines. The working paper presents examples using data from the Bradford district and outlines data sources that could be used to assist in the creation of mortality rates for Bradford wards and/or ethnic groups (including country of birth (COB)). While no ethnic-specific age-specific mortality rates (ASMRs) are created, two strategies of using relational-type approaches are presented for the Bradford wards.

Background information

The working paper originally formed part of the authors PhD research training as a review of methods and previous research on mortality rates at the subnational-level and by ethnic group (and country of birth). Basic strategies are presented on how to deal with small area mortality rates; however, there is no actual modelling/smoothing of mortality rates. Further to this, while ethnic-specific ward-level death data was obtained from Bradford Health Authority (HA) (based on name analysis for South Asian/non-South Asian names), the form of the data meant that it would require complex data calibration in order to create any useful rates using it and this was subsequently not pursued (this is all described in the working paper).

The authors PhD was CASE (Co-operative Awards in Science and Engineering) sponsored by Bradford Metropolitan District Council (MDC) and therefore demographic rates were created for the thirty electoral wards within the Bradford district, with the aim that these would then be used in population projections.

Introduction

This aim of this working paper on developing strategies for deriving small population mortality rates is to provide useful resources, such as methods and data sources, that could potentially be used to create the ethnic-specific small area mortality rates which may then be used in population projections. Where the population of small area will be taken, in line with work by Simpson *et al*. (1997:266), Rees *et al*. (2004:7) and Williamson (2006), to be approximately 5,000-20,000 people, around the size of an electoral ward. There are various examples given in this working paper, where appropriate, using data from the Bradford district, which comprises thirty electoral wards, illustrating the application of various proposed strategy options to estimate the mortality rates. Identification of potential users of the information presented here, for example, people working in Local Authorities (LAs), is considered throughout.

Further, the aims are to provide an overview of the different ways that death/mortality rates are researched or analysed, with a focus on mortality laws and smoothing, or graduation, of mortality rates. This working paper then goes on to describe the collection of data used to create mortality rates for subnational small areas and, finally, tries to address mortality according to ethnic groups. While there is difficulty in estimating mortality according to ethnic group, due to ethnic group not actually being recorded at death registration, it can be crudely approximated by country of birth (COB) for first generation immigrants, or, more recently, albeit only for those people of South Asian ethnic origins, name analysis can be carried out on death certificates. One such name analysis package, called Nam Pehchan, has previously been executed on Bradford electoral registers and on death records (a discussion of the study using the death records for Bradford district is provided later in the paper).

Later sections of the working paper on small area and ethnic group mortality present the examples using data from the Bradford wards, with information on obtaining deaths data after undergoing name analysis from Bradford Health Authority (HA). Overall, it is intended that this working paper will be a useful resource, presenting methods, data and various strategies for creating mortality rates for use in projections.

1.1.1 Terminology for the operators of mortality

At the outset, it should be noted that there may be some confusion over the different aspects of mortality namely m_x the central death rate, q_x the initial rate, life table probabilities, or mortality rate and μ_x the force of mortality. Full definitions of these operators will not be given here; further details can be found in any standard demography textbook such as by Hinde (1998) or Preston *et al*. (2001). Users making population projections will be familiar with m_x the central death rate, so called since 'the exposed-to-risk is an estimate of the number of persons aged x last birthday at the time the events took place. The average age of these persons is x + ½: they are half-way through (in the 'centre of') the year in question' (Hinde 1998: 11).

The *m*-type death rates are calculated from the deaths in any particular year divided by the mid-year population for that year, which is also referred to as the 'crude death rates' in actuarial journals (McCutcheon 1985-1987). In general, the *m*-type death-rates violate the principle of correspondence, since the people in the numerator could have died at the beginning of the year, and hence would not be in the denominator (mid-year population estimate) for the rate calculation. For more details and a fuller explanation of the principle of correspondence and *m*-type or *q*-type rates, as well as their advantages and disadvantages see Hinde (1998).

The q_x rates, or initial rates so called since the 'exposed-to-risk is defined at the start, or initiation, of the year of age under investigation (that is, when the members of the exposed-to-risk celebrate their *x*th birthday)' (Hinde 1998: 11). The relationship between the *m*-type and *q*-type rates was originally proposed by Reed and Merrell in 1939, where they changed *m*-type death rates into *q*-type lift table probabilities. The basic formula is given by:

$$q_x \cong \frac{m_x}{1+0.5m_x}$$ (Smith and Keyfitz 1977; GAD 2004a:58)

$$q_x = \frac{2m_x}{2+m_x}$$ (Hinde 1998:15; GAD 2004b)

This applies to the deaths for people aged one year and over; deaths are treated differently when occurring at less than one year of age, due to the known distribution of the deaths at the beginning of life. For details on the formulas used to account for deaths under one year of age, see the English Life Tables (ELT) no.15 (ONS 1997), the Government Actuary Department ((GAD) 2004b) and the General Register Office for Scotland (GRO(S)) (GRO 2004).

Finally, while *m*-type death rates are investigated by agencies making population projections (GAD 2004a), it is the *q*-type mortality rates that are used in the projections.

1.2 Methodologies for modelling mortality

This section reviews the different ways of modelling death/mortality rates (which could subsequently be used in population projections). A general grouping of these:

- mortality laws or mathematical models (curve fitting)
- standard or model life tables
- relational methods

These approaches will be summarised by referring to previous studies, and examples of their practical application emphasised where relevant. While it may be unrealistic to expect producers of population projections in LAs, for example, to administer all these different methodologies, they are included in order that a comprehensive methods review is given.

1.2.1 Graduation of mortality rates, including Mortality Laws

This section will describe mortality laws and more broadly cover the smoothing, or graduation, of mortality rates by mathematical formulae and by splines. Where 'a mortality law is a mathematical expression that describes mortality as a function of age' (Hannerz 2001b:338). The most famous representation of mortality, that is described in most standard demography textbooks, is the Gompertz (1825) 2-parameter function for the force of mortality, given by $\mu(x) = Bc^x$. This was later modified by adding in a constant, yielding 'the constant adapted Gompertz', the 'Makeham' (in 1860): $\mu(x) = A + Bc^x$. These functions are not, however, suitable for mortality across all ages; they are not, for instance, appropriate for mortality at young ages (Pollard 1975). Moreover, μ_x the force of mortality, as it is known, as opposed to the other operators introduced in this introductory section, will rarely be used in the projection process and will not be discussed further. Although, it is sufficient to say here that in actuarial science, all three operators (m_x, q_x and μ_x) are modelled. In a paper on '*Graduation by mathematical formula*' Forfar *et al*. (1988) cover the methods for modelling each of these operators. Although it is now slightly out of date, it includes the method of splines that has, and is, currently used by GAD and GRO(S) (McCutcheon and Eilbeck 1975-1977; McCutcheon 1985-87; OPCS 1987; ONS 1997; GRO 2002, 2004).

Further, when making reference to the mathematical models, models will be described and examples given of their application on mortality data. A full discussion of weights used for the mathematical functions will not be attempted. For more details on the weighting used for mortality rates (and the numbers of knots for splines) see McCutcheon and Eilbeck (1975-1977), McCutcheon (1979-1981), McCutcheon (1985-1987) and Congdon (1993), all of which give specific examples of their use in relation to mortality rates[1]. In addition, more practically, the aim is to review methods used by agencies, such as GAD and GRO(S), not to reproduce their work.

In actuarial terms, the smoothing of mortality rates is referred to as graduation, although the definition of graduation can change:

> In the course of time the term 'graduation' has been both extended and narrowed in its application, according to the particular interests of the writers. (Benjamin and Pollard 1980:239)

where

> Graduation may be regarded as the principles and methods by which a set of observed (or crude) probabilities are adjusted to provide a suitable basis for inferences to be made and for other practical calculations to be made. One of the principle applications of the graduation is the construction of a survival model, usually presented in the form of a life table.
> (Haberman and Renshaw 1996:411)

In addition:

> For the English Life Tables [ELT], the intention of this smoothing (or 'graduation', as it is called by mathematicians) is to replace the crude rates by a series of graduated rates which, while forming a smooth progression over the whole age range covered, still preserve the general shape of the mortality curve. (ONS 1997:1)

Further to this, the different types of processes that are termed as graduation processes vary. The main forms of graduation are:

- graphic method
- graduation by mathematical formula
- graduation by reference to a standard table
- oscillatory and spline functions (Benjamin and Pollard 1980:241)

Caution should be taken when using these graduation techniques, as Benjamin and Pollard (1980:31) warn when commenting on graduation methods. The data must be fully explored, otherwise what they refer to as 'true irregularities' will generally be

[1] Regarding knots for splines see McCutcheon (1979-1981:435) or Papaleonidas (1995:15)

smoothed over. For example, McCutheon and Eilbeck (1975-1977), in investigating the ELTs, found that at age 17 the *m*-type death rate was bigger than that for the surrounding ages. They suggest that this could be due to deaths occurring as a result of road accidents and further comment that it may well prove necessary to produce these rates by a ‘hand polishing’ process applied to the crude data (McCutheon and Eilbeck 1975-1977:287). More recently, the increase in mortality for young adult males, now known as the ‘accident hump’, has been identified by Heligman and Pollard (1980), Congdon (1993, 2004), McNown and Rogers (1989) and Hannerz (2001a, 2001b). This is now accepted as a feature of the male mortality experience. However, it illustrates that, despite being originally exposed over 25 years ago at the country level (from the ELTs), there are still problems with this data, which means that users may have to rely on unsophisticated ways to deal with the ragged death-rates[2]. An option which could be used to capture these unusual irregular features are kernel methods, they were successfully fitted to mortality data by Congdon (1993).

The following sections will describe two principal methodologies that can be used to fit mortality rates over the whole age range. It was mentioned earlier that the Gompertz and the Gompertz-Makeham functions are the most ‘famous’; however, they are not suitable to model the morality rates for all ages.

1.2.1.1 Curve Fitting Mortality rates

Mortality can, due to its known regularities, be modelled using a curve. In general, the larger the number of parameters can make the process less stable, or over-parameterised (Congdon 1993). While there are many facets of mortality, when it comes to curve fitting, most of the actuarial literature has taken the form of fitting the q_x rates; although there are exceptions and adaptations. For example (Rogers 1986), uses m_x death rates fitted by a modified Heligman-Pollard (H-P) curve (Heligman and Pollard 1980) - adapted originally by Brooks *et al.* (1980 cited in Rogers (1986)). Moreover, due to the regularities in mortality, human mortality can be fitted using functions with between 2 and 9 parameters. Tabeau (2001:2-13) gives a comprehensive review of every possible model that can be used to model all aspects of mortality from

[2] The term ‘ragged’ (as described by Congdon (1993:238) or ‘rugged’ by Gilje (1969:118) describes cases where the age-specific rates fluctuate when progressing from one age to the next, as oppose to a smooth progression from the youngest to the oldest.

the old non-polynomials of De Moive (1725), to recent non-polynomials and partial age functions, along with all the polynomial functions, including the H-P curve.

With regards to fitting q_x mortality rates, the most common is using the H-P curve, which is standard in the UN computer package MORTPAK. Heligman and Pollard (1980) introduced the curve to investigate the age-pattern of mortality, their primary curves are:

$$\frac{q(x)}{p(x)} = A^{(x+B)^C} + D\exp[-E(\ln x - \ln F)^2] + GH^x$$

8-parameters

$$q(x) = A^{(x+B)^C} + D\exp[-E(\ln x - \ln F)^2] + \frac{GH^x}{(1+GH^x)}$$

or

9-parameters

$$q(x) = A^{(x+B)^C} + D\exp[-E(\ln x - \ln F)^2] + \frac{GH^x}{(1+KGH^x)}$$

The H-P curve boasts that the parameters are easily interpreted in demographic terms. The parameters, as described and used by Congdon (1991, 1993), when fitting it to London borough data, breaks down into three curves, these are shown in figure.1:

Figure 1 'The age pattern of mortality'

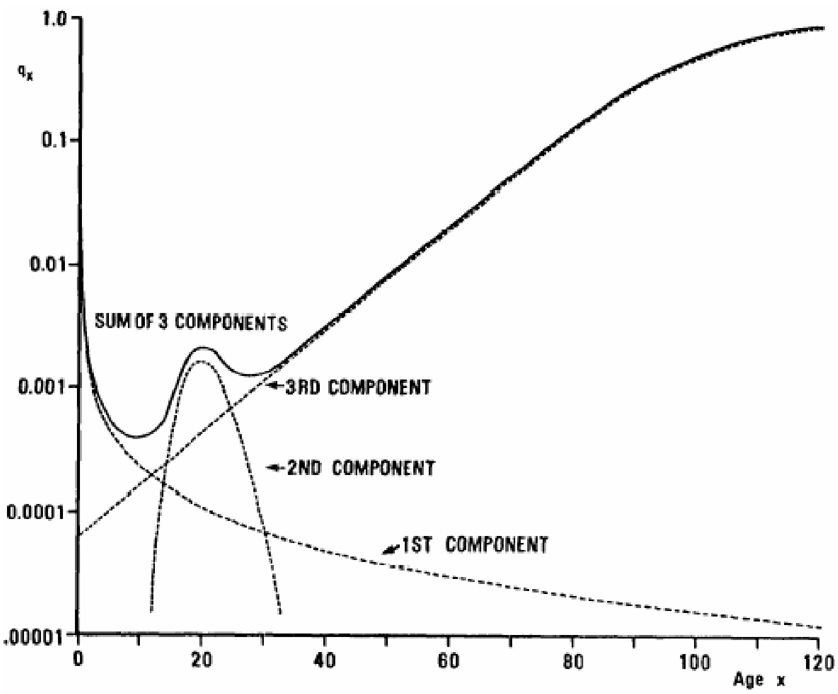


Figure 1. The graduated q_x curve and its three components: Australian national mortality, 1970–72 (males).

Source: 'The age pattern of mortality' by Heligman and Pollard (1980:51)

The first component corresponds to childhood and infant mortality; this is for parameters A and B relating to mortality for ages under 1 and C for childhood mortality (McNown and Rogers 1989). The second curve, given by parameters D, E and F, represents adult mortality, specifically accounting for an accident hump for ages around 20, with 'D related to the severity of the accident hump, E related to its spread and F indicating the location of that hump' (Karlis and Kostaki 2002:852). The third component, for parameters G and H, is the general increasing mortality for older ages using the Gompertz exponential (Heligman and Pollard 1980). For a more detailed description of each parameter and their changes over the 1990-1985 period using US data, see McNown and Rogers (1989).

The original mortality data used to fit the H-P curve was Australian male and female death-rates for the 3 years 1946-48, 1960-62 and 1970-72. Through least squares Gaussin-Newton iteration, the parameters were determined (Heligman and Pollard 1980). The results, similar to Congdon's (1991, 1993), uncovered that the accident hump is not always properly represented. More recent than the original Heligman-Pollard paper, Kostaki (1992) modified the curve to include two parameters E1 and E2 to deal with the 'spread of the accident hump to the left and right of its peak, respectively' (cited in Tabeau (2001:11)). Otherwise, the fit is reasonable and the parameters easy to interpret. Originally, in an attempt to remedy the problem of the accident hump, Heligman and Pollard (1980) modified the original formula. The two resulting modifications provided curves that fit the data better at all ages (the formula for the 9-parameter version is given above). They do however note that the changes to the curve may result in the parameters losing their demographic interpretation.

In an attempt to overcome the problem of over-parameterisation, Congdon (1993) - in investigating mortality data from London - proposed setting some of the 9 parameters to a set sensible value, as outlined by Rogers (1986). Initial results indicate that there is little loss of fit, given by R^2, from that of the fuller 9-parameter model (Congdon 1993). On the other hand, McNown and Rogers (1989:648) comment that 'models are complex and involve a large number of parameters, their flexibility is necessary to capture the full curvature of the mortality profile'.

To give this problem some context and relating this to practical application of mortality curves, Congdon (1993) from his mortality modelling using data from all

London boroughs over three year periods, concluded that the process was unstable. Based on this assumption of over-parameterisation and instability when fitting the data to London boroughs, the process will not be attempted in this working paper for the Bradford wards owing to the small numbers of deaths involved compared to a London borough (even if the Bradford ward-level deaths data had been available in, or were changed to, single-year of age (SYOA)).

Given the general acceptability of the H-P curve for *q*-type rates, for any users wanting to use the curve to model mortality data there are two options. First, either to transform *m*-rates in to life table *q*-type rates; this can readily be done by adapted formulae[3]. Second, it may be easier for users aiming to smooth death rates to create the *m*-type rates, since these can be calculated directly from deaths registrations and MYEs. Rogers (1986:49) uses the modified Heligman-Pollard 8-parameter curve for annual death rates - which was by modified by Brooks *et al*. (1980, cited in Rogers (1986)) - given by the formula:

$$d(x) = d_I(x) + d_A(x) + d_S(x),$$

Where $d_I(x) = Q_0$ for $x = 0$

$$= Q_1^{x^\gamma} \text{ for } x > 0,$$

$$d_A(x) = Q_A \exp\left[-\left(\left(\ln x - \ln X_A\right)/\sigma\right)^2\right] \text{ for } x \geq 0,$$

$$d_S(x) = Q_S\left[e^{x/X_S} \Big/ \left(1 + Q_s e^{x/X_S}\right)\right] \text{for } x \geq 0.$$

Here, as with the original parameterisation, the terms characterise:

- I infant and childhood mortality
- A deaths due to accidents and, among females, maternal mortality
- S a senescent mortality component, which reflects mortality due to ageing (Rogers 1986:50)

A more detailed description of each parameter in turn is provided by Rogers (1986:50). The interpretation is essentiality the same as for the original H-P curve for the *q*-rates. Furthermore, considering the use of parameterised curves, Kostaki and Panousis (2001) acknowledge that there is an advantage in using these types of parametric techniques in terms of interpretability, over that of splines for example. The point is that there may be preference for mathematical models in which the

[3] formula given in the introductory section

parameters can often be related to specific demographic features. This was also discussed when comparing curve fitting option with using splines in relation to age-specific fertility rates (ASFRs) by Congdon (1993:238) and Hoem *et al*. (1981:235).

1.2.1.2 Spline functions for mortality rates

Splines are not a mortality law as such; however, splines have been used in the graduation of mortality rates for some time now. McCutheon and Eilbeck (1975-1977) in the '*experiments in graduation of the English life tables (ELT) no. 13 data*' use splines to create the England mortality rates for the life tables using data from 1970-1972. This was carried on to the ELT number 14 for data from 1980-1982 (McCutcheon 1985-87; OPCS 1987), and to the most recently published ELT number 15 for the ELTs based on mortality from 1990-1992 (ONS 1997). Moreover, GRO(S) in the construction of the 1990-92 Scottish life tables essentially used the same method, with the only difference resulting in the construction of rates for the oldest old (GRO 2002). A definition of a spline is not given in this working paper since they are described in the PhD by the author (Williamson 2006), definitions are also given by Greville (1969:1), Kostaki and Panousis (2001:4) and Papaleonidas (1995:15-16).

The use of splines in construction of mortality rates has proved useful as a way to smooth out the *m*-type central death rates before changing them into initial rates. This is achieved using the formula (provided previously) for use in a life table for making projections and other interesting features - such as calculating the expectation of life at any given age. This is common practice by GAD and GRO(S) through calculating the national-level life tables (GAD 2004a; GRO 2002). With the slight exception that these agencies do not graduate the rates for *q0* and the oldest ages, these are estimated from births and deaths for age zero, and for the oldest ages by extrapolating the *q*-values (ONS 1997).

Despite the purpose of splines being used to smooth ragged rates, caution will be given to anyone attempting to do so for the small areas/groups, for the main reason that the mortality rates could simply be too ragged and that they may be smoothed to unrealistically. In other words, the data may be smoothed into an age-pattern not actually realised in the underlying data. For example, in creating the ELT number 15, GAD comment with reference to ungraduated country-level rates that these '…are not suitable as the basis for a standard life table because they tend to vary erratically from age to age owing to the small numbers of deaths involved' (ONS 1997:1). They are referring to three years worth of deaths, those that are used in the mortality rate calculation, as a small amount of deaths.

1.2.1.3 Conclusions on the graduation of mortality rates

In conclusion, the methods outlined in this section have all been used on large numbers of deaths at the national-level. For this reason they will not be attempted at the subnational Bradford ward-level, considering the quote from the creation of the ELT, using three years worth of deaths, the rates created are ragged. Finally, Preston *et al*. (2001) comment

> Fitting mathematical functions to age patterns of mortality requires accurate data on morality at certain ages. […] Tabular representations of model age patterns were developed to deal with situations of missing, inadequate, or in accurate data. (Preston *et al*. 2001:194)

The next section covers these model life tables, along with relational methods.

1.2.2 Model Life Tables and Relational Methods

A concise description of what a standard, or model, life table is and why it is useful is:

> A model life table is simply a numerical table that gives death rates, probabilities for survival, expectations of life and other related information as a function of age. The model life tables are normally based on empirical observations in large populations believed to have reliable population and mortality data. The purpose of the model life table is to substitute for unknown aspects of mortality in a population with limited or unreliable mortality data. (Hannerz 2001b:338)

The most famous collection of model life tables were produced by Coale and Demeny for the UN in 1966. Their collection was for four different set of types, that is East, West, North and South (Shryock and Siegel 1976; Coale *et al*. 1983; Preston *et al*. 2001). Where these standard tables can be used as a proxy when there are unreliable sources for making estimates of the mortality of the underlying population.

In terms of application, in the context of this working paper for the specific problem of small areas, model life tables could be used that have been created based on a larger areas death data. Using GAD mortality rates for England in projections is one example of using a model life table for projections, since they are widely available and free from the GAD website. Alternatively, given a higher degree of expertise, a user could attempt a relational approach as demonstrated by Congdon (1991, 1993) for London boroughs, using the overall London life table as a standard.

Relational methods in their classic form can be complex and requiring extra data and specialist knowledge. ‘A relational method is a mathematical expression, which relates to mortality in one population to that in others’ (Hannerz 2001b:338). The most well know demographer who developed these is William Brass (for example, his 1971 chapter ‘On the scale of mortality’). Zaba (1979) also used relational methods, as has Ewbank *et al*. (1983), Congdon (1991, 1993) Hannerz (2001b), Murray *et al*. (2003), Bhat (2004) and Woods *et al*. (2005). In addition, Hannerz (2001b) combines relational methods, mortality laws and model life tables to model Swedish data.

Relational methods are useful as they have many benefits in relating mortality behaviour in a population where there is poor mortality data, to that of a population that is known to have sufficient or reliable data, and require less parameters (Preston *et al*. 2001). Despite being useful, relational methods in their true statistical form will not be fully described, since they may be beyond the scope of creating mortality rates by LAs, for example. Nevertheless, as discussed, relational methods have previously been used subnationally by Congdon (1991, 1993) on London boroughs, relating the boroughs to the overall London life table as a standard. More recently, Woods *et al*. (2005) used a relational approach when exploring mortality by Government Office Region (GOR), deprivation quintiles, and both, in England and Wales, where the deprivation measure was the income domain of the 2000 (Index of Multiple Deprivation) IMD and the standard was from the GAD ELTs. Moreover, the latter are at present developing a program that will run in the statistical software Stata for users to download. Options will vary from the simple Brass 2-parameter version to the Ewbank *et al*. (1983) 4-parameter model (Woods, L. personal communication). This means that the relational methods may become more accessible to users that are required to estimate mortality rates when faced with less reliable mortality data.

The intention of including this section on model life tables (and relational methods) is to emphasise that over the past fifty years other countries have been using standard life tables in the absence of reliable data. It is not a new concept to use life tables that are based on similar conditions; and it will be recommended (and demonstrated later in the chapter) that users intending to make population projections for small areas/groups could use national-level data as a standard, combined with local death counts to constrain the mortality level.

1.3 Subnational Mortality

Given that the emphasis is on small areas/groups, in an attempt to include rates below the national-level, the methodology that Office for National Statistics (ONS) uses for subnational population projections (SNPP) is briefly detailed. The method comprises:

> using the past few years' death registrations to provide subnational age-specific rates, given for each year of age to apply to the population (ONS 1999, 2004a, 2004b). The deaths used to create the mortality rates in the 1996-based SNPPs were taken from the years 1993-1996. More recently, deaths from the period 1997-2000 were used for the 2000-based short-term SNPP, and from years 1999-2003 for the 2003-based long-term SNPPs (ONS 2004a, 2006a). It has been projected (in 1996-based SNPPs) that from the pattern of deaths there will be a reduction until 2011, and then an increase to 2021. This reflects or anticipates the post-war baby boom cohort, which will by then be at a high-risk (older) age (ONS, 1999). Overall, SNPP deaths are constrained by the national-level projected deaths (Wood *et al*. 1999:22; ONS 1999, 2004a).
> (adapted from Williamson 2006)

However, these sets of ASMRs for subnational areas 'building bricks' (essentially districts) are not offered for wider use, meaning that users from LAs would still have to use the GAD standard rates for projections in the absence of adequate data to create their own (or use another approach). In previous sections some subnational mortality analysis has been mentioned, Congdon (1991, 1993, 2004) carried this out. This section details easily accessible data sources that could be exploited to create subnational mortality rates, gives potential data problems and provides examples of ways that can be termed relational approaches. These practical examples are applied to the Bradford wards and could readily be reproduced for any other small areas.

1.3.1 Data and potential problems in creating subnational mortality rates

The data most LAs will have readily available in order to create mortality rates will come from the vital statistics (VS) data. These are available from the ONS and are tables of deaths by five years of age for each calendar year. The exception is those of the infant deaths, the 1-4 year olds, and then the highest age group - which is banded. The lowest output area available at present is the ward-level. Thus, at best one could create age-specific death rates by 5-year age-groups using an average of a few years of deaths in order to smooth over fluctuations in the distribution of deaths that occur in a single year. GAD and GRO(S) use an average of three years worth of deaths and then a formula to weight out three years of populations (the denominator) to create the age-specific mortality rates[4]. Further, there are others problems when trying to operationalise the deaths and the population data into rates. The problems of five-year data and of how to handle the central *m*-type death rates are considered in turn.

1.3.1.1 Base population for the ASMRs

For the most part in trying to create rates for small areas, such as wards, it was not possible to have access to accurate population counts (to make MYEs from) for years other than Census years. Previously ONS only produced subnational estimates at the 'building brick' level (ONS 2002), however recently (April 2005) the SAPE team at ONS produced mid-year ward-level population estimates (termed 'experimental statistics') (ONS 2005). The ward-level population estimates are by quinary age-groups for England and Wales for the 2003 ward boundaries. Nevertheless, previously, there were other ward-level population estimates these

> were the 1991 EWCPOP estimates (from the Estimating with Confidence) EwC project) and the 1998 DETR (now ODPM) estimates (created by the Social Disadvantage Research Group at Oxford University). The 1991 EWCPOP estimates are by quinary age-groups and sex, and are adjusted for undercount (CDU 1999). The 1998 ward-level estimates are not by quinary age-groups, they are only broken down by age-groups: under 16, 16-59 and 60+ (ONS 2006b, 2006c). (Williamson 2006)

In addition, Congdon (2004) did manage to gain access to ward-level data for the years 1986-2000, by five-year of age-groups, based on data from the Compendium of Clinical and Health Indicators and GP registrations. Even though it is impressive to have access to year on year ward estimates, Congdon (2004) found however, that not all population estimates were in line with the 2001 Census results.

1.3.1.2 Five-year mortality rates to ASMRs

If there is sufficient interest in expanding the five-year rates into age-specific rates there has been work carried out over the years by Kostaki (1987 and 2000 cited in Kostaki and Panousis (2001)) and by Kostaki and Panousis (2001), where Kostaki and

---

[4] Noting the formula used is not given or discussed, for more details see ELT no.15 (ONS 1997) and GRO(S) website, and as Hinde (1998:39) comments that it is calculated 'using a rather complex procedure'.

Panousis (2001) state there is also a feature to do so in the UN package MORTPAK. To achieve this firstly a life table has to be constructed and only then it is possible to 'expand' the abridged life table to age-specific rates. Kostaki and Panousis (2001) offer reviews of different methodologies that can be used to 'expand the life table' and these include:

- Heilgman and Pollard 8 parameter curve
- Natural spline interpolation
- Complete spline interpolation
- Langrange interpolation
- Non-parametric relational technique
- An Heilgman and Pollard 8 adjusted method

(Kostaki and Panousis 2001)

These are possible ways of handling five-year mortality rates in order to create age-specific rates; however, in practice they may not be a realistic option. Kostaki and Panousis (2001:12) comment that despite the parametric methods giving the better outcomes they 'require advanced software', yet they comment that advanced software problem is also true for spline techniques.

1.3.1.3 The use of m-type death rates in projections

Concerning the use of *m*-type central death rates in population projections, it is usual practice to use *q*-type mortality rates, or probabilities, in population projections. The way in which agencies exploit these rates in projections varies. For example, ONS do not use a straightforward population multiplied by the death rates calculation, it also takes into account some of the migrants for the year (GAD 2004a).

The forecasting package developed by Bradford Council and Andelin Associates called POPGROUP has been used to assess rates created through the different strategies (for both rates created in this working paper and the author's PhD research). The projection calculations/methodology used in the POPGROUP package differs slightly from that used by GAD. With respect to mortality the POPGROUP package does not apply the mortality rates to any of the migration taking place, the full calculations used to create the population projections can be found in Andelin and Simpson (2005:69) or summarised by Williamson (2006).

In trying to decide what set of rates to apply to a population, ONS in the methodology section offers an explanation of how it creates and why it uses the *q*-rates:

> The q_x rates in the projections are the result of two interpolations. The first interpolation takes place between the q_x rates for adjacent calendar years and produces rates on a mid-year to mid-year basis. The second interpolation is between adjacent ages and gives a set of q_x rates that, in life table terms, relate to exact age x+½ on a mid-year basis. These are assumed to be applicable to the mid-year population at age last birthday. (GAD 2004a:58)

In practice this process may be far too data intensive and time consuming to perform in order to create small area/group mortality rates in LA research departments.

This is in contrast to recommendations in the *POPGROUP users guide worked examples*, where it is suggested that users could create their own ASMRs for use in a population projection from the VS data at the district-level (Simpson 2001). This would mean using the *m*-type death rates in the projection rather than the *q*-type mortality rates if these were calculated. However, Simpson (2001:36) does state that the actual numbers of deaths occurring at this district-level will probably be small, which could be hinting at the reliability of doing this.

1.3.2 Examples of two relational-type options for small group mortality rates

This section details relational-type approaches for small area mortality rates that can be achieved using the projection package POPGROUP. The examples are for the Bradford wards. Before the examples are described, the average number of deaths for the Bradford district for ages under 75 are given in the table 42. These are for the four Bradford Primary Care Trusts (PCTs), information on the electoral wards that are included in the PCTs are in appendix 1[5].

Table 1 Average of 1995-2000 deaths in the Bradford disrict for all causes (for ages under 75)

PCT	Average Annual deaths			SMR		
	Male	Female	Both	Male	Female	Both
Airedale	262	194	456	93	98	95
Bradford City	333	200	533	139	140	139
North Bradford	255	181	436	113	111	112
Bradford South & West	379	252	631	118	113	116
Total	1229	827	2055	115	114	115

Source: Bradford Health Authority (2001)

[5] The thirty Bradford wards are the 1991 based Bradford wards and not the new electoral wards following the recent boundary changes.

From table 1 it can be acknowledged that Bradford district overall has higher mortality than at the national-level, given by the Standard Mortality Ratio (SMR)[6] of over 100. However, this table was included for another reason; this is the first time actual number of deaths has been presented. This is to illustrate exactly how small the average death counts for Bradford as a district and the four PCT are, where obviously going below this to the ward-level means even smaller numbers of death, especially when considered by SYOA (in addition, if also considered by ethnic group these counts would be even smaller). The counts of deaths for both sexes range from 47 in the ward Ilkley (SMR of 72), to 85 deaths (SMR of 147) in the University ward[7]. While Bradford HA does produce SMRs by ward, a strategy of how to indirectly estimate ward-specific SMRs using the POPGROUP package are detailed next.

1.3.2.1 Using VS data and SMRs in the Bradford ward-level projections

The basis for this strategy is to use the national-level sets of *q*-rates from a standard life table and to adjust these for the area/group of interest. This can be carried out in POPGROUP. When the latest GAD national-level *q*-rates are made available, users can simply either add in mortality counts for any ‘group’ of interest or add in a pre-calculated SMR for the areas/groups. The effect will change the overall level of mortality, however, the ‘group’ still retains the age-structure of the original GAD ASMRs specified in the POPGROUP package, hence a relational-type approach. In the case of the thirty Bradford wards, users of POPGROUP could enter pre-calculated SMRs which will be more detailed than using only the GAD England mortality rates. As shown in table 1, below the district-level, the corresponding SMRs for the four Bradford PCTs could be entered for each ward.

Another more detailed example would be to use death counts from the VS tables and enter these for each ward; the process of how the natural increase ward-level Bradford projection (no migration) was constructed was detailed in Chapter 4 (section 4.4.3.1 including data sources and assumptions) of the PhD research by the author

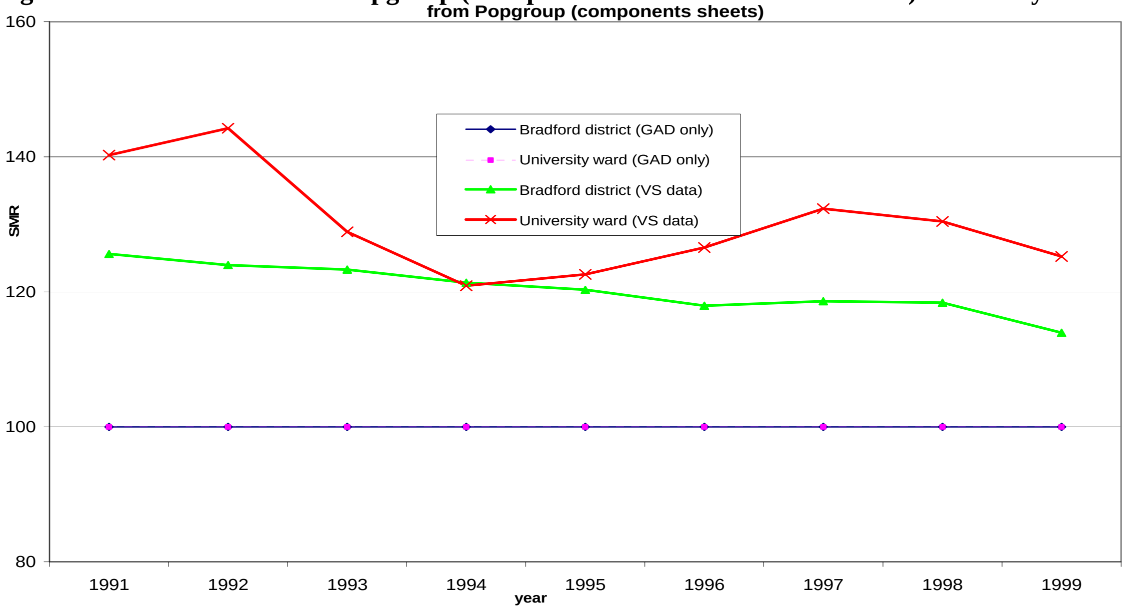
---

[6] ‘The SMR is the ratio of the actual number of deaths in population *A* [for example, the Bradford district or wards] to the number of deaths that would be expected to in population *A* if it experienced the age-specific death rates of the standard population [for example, GAD England rates]’. (Hinde 1998:23)

[7] Based on average of deaths for years 1995-2000 for ages under 75

(Williamson 2006). Essentially, this projection is using GAD rates as the assumptions for fertility and mortality, and then entering VS data to make the projection more localised. For the mortality component, this would be to enter the death data as mid-year deaths, which will change (raise or lower) the level of mortality in each of the wards while retaining the age-specific pattern of it as given by the 'standard' schedule of (GAD England) rates. Therefore the SMR for each ward will be an adjusted based on the localised death counts from the VS data (the adjusted SMRs are provided in the 'components' summary POPGROUP sheet). Figure 2 displays how this works; for simplicity only the Bradford district and one ward, University are shown in the chart.

Figure 2 Chart of SMRs from Popgroup (over period when VS data was entered) males only



Source: The data displayed in the chart was from a previous population projection for Bradford using a base population that was not scaled to the most recent revision of the ONS 1991 MYEs. Nevertheless, all the methods to create it, outlined in chapter 4 of the author's PhD (Williamson 2006), are the same and it will still serve as valid test data for the strategies demonstrated here.

From figure 2 it is shown that for the projection using the only the GAD England rates the SMRs are 100 (this means it is the same as the 'standard'). However, when VS data are included in the projection this changes the mortality level. Similar to the SMRs displayed in table 1, Bradford's overall mortality level is above England's.

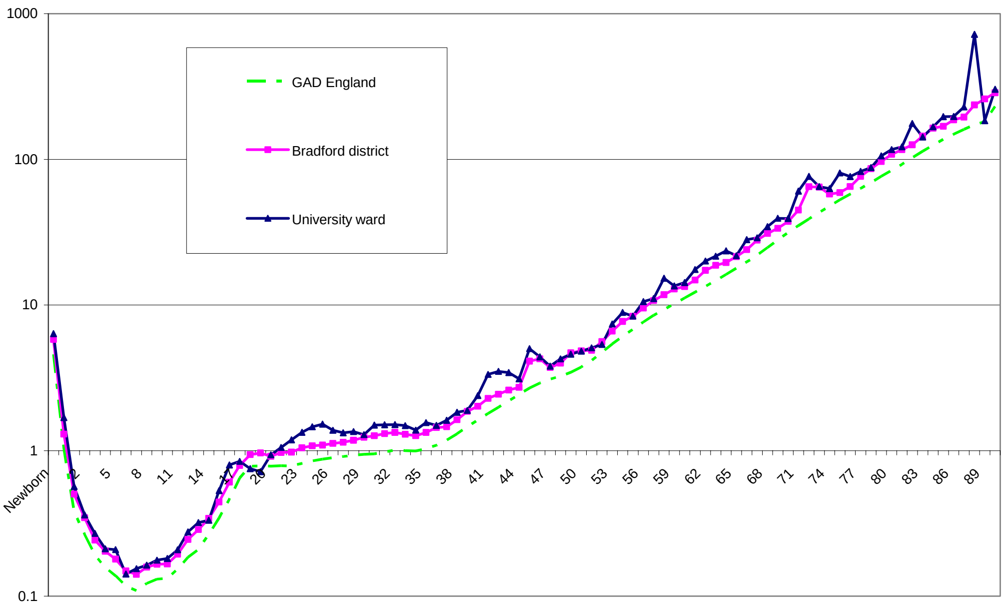
Furthermore, the implied SMR for the years in which the VS data has been entered for use could be used to create a SMR for use in projecting into the future. In Williamson (2006: section 4.4.3.2) there was an example of this using the Bradford ward-level data and it was suggested, based on an investigation of various years worth of SMRs, to use an average of the last three years SMRs for which VS data was entered to produce a future ward-specific mortality differential for the projection.

1.3.2.2 Using POPGROUP to create ASMRs for the Bradford ward-level projections

Another relational-type strategy in order to obtain subgroups' actual ASMRs, as oppose to SMRs, would be to exploit the 'dump-file' created in POPGROUP. The dump-file holds the information for all years and it is what is re-packaged to create all projections in the output files. It is the behind the scene part of the POPGROUP projection process. For example, if VS data were entered for wards, ASMRs for the Bradford district and all the wards could be indirectly created. This would be to use the GAD England ASMRs as the 'standard' to which their mortality was related to.

To give an example, ASMRs for males in 1991 for both the Bradford district, and for the ward University, produced by taking the 'workings' from the projection process to create the ASMRs. This was achieved by using the 1991 population as the base population (and the projected male births as the newborns). Then the deaths were then averaged over the three years 1991-1993 to create the 1991 ASMRs (three years were used to reduce the degree of raggedness of the rates), these are shown in figure 3:

Figure 3 Males only ASMRs for 1991 Bradford district and Univeristy from POPGROUP (log scale)



Source: Same as figure 2, this is taken from an older population projection for Bradford using a base population that was not scaled to the most recent revision of the ONS 1991 MYEs, where the purpose is to demonstrated the strategy here.

From figure 3 (on the log scale) it can be seen, as was pointed out earlier, that for males, in general, the Bradford district has higher levels of mortality than England. Moreover, it may be observed that the implied age-specific deaths (originally from using the GAD rates with level defined by the ward births), even when averaged across three years to create the ASMRs, that these are much more ragged than the GAD England ASMRs. This is most evident for the ASMRs for the ward University; this is the small numbers problem and highlights the potential for smoothing rates, the problem is also stressed/shown throughout the PhD by the author (Williamson 2006).

1.3.2.3 Conclusions on relational-type approaches

The strategies described in this subsection, while they are not true relational approaches to mortality data, are options that could be easily implemented in the projection of small area/group mortality rates, by LAs for example, where access to deaths data is limited. Moreover, since these strategies are based around POPGROUP any POPGROUP user could create them. In addition, in obtaining ASMRs from using the 'dump file' an H-P curve could be fitted to them to smooth the ragged rates.

1.4 Mortality data sources according to ethnic groups (including COB)

The literature that has been covered so far has been a methods review of modelling mortality rates at the national-level. Then subnationally, after some general issues were addressed, two strategies of how using a relational-type approach for estimating small area mortality were presented. In addition, it was reported there has been some research on mortality for wards and boroughs in London (Congdon 1991, 1993, 2004).

The focus of the working paper is on small areas and ethnic groups; it is at this point that ethnic group mortality is considered. Before engaging in the literature, it must be acknowledged that, similar to creating ethnic-specific fertility rates (as discussed in detail in Williamson (2006)), ethnic group is not recorded in the registration data, which makes the process of estimating mortality rates by ethnic group difficult. Therefore, this section also considers previous mortality research using country of birth (COB) and ethnicity. Traditionally, country of birth has been asked at death registration, and this has typically been used as a proxy for ethnic group mortality.

Previous research on mortality by migrants and COB has been carried out by Marmot *et al*. (1984), Balajaran *et al*. (1984), Balajaran and Bulusu (1990), Harding and Maxwell (1997), Wild and McKeigue (1997), Harding and Balajaran (2001, 2002) and Storkey (2002), Harding (2003, 2004). Of which Harding and Balajaran (2002) in their chapter '*Mortality data on migrant groups living in England and Wales: issues of adequacy and interpretation of death rates*', provide a review of all the pre-2000 work mentioned above using COB, detailing when the research took place and what data was used in the study (with the exception of Balajaran and Bulusu (1990), which is reviewed by Storkey, (2002)).

This section gives details of recent data sources that could be exploited to estimate mortality by COB and ethnicity and provides information from previous research on mortality, with the aim of using the information to set mortality assumptions for use in ethnic group population projections.

1.4.1 Data sources for estimating mortality at the national-level

This section considers mortality at the national-level. The process of estimating mortality by ethnic group and COB even at the national-level proves difficult. Unlike fertility, for example, other data sources cannot give any information or intentions on mortality. Using data from surveys, such as the Labour Force Survey and General Household Survey, these obviously do not account for deaths. Previously, in trying to determine mortality rates by ethnic origins due to the lack of data, in the 1970s for example, resulted in mortality rates for the UK applied to the people of NCWP origins (an old rather crude categorisation of ethnicity based on people of the New Commonwealth and Pakistan), as they were assumed to be the same (OPCS 1979:22).

Recently things have improved. When attempting to establish mortality rates according to COB and ethnic group cross-sectional data from the Census and death registration data and could be used to build up an age distribution, by COB (as a proxy for ethnicity). This was first carried out by Marmot *et al*. ((1984), cited in Harding and Balarajan, 2002:115) in using 1971 Census data. The other main data source which could be used is the Longitudinal Survey (LS), which can be linked to death registration data, although numbers for minority groups is relatively small. To aid understanding of these sources and potential weaknesses in operationalising them, a table created by Harding and Balarajan (2002) is reproduced:

Table 2 Data Sources which can be used to calculate mortality rates in England and Wales

Sources	**Country of birth (COB)**	**Ethnic Origin**	C.O.B. grand-parents or parents	Ethnic Origin of grand-parents or parents	**Cover age**	**Quality of data**
Unlinked						
Death registrations	Yes	No	No	No	100%	Country of birth reported by relative of deceased maybe different from that reported in the Census
Census	Yes	1991/ 2001	Question asked in 1971 Census	No ^	100%	1. Under-enumeration of some groups in specific geographical areas 2. May contain migrations who return to home countries, leading to inflation of denominators.
Linked						
ONS Longitudinal Survey	Yes	1971 (South Asians)~ 1991/ 2001	Data in non-members' files, could be linked to memebers' data	Data in non-members' files, could be linked to memebers' data	1%	1. Loss to follow-up at older ages can be up to 30% for some groups 2.Small sample size

~ Question not asked in 1981and 1991 Censuses but if the person was usually resident in the household it could be derived using relationship to head of household

^ Question not asked but if the person was usually resident in the household, ethnic origin could be derived using relationship to head of household. ~ Special exercise to identify South Asian names

Source: Harding and Balarajan (2002:116,table1)

In attempting to appreciate ethnic group mortality by using COB as a proxy for the death registration information, as with most of the data sources, can often be flawed. Harding and Balarajan (illustrated in the table 2), uncover another problem when filling in the death certificate. They suggest that there may be wrong mortality calculations for people from Pakistan. Pakistan used to be part of India, and especially if estimated wrongly this will affect the mortality rates. Moreover, Harding and Balarajan (2002) identify a reliability problem in using these data sources, due to the few deaths in the younger groups, even when combined over a few years.

In addition, if subsequent analysis was to be carried out on death registration data, there could be further problems. This is due to family members, when filling in their occupation may fill them in subjectively, making the deceased more important or of better status than they really were (Townsend and Davidson (1982) and Nazroo (1998) cited in Aspinall (2001:852)). This is of course true for a deaths registered of any ethnic group, however if conducting research into social class and ethnic group/COB mortality this is another important consideration. In considering infant mortality, for example, (based on deaths where the father was on the registration form, using his occupation), ONS report variations according to social class. ONS show that the infant mortality rate for social class I is 4.5 deaths per 1,000, while this rises to 8.8 for social class V (for England and Wales 1993-1997 (Cooper, 2001:189)).

Refocusing on creating ethnic group mortality rates, Harding and Balarajan (2002) outline six main areas that are problematic when trying to calculate mortality rates for minority groups, their headings are given below:

- Misclassification at Census and death registration
- Heterogeneity within country of birth grouping
- Children born in Indian subcontinent whose parents were British
- Return migration of older migrants
- Generational differences in mortality
- Changes in Mortality rates with increased length of residence in the UK (Harding and Balarajan 2002:119-124)

Storkey (2002:108) likewise discussed that countries can have people born to different ethnicities within them, which causes difficulties if trying to estimate mortality rates, and provides an example from of people of ethnic group White who were born in India. Moreover, previously the Office for Population Censuses and Surveys (OPCS, now ONS) in creating the estimates made allowances for people who were White in the deaths by COB accounting for NCWP origins (Shaw 1988:29).

The issues listed above are becoming all too evident. Mortality rates are being produced, with questions over how to account for or to deal with the problems that impact upon the rates. Most of these difficulties will hopefully be remedied in the future by the availability of more reliable data to challenge these problems. Moreover, these problems listed are not restricted to creating mortality rates, with the exception of the ‘return migration of older migrants’, all the problems are also applicable when attempting to create ethnic-specific fertility rates.

Notwithstanding, Harding and Balajaran’s work considers COB and also ethnic group data from the LS, the suitability of the findings based on small sample sizes, and they extend this to consider people from the first and second generation for some of the main ethnic groups by certain age-groups, where numbers are small (Harding and Balajaran 2002). Their main measure of mortality being the SMR, although broad banded age-specific death rates by country of birth are presented for 1988-92 using the LS for ages 20-70+.

Since the emphasis of the working paper is on ethnic group and not COB, the previous research on COB will not be covered in detail. In addition, given that time has passed, many of the people in the ethnic group categories are now not readily identifiable by using the country of birth as a proxy. Table 3 adapted from Harding and Balajaran (2002) using the LS that shows by ethnic group ratio of those first and second generation ethnic groups.

Table 3 Number of persons amongst first and second generations ethnic groups in the 1991 Census

Number of persons amongst first and second generations ethnic groups aged under 65 years in the 1991 Census and living in England and Wales.		
Ethnic origin at the 1991 Census	Sample size	ratio of second to first generation
First generation South Asians (total)	12,391	0.59
Second generation South Asians (total)	7,261	
First generation Indian	6,968	0.55
Second generation Indian	3,801	
First generation Pakistani	3,557	0.80
Second generation Pakistani	2,847	
First generation Bangladeshi	1,866	0.33
Second generation Bangladeshi	613	
First generation Black Caribbean	2,306	1.18
Second generation Black Caribbean	2,728	
First generation Black African	1,562	0.51
Second generation Black African	794	
First generation Chinese	1,213	0.36
Second generation Chinese	438	

Source: Adapted table no 5 from Harding and Balajaran (2002:119) from ONS LS, 1991 cohort followed up to 1991-7 [adapted into include a ratio]

Admittedly, this ratio only can represent these generational differences based on the sample of people included in the LS. The LS is intended to be a 1.1% sample of the whole population (ONS, 2004c) and could be an indication that for some ethnic groups, for example, for general South Asian ethnicity, there is over half of the population who were classified as South Asians are second generation, which would render mortality rates by COB wrong when applied to all South Asians.

For the projection process, it is may also be useful to consider mortality at the youngest ages. For despite the vast improvements in infant mortality in the last fifty years there are still vast differences between different countries and within countries (and subnationally, as was shown for Bradford in figure 3 previously). Storkey (2002), Harding and Balajaran (2002) and Cooper (2001) have explored literature on infant mortality by COB for the mother. Their findings are condensed in table 4:

Table 4 Mortality rates for under 1 year olds - various sources based on mothers COB

<table>
<tr><td colspan="4">(1)</td><td colspan="2">(2)</td><td>(3)</td></tr>
<tr><td>Source: Presented in table 5.4 Storkey (2002:112)</td><td colspan="3">Perinatal, neonatal and postneonatal mortality (deaths per 1000 births) by mother’s country of birth, England and Wales, 1989-91</td><td>Source: Presented in table 11 Harding and Balajaran (2002)</td><td>Infant mort. rates, 1996, by mother’s country of birth, England and Wales</td><td>Infant mortality rates, 1993-1997 For England and Wales (not UK) produced by ONS (Cooper 2001)</td></tr>
<tr><td>Country of Birth</td><td>Perinatal mortality</td><td>Neonatal mortality</td><td>Postneonatal mortality</td><td>Country of Birth of mothers</td><td>Infant mort. rate per 1000 births</td><td>Infant mortality rate per 1000 births</td></tr>
<tr><td>UK</td><td>7.9</td><td>4.4</td><td>3.2</td><td>UK</td><td>5.8</td><td>5.9 (England & Wales)</td></tr>
<tr><td>India</td><td>9.9</td><td>5.9</td><td>2.7</td><td>India</td><td>5.4</td><td>5.9</td></tr>
<tr><td>Pakistan</td><td>14.2</td><td>7.4</td><td>5.5</td><td>Pakistan</td><td>10.1</td><td>11.7</td></tr>
<tr><td>Bangladesh</td><td>10.9</td><td>3.8</td><td>2.3</td><td>Bangladesh</td><td>6.3</td><td>6.7</td></tr>
<tr><td>East Africa</td><td>10.0</td><td>5.1</td><td>2.3</td><td>East Africa</td><td>6.1</td><td rowspan="3">7.1 (all non-England & Wales born mothers)</td></tr>
<tr><td>Caribbean</td><td>13.3</td><td>7.5</td><td>4.3</td><td rowspan="2">Caribbean Commonwealth</td><td rowspan="2">8.4</td></tr>
<tr><td>Rest of Africa</td><td>13.2</td><td>7.6</td><td>3.5</td></tr>
</table>

(1)Original source Balarajan and Raleigh (1995) from Bardsley and Lowdell (1999) (2)Original source Davey Smith *et al* (2000) (3)Cooper (2001:192)

These infant mortality rates could be considered for use in a population projection, since they are at least still capturing first generation general behaviour for the non-UK born populations (since they are by the mother’s COB). Although, admittedly these rates may be different from the UK born ethnic group populations that the COB rates are being used to approximate, which again may render these rates incorrect.

Considering ethnic group infant mortality directly, Simpson (2002:58) reports that research by Bradford MDC and ONS found that for ethnic groups Pakistani, Bangladeshi and Black, the infant mortality rate is 1.6 higher than for other groups.

1.4.2 Data sources for estimating mortality at the subnational-level

Storkey (2002) in her unpublished thesis discusses both the mortality by COB and the possibility of mortality rates by ethnic group either at the borough-level or for London as a whole. Specifically in relation to projections, Storkey (2002) reviews and proposes many different methodologies for mortality by COB for London that could potentially be used in projections. One method was carried out by the London Research Council (LRC 1995) for the period 1998-1990 and another on which she was involved in by Bardsley *et al*. (2000) investigating mortality for the period 1996-1998. The first method considers mortality for ages 46-64 only for nine different COBs. Storkey has concerns over the results however, based on the base populations that were used to calculate the rates, age structure and sample sizes (Storkey 2002:114-115). The second method involved four differing methods to create a base, or ‘exposed to risk’ population in order to estimate the mortality rates by COB based on average deaths from death records (Storkey 2002:115-119).

Since the wider PhD research by the author focuses on using data from the Bradford district, an investigation of country of birth and ethnicity was undertaken, displayed in table 5 below. This was an attempt, similar to that of table 3 using the LS, to gauge an understanding of the UK-born generations for non-White ethnic groups; to determine if COB rates will be useful in population projections. Unfortunately, since the data was taken from the standard 2001 Census tables there could not be a true first or second generation breakdown. The distinction used was born in UK or not, and to account for generation, only taking people aged 0-19.

Table 5 Country of birth UK, for region Yorkshire and the Humber (2001 Census table C0004)

Ages	All Ethnic Groups	All non-white ethnic groups	Black or Black British: Total	Asian or Asian British: Indian	Asian or Asian British: Pakistani	Asian or Asian British: Bangladeshi	Chinese, Other, Mixed and Other Asian
All ages	94.75%	50.54%	48.64%	53.64%	57.39%	48.84%	61.10%
ages 0-19	97.41%	84.82%	73.03%	93.49%	90.68%	82.99%	84.84%

Source: calculations from the 2001 Census table C0004 AGE AND COUNTRY OF BIRTH BY ETHNICITY

Table 5 presents for the Government Office Region (GOR) of Yorkshire and Humber (which includes the Bradford district) the percentage overall of people born in the UK and also for ages 0-19. This is for all people and for some ethnic groups. Again, it can be appreciated that especially for Asian ethnic groups that to use mortality rates (and indeed other demographic rates) based on COB would not be capturing the full ethnic differences in the population, due to second and third generations.

It can be seen that when infant mortality is considered, for example, subnationally, even at the GORs, that there are differences, presented in table 6 below:

Table 6 Infant mortality rates (per 1000 live births 1993-1997) by mothers country of birth

Mother’s country of Birth	England or Wales	Not born in England or Wales	Scotland	Other European Union	Bangladesh	Indian	Pakistan
England and Wales	5.9	7.1	5.8	5.5	6.7	5.9	11.7
Yorkshire and Humber	*6.6	*9.3	6.3	-	-	-	13.6

- the rates for areas with less than 20 infant deaths for the combined 1993-1997 period have not been calculated because of the small numbers involved
Source: ONS publication (Cooper 2001:192)

Table 6 shows that overall, similar to what was shown at the national-level, mothers who were not born in England and Wales have higher rates of infant mortality.

However, when broken down this is not the case for all countries; for example, for the mothers COB of Scotland, Other European Union and Indian the rate is not higher. This suggests caution in using a ‘fits all’ non-England COB mortality rate. Table 6 also stresses the importance of using subnational mortality differentials, and the difficulty in creating these due to small numbers. For example, the GOR, Yorkshire and Humber, has in all cases higher infant mortality than for England and Wales (where there is sufficient data). Further to this, when considering mortality differentials at the GOR, there are also marked differences by social class. For example, Cooper (2001:189) notes that the largest differences between social classes V and I for infant mortality occur in the GOR of Yorkshire and Humber. This is from under four deaths per 1,000 for social class I to over nine for social class V.

1.4.3 The Bradford mortality data

The previous research presented was either for England and Wales, or by GOR, which may not be entirely relevant to Bradford wards. Given the socio-economics of the country as a whole, or even at GOR, these characteristics will be different to people resident in the Bradford district and the wards within it (this is highlighted in the author’s PhD (Williamson 2006). Overall, this section tries to assess potential uses of mortality data related to Bradford with a focus on mortality according to ethnic group.

1.4.3.1 Deaths and mortality by South Asian and Non-South Asian

This section initially considers mortality in the Bradford district by COB, and then presents research using COB and name analysis. The aim here is to present sources of previous mortality research in Bradford, with the potential for use in projections.

Previously research has been undertaken by Bradford MDC considering COB and ethnicity broadly by South Asian and non-South Asian origins. For example, Storkey (2002) presented a table produced by Bradford HA for Bradford MDC (1995), which gives SMRs by COB this is shown below in table 7.

Table 7 SMR by Country of birth for Bradford from 1998-1992

Country of Birth	No. of deaths (all five years)	**SMR**	SMR 95% Lower Confidence Limit	SMR 95% Upper Confidence Limit
Asian	521	100.1	91.7	109.1
Non-Asian	10,326	112.5	110.4	114.7
India	171	135.9	116.3	157.9
Pakistan	315	96.5	86.1	107.7
Bangladesh	35	129.4	90.1	179.9
All groups	10,847	111.9	109.8	114.0

Source: Table 5.5 presented by Storkey (2002:113)

From table 7 it is shown that overall taking account of mortality by country of birth, Asians have a SMRs (of 100.1) which is closer to that of national-level, than for Bradford district, which has a higher SMR (of 111.9). Although, considering Asian countries of birth as one large group masks the high SMRs of people born in India and Bangladesh (even though there are a very small number of deaths for Bangladesh, as country of birth, for the five year period). Again, as with the research presented in the previous section, since these are by COB then the SMR will not be accounting for all people of that ethnic group. Notwithstanding, for use in population projections, the SMR could be used if it was fully justified, or perhaps using a mid-way between the SMR for the ward, or district, and the SMRs from the table could be used to provide an ethnic differential (albeit based on COB).

Around the same time, a small sample study was carried out by Bradford MDC which was based on COB and also incorporating name origin of the deceased. In general, name analysis should remedy the problem of whether any ethnic group has a mixture of UK and non-UK born population. These SMRs based on the name analysis are shown in table 8 (only for the ages 35-74). Even though the name analysis is only identifying names of South Asian origins, these SMRs may be more reliable than using the COB SMRs alone in a projection, since adding in the name analysis is giving mortality by ethnic group.

Table 8 SMRs from small sample study based on COB and name analysis of deceased (1988-92 ages 35-74)

	Males	Females
Indian	147	116
Pakistani	102	84
Bangladeshi	152	109
All others	115	109

Source: SMRs from Table 2 presented by Simpson (2002:58)

Table 8 shows overall male SMRs are higher than for females. However, it is not possible to generalise for South Asian and non-South Asian mortality.

More recently in the Bradford HA produced SMRs based on name analysis from the Bradford MDC name identification package Nam Pehchan, these are given in table 9 below. Testing using the Nam Pehchan package has proved accuracy of around 90% (Cummins *et al*. 1999; Harding *et al*. 1999; Horner 2004). Table 9 shows the results from using Nam Pehchan from the publication '*2001 – A public health odyssey the journey towards better health in Bradford*'. Noting that in the publication there are charts of the SMRs with confidence intervals (although these are not explicitly given).

Table 9 Bradford Deaths for all causes for ages 35-74 (average over the years 1995-2000)

Ethnicity	Average Annual deaths			SMR		
	Male	Female	Both	Male	Female	Both
South Asian	92	46	138	124	131	126
Non-South Asian	1029	717	1746	115	112	113
Total	1121	763	1883	115	113	114

Source: Bradford Health Authority (2001)

From the table it can be noted that the SMRs for the South Asians are based on a small number of average deaths per year. Given the size of the South Asian population in Bradford and its youthful age structure - hence the small amount of deaths - the impact of the higher SMRs on the Bradford district SMR is small. These SMRs in table 9, even though for only two ethnic groups, may be better to use in a population projection, than SMRs by COB, given that they are directly accounting for ethnicity. It is acknowledged that there could be problems in the accuracy of Nam Pehchan or if South Asian names have been anglicised over time, however, to use name analysis is a step towards uncovering ethnic mortality behaviour that cannot presently be obtained from any other data source.

1.4.3.2 Preliminary attempt to create age-specific South Asian/Non-South Asian mortality rates

Given that Bradford MDC has the ability to use the name identification package Nam Pehchan and has already run such an analysis on recent deaths data, the research proceeded by investigating mortality according to the two broad ethnic groups South Asian and non-South Asian at the Bradford ward-level. Using methods already outlined in this chapter, the intention was to perform some kind of smoothing/graduation of the ethnic-specific, age-specific deaths data by ward in order to create rates, which could then be incorporated into population projections.

The necessary data was requested from Bradford HA/Primary Care Information Team. This resulted in a lengthy process involving meeting the holders of the Bradford deaths data to gain access the death data after the name analysis had been already completed. There were problems in gaining access to the data even though the PhD research by the author was at the time CASE sponsored by Bradford MDC. For example, the data request had to pass an ethics board, which meant that, despite having requested the data in July 2003, it was not until January 2004 that the data was provided. When the data arrived, the deaths were in 5-year age groups according to the two broad ethnic groups. These categories had been totalled up for a five-year period (1995-1999). In addition, many relevant data fields were suppressed due to the counts being under five.

Originally, the intention was to create ASMRs for each ward according to the two ethnic groups, South Asian and Non-South Asian. This was not possible, not because the data was in 5-year age groups, but owing to the presence of empty cells where the number had been suppressed (because there were fewer than five deaths over the period). Finally, it was decided that this part of the research was deemed unfeasible, given the effort needed to get the data into a useable form and due to the very small number of events (further reasons are given below). In addition, the long waiting time for the data in relation to the time constraints of the PhD research timetable made it unrealistic to fully exploit this deaths data.

The suppression of data was anticipated. To overcome this in the data request, many different sets/combinations of the data were asked for. The intention was then to incorporate VS data and use the method of Iterative Proportional Fitting (IPF) to uncover some of the suppressed counts, then apportion out the 5-year rates (for 1995-1999) by SYOA to 1-year rates. In the end, however, it was apparent that at the ward-level the counts were too small to be modelled by SYOA. The next option was to consider mortality within the four PCTs, since some of the data obtained was at the PCT level. It soon became obvious that the difficulties of pursuing this approach would outweigh the benefits. First, the correct number of deaths at the PCT-level would be needed (using IPF) for the numerator of the rate. Second, a reliable population base for the 1995-1999 period for to the two ethnic groups would have to be determined. This denominator would have to be approximated using the 1991 and 2001 Census data. All of which was too time consuming for the PhD research.

1.5 Chapter summary and concluding remarks on strategies for creating small area and ethnic group mortality rates

The aim of this working paper was to provide an overview of different strategies and sources of data that can be used to create mortality rates for use in population projections. Initially, a review of methods from the literature on modelling ASMRs was provided, including curve fitting and splines, which are, for the most, applied at the national-level. Then age-specific mortality considerations were addressed at the subnational-level. This section focused more on problem solving and data sources than on providing evidence from previous subnational research on mortality research. It included two possible strategies using Bradford ward-level data, using a relational-type approach to deal with mortality for projections. This was to demonstrate the use of different strategies to estimate mortality behaviour for small areas, when faced with data availability problems. The final section covered possible data sources and previous research that could be used for projection purposes when attempting to account for ethnic group mortality, which also included a review of previous research according to COB. It was also stated that while ethnic-specific mortality, from performing name analysis to identify South Asian and non-South Asian names, had been considered, once the data was obtained many cells had been suppressed due to the small numbers of deaths in the five-year period.

Despite the fact that the goal of the author's PhD research was to estimate demographic rates for ethnic groups within small areas (the Bradford wards), the decision was taken not to create ethnic-specific ASMRs - even where suitable data was made available (see also Storkey (2002) who made a similar decision). While the research that could be carried out using the ethnic-specific mortality data (based on name analysis by South Asian/non-South Asian) would be both interesting and useful, due to the time restrictions of the PhD this was not feasible. Nevertheless, given that the data has already been obtained, creating the ethnic-specific ASMRs for either Bradford district, the four PCTs in the Bradford districts or even the 'Bradford fertility areas' created by Williamson (2006) may have potential for future research.

Furthermore, Storkey (2002) considers creating SMRs for each ethnic group in London, then scaling each of the borough life tables. However, after much deliberation about many different approaches to deal with COB and ethnic group mortality behaviour, she decided against this because of the socio-economic differences between the boroughs. This meant that using an ethnic group only SMR for London would not be the most suitable approach to mortality at the borough-level. In her section on mortality rates for use in the London projections, she settles on using the borough-level rates. This decision was based on the small numbers and:

> Given the difficulties with the mortality data and the fact that differences between ethnic groups seemed small compared with differences in mortality by borough across London, the decision was taken to use borough specific survivorship rates by age and gender but not by ethnicity in the projections model (Storkey 2002:126).

This point should be strongly considered when attempting projections according to ethnic groups as Storkey's findings suggest that area differences may be more important than ethnic group differences.

Storkey's conclusion conflicts with research findings presented in the author's PhD, where ethnic-specific ASFRs were found to give more accurate projected births than ward-level ASFRs for ethnic groups at the ward-level (rather than at the ward-level only) (Williamson 2006). Despite this, Storkey's point that differences in mortality between groups may be small should be recognised. Since the differences in ethnic group mortality rates may be small, these differences will have only a minor impact on the number of deaths. Given that many of the minority ethnic groups have younger age structures, there will be fewer people at ages where mortality rates are high.

It should be stressed that Storkey (2002) was comparing differences in using ethnic-specific and borough-specific morality rates for London and not ethnic-specific rates at the borough-level, which would be more detailed. Storkey does consider borough-level, ethnic-specific rates estimated from COB; however, after considering the data used to produce this, and the reliability of these if they were produced, she decided against producing the rates. In such a case, as was suggested for the Bradford wards, it may be an option to use the area-specific mortality rates and scale them depending on the differences between ethnic group mortality behaviour if it is available; even though it is acknowledged that to use an area-specific rate should partly account for ethnic differences.

Nonetheless, even when differences in mortality is artificially accentuated for non-White groups, this may not impact greatly on their population size, due to their young age structures. This finding was realised in research carried out at the national-level, for England only by ten ethnic groups, in sensitivity testing of the projection process using two alternative options (Williamson 2003). The options were compared to a baseline population projection, using GAD England mortality rates equally across each ethnic group. The first option was to use an SMR by COB for people aged 20-69 in the groups: Caribbean Commonwealth to account for Black Caribbean ethnic group and India, Pakistan and Bangladesh to account for Indian, Pakistani and Banglandeshi ethnic groups (estimated by Harding and Balarajan 2002:124). The second option was to crudely increase the mortality level for non-White ethnic groups by twenty per cent. The consequence in the percentage population change from 1991 to 2001 showed that using the SMRs from Harding and Balarajan’s work would only result in differences of only 0.1%, 0.1% and 0.3% in the population sizes from the baseline projection for these three ethnic groups. On the other hand, using the crude twenty percent increase from the GAD England SMR, results in a maximum reduction in the population size for any of the nine non-White ethnic groups of 0.7%. For example, the greatest change found, 0.7%, was for ethnic group Black Caribbean. The baseline projection suggested that the Black Caribbean population would increase by 10% over the period 1991-2001, whereas using the crude 20% increase in the SMR yielded a 9.3% population increase. Hence, at the national-level crudely artificially increasing SMRs by 20% for non-White groups does not greatly impact on overall population size.

In conclusion, however, drawing on all the information presented in the chapters on fertility rates in the author's PhD (Williamson 2006), it is important to include as much detailed information as possible on the population of interest when trying to create age-specific rates. Thus, despite the differences in ethnic-specific mortality rates being small, described by Storkey (2002), and the example of the impact of different mortality levels at the national-level, it will still be recommended to base rates on the most detailed information available. For example, if the projections are long-range then ethnic mortality differences will be more obvious, as minority ethnic group age structures will change over time. Moreover, in trying to recommend the best method it is useful to take account of work by Storkey (2002), she compares four different methods to estimate the base, or ‘exposed to risk’, populations to create mortality rates according to COB for the period 1996-98:

> There are clearly no ‘correct’ answers from these different models. What is important is to be aware of the nature and the extent of difference in mortality estimates that arise through these different approaches.
> (Storkey 2002:118)

1.6 Potential future research

The review of data sources and methods has uncovered the possibility of extending work presented in the PhD research by the author (Williamson 2006) on ASFRs, whereby ‘Bradford fertility areas’ were created and then ethnic-specific fertility rates were created for these ‘fertility areas’ (which were groupings of wards based on cluster analysis and the 1991 ONS classification of wards). These ‘fertility areas’, or even the four Bradford PCT’s, could be used to create ethnic-specific ASMRs and the rates could be then smoothed by using the H-P curve (Heligman and Pollard 1980) or by using a relational approach (either using the ASMRs from the Bradford district, the GOR of Yorkshire and Humber or GAD England as a standard).

References

Andelin, J. and Simpson, L. (2005). *POPGROUP Population estimation and forecasting system. Reference manual*. version 3 (available from www.ccsr.ac.uk/ popgroup/useful/pg3RefManual April05.pdf).

Aspinall, P.J. (2001). Operationalising the concept of ethnicity data in studies of sociology of health and illness. *Journal of Health and Illness*. Vol.23(6), pp829-862.

Balarajan, R., Bulusu, L, Adelstein, A.M. and Shukla, V. (1984). Patterns of mortality among migrants to England and Wales from the Indian subcontinent. *British Medical Journal (BMJ)*. Vol.289, pp1185-7.

Balajaran, R. and Bulusu, L. (1990). Mortality among Immigrants in England and Wales 1979-1983, in Britton, M. *Mortality and geography: a review in the mid-1980s: the Registrar General's decennial supplement for England and Wales*. OPCS Series DS no.9, pp103-121. HMSO, London.

Bardsley, Hamm *et al* (2000) [cited in Storkey (2002)].

Benjamin, B. and Pollard, J.H. (1980). *Analysis of Mortality and other Actuarial Statistics*. Heinemann, London.

Bhat, M.P.N (2004). *A Generalized Relational Model for Demographic Analysis*. Adult Mortality in the Developing World Preliminary Papers, Center on the Economics and Demography of Aging, University of California, Demography Department, Berkeley (available from www.ceda.berkeley.edu/events/AMDC_Papers /Bhat-amdc.pdf).

Bradford Health Authority (2001). *2001 - A Public Health Odyssey the journey towards better Health in Bradford*. Annual Health Report.

Brass, W. (1971). On the scale of mortality, in W. Brass (ed) *Biological Aspects of Demography*. Taylor and Francis Ltd, London, pp69-110.

Brooks, C., Sams, D. and Williams, P. (1980). *A Time Series of Smooth Approximations for Age, Sex, and Marital Status Specific Death Rates in Australia, 1950/51 to 1975/76 With Projections to the Year 2000*. Research memorandum, Melbourne, Australia: Impact Project Research Centre [cited in (Rogers 1986), the version of H-P curve for *m*-type rates].

CDU (1999). *Undercount adjusted Census data: EWCPOP, SOCPOP & MIGPOP*. http://census.ac.uk/cdu/Datasets/1991_Census_datasets/Area_Stats/Adjusted_data/Undercount_adjusted_census_data/ [accessed 15/12/05].

Coale, A.J. and Demeny, P. (1966). *Regional Model Life Tables and Stable Populations*. Princeton University Press, New Jersey.

Coale, A.J, Demeny, P. and Vaughan, B. (1983). *Regional Model Life Tables and Stable Populations*. (2nd edition) Academic Press, New York.

Congdon, P. (1991). *Uses of Demographic Graduation in Local Projections*. BSPS working paper (Peter Congdon, LRC).

Congdon, P. (1993). Statistical Graduation in Local Demographic-Analysis and Projection. *Journal of the Royal Statistical Society Series (RSS) A*. 156(2), pp237-270.

Congdon, P. (2004). Modelling Trends and Inequality in Small Area Mortality. *Journal of Applied Statistics*. 31(6), pp603-622.

Cooper, N. (2001). Analysis of infant mortality by risk factors and by cause of death in England and Wales, in Griffiths, C and Fitzpatrick, J. *Geographic Variations in Health - Decennial Supplement DS16*. TSO, London.

Cummins, C., Winter, H., Cheng, K-K., Maric, R., Silcocks, P. and Varghese, C. (1999). An assessment of the Nam Pehchan computer program for the identification of names of south Asian origin. *Journal Public Health*. 21, pp401-6.

Ewbank, D.C, Gomez de Leon, J.C., Stoto, M.A. (1983). A reducible four-parameter system of model life tables. *Population Studies*. 37, pp105-127.

Forfar, D.O., McCutcheon, J.J. and Wilkie, A.D. (1988). On Graduation by Mathematical Formula. *Institute of Actuaries*. 115, pp1-149.

GAD (2004a). *National Population Projections 2002-based*. Series PP2, No.24. HMSO, TSO. London.

GAD (2004b). *Interim life table methodology*. www.gad.gov.uk/Life_Tables/ Methodlogy.htm [accessed 27/09/04].

Gilje, E. (1969). Fitting curves to age-specific fertility rates: Some examples. *Statistisk tidskrift (Statistical Review of the Swedish National Central Bureau of Statistics)*. Series III, vol.7, pp118-134.

Greville, T.N.E. (1969). *Theory and Applications of Spline Functions*. Academic Press, Inc., London.

GRO (2002). *Life Tables 1990-92. First Supplement to the Annual Repost of the Registrar General for Scotland 1996*. www.gro-scotland.gov.uk/statistics/library/ annrep/lifetabs/index.html [accessed 23/10/02].

GRO (2004). *Life Tables 1990-92, Appendix B.* www.gro-scotland.gov.uk/grosweb/ grosweb.nsf/pages/sl_appb [accessed 13/10/04].

Haberman, S. and Renshaw, A.E. (1996). Generalized Linear Models and Actuarial Science. *The Statistician*. Vol.45(4), pp407-436.

Hannerz, H. (2001a). Manhood trials and the law of mortality. *Demographic Research*. Vol.4, article 7 (available from www.demographic-research.org).

Hannerz, H. (2001b). An extension of relational methods in mortality estimation. *Demographic Research*. Vol.4, article 10 (available from www.demographic-research.org).

Harding, S. (2003). Mortality of Migrants from the Indian Subcontinent to England and Wales: Effect of Duration of Residence. *Epidemiology*. Vol.14(3), pp 287-292.

Harding, S. (2004). Mortality of migrants from the Caribbean to England and Wales: effect of duration of residence. *International Journal of Epidemiology*. Vol.33(2), pp382-386.

Harding, S. and Balarajan, R. (2001) Longitudinal study of socio-economic differences in mortality among South Asian and West Indian migrants. *Ethnicity and Health*. 6, pp121-128.

Harding, S. and Balarajan, R. (2002) Mortality data on migrant groups living in England and Wales: issues of adequacy and of interpretation of death rates, in Haskey, J. *Population projections by ethnic group a feasibility study*. HMSO. No.67, pp115-127.

Harding, S., Dews, H. and Simpson, L. (1999). The potential to identify South Asians using a computerised algorithm to classify names. *Population Trends*. 97, pp46-49.

Harding, S. and Maxwell, R. (1997). Differences of mortality of migrants, in Drever, F. and Whitehead, M. (eds). *Health inequalities*. OPCS series DS no.15. TSO, London.

Heligman, L. and Pollard, J.H. (1980). The age pattern of mortality. *Journal of the Institute of Actuaries*. 107, pp49-75.

Hinde, A. (1998). *Demographic Methods*. London, Arnold Publishers.

Hoem, J.M., Madsen, D., Nielsen, J.L., Ohlsen, E.-M., Hansen, H.O., and Rennermaln, B. (1981). Experiments in Modelling Recent Danish Fertility Curves. *Demography*. 18, pp231-244.

Horner, D. (2004). *Identifying Ethnicity: A comparison of two computer programmes designed to identify names of South Asian ethnic origin*. 2005.

Karlis, D. and Kostaki, A. (2002). Bootstrap Techniques for Mortality Models. *Biometrical Journal*. 44(7), pp850-866.

Kostaki, A. (1992) [cited in Tabeau, Van den Berg Jeths *et al* (2001)].

Kostaki, A. and Panousis, V. (2001). Expanding an abridged life table. *Demographic Research*. 5(1), pp139-144 (available from www.demographic-research.org).

LRC (1995) [cited in Storkey (2002b)].

Marmot, M.G., Adelstein, A.M. and Bulusu, L. (1984). *Immigrant mortality in England and Wales 1970–78 Causes of death by country of birth*. OPCS series SMPS no47. HMSO, London.

McCutcheon, J.J. (1979-1981). Some remarks on splines. *Transaction of the Faculty of Actuaries (TFA)*. 37, pp421-438.

McCutcheon, J.J. (1985-1987). Experiments in the graduation of the English life tables (No.14). *TFA*. 40, pp135-147.

McCutcheon, J.J. and Eilbeck, J.C. (1975-1977). Experiments in the graduation of the English life tables (No.13) data. *TFA*. 35, pp281-296.

McNown, R. and Rogers, A. (1989). Forecasting Mortality: A Parameterized Time Series Approach. *Demography*. 26(4), pp645-660.

Murray, C.J.L., Ferguson, B.D., Lopez, A.D., Guillot, M., Salomon, J.A. and Ahmad, O. (2003). Modified logit life table system: principles, empirical validation, and application. *Population Studies*. Vol.57(2), pp165-182.

ONS (1997). *English Life Tables No.15: 1990–92*. ONS Series DS No.14. TSO, London.

ONS (1999). *Subnational population projections*. PP3, No10. TSO, London.

ONS (2002). *Making a population estimate in England and Wales*. Occasional Paper 37, version 4. October 2002. ONS, Titchfield. www.statistics.gov.uk/about/methodology_by_theme/making_population_estimates/downloads/pap37v4.pdf [accessed 11/05/03].

ONS (2004a). *2003-based Subnational Population Projections for England: Methodology Guide*. www.statistics.gov.uk/downloads/theme_population/pp2003methodology.pdf [accessed 08/02/06].

ONS (2004b). *2003-based SNPP for England: Data and Methodology Issues*. (November 2004). www.statistics.gov.uk/downloads/theme_population/2003snppissues.pdf [accessed 08/02/06].

ONS (2004c). *Population and migration methodology / Longitudinal Study / Data quality*. www.statistics.gov.uk/about/data/methodology/specific/population/LS/quality/default.asp [accessed 29/10/04].

ONS (2005). *Mid-2001 and mid-2002 CAS Ward Population Estimates (Experimental Statistics)*.www.statistics.gov.uk/about/consultations/downloads/Methodology_Notes.pdf [accessed 06/12/05].

ONS (2006a). *Health authorities: 2000-based England SNPP*. www.statistics.gov.uk/statbase/xsdataset.asp?vlnk=1504&More=Y [accessed 08/02/06].

ONS (2006b). *Oxford University Population Estimates for Wards, Mid-1998*. http://neighbourhood.statistics.gov.uk/dissemination/MetadataDownloadPDF.do?downloadId=7656 [accessed 27/06/06].

ONS (2006c). *Information On Oxford University Population Estimates for Wards*. http://neighbourhood.statistics.gov.uk/dissemination/MetadataDataset.do?adminCompId=7631&metadataType=DatasetFamily&bhcp=1 [accessed 27/06/06].

OPCS (1979). Population of New Commonwealth and Pakistani ethnic origin: new projections. *Population Trends*. 16, pp22-7.

OPCS (1987). *English life tables no.14: 1980–82 the report by the Government Actuary for the Registrar General for England and Wales*. OPCS Series DS; no.7. HMSO, London.

Papaleonidas, P. (1995). *Graduation by smoothing splines. Experiments with English and Greek fertility rates*. unpublished Masters thesis, Department of Actuarial Science and Statistics, City University, December 2005.

Pollard, J.H. (1975). *Mathematical models for the growth of human populations*. Cambridge University Press, Cambridge.

Preston, S.H., Heuveline, P. and Guillot, M. (2001). *Demography: Measuring and Modeling Population Processes*. Blackwell Publishing Ltd, Oxford.

Rees, P., Norman, P. and Brown D. (2004). A framework for progressively improving small area population estimates. *Journal of the RSS A*. Vol.167(1), pp5-36.

Rogers, A. (1986). Parameterized Multistate Population Dynamics and Projections. *Journal of the American Statistical Association.* 84(393), pp48-61.

Simpson, L. (2001). *POPGROUP users guide worked examples*, Bradford Council. December 2001.

Simpson, L. (2002). Estimation methodologies and assumptions for projections: current practice for population projections by ethnic group, in Haskey, J. *Population projections by ethnic group a feasibility study*. HMSO. No.67, pp53-72.

Simpson, S., Middleton, L., Diamond, I. and Lunn, D. (1997). Small-area population estimates: a review of methods used in Britain in the 1990s. *International Journal of Population Geography* 3, pp265-280.

Shaw, C. (1988). Components of growth in the ethnic minority population. *Population Trends*. 54, pp26-30.

Shryock, H.S., Siegel, J.S. and Associates (1976). *The methods and materials of demography* (condensed edition by E.G. Stockwell). Academic Press, London.

Smith, D. and Keyfitz, N. (1977). *Mathematical demography: selected papers*. Springer-Verlag, Berlin. [citing the Reed and Merrell equation (1939)].

Storkey, M.E. (2002). *Population projections of different ethnic groups in London 1991 to 2011*, unpublished PhD thesis, Department of Social Statistics Faculty of Social Sciences, University of Southampton, January 2002.

Tabeau, E (2001). A Review of Demographic Forecasting Models for Mortality, in Tabeau, E; Van den Berg Jeths, A. and Heathcote, C. *Forecasting Mortality in Developed Countries: Insights from a Statistical, Demographic and Epidemiological Perspective*. NIDI, Kluwer Academic Press European Studies of Population. Vol.9.

Wild, S. and McKeigue, P. (1997). Cross sectional analysis of mortality by country of birth in England and Wales. *BMJ*. 314, 705 (8 March).

Williamson, L.E.P. (2003). *Population projections of minority ethnic groups in England: What can we conclude from the differences with the 2001 Census?* BSPS Conference, September 2003 Bristol.

Williamson, L.E.P. (2006). *Population projections for small areas and ethnic groups - developing strategies for the estimation of demographic rates*, unpublished PhD thesis, University of Manchester, School of Social Science, Faculty of Humanities, *forthcoming*.

Wood, J., Horsfield, G. and Vickers, L. (1999). The new subnational population projections model: methodology and projection scenarios. *Population Trends*. 98 Winter, pp21-28.

Woods, L.M. (Laura.Woods@lshtm.ac.uk), Wed 1 Sept 2004. Re: Popfest 2003 / mortality smoothing technique. e-Mail to L. Williamson (Lee.Williamson@stud.man.ac.uk).

Woods, L.M., Rachet, B., Riga, M., Stone, N., Shah, A. and Coleman, M.P. (2005). *Journal of Epidemiology and Community Health*. 59, pp115-120.

Zaba, B. (1979). The four-parameter logit life table system, *Population Studies*. 33, pp79-1.

Appendix 1

Information on the wards included in each of the 4 Bradford Primary Care Trusts

Ward	Primary Care Trust (PCT) Description
Baildon	North
Bingley	Airedale
Bingley Rural	South & West
Bolton	North
Bowling	City
Bradford Moor	City
Clayton	South & West
Craven	Airedale
Eccleshill	North
Great Horton	South & West
Heaton	City
Idle	North
Ilkley	Airedale
Keighley North	Airedale
Keighley South	Airedale
Keighley West	Airedale
Little Horton	City
Odsal	South & West
Queensbury	South & West
Rombalds	Airedale
Shipley East	North
Shipley West	North
Thornton	South & West
Toller	City
Tong	South & West
Undercliffe	City
University	City
Wibsey	South & West
Worth Valley	Airedale
Wyke	South & West