The Survivor Ratio Method for Estimating Numbers at High Ages

Roger Thatcher

Vaino Kannisto

Kirill Andreev

Table of Contents

The Survivor Ratio Method for Estimating Numbers at High Ages

Roger Thatcher [1], Vaino Kannisto [2], and Kirill Andreev [3]

Abstract:

Only a few countries have official population statistics which are sufficiently accurate to produce reliable estimates of death rates at high ages. For other countries, there are several methods which can be used to produce improved estimates. The choice is important for research on old age mortality. In 1999 the Max Planck Institute for Demographic Research undertook a research project to compare the performance of the three leading methods, using data for nine countries over 35 years. This paper describes the research and the results, which were unexpectedly simple. It also gives an authoritative account of the most successful method.

[1]   129 Thetford Road, New Malden, Surrey, KT3 5DS, United Kingom.

[2]   Campo Grande 1, 6-D, P-1700-086 Lisbon, Portugal.

[3]   Department of Community Health & Epidemiology, Abramsky Hall, Queen's University, Kingston, Ontario, Canada K7L 3N6. Email: andreevk@post.queensu.ca OR kirill_andreev@hotmail.com. Homepage: http://post.queensu.ca/~andreevk/.

1. Introduction

For research on trends in mortality at high ages, it is necessary to know not only the numbers of deaths at each year of age, but also the corresponding numbers who were at risk, because they were living in the same population at the same time. Many countries have reliable data on deaths, but at present only a few have equally reliable statistics on the numbers at risk at high ages. These few countries, which all derive their statistics from population registers, are Denmark, Finland, Iceland, the Netherlands, Norway, Sweden and Switzerland.

For the other countries, several methods have been proposed to make improved estimates of the numbers at risk. The first, historically, was the "method of extinct generations", also known as the "method of extinct cohorts", which was proposed by Vincent (1951) and is described in the next Section. This method can only be applied strictly to "extinct" cohorts in which all the members have already died. However, in order to extend their work, Vincent and his assistant Depoid sometimes made estimates of the numbers of deaths which were still to come in the cohorts which were <u>almost</u> extinct. It was they who first used survivor ratios. Other methods of estimation were later developed and are described in a review by Thatcher (1993). A further method has since been added by Andreev (1999).

It should be noted that the work in this paper is not confined to cohorts which are almost extinct. Some of the estimates extend down to age 90. As 20 per cent of females live to reach age 90, the cohorts concerned still have some way to go.

However, these methods cannot all be applied in all countries at all times, due to data limitations and other difficulties. The present paper therefore concentrates on the three main methods. In 1999 the Max Planck Institute for Demographic Research decided to initiate a research project, to see how the performance of these three main methods would have compared, if they had been applied at the time to the data for nine countries in each of the years from 1960 to 1995. The nine countries were Denmark, England and Wales, Finland, France, Japan, the Netherlands, Norway, Sweden and Switzerland. The data were taken from the Kannisto-Thatcher database on old age mortality, currently maintained by the Max Planck Institute at Rostock.

The outcome was unexpectedly simple. When the results were ranked, it was found that one particular variant of one of the main methods performed consistently better than all the others. This was a very fortunate conclusion. If the tests had shown that one method was better in some cases and that another method was better in other cases, it would have been much more difficult to decide how to update the database. The present paper is designed to describe this research and to put the results on record.

The most successful method was a variant of the Survivor Ratio (SR) method, which is a modern development of a method which was first used by Vincent (1951).

However, much has changed since 1951 and the variant includes a correction to allow for declining mortality at high ages. An up-to-date account of the method is therefore needed and this is given in Section 2, contributed by Thatcher and Kannisto. The research itself and the results are given in Section 3, contributed by Andreev.

For those who are interested in the data, the book by Kannisto (1994) describes how the Kannisto-Thatcher (K-T) database was set up. It gives the sources, briefly appraises the reliability of the data and also deals with the problem of splitting numbers of deaths between the triangles on the Lexis diagram. Some further information relating to England and Wales will be found in Thatcher (1992). A full description of the current methods is given in the documentation of the database which is available from the Max Planck Institute.

2. The Survivor Ratio Method

2.1 Vincent's method

In some countries, there are reliable population registers and there is no difficulty in calculating accurate death rates at each year of age. In other countries, death rates are calculated by the traditional method of dividing the numbers of deaths (found from civil registrations) by the numbers at risk as given by the official population statistics, which are based, ultimately, on the Censuses of Population. Unfortunately it is known that some of the very high ages recorded in the censuses are not always accurate. At such very high ages, death rates calculated from census data are not reliable.

A different method was proposed by Vincent (1951). At high ages, and for dates sufficiently far in the past, the population at risk can be estimated directly from the death registrations, without using the censuses of population at all. Once all the members of a given cohort have died, their dates of birth and death give enough information to reconstruct the numbers who were alive at earlier dates, provided that we confine ourselves to ages which are high enough for international migration to be negligible. This is the "method of extinct generations" or the "method of extinct cohorts".

However, it is not really necessary to wait until all the members of the cohorts concerned have died, before one can start to make provisional estimates of the numbers at risk. By the time that the members of a cohort have reached (say) age 100, only a small proportion of its original members will still be surviving and this proportion - or, more precisely, the ratio of the number of survivors who are still living to the numbers in the cohort who died during (say) the last five years - can be estimated from the experience of previous cohorts. On multiplying this estimated "survivor ratio" by the

known number of deaths which have occurred in the given cohort during the last five years, we obtain a provisional estimate of the number of survivors who are still alive. We can then reconstruct the past population by adding the estimated number of survivors to the known number of past deaths, cohort by cohort. This illustrates the method which was used by Vincent and Depoid to extend the method of extinct cohort so that it could be applied to cohorts which are <u>almost</u> extinct.

In order to see how this method works, we need a notation. Consider the cohort born in the calendar year b and let d(b,y) be the number of deaths in this cohort during the calendar year y. Let s(b,Y) be the number of survivors in this cohort who were alive at the end of the calendar year Y, or on 1st January of the year Y+1. Then the "survivor ratio" for this cohort at the end of year Y, when the denominator of the ratio is taken to include deaths in the last k years of age, will be given by

$$R(b,Y) = s(b,Y)/[d(b,Y) + d(b,Y-1) + ... + d(b,Y-k+1)] \qquad (1)$$

In the immediately preceding cohort, born in the year b-1, the corresponding ratio at the same age was

$$R(b-1,Y-1) = s(b-1,Y-1)/[d(b-1,Y-1) + ... + d(b-1,Y-k)] \qquad (2)$$

Also, as an identity, we have

$$s(b-1,Y-1) = s(b-1,Y) + d(b-1,Y) \qquad (3)$$

All the d's are known. Thus if we already have an estimate of s(b-1,Y), we can use (3) to find s(b-1,Y-1) and hence (2) to find R(b-1,Y-1). If we are now prepared to assume that the cohort born in the year b has the same survivor ratio as the cohort born in year b-1, as should be the case if mortality rates are stable, then we shall have

$$R(b,Y) = R(b-1,Y-1) \qquad (4)$$

and we can now use (1) to estimate s(b,Y). In this way, we can derive an estimate of s(b,Y) from an estimate of s(b-1,Y). If we know the highest age ω where there is expected to be only one survivor, so that s(Y- ω ,Y) = 1, then we can start the process at age ω and work down the cohorts, by iteration, and so estimate the numbers of survivors at lower ages at the end of the year Y.

This description is slightly simplified because it bases the estimate of R(b,Y) on the assumption that this is the same as in the single immediately preceding cohort. It is

possible to damp down fluctuations in the ratios by using the average ratios over the last few cohorts, say m cohorts, instead of just a single cohort.

This original version of the SR method works well when mortality rates are stable but the sizes of successive cohorts are changing. In effect, it then uses the deaths in the last k years as a measure of the relative sizes of the cohorts. These were the conditions under which Vincent first devised the method and applied it to centenarians. However, when the method was applied in different circumstances, where mortality rates were falling quite rapidly and where it was sometimes necessary to extend the estimates to ages as low as 90, it was found that modifications are needed. The main work on this was done by Kannisto and his main findings can be summarised as follows:

(i) When mortality rates are falling, the survivor ratios are not constant from one cohort to the next and the simple assumption (4) is no longer valid. Kannisto proposed the introduction of a correction factor c, so that (4) is replaced by

$$R(b,Y) = c.R(b-1,Y-1) \tag{5}$$

When mortality rates are falling, we expect c to be slightly greater than 1.

(ii) The method starts at the highest age and works downwards. Kannisto devised an algorithm for estimating the highest age ω . This is taken as the lowest age at which the latest five-year average of the total number of deaths at still higher ages is less than 0.5.

(iii) As the SR estimates are produced one by one, starting from the highest age and working downwards, they gradually become less and less reliable. This is mainly because they become more and more dependent on the assumptions. (They also ignore migration). Eventually there will come a point, typically somewhere between ages 90 and 100, where the SR estimates need to join on to the official statistics, at the lower ages where the official statistics can be accepted as reliable. The condition that the SR estimates should join on to the official statistics in an acceptable way imposes a constraint. In fact, this constraint has a distinct benefit. Provided that it can be specified sufficiently closely, there will be only one value of c which will produce SR estimates which are consistent with the constraint, and this makes it possible to determine a definite numerical value for the parameter c.

2.2 Practical problems

In order to apply the SR method in practice, we need to have values for the parameters k, m and c. Here k is the number of years of age in the denominator of the survivor ratio, and this needs to be at least five. Also m is the number of cohorts which are used

to smooth the survivor ratio and this is also commonly taken as five. These are not critical assumptions, but the next Section reports on some research which was designed to see whether they can be improved.

In order to determine the parameter c, we need to specify the way in which the SR estimates should join on to the official statistics. In some countries, there is not much choice. For example, if there are no official statistics for individual ages above 90, but only a total for the age group "90 and over", then the only practical choice is to link the SR estimates either to the official number at age 90, or to the official total for the age group "90 and over". It is important to know which of these choices performs best, and this is the reason for some of the variants which are investigated in the next Section.

There are other countries where the choice is not quite so simple, because the official statistics are believed to be reliable for a few years of age above 90, but not as far as 100. In principle, one would like to know the age at which the official statistics cease to be reliable. Unfortunately, the only sure way to determine this is to compare the official statistics with the estimates found by the method of extinct generations, but this involves waiting until the cohorts are extinct. The Max Planck Institute is hoping to discover a more rapid mechanical rule. In the meantime, there is a more traditional method which can be used. This is to place the official statistics for ages 90 and over alongside the corresponding SR estimates, to see where they diverge. A judgement can then be made on how they should be joined.

3. Research

3.1 Alternative methods

Of course, the SR method is not the only way to produce estimates of the number of survivors in each cohort. The main alternatives are the Das Gupta method (DG), which was applied to the U.S. Census of Population in 1980 and has been fully described by Das Gupta (1990); and the more recent Mortality Decline method (MD), proposed by Andreev (1999).

The SR, DG and MD methods all operate by using ratios calculated from recent data. The SR method uses the "survivor ratio", as described in Section 2. The DG method uses the "deaths ratio", which is the ratio D_{x+1}/D_x of the number of deaths at age x+1 to the number of deaths at age x in the same cohort. The MD method uses death rates, estimated by a linear statistical model.

Although there is a broad similarity between these approaches, in that they all use ratios derived from recent data on deaths, they differ considerably in the details. There are also variants of each method. In 1999 the Max Planck Institute for Demographic

Research decided to make a systematic comparison between these methods and their variants, to see how well they would have performed if they had been applied at the time in each of the years 1960 to 1995, in each of the nine countries listed in Section 1.

3.2 Further details, variants and notation

The SR method has already been described fully in Section 2. It has two parameters which can be chosen at will, namely k (which is the number of years of age in the denominator of the survivor ratio) and m (which is the number of cohorts used to smooth the survivor ratios). The SR estimates found by using just these two parameters, and the equation (4) above without any constraint, are denoted by the abbreviation SR(k,m). However, equation (5) introduces a parameter c which is later determined by a constraint. The variants of particular interest are those in which the SR estimates are constrained either to agree with the official statistics (or other independent estimates) at age 90, or to give the same total population at ages 90 and over. These variants are denoted by SR(k,m,90) and (k,m,90+), respectively.

In the DG method, the ratios D_{x+1}/D_x are calculated from the data for the last k calendar years. These ratios are then used to "predict" the numbers of future deaths in each future year in each cohort, until the cohorts are extinct. Estimates of the numbers living are then found by the method of extinct generations. Finally, the results are adjusted so that the total is consistent with an independent estimate. For the purpose of the present comparisons, the DG totals were adjusted to agree with the same independent estimates as were used for the SR method, at ages 90 and 90+. The results are denoted by DG(k) when unadjusted and by DG(k,90) or DG(k,90+) when adjusted.

In the MD method, current death rates were estimated by extrapolating the death rates in the preceding k cohorts. This was done by fitting linear statistical models to the death rates at each age and then using the predictions given by the model for the current cohorts. For full details see Andreev (1999). Again, the totals may or may not be constrained to agree with independent estimates at ages 90 or 90+. The resulting estimates are denoted by MD(k), MD(k,90) and MD(k,90+), respectively.

3.3 Method of comparison

In order to investigate the relative merits of these three methods and their more obvious variants, it was decided to apply them to nine of the countries for which reliable data are available in the Kannisto-Thatcher (K-T) database for old age mortality (Kannisto, 1994). The calculations were made for each of these nine countries, separately for each

of the years from 1960 to 1995, and separately for each selected variant of each method. These included three variants of the SR method, six of the DG method and three of the MD method. However, because the MD method required a longer run of years than the others, it was applied only from 1980 onwards. The results are accordingly presented in two parts, for 1960-1979 showing the SR and DG methods only, and for 1980-1995 showing all three methods.

For the constrained calculations which require independent estimates of the number aged 90 and the total number aged 90 and over, these were taken from the K-T database. For the earlier period 1960-1979, the independent figures were calculated entirely by the method of extinct generations, using only the deaths after the year concerned. They are therefore completely independent of the SR and DG estimates, which are based entirely on the deaths <u>before</u> the year concerned. They therefore did not depend at all on the SR or DG methods, or on the population registers. They were also completely independent in the earlier part of the period 1980-95.

However, in the latter part of 1980-95 they were no longer completely independent. For example, the K-T estimate for the number aged 95 in 1990 is found by adding the known number of deaths in this cohort between 1990 and 1995 to the estimated number of survivors in the cohort who were aged 100 in 1995. Thus the K-T estimate for 1990 will depend, at least partly, on the method of estimation which was used in 1995. However, there is not much alternative to using the K-T Database in the comparisons, because it contains all the best information which is available at present. Improved comparisons will become possible in time, when the cohorts concerned are nearer to extinction.

All estimates were made for males and females separately. Thus for any given age, there were separate estimates for 9 x 9 x 20 x 2 = 3,240 countries/variants/years/sexes in the years 1960-1979 and for a further 9 x 12 x 16 x 2 = 3,456 in the years 1980-1995. These 6,696 estimates, which each involved a great deal of calculation, were made for each separate year of age from 90 upwards.

Because the amount of information produced by these calculations was so huge, it is desirable to find a condensed method of presentation and a simple procedure for discriminating between the different methods and variants. For presentation, it was decided to concentrate on the relative success of each method and variant in estimating (a) the total number in the population at ages 100 and over: and (b) the total number in the population at ages 95 and over. Here the estimates are all based on the numbers of deaths before the event. The observed populations, with which they are compared, are taken from the K-T database.

The procedure ranks the estimates given by the various methods and variants and assigns them a score between 0 and 1,000, based on how close the estimate falls to the observed population in the year concerned. For example, nine methods were used to

produce estimates for the year 1960. For each method, we can compare the estimate of the population aged 100 and over with the corresponding observed population, in each country. The method which gives the smallest absolute difference is given the highest rank. The ranks are then summed over countries and years, divided by the total sum of all ranks, and finally multiplied by 1,000. The most successful methods are those with the highest scores.

3.4 Results in Table 1

The relative scores produced by this procedure are summarised in Table 1. This table shows a remarkable result which leaps to the eye: on each line, the highest score is in the last column. This means that a particular variant of the SR method, namely the variant SR(5,5,90+), scored better than all the other methods in every case. It scored better both for ages 100 and over and for ages 95 and over; it scored better for both males and females; and it scored better in both the earlier years 1960-79 and in the later years 1980-95. Such a simple finding was totally unexpected.

This highly successful variant of the SR method is constrained to agree with the independent estimate of the total population at ages 90 and over. This variant performed better than the unconstrained version of the SR method, and also better than the variant constrained to agree with the independent estimate for age 90 (as opposed to ages 90+). Table 1 shows that this was also found for the DG and MD methods. In all cases, the variant constrained to agree at ages 90+ performed better than the unconstrained version and also better than the variant constrained to agree at age 90.

It is not surprising that the constrained variants of the SR and DG methods perform better than the unconstrained versions. The reason is that the unconstrained versions do not make adequate allowance for the fall in mortality rates at high ages. Indeed, the constraints were proposed by the authors of the methods themselves, on the basis of previous experience.

It is not so obvious why the constraint to the total at ages 90+ should perform better than the constraints to the single age 90, but there are some points which may be relevant. The official estimate for the total at ages 90+ may or may not be completely accurate, but its percentage error will certainly be far less than the percentage error at higher ages. Also, it will produce an estimate of the <u>change</u> between successive years which can certainly be accepted as reliable. This is because the change in the total depends only on the total deaths above age 90 and on the numbers reaching age 90. It does not depend at all on the accuracy of the numbers in the population at individual ages over 90. A reliable estimate of the change is an important piece of information which needs to be taken into account in the estimation process.

Table 1:   *Relative rank distributions for all methods*

	MD(10)	MD(10,90)	MD(10,90+)	DG(5,90+)	DG(3,90+)	DG(5,90)	DG(3,90)	DG(5)	DG(3)	SR(5,5)	SR(5,5,90)	SR(5,5,90+)
Ages 100 and over												
						Years 1960-1979						
Males				120	107	114	97	112	91	107	121	**132**
Females				115	92	109	86	104	75	121	148	**151**
Both sexes				117	99	111	92	108	83	114	135	**141**
						Years 1980-95						
Males	76	70	76	85	75	85	71	81	65	93	110	**113**
Females	76	78	89	93	78	87	73	73	59	72	108	**115**
Both sexes	76	74	83	89	76	86	72	77	62	83	109	**114**
						Years 1960-1995						
Males				117	102	113	94	108	86	111	130	**138**
Females				119	97	111	90	100	75	110	146	**152**
Both sexes				118	99	112	92	104	80	111	138	**145**
Ages 95 and over												
						Years 1960-79						
Males				135	116	103	93	92	73	97	132	**159**
Females				141	124	113	98	75	61	86	137	**166**
Both sexes				138	120	108	95	84	67	91	134	**162**
						Years 1980-95						
Males	83	71	95	100	86	81	67	64	56	61	108	**127**
Females	69	87	110	107	93	89	74	59	51	46	97	**120**
Both sexes	76	79	103	103	89	85	70	62	54	53	103	**123**
						Years 1960-1995						
Males				136	115	105	90	89	74	90	137	**163**
Females				145	126	117	99	78	64	76	133	**164**
Both sexes				140	120	111	95	84	69	83	135	**164**

Bold items shows the maximal rank (best results)

Countries included in calculations: Denmark, England and Wales, Finland, France, Japan, Netherlands, Norway, Sweden, Switzerland

Table 2   *Relative rank distributions for unconstrained methods.*

	Ages 100 and over				Ages 95 and over				Ages 90 and over			
	MD(10)	DG(5)	DG(3)	SR(5,5)	MD(10)	DG(5)	DG(3)	SR(5,5)	MD(10)	DG(5)	DG(3)	SR(5,5)
Years 1960-79												
Males		360	274	**365**		352	275	**373**		343	298	**360**
Females		339	232	**429**		354	270	**376**		**373**	303	324
Both sexes		350	253	**397**		353	272	**375**		**358**	301	342
Years 1980-95												
Males	227	263	198	**313**	**309**	257	203	231	**323**	258	207	212
Females	**271**	265	209	255	**314**	262	240	185	**324**	272	245	159
Both sexes	249	264	204	**284**	**311**	259	221	208	**323**	265	226	185
Years 1960-95												
Males		356	261	**383**		**365**	283	352		**364**	295	341
Females		350	255	**395**		**370**	302	328		**389**	328	283
Both sexes		353	258	**389**		**367**	292	340		**377**	312	312

Bold items shows the maximal rank (best results)
Countries included in calculations: Denmark, England and Wales, Finland, France, Japan, Netherlands, Norway, Sweden,
   Switzerland

3.5 Results in Table 2

The results in Table 1 clearly show that constrained methods perform better than unconstrained methods, at least when independent estimates (such as the total population at ages 90 and over) are available to provide the constraint. However, it is still of interest to see how well the unconstrained methods compare with each other when no independent estimates are available, and also how well they succeed in producing estimates of the total at ages 90 and over.

    The resulting scores are shown in Table 2. It will be seen that in the years 1980-95, when all three methods can be compared, the unconstrained MD method performed better than the unconstrained SR and DG methods, for both males and females at ages 90+ and 95+, and also for females at ages 100 and over.

3.6 Relative error of survivor estimates (Figures 1a, 1b, 1c)

Tables 1 and 2 show the relative rankings of the various methods and variants, but these do not tell us the accuracy of the estimates, as measured by the percentage differences between the estimates and the observed population values. These percentage errors from the pooled data for the nine countries were calculated for each year, and for each of the ages 90, 91, ..., 100+. The resulting percentage errors are illustrated in the form of Lexis maps in Figures 1a, 1b and 1c.

Each rectangle on the map corresponds to a deviation at a certain age and year. The scale on the right separates the deviations into groups and each group is displayed with a certain colour on the map. For example, all deviations in which the observed population is underestimated by more than 15% are displayed in the deepest blue. All deviations in which the observed population is overestimated by more than 15% are displayed in the deepest magenta.

As expected, the methods which are constrained to agree with the observed population at ages 90 or 90+ perform better than the unconstrained methods. The constrained methods still tend to overestimate the observed population, but the error is typically only of the order of 1%-5%. The unconstrained methods, however, tend to underestimate the observed population and have larger errors, typically of the order of 5%-15%. For the SR and DG methods this is simply due to the fact that, when unconstrained, they do not allow for the fall in mortality rates at high ages. The unconstrained MD method allows for the mortality decline but at certain ages the estimates are still too low. This appears to be due to the fact that death rates have not declined at a constant rate over time.

The errors illustrated in Figure 1 are the averages found when the errors in the nine countries are pooled. For individual countries, the errors can be larger or smaller. As an illustration, the method SR(5,5,90+), when applied to estimate the total number of centenarians in England and Wales, had errors which in about half the years between 1980 and 1995 were as large as 5%. The highest error was 9%. Percentage errors were even higher in the smaller countries.

3.7 The parameters in the SR method

We recall that the SR parameters has the parameters k (the number of ages in the denominator of the survivor ratio) and m (the number of cohorts in the moving average). The most commonly used values are five in each case and these were the values used in the comparisons described above. Since then, some further work has been done to see whether k = 5, m = 5 is an optimal combination.

For the case of England and Wales, estimates were made of the total number of centrenarians (aged 100 and over) in each year from 1970 to 1995, using both the unconstrained method SR(k,m) and the constrained method SR(k,m,90+), for all possible values of k and m from 1 to 10. This made 100 combinations altogether, for each year. The results were then ranked, in the same way as for Figure 1, and the results are plotted on the Lexis map in Figure 2. The lightest entries show the best fits. Relatively small values of k (up to 6 ages) and slightly larger values of m (4 cohorts or more for the unconstrained SR method and 6 or more for the constrained method) seem to produce the better rankings, but there is no clear optimal combination.

A point to be remembered is that the small cohorts which were born when the birth rate was low during the First World War are now entering the old age field. In order to maintain the sizes of the denominators in the survivor ratios, higher values of k (perhaps up to 10) may be needed in the coming years.

4. Summary

In circumstances where existing estimates of the numbers living in the population at very high ages are not reliable for individual years of age, there are several methods which can be used to produce improved estimates of the numbers living, by making use of the data on deaths. The aim of the research was to find out which of the available methods performed best, when applied to the data for nine countries in each of the years 1960-1995.

The research has produced a very simple and totally unexpected result. There was a single particular method which, in all the cases tested, performed better than all the other methods included in the comparisons. It gave better estimates, not only for the total population aged 100 and over, but also for the total aged 95 and over, for both males and females, and in both the earlier years 1960-79 and in the later years 1980-95. This unexpectedly successful method was a particular variant of the survivor ratio method, constrained to agree with independent estimates of the total population aged 90 and over.

The research also showed that the other two main methods, the Das Gupta and the Mortality Decline methods, also performed better when constrained to agree with independent estimates of the total population aged 90 and over.

Results are also given to show the sizes of the errors in the estimates given by the various methods. Other topics discussed in the paper include the choice of parameters for the survivor ratio method and the practical problem of "switching" from official statistics to improved estimates in certain countries.

References

Andreev, K.F. Demographic surfaces: Estimation, Assessment and Presentation, with application to Danish Mortality, 1835-1995. Ph.D. thesis, University of Southern Denmark (1999).

Das Gupta, P. Reconstruction of the age distribution of the extreme aged in the 1980 census by the method of extinct generations. Washington D.C.20233: Population Division, U.S.Bureau of the Census, 1990.

Kannisto, Vaino. Development of oldest-old mortality, 1950-1990: Evidence from 28 developed countries. Odense University Press, 1994. ISBN 87 7838 015 4

Kannisto, V. On the survival of centenarians and the span of life. Population Studies, 42 (1988) 389-406.

Thatcher, A.R. Trends in numbers and mortality at high ages in England and Wales. Population Studies, 46 (1992), 411-426.

Thatcher, A.R. Overview of methods for estimating population numbers at high ages from data on deaths. Paper presented at the workshop on old age mortality, Duke University, February 1993.

Vincent, P. La mortalité des vieillards. Population, 6 (1951) 181-204.

Figure 1a: *Relative error of survivor estimates (%), Males*

Populations included: Denmark, England & Wales, Finland, France, Japan, Netherlands, Norway, Sweden, Switzerland

Figure 1b: *Relative error of survivor estimates (%), Females*

Populations included: Denmark, England & Wales, Finland, France, Japan, Netherlands, Norway, Sweden, Switzerland

Figure 1c: *Relative error of survivor estimates (%), Both sexes*

Populations included: Denmark, England & Wales, Finland, France, Japan, Netherlands, Norway, Sweden, Switzerland

Figure 2:   *Relative ranks of SR method with different number of ages and cohorts*

Populations included: England & Wales, Sweden, years 1970-1995.
Population at age 100 and over was used between methods. The lighter colours show the highest rank.