# How Many Americans Could Be Alive? 

Magdalena Muszyńska ${ }^{1,2 *} \quad$ Roland Rau ${ }^{1,2+}$<br>${ }^{1}$ Duke Population Research Institute (DuPRI)<br>${ }^{2}$ Population, Policy, and Aging Research Center (PPARC)<br>Terry Sanford Institute of Public Policy; Duke University

April 15, 2008

## 1 Introduction

The extraordinary improvements in human longevity over the last century and a half resulted in a steady increase of the record life-expectancy, by almost 3 months per year for females, and 2.7 months for males (Oeppen and Vaupel, 2002). In only the last half a century, life-expectancy in record-holding countries increased annually by 2.3 months for females and 2 months for males. At the same time, the slower than the best-practice increase in the average life duration in the United States ( 2 months per year for males and 1.7 months for females) resulted in a growing gap between life-expectancy in this country and its record value. In 1950, for example, the gap between the United States and Norway, the country with highest female life expectancy at that time, was 2.2 years. With Japan as the record holder in 2003 with life expectancy at birth of 85.3 years for females, the corresponding value for the United States was 80 years, resulting in a gap of 5.3 years. Various explanations are discussed for the not optimal improvements in life-expectancy in the US, such as non-universal health-care coverage (Boyle Torrey and Haub, 2004), obesity epidemic (Boyle Torrey and Haub, 2004; Preston, 2005), emergence of new diseases related to mental health (Meslé and Vallin, 2006). In addition, some authors claim that the value of life-expectancy before 1980s, when the slow down started, was overestimated as a result of the undercount of deaths at older ages (Meslé and Vallin,

[^0]2006). This would mean that the growing disadvantage of the US is not only the recent phenomenon and to understand the reasons for the poor performance of the country we should concentrate not only on the last decades.
In this paper, our aim is to estimate how this mortality disadvantage in the last fifty years had influenced the size and structure of the US population at the beginning of the XXI century. Results obtained in this study are complementary to those presented by White and Preston (1996), where the authors, asking "How many Americans are alive because of Twentieth-Century improvements in mortality?", compared the actual population size and structure in the year 2000 with one projected with no mortality improvements since 1900. According to this study, the population in the year 2000 would be half as large as the observed 276 million if death rates had remained constant on mortality levels from 1900 with fertility and migration at the actual rates. The largest share of this difference can be attributed to improvements in infant-, child- and young-adult-mortality throughout the first decades of the twentieth century. Nevertheless, also survival advancements during recent decades translate into considerable numbers of lives saved. White and Preston (1996) expect that the population of the United States would have been smaller by 12 million people in 2000 - approximately the population of Illinois, Ohio, or Pennsylvania — if no further improvements in mortality had been made since 1960. The advancement in mortality reduction in the United States, however, could have been even greater if we consider the best practice countries. In this study we adress the question how many Americans could be alive if mortality in the United States over the last half a century had been the lowest possible, and how the structure by age and sex of this hypothetical population would differ from the one observed.
As the lowest-possible mortality here we understand the lowest age-specific death rates. The following two assumptions concerning these rates are made. First, we assume that in each year the lowest mortality pertains to a single country that is the record holder in lifeexpectancy. As this study is only a hypothetical scenario, the lowest possible mortality in every age group that could have been achieved in a given year does not necessarily belong to a single country. As the aim of our study is to assess the number of lives lost in the US as a result of non-optimal developments in mortality, we introduce a second set of rates. It is estimated for each sex, year and age as a minimum from smoothed national mortality surfaces over age and time for 15 low-mortality countries. The reason for estimating rates from smoothed surfaces is random variation in age-specific mortality rates resulting from small population sizes in some countries. When the number of people exposed to the risk is small the age-specific death rates are seriously distorted as a result of fluctations in the observed realization of the underlying risk of dying. This effect is particularly strong at older ages, where the number of surivors is small. The same problem exists at young ages
in countries with small populations. For example, with the very low probability of dying, and small population size there might be years when no event occured. As a result the observed death rates for single countries, as well as record life-expectancy, are distorted and does not necessary reflect the underlying pattern of mortality.

## 2 Methods and Data

### 2.1 Method

The number of Americans who could be alive, given that the level of mortality in the last half a century had been the lowest possible, was estimated in a female-based projection using the component method in one-year intervals of age and time (Preston et al., 2001; Shryock and Siegel, 1976; Smith et al., 2002). For single year and age, separately for both sexes, we projected the hypothetical population at the beginning of the year applying the best practice probabilities of dying, fertility rates and estimated net migration to the population on January 1st of the previous year.
Two measures of best-practice probability of dying have been applied in the projections. The first one was derived for each calendar year and sex from age-specific death rates observed in the country with the record life-expectancy. The second measure is a minimum value of the probability of dying across the world. These values were estimated from best practice death rates for each calendar year and single age group (compare Section 2.2 on Mortality).
In the projections, we assume that the events occur in the following order: births occur at the beginning of a calendar year, deaths in the middle and migration at the end. As a result, the number of newborns was estimated by applying the single-age-specific fertility rates to the number of women in a given age group at the beginning of a given calendar year. The number of children under 1 year of age as on January 1st 2004 was estimated adjusting the number of births in the previous year using the probability of dying for babies born in the previous year aged zero (the lower right triangle in a Lexis-square). As we assume that death and childbirth occur before the migration event, immigrants start to be exposed to the risk of dying and giving birth to a child only the next year after they move into the country.

Before comparing the total size of the actual population with the total number of people alive in the hypothetical population, we added to the latter the number of men and women aged 101 and older in the actual US population (alive on January 1st 2004). We argue that the error of applying observed US death rates to our hypothetical population at ages above 100 instead of best-practice values is negligible: For men the remaining
live-expectancy at age 101 in 2003 in the US was highest in the world, which means that for these ages mortality can be considered best-practice. For women the mistake resulting from this approximation is rather small: in Japan, which in 2003 was the leading country in mortality at elderly ages, the remaining life-expectancy at age 101 is only one month longer than in the United States ( 2.59 and 2.47 respectively).

The number of people who could be alive in 2004, given the best practice mortality regime, consists of those who would not have died in the last half a century, but also of those who would have been born otherwise: whose parents would not have died before giving birth to them, or those whose grandparents would not have died before giving birth to their parents, and so on. In order to recognise this cumulative effect of lost reproductive resources, in addition to quoting the number of lives that could have been saved, we distinguish in this paper between (a) direct deaths (first generation) and (b) indirect deaths (second and higher generation). To do so, we apply age-specific fertility rates to the number of women who could have been otherwise alive in each year over the period between 1954 and 2003 and age those babies by applying best practice probabilities of surviving until January 1st 2004.

### 2.2 Data

Population The US population counts in single age groups for ages 0 to 100, and aggregate number of those aged 101+ as on January 1st 1954-2004 were obtained from the Human Mortality Database (2007).

Fertility The data on female age specific fertility rates for five-year groups in 1954-2003 were derived from Vital Statistics (National Center for Health Statistics). The sex-ratio at birth according to the age of a mother in five-year age groups in the years 1954-2003 was derived from the same data-source. Both were redistributed to one-year age-group rates using the Karup-King third-difference osculatory interpolation formula (Shryock and Siegel, 1976).

Mortality To obtain best practice death rates, we retrieved data for the following low mortality countries from the Human Mortality Database (2007): Australia, Canada, Denmark, England and Wales, Finland, France, Iceland, Italy, Japan, Netherlands, New Zealand (Non-Maori population), Norway, Sweden, Switzerland, United States. First, for each single year we estimated record life-expectancy and age-specific death rates for the record holder constituted the best-practice values for a given year.
Random fluctuations in the data, however, might lead to estimates of best-practice mortality which are too low. For example in Iceland, where the population is small and
the probability of dying at young ages is very low, in one or more of the last years in some age-groups no death occured. This does not mean that the underlying mortality is zero at these ages; it is rather the outcome of a small numbers of exposures causing large random fluctuations in the number of events (i.e., deaths). To avoid these artefacts, we smoothed mortality surfaces for each of those 15 countries for ages 0-100 and years 1900-2003 separately for females and males. The smoothing itself of the two-dimensional surfaces has been conducted as described in Currie et al. (2004). Their non-parametric approach, which is based on P-Splines (Eilers and Marx, 1996), is rather flexible by not making any assumptions about the functional form of mortality over age or time. In order to estimate the best-practice surface of mortality, minimum values were taken from the 15 smoothed surfaces for each age and year. In our projections we restricted the range to the years 1954-2003.

Migration The net number of international migrants was derived comparing the actual US population as of January 1st each year (1955-2004) to the population projected by applying age-specific fertility rates and death probabilities to the actual population of a previous year. We assumed that migration occurs at the end of the year. Migrants aged $i$ in a given year are calculated as the difference between the actual and projected population aged $i+1$ on the January 1st next year.

### 2.3 Empirical Tests of Methodology

Our paper represents a classic "what-if" scenario. In such cases, various assumptions about reality have to be made. The final outcome of these scenarios can be highly dependent on this input. Therefore we conducted two empirical checks of our methodology:

- First, we tested our method by repeating the projections of White and Preston (1996) for the years 1960-2000. According to our estimates, if there was no health improvement since 1960, the population in 2000 would be $4.2 \%$ smaller than its actual size, which compares to $4 \%$ as reported by White and Preston (1996) or $4.3 \%$ if we divide the size of the population in 2000 given that the mortality improvements stopped in 1960 by the forecasted population in 2000 (White and Preston, 1996, the numbers also reported by). Taking into account the fact that in the original paper the size and structure of population in 2000 were also projected by the authors (the paper was published in 1996) and the fact that the precision reported is up to full-percentage points only (or total millions in the second case), and we could not find information concerning the assumption on the sequence of vital events in the original paper, the difference between these numbers is negligible. This suggests that our method
yields results which are comparable to the study of White and Preston (1996).
- Secondly, applying the rates from the best-practice mortality surfaces, we compared how sensitive our results are to the assumption concerning the sequence of events. Originally, we assumed the following order: 1) births, 2) deaths, 3) migration. Another of many possible scenarios is that the death and migration are evenly spread in a year and the births occur in the middle of the year. This assumption was applied both to the final projections as well as the ones in which we derived the number of net migrants. In Table 3, we compare the result of the two sequences of events on the size and structure of the projected population. The total size of the population projected, assuming the second sequence of events, is larger by 147 thousand than in our original projection. This difference, however, equals to only $0.05 \%$ of the total population projected with the original sequence of events. In addition, there are no significant differences in the sex- or age-structure of the two populations. Hence, comparing results of the two scenarios, we claim that the assumption concerning the sequence of events does not influence significantly the size or structure of the projected population. We could expect the largest difference between the two scenarios in the number of births, as a result of the timing the immigrants are exposed to childbirth. The reason for small differences, however, is that both assumptions concerning the sequence of events are also applied to determine the number of net migrants in the studied period. The estimated number of immigrants is significantly smaller in the case of the second assumption and, as a result of that, the difference in the population size between the two scenarios, resulting from higher number of births in the second projection, is reduced by the smaller number of immigrants in the second one.


## 3 Results

### 3.1 Pattern of US Mortality Disadvantage

Over the last half a century, United States has never been the record holder in respect to life expectancy (compare, Figure 1). In 1954 the gap between mean duration of life in this country and its best-practice value was almost 5 years for men and 3 years for women. The gap was decreasing until the early 1980s. Since the early 1980s, when the gap was minimal, we observe the remarkable development in female mortality in Japan, becoming the record holder for females. The speed of increase of female life expectancy at birth in Japan over the years 1982-2003 was of about 3 months a year, while in the US it accounted to only about a month each year. For males the gap in life expectancy betweeen the US
and the best-practice values has been realtively stable since the beginning of the 1980s and equalled to almost 5 years in 2003, the same value as in the 1950s.
Figures 2 and 3 present the ratio of US death rates to best practice values from smoothed surfaces for each single age (between 0 and 100 years) and calendar year (years 19542003), for females (Fig. 2) and males (Fig. 3) respectively. White areas in the figures correspond to ages and years when mortality in the US was best. Blue colors indicate that US death rates were less than $80 \%$ higher than in the best-practice country. Areas in green denote death rates which were 1.8 to 2.2 times as high as in the US. Even brighter colors point to larger US mortality disadvantages.
For long time periods, the United States used to be the country with the best survival chances at advanced ages (Manton and Vaupel, 1995). Males still belong to the group of 'top-performers'. Women have been losing their advantage at the elderly ages, though, since the late 1980s. Most dramatic are the disadvantages in mortality at adult ages below 60. During the last 15 years, death rates were at least twice as high as in the record holding countries. In particular for women at reproductive ages, an unfavorable trend began in the 1990s with death rates being three and four times larger than the best-practice values. We can already now conjecture that this elevated female mortality at young adult ages implies a large cumulative effect on lives lost in a long time perspective: if a woman dies before giving birth, not only the future babies of her are not born, but also they will not give birth to her grandchildren, and so on. As the minimum age of childbearing is assumed to be 15 years, over 50 years this indirect effect accumulates to a maximum of 4 generations of women.

### 3.2 How Many Americans Could Be Alive?

The differential from optimal development of mortality in the United States in the last fifty years had an impact on the size and structure of the population at the beginning of this century. If mortality in the United States in the last half a century had been as low as in the the countries with the highest life-expectancy, the population in 2004 would be $4.4 \%$ larger than its actual size. When the best-practice age-specific death rates from mortality surfaces are applied in our projections this difference increases to $5 \%$ of the actual population size (compare, Table 1). The number of people who would be alive in January 2004 in the United States according to the first projection equals 305 million and in case of the estimations based on the best-practice surfaces - 306.6 million. That means that 12.8 million lives could have been saved over the last 50 years given the mortality experience in the record-holding country, and 14.3 million with the assumption of minimum mortality experience separately in each year and age group.
The total number of people in 2004, whose lives were lost in the last half century, is similar
to the number of lives who have been actually saved (as in 2004) due to improvements in mortality during the same period. We estimated that if death rates in the years 1954-2003 were fixed at the level of 1954, the population in 2004 would be smaller by 15.5 million. Comparable results have been reported by White and Preston (1996) who kept death rates constant from 1960 to 2000 (compare, Section 2.3).
The number of women who would be alive in January 2004 is 155 million if best-practice mortality conditions prevailed in the United States, disregarding of the assumptions on mortality. That figure, compared to the actual population, means that almost 7 million female lives could have been saved during the last 50 years ( 6.3 million in the first set of projections and 6.9 million in the second). For men, the absolute number of lives lost is slightly higher: With the assumption based on record life-expactancy it equals 6.4 million. Once the death rates from best-practice surfaces were applied, the number of lives, which could have been saved for men, equalled 7.4 million. For both sexes, the relative difference between the size of the actual population and the one projected with best practice mortality grows with age (for example, for the age-distribution according to the second set of projections compare Table 1). However, the distribution of the number of lives saved below 65 and at age 65 and older differs between the sexes. The number of women aged 0-64 who could be alive equals 2.6 million in the projection based on rates from the record life-expectancy ( 2.7 million in case of the second assumption) and those at age 65 and older 3.7 million ( 4.2 million). This means that $41 \%$ ( $39 \%$ ) of women, who could be alive given the best-practice mortality conditions, are below age 65. Among men, approximately the reciprocal distribution by age can be observed. Given the assumption that mortality patterns in the last half a century in the United States equalled the one resulting in record life-expectancy, $33 \%$ of those who could be alive in January 1st 2004 were at ages 65 and higher ( 4.2 million aged $0-64$ and 2.1 aged 65 and older). Under the assumption of best-practice death rates $41 \%$ of the loss of men's lives would be at ages 65 and higher ( 4.4 million aged 0-64 and 3 million men aged 65 and older). In both sets of projections, irrespective of the assumption on the optimal pattern of mortality, the number of lives that could have been saved under age 65 is similar for both sexes. The different assumptions in the two projections influence the number of lives lost at the ages 65 and above.
The next part of the analysis refers to ages 0-49 (reproductive ages or below). As the difference between the number of people whose lives could have been saved and who in 2004 would be below age 65 was similar in the two sets of projections, we decided to present below only the results from one of them. These estimations were made based on the assumption that mortality was as low as in the best-practice mortality surfaces. Among women whose lives could have been saved during the last 50 years, 1.9 million
women would be at reproductive ages or below. The number of female lives aged 0-49 that were actually saved due to the observed improvements in mortality since 1954 is of similar magnitude according to our estimations ( 2.2 million, result not shown in tables). The loss of women at reproductive ages generates large number of lives lost in the second and successive generations. These are the babies which were not born because their mothers and grandmothers died before giving birth to them.
In Table 2, we compare our (hypothetical) population with the US population on January 1 st 2004 in 5-year age-groups, separately for direct (1st generation) and indirect (2nd and higher generations) lives lost for ages below 50. Age 50 was chosen not because it could mark the end of the reproductive period of women but rather because we only made a projection for 50 years. The cumulative effect of lives lost in the second or higher generation can therefore have no impact on higher ages. Additionally, we provide information on the proportion of indirect lives lost vs. all lives lost.
The total number of lives lost as of January 1st 2004, because their mothers, grandmothers, etc., died before giving birth to them, equals 1.6 million which is approximately one third of all lives lost in the respective age range (0-49). As a result of the reproductive loss, over the last 50 years, the 2 nd+ generation female lives lost accumulated to almost 760 thousand. It means that $39 \%$ of all female lives lost at the reproductive ages and below in 2004 are those whose mothers and grandmothers died before giving birth to them. While the proportion of male lives lost in the second and higher generations is smaller ( $29 \%$ ), the absolute number of 855 thousand is larger by about 100 thousand than among females. One of the reasons, for this absolute difference is that the sex-ratio at birth is in favor of male babies. The relative difference, however, might be smaller as at the age 50 and below the number of males is larger than the number of females.
As demonstrated above, the number of lives lost accumulates fast with every new generation. This brings us to the question concerning the future consequences of the mortality disadvantage at the young ages in the United States. For example, in 2004 alone - taking into account only those women in reproductive ages whose lives could have been saved and applying age-specific fertility rates for 2004 (the 2004 Total Fertility Rate was redistributed to age-specific fertility rates by applying the 2003 distribution of rates) - these women would give birth to more than 68 thousand children, which makes almost $2 \%$ of all births that actually occurred in that year (National Center for Health Statistics, various years).

## 4 Summary

In this study we estimated the number of Americans who could be alive at the beginning of the XXI century, given the mortality level in the preceding fifty years was as low as its best-practice values. In order to achieve this result, we assumed that fertility and migration would remain at their real level over the period of the projection and hence this new health regime would not have an effect on fertility rates or net migration. Two sets of projections were estimated with different assumptions concerning the best-practice mortality. In the first set, in each calendar year we applied those death rates that constituted the record life-expectancy for that year. As the second, the best-practice mortality was estimated from smoothed surfaces of 15 low mortality countries, for each single year of age and calendar year, separately for men and women.
The disadvantage of the United States, as far as the absolute difference in death rates is concerned, concentrates at older ages, however the mortality is much higher in relative terms at younger ages. These higher death rates below age 65 are of crucial significance for the results of this study: women who die before or during the reproductive ages do not give births to as many children as they would do otherwise. In return, these babies who have not been born can also not have children. This is what we called indirect deaths or lives lost in the second and higher generations. This effect accumulates very quickly in successive generations.

As a result of the actual mortality improvement in the United States in the last half a century, there were 15.5 million people in 2004 who otherwise would not be alive. The mortality improvement in the United States in the last half a century, however, was only half as effective as the improvement in selected other developed countries, if we consider together the number of lives saved and those who could have been saved. In the same year 2004, the number of people whose lives could have been saved, given the mortality level in the United States was as low as the best-practice, amounted to 12.8 million in the estimations based on death rates from the record life-expectancy, and 14.3 million in the part applying the best-practice surfaces of mortality. While for men, most of the lives which could be saved are below age 65 , for women the opposite is true. The difference in the assumption concerning the best-practice mortality had an influence on the number of saved lives aged 65-100 in 2004, and not on the younger age groups. Despite the fact that the majority of female lives could have been saved above age 65, there was a significant loss of female lives at reproductive ages, leading to 1.6 indirect lives lost. It means that 1.6 million children could have been born and survived to January 2004 if female death rates at ages below the upper reproductive age were as low as the best practice value.

## References

Boyle Torrey, B. and C. Haub (2004). A comparison of US and Canadian mortality in 1998. Population and Development Review 30, 519-530.

Currie, I., M. Durban, and P. H. C. Eilers (2004). Smoothing and forecasting mortality rates. Statistical Modelling 4, 279-298.

Eilers, P. H. C. and B. D. Marx (1996). Flexible smoothing with b-splines and penalties. Statistical Science 11, 89-121.

Human Mortality Database (2007). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org (data downloaded on 08.01.2007).

Manton, K. G. and J. W. Vaupel (1995). Survival after age 80 in the United states, Sweden, France, England, and Japan. New England Journal of Medicine 333, 1232-1235.

Meslé, F. and J. Vallin (2006). Diverging Trends in Female Old-Age Mortality: The United States and the Netherlands versus France and Japan. Population and Development Review 32, 123-145.

National Center for Health Statistics (2007). Vital Statistics of the United States, Part I: Natality. Washington, D.C.: Government Printing Office, Various Years.

Oeppen, J. and J. W. Vaupel (2002). Broken Limits to Life Expectancy. Science 296, 10291031.

Preston, S. H. (2005). Deadweight?- the influence of obesity on longevity. The New England Journal of Medicine 352, 1135-1137.

Preston, S. H., P. Heuveline, and M. Guillot (2001). Demography. Measuring and Modelling Population Processes. Blackwell Publishing.

Shryock, H. S. and J. S. Siegel (1976). The Methods and Materials of Demography. Academic Press, San Diego.

Smith, S. K., J. Tayman, and D. A. Swanson (2002). State and Local Population Projections. The Springer Series on Demographic Methods and Population Analysis, Springer, Netherlands.

White, K. M. and S. H. Preston (1996). How many americans are alive because of twentieth-centrury improvements in mortality? Population and Development Review 22, 415-429.

Figure 1: Female and Male Life-expectancy at Birth, United States and Best-practice, 19542003


Source:HMD

Figure 2: Ratio of USA Death Rates (smoothed) to the Lowest Rates from Best Practice Surface, Females


Source: Authors' Estimations Based on Data Derived from the HMD

Figure 3: Ratio of USA Death Rates (smoothed) to the Lowest Rates from Best Practice Surface, Males


Source: Authors' Estimations Based on Data Derived from the HMD
Table 1: Hypothetical and Actual US Population in 2004 (in thousands): Estimated with Mortality from Best Practice Surface

| $\left\|\begin{array}{l} \vec{\pi} \\ \hat{0} \end{array}\right\|$ |  |  <br>  <br>  <br>  <br>  <br>  |  |
| :---: | :---: | :---: | :---: |
| $\begin{array}{\|c\|c\|c\|} \frac{0}{\pi} \\ n^{\pi} \end{array}$ |  |  - ri ri ri ri ri ri ri ri ri ri ri ri -i ず <br>  <br>  <br>  <br>  |  |
|  |  |  <br>  <br>  <br>  <br> N <br>  <br>  |  |
|  |  |  <br>  | $\stackrel{+}{\circ}$ |

Table 2: Lives lost as of January 1st 2004, All Deaths and 2nd and Further Generation Deaths in Reproductive Ages


A Appendix
Table 3: Comparison of projected US population in 2004 (in thousands) with mortality from best practice surface with two

$\ddagger$ Seq. 2: Deaths and Migration evenly spread throughout the year.


[^0]:    *Corresponding Author; Duke University; Terry Sanford Institute of Public Policy; Population, Policy, and Aging Research Center; Box 90309; (for express mail, add: 302 Towerview Road); Durham, NC 277080309; USA; Email: mm135@duke. edu
    ${ }^{\dagger}$ roland.rau@duke.edu

