ISSN 1440-771X

MONASH University

Department of Econometrics and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Forecasting age-related changes in breast cancer mortality among white and black US women: A functional data approach

Farah Yasmeen, Rob J Hyndman and Bircan Erbas

April 2010

Working Paper 09/10

Forecasting age-related changes in breast cancer mortality among white and black US women: A functional data approach

Farah Yasmeen
Department of Econometrics and Business Statistics
Monash University, VIC 3800, Australia.
Email: fyas2@student.monash.edu.au

Rob J Hyndman
Department of Econometrics and Business Statistics,
Monash University, VIC 3800, Australia.
Email: Rob.Hyndman@buseco.monash.edu.au

Bircan Erbas*
Department of Health Sciences,
School of Public Health,
La Trobe University, VIC 3800
Australia.
Email:B.Erbas@latrobe.edu.au
Tel: +61 3 9479 5657
Fax: +61 3 9479 1783

*Corresponding author

22 April 2010

JEL classification: C14,C23,J11

Forecasting age-related changes in breast cancer mortality among white and black US women: A functional data approach

Abstract

The disparity in breast cancer mortality rates among white and black US women is widening, with higher mortality rates among black women. We apply functional time series models on age-specific breast cancer mortality rates for each group of women, and forecast their mortality curves using exponential smoothing state-space models with damping.

The data were obtained from the Surveillance, Epidemiology and End Results (SEER) program of the US (SEER, 2007). Mortality data were obtained from the National Centre for Health Statistics (NCHS) available on the SEER*Stat database. We use annual unadjusted breast cancer mortality rates from 1969 to 2004 in 5-year age groups (45–49, 50–54, 55–59, 60–64, 65–69, 70–74, 75–79, 80–84). Age-specific mortality curves were obtained using nonparametric smoothing methods. The curves are then decomposed using functional principal components and we fit functional time series models with four basis functions for each population separately. The curves from each population are forecast and prediction intervals are calculated.

Twenty-year forecasts indicate an over-all decline in future breast cancer mortality rates for both groups of women. This decline is steeper among white women aged 55–73 and black women aged 60–84. For black women under 55 years of age, the forecast rates are relatively stable indicating no significant change in future breast cancer mortality rates among young black women in the next 20 years.

Keywords: Breast cancer mortality, racial and ethnic disparities, screening, trends, forecasting, functional data analysis

1 Introduction

Breast Cancer is the most common cancer among women in the United States excluding skin cancers (American Cancer Society, 2007). It accounts for about 33% of all cancers diagnosed in women (Smigal, Jemal, Ward, Cokkinides, Smith, Howe, and Thun, 2006). After cancer of the lung, it is the second leading cause of cancer deaths (Jemal, Thomas, Murray, and Thun, 2002) with approximately 40 480 women expected to die in 2008 (www.cancer.gov). Hence, it is necessary and important to get accurate projections of age-specific breast cancer mortality rates. These projections are of major public health interest. Their interpretations, however, may be complex because they reflect the combined effects of screening usage, trends in risk factors and effectiveness of treatments (Cheverley and White, 1997).

In the United States there are striking differences in breast cancer mortality in black and white women and in other ethnicities such as Hispanics, but there is very little historical data on these populations (see www.seer.cancer.gov). These differences are causing a long term divergence in mortality trends between African American and white women. Many studies are showing that the disparity in breast cancer death rates between whites and blacks continues to grow with mortality rates higher among black women compared to white women (see, e.g., Joslyn, 2002; Marbella and Layde, 2001; Tarone, Chu, and Gaudette, 1997).

In addition, there are major differences in breast cancer mortality in black and white women in the US by age. In a study of age- and race-specific breast cancer mortality, Cheverley and White (1997) reported an apparent decline in breast cancer mortality rates since 1992 among white women across all ages. However, there was no significant decline among white women aged 80 years and over or among black women of all ages. Smigal, Jemal, Ward, Cokkinides, Smith, Howe, and Thun (2006) found a decreasing trend in breast cancer death rates of white women at an average annual rate of 2.4% since 1990 and by 1.1% per year among black women since 1991.

However, few studies have predicted age-related changes in black/white breast cancer mortality. Jatoi, Anderson, Rao, and Devesa (2005) used age-period-cohort (APC) models based on Poisson regression and observed the changes occurring in mortality trends among white and black women in the United States according to age, calendar period (year of death) and birth cohort (year of birth). They observed that, for both races, the overall mortality rates increased with age, but black women have higher rates at ages younger than 57.

Tyczynski, Hill, and Berkel (2006) observed breast cancer mortality rates by race and age in Ohio and reported an overall decline in mortality since 1990. They used joinpoint regression

and an APC model to evaluate temporal changes in mortality. Hanayama (2007) proposed an extended version of the APC model for analysing (age, period) tabulated data on mortality and fitted the proposed model to breast cancer mortality data for women in the United States.

However, a limitation in these studies is that neither of them projected future breast cancer mortality rates for white and black women taking account of age-related changes. These changes need to be allowed for rather than assuming constant change in the age-mortality associations for future years.

In this paper, we use functional data analysis (Ramsay and Silverman, 2005) to model age-specific breast cancer mortality tends among white and black women in the United States. We also use these models to predict future breast cancer mortality rates by age separately for white and black women. These predictions are important for public health. To the best of our knowledge, this is the first study to show age-related predictions of black and white breast cancer mortality.

2 Methods

2.1 Data

The data were obtained from the Surveillance, Epidemiology and End Results (SEER) program of the United States (SEER, 2007). Mortality data were obtained from the National Centre for Health Statistics (NCHS) available on the SEER*Stat database. Annual age-specific breast cancer mortality data are designated by ICD 8 & 9 (1979–1998) code 174 and ICD 10 (1999+) code C50 and are available for the two racial groups, whites and blacks, and all races combined since 1969 in nineteen 5-year age-groups. However in this study, we analysed the data for eight age groups: 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, 75–79, 80–84.

2.2 Statistical Methods

To see the association between age and breast cancer mortality rates, we plot mortality-age curves. These curves are plotted by treating the annual mortality rate as a function of the mid-point of the age groups. Log mortality rates for all age groups were obtained separately for white and black women.

From the original log mortality data, the first step is to obtain smooth functions which we observe with error (Hyndman and Ullah, 2007) (see the appendix for details). These are obtained using

nonparametric smoothing methods. We treat the smooth curves as our functional observations. Hence, we have two functional time series for white and black women.

To analyse these curves, we use functional time series models introduced by Hyndman and Ullah (2007). They decomposed the smoothed curves via a basis function expansion using functional principal components (Ramsay and Silverman, 2005). This approach gives a small number of basis functions (functions of age) and gives time series coefficients which are uncorrelated with each other. We follow the same methodology as Erbas, Hyndman, and Gertig (2007) and use functional regression models separately for white and black women. An adequate fit for both white and black women is obtained using a functional regression model with four basis functions.

The time series coefficients are forecast using univariate time series models. Here we use exponential smoothing state-space models based on Hyndman, Koehler, Ord, and Snyder (2008), which provide a statistical framework for automatic forecasting. These forecasts are then multiplied with the estimated basis functions resulting in forecasts of the entire mortality age curve.

This technique also provides computation of prediction intervals for the forecasts using the estimated variance of error terms. Finally, we construct prediction intervals around our predictions assuming the forecast errors are normally distributed.

We estimated twenty-year predictions of breast cancer mortality rates for white and black women in the United States. All statistical analyses were performed in R version 2.10.1.

3 Results

Figure 1 displays age-specific mortality rates for breast cancer among white and black US women during the study period (1969–2004). These graphs show that breast cancer mortality rates among white women under 55 years declined slowly from 1969 and then quite sharply after 1990. A similar decline started in 1988 for age-group 55–59.

The patterns for older white women (aged 60–79 years) are slightly different. The rates first increased from 1970 to 1990 and then started to decline very sharply after that period. However, for women aged 80–84 years, the rates increased from 1970, stabilized between 1988–1994 and then declined after 1994.

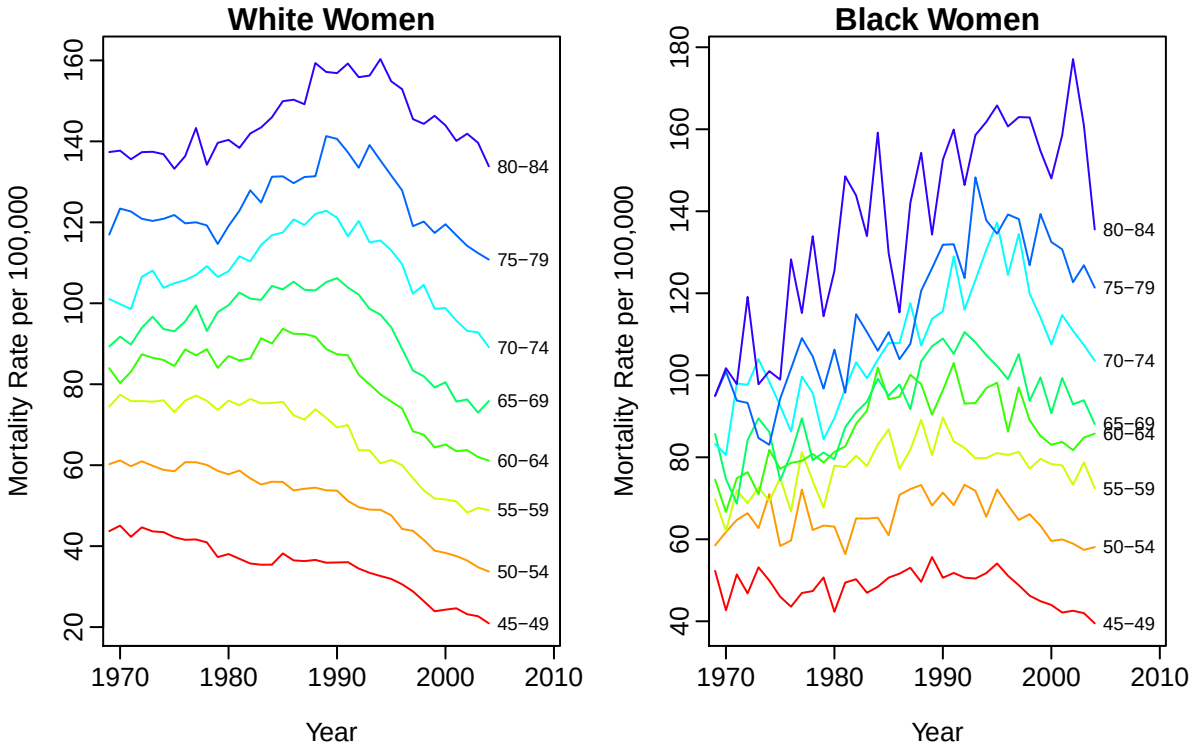

Figure 1: *Breast cancer mortality rates by age-groups among white and black US women (1969–2004).*

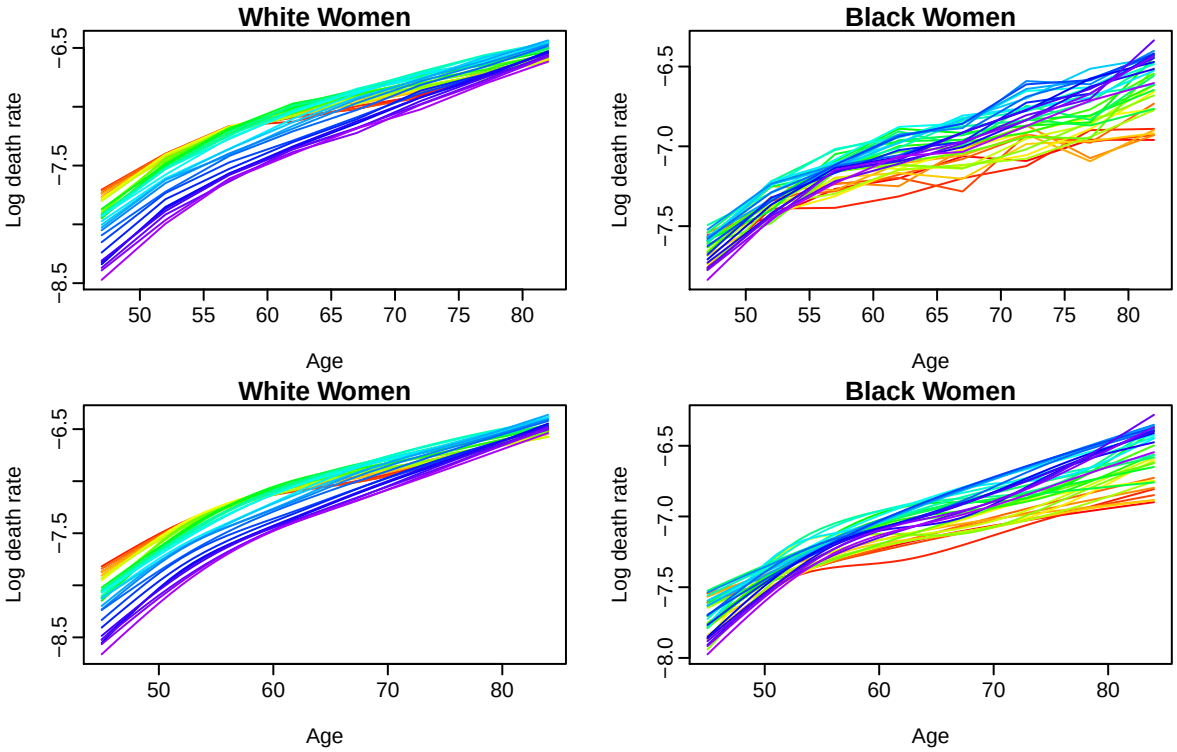

Figure 2: *Log mortality rates for white and black US women (1969–2004). In the upper panel original log rates are plotted whereas in the lower panel, log rates after smoothing are plotted. Penalized regression splines are used to smooth the data. The curves are ordered in time using rainbow colors. The oldest data (from 1969) is shown in red, with the most recent data (from 2004) shown in purple.*

The patterns for breast cancer mortality rates among black women are more variable, especially for women aged 65 years and above. Among African Americans, the rates slightly increased for women aged 45–54 years and then started to decline from 1994. However, for age-group 55–59, the decline started from 1990.

For black women of middle age (60–69) years, breast cancer mortality rates first increased very sharply after 1970 and then declined from 1990. There is a clear and sharp increasing trend for black women aged 70 years and above until 1992. The rates started to decline from 1996 for age group 70–74 years and since 1992 for women aged 75–79 years. However, the mortality rates for black women aged 80–84 years have high variability compared to the other age groups, and have declined since 2000.

The first step in using functional time series models is to obtain smoothed log mortality rates. We use penalized regression spline (Wood, 1994) for smoothing. The observed log mortality rates and smoothed rates are given in Figure 2 using rainbow plots (Hyndman and Shang, 2010). The first row of Figure 2 shows exactly the same data as in Figure 1 but as functional data instead of age-specific time series.

By looking at these graphs, we can see that before 1990, the mortality rates among black women under 55 years of age were higher than those for the white women of same age. However for older black women (65 and above) the rates were much smaller than the rates for white women.

For whites, the first four basis functions in the functional model explain 86.6%, 10.2%, 2.9% and 0.2% of the total variability in mortality rates, while in blacks they explain 77.6%, 14.3%, 5.4% and 1.2% of the total variability. Hence only four basis functions are sufficient to explain about 99% of the total variability in breast cancer mortality rates of white and black women.

Figure 3 depicts the plots of the first basis functions and corresponding coefficients for the models for each race. It is interesting to note that the first basis function for white women decreases with age while the corresponding function for black women increases with age. That is, the major changes in breast cancer rates amongst white women occur at the younger ages, while the major changes amongst black women occur at the older ages. The plot of the first time series coefficient for white women shows that the mortality rates were almost unchanging from 1969 to 1988, and have since decreased. In contrast, the first coefficient for black women shows that breast cancer mortality rates increased from 1969 to about 1992, were relatively stable between 1992-1996, and have started to decline since 1997. We do not attempt to interpret the other basis functions as they involve second-and higher-order effects.

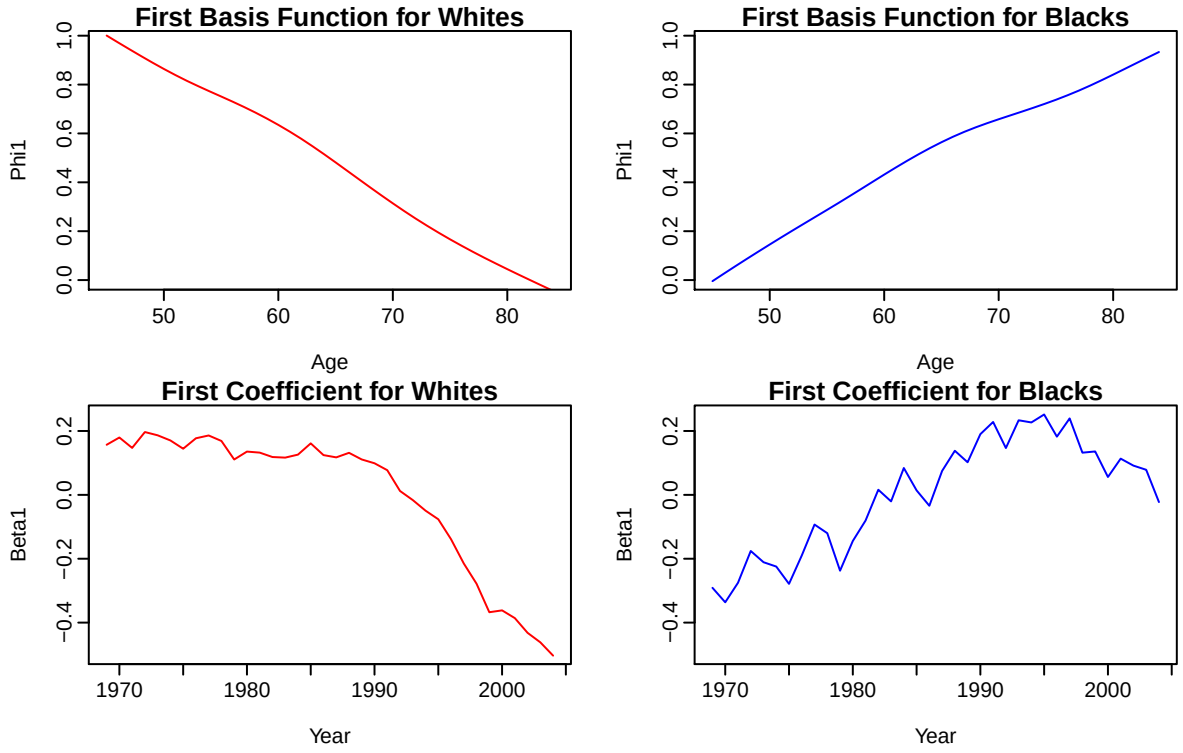

Figure 3: *First basis functions and first coefficients for whites and blacks with red color is used for whites and blue color for blacks.*

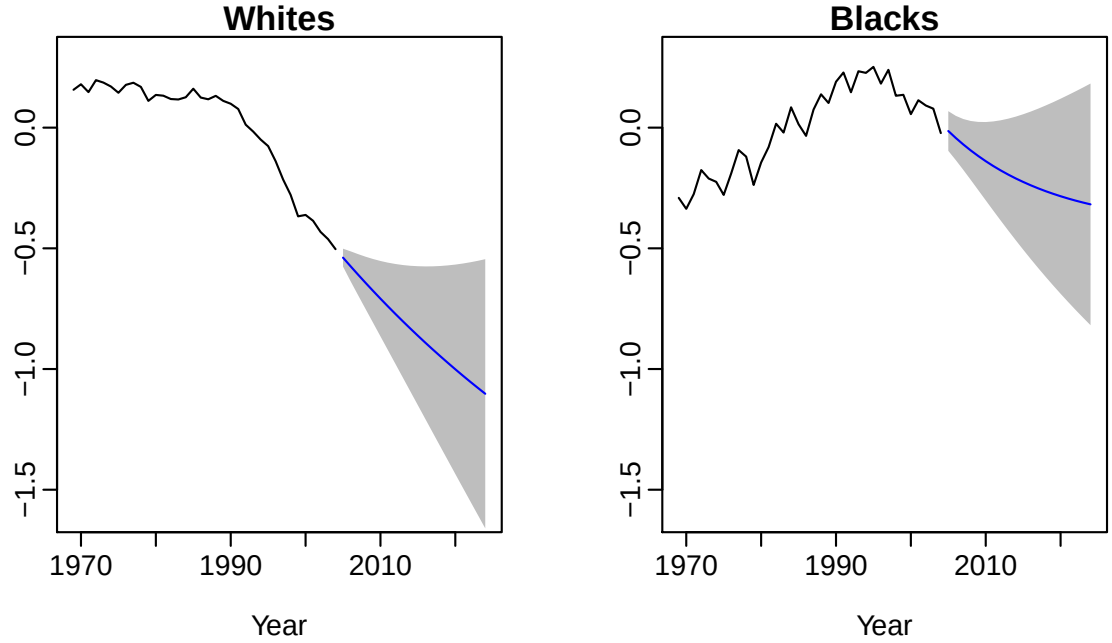

Figure 4: *20-year mortality forecasts for the first coefficients of each model.*

We obtain 20-year predictions of the future mortality rates using exponential smoothing state-space models of Hyndman, Koehler, Ord, and Snyder (2008) with damping. Figure 4 displays the predictions of the first coefficient of future mortality rates. The forecasts are plotted separately for white and black women along with 80% prediction intervals. From these graphs, it is clear that the future mortality trends are decreasing for both white and black women.

By multiplying the forecasts of the coefficients with the basis functions and summing the results, we obtain forecasts of the entire mortality curves. Figures 5 and 6 represent 20-year forecasts for the entire curves. These plots exhibit that the future mortality rates will decline for all ages, but that the decline will be greatest for the oldest black women, and smallest for the youngest black women. The changes for white women are more similar across the age range.

In Figures 7 and 8, we convert the functional forecasts back into age-group forecasts for comparison with the historical rates.

To check the forecast uncertainty, we also plot the mortality forecasts for forecast horizons of one and twenty years, along with 80% prediction intervals for the two races. These are given in figures 9 and 10 respectively. The estimated predictions have narrow bands for white women and young black women under 60 years. However, they are much wider for older black women due to more variability among the historical mortality rates. Some forecasts at selected ages are shown in Table 1, along with 80% prediction intervals.

	White Women				**Black Women**			
Age	Forecast 2005		Forecast2024		Forecast 2005		Forecast 2024	
	Mean	80% PI	Mean	80% PI	Mean	80% PI	Mean	80% PI
45	16.6	[15.5, 17.8]	8.7	[4.9, 15.6]	35.3	[30.4, 41.0]	33.7	[21.1, 53.6]
50	27.6	[26.7, 28.4]	16.9	[12.8, 22.3]	50.8	[47.9, 53.9]	47.2	[37.6, 59.1]
55	41.6	[40.5, 42.7]	28.2	[22.3, 35.9]	67.0	[62.9, 71.3]	60.5	[49.7, 73.5]
60	55.2	[53.5, 56.9]	39.9	[30.7, 51.8]	77.6	[72.8, 82.8]	67.4	[52.9, 85.8]
65	67.8	[66.0, 69.6]	52.1	[41.4, 65.5]	84.5	[79.1, 90.3]	70.4	[51.6, 95.8]
70	82.3	[80.1, 84.6]	67.2	[53.6, 84.2]	95.3	[89.7, 101.3]	77.5	[57.2, 105.0]
75	100.8	[98.2, 103.5]	87.1	[70.8, 107.1]	113.4	[106.7, 120.6]	91.2	[67.8, 122.7]
80	123.4	[120.6, 126.4]	112.9	[94.5, 134.9]	136.8	[126.0, 148.5]	107.7	[70.8, 163.9]

Table 1: *Forecasts and 80% prediction intervals of mortality rates for selected ages in 2005 and 2024*

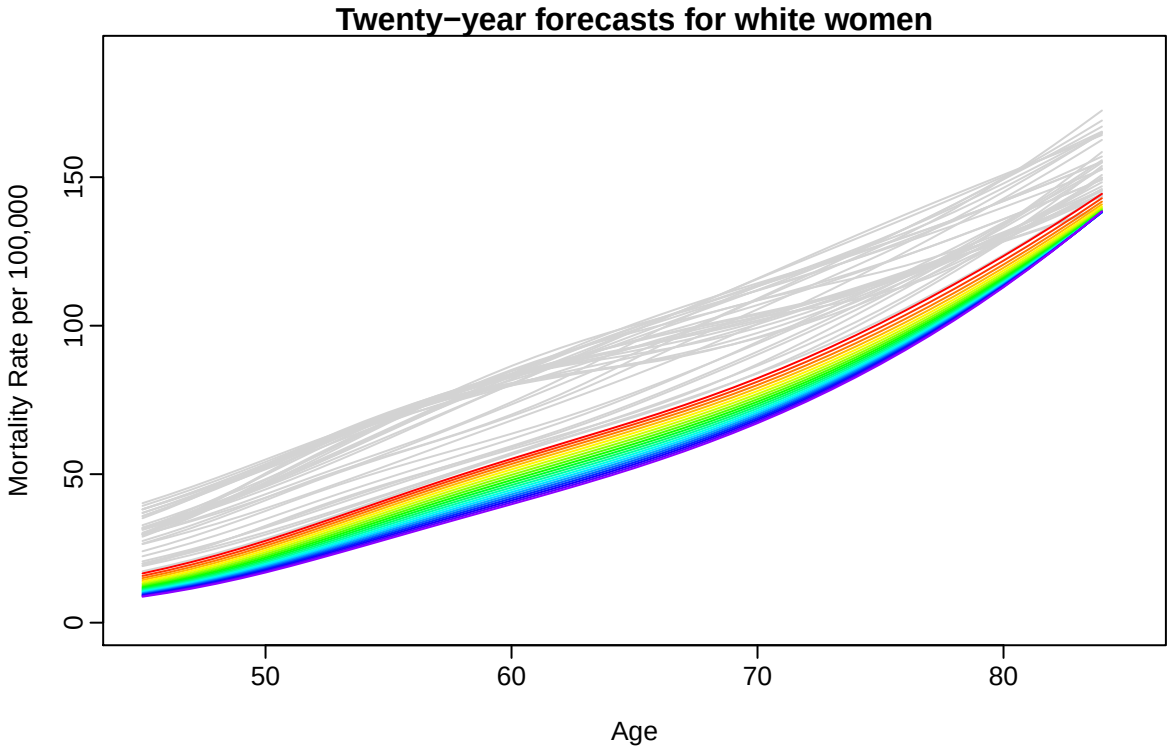

Figure 5: *20-year forecasts (2005–2024) for breast cancer mortality rates for white women. The years are plotted in "rainbow" order with the later years in violet.*

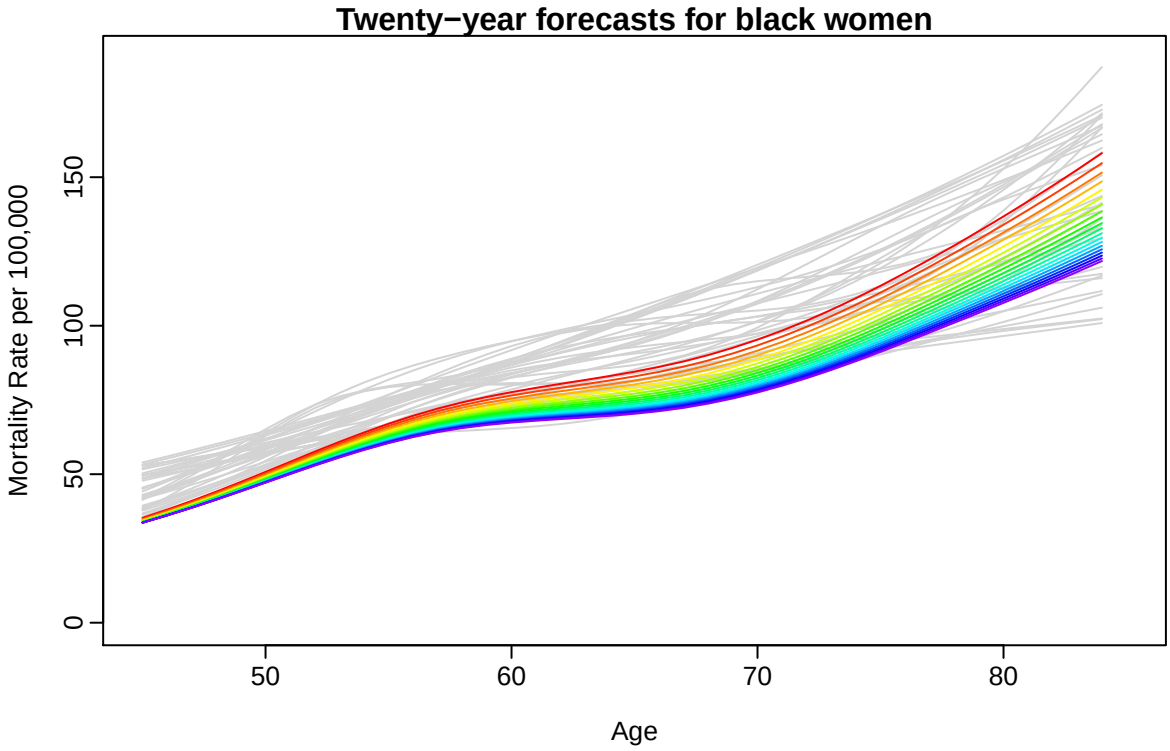

Figure 6: *20-year forecasts (2005–2024) for breast cancer mortality rates for black women. The years are plotted in "rainbow" order with the later years in violet.*

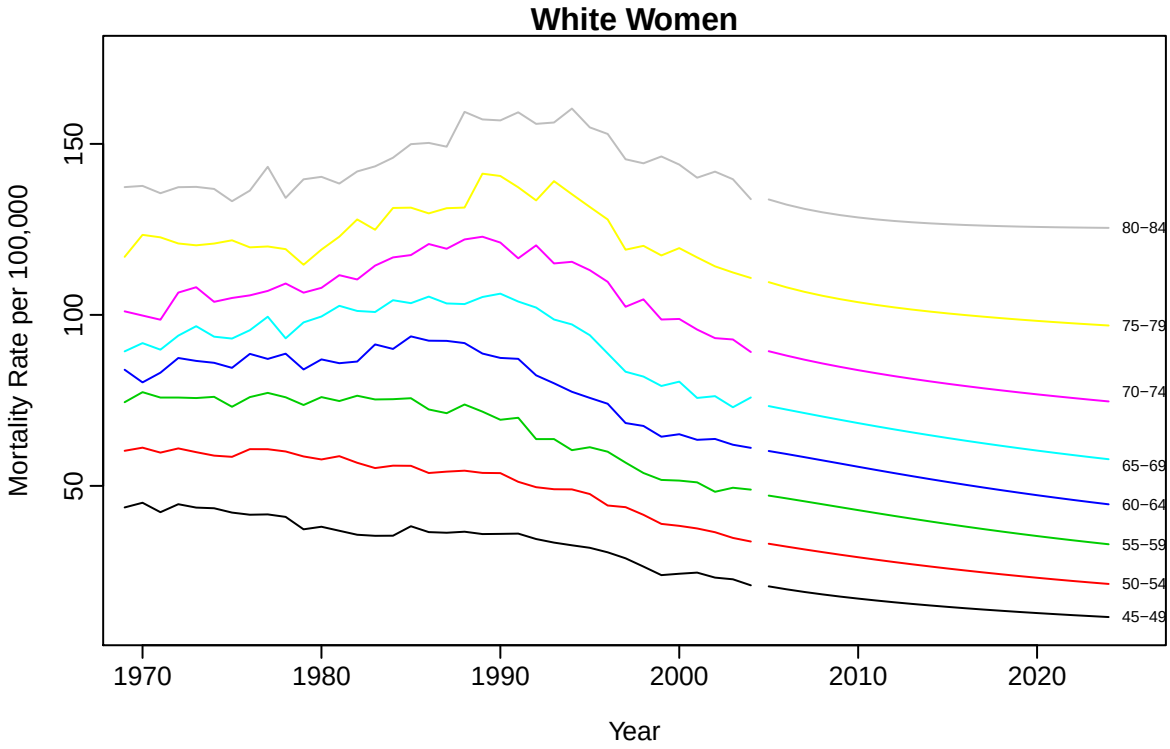

Figure 7: *20-year forecasts (2005–2024) for breast cancer mortality rates for white women by age-group.*

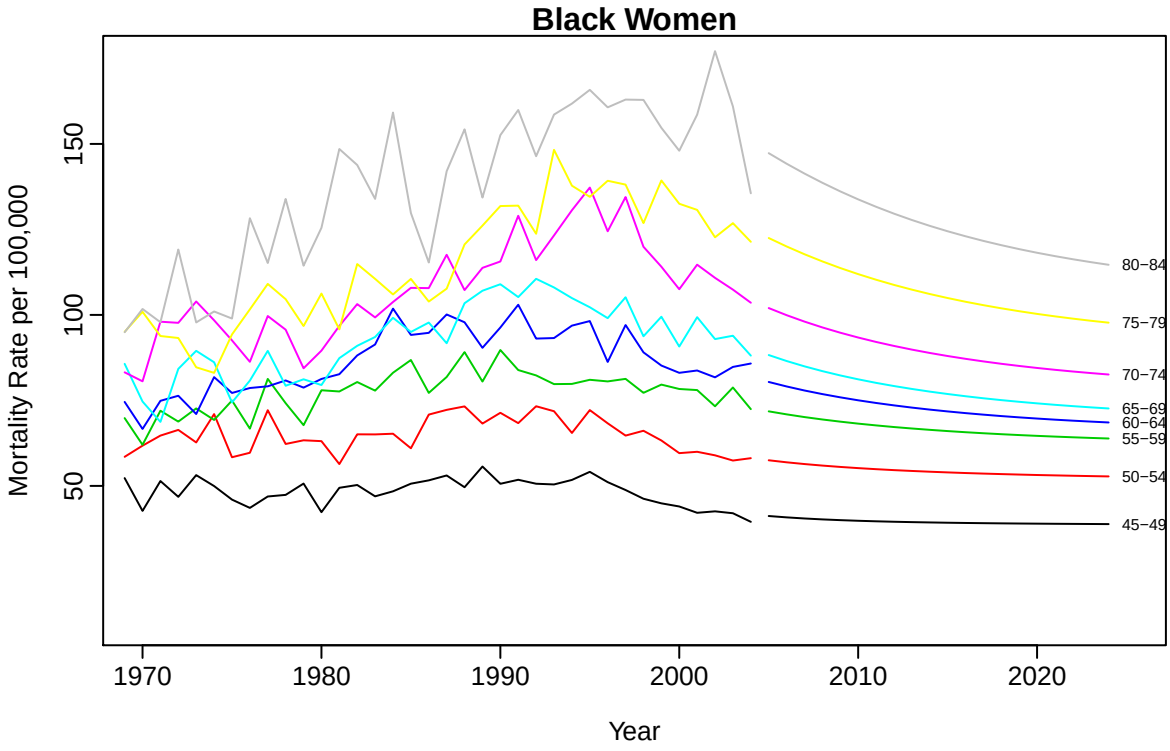

Figure 8: *20-year forecasts (2005–2024) for breast cancer mortality rates for black women by age-group.*

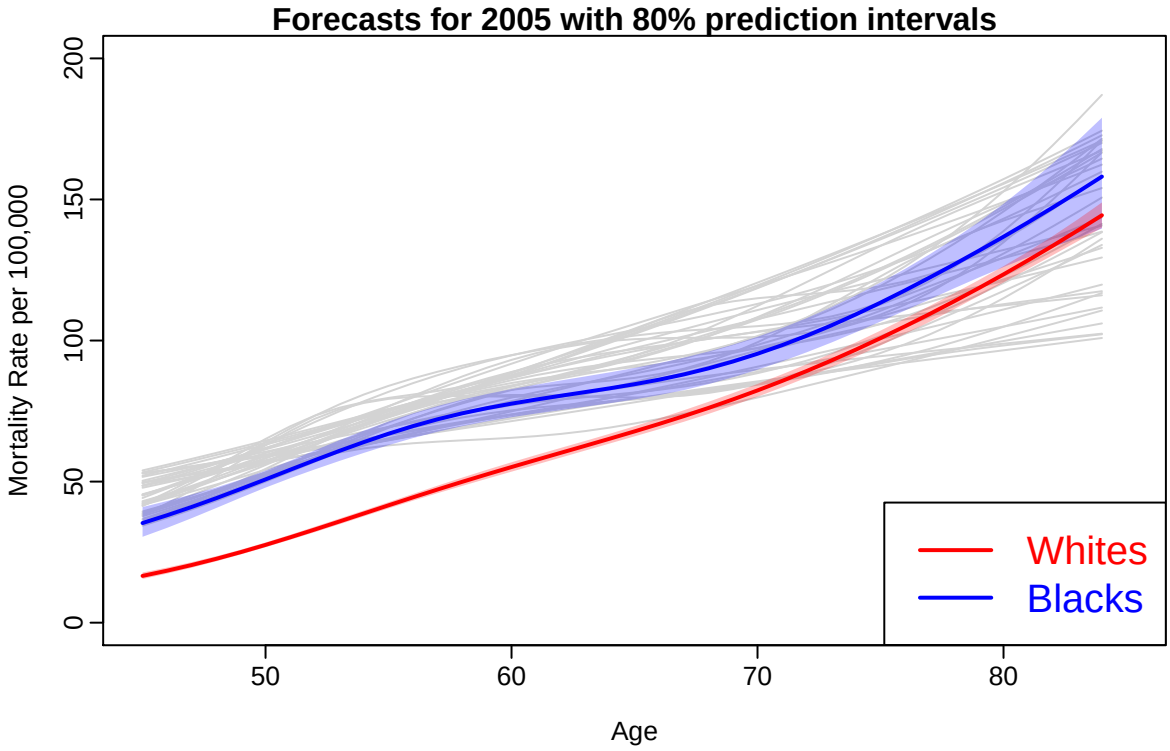

Figure 9: *One-year forecast (for 2005) for breast cancer mortality rate along with 80% prediction intervals.*

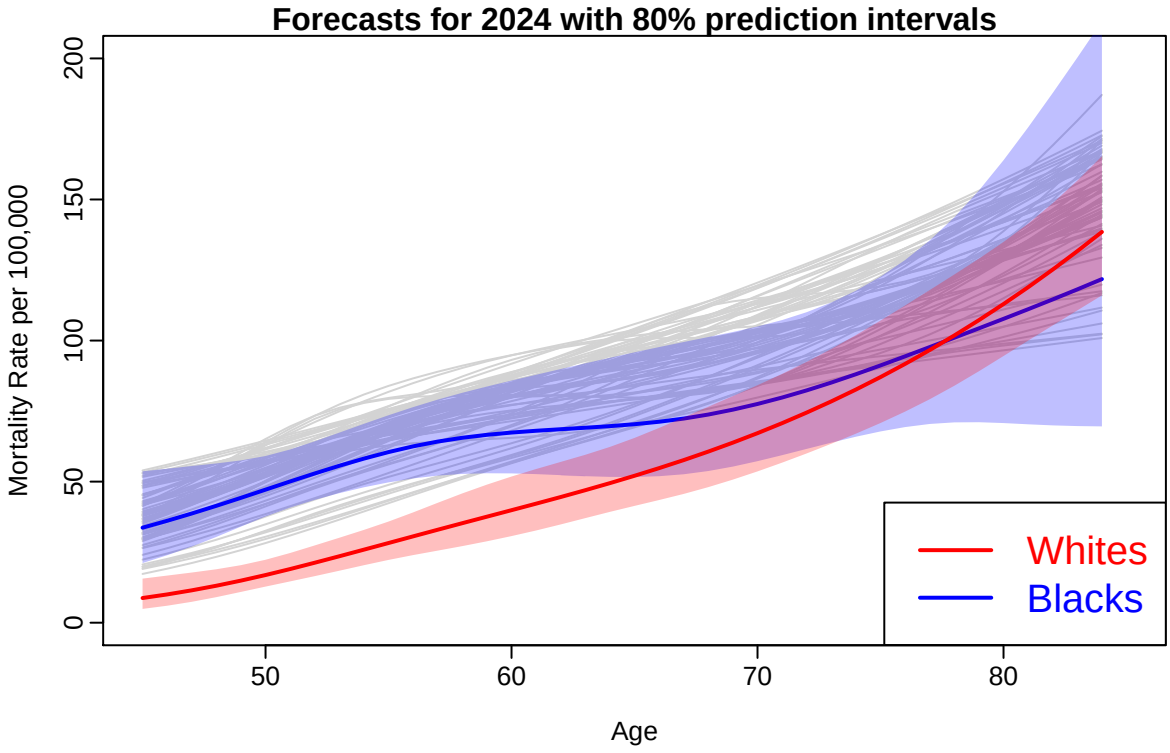

Figure 10: *20-year forecast (for 2024) for breast cancer mortality rate along with 80% prediction intervals.*

4 Discussion

In this paper, we have presented the functional forecasting method for age-related predictions of black and white breast cancer mortality in the United States. We have shown that white women have smooth and consistent patterns in breast cancer mortality rates for all age-groups whereas the mortality curves for black women are much more variable. During the most recent years, mortality rates for black women have been higher than those from white women for all age groups under 80 years of age. Our analysis suggests black American women do not bene.t equally from the overall decline in breast cancer mortality in the United States.

Our models show that the most important source of variability in mortality rates for white women is in the younger ages, while black women show greatest variability amongst older women. All breast cancer mortality rates are expected to decline. The greatest rate of decline in white women is expected for ages between 55–74 years (see Figures 5 and 6). For black women, the plot of the forecasts shows that there the rates will decline very slowly for women aged under 60 years. However, they will decline significantly for the older black women (age 70 and above).

Our results are consistent with results found by other studies such as Smigal, Jemal, Ward, Cokkinides, Smith, Howe, and Thun (2006), Tyczynski, Hill, and Berkel (2006) and Chu, Tarone, and Brawley (1999) in the sense that they showed that the over-all breast cancer mortality rates in US have declined in the last two decades. This decrease is due to the advances in detection and diagnosis and improvements in the treatments of breast cancer (Canto, Anderson, and Brawley, 2001; Cheverley and White, 1997; Chu, Tarone, and Brawley, 1999). However, the mortality rates for white and black women are significantly different. This disparity started in the 1990s as a result of a sharp improvement in mortality among white women with no improvement for black women (Hirschman, Whitman, and Ansell, 2007). Studies have also shown that there are racial disparities in almost every step of the breast cancer process, including detection, treatment and survival (Chelbowski, Chen, Anderson, Rohan, Aragasi, Lane, Dolan, Paskett, McTiernan, Hubbel, Adams-Campbell, and Prentice, 2005; Li, Malone, and Daling, 2003).

Tarone, Chu, and Gaudette (1997) found an increase in mortality calendar period slope in the 1980s compared with the 1970s for all women, and then a decreasing mortality trend from 1991 only for white women. During the 1980s, the larger increase in the mortality calendar period slope for black women could reflect the higher stage at diagnosis for them (Tarone, Chu, and Gaudette, 1997). The significant decrease during the 1990s in calendar period mortality slope for US white women supports the conclusion that this decrease in mortality is due to

medical interventions (Chu, Tarone, Kessler, Ries, Hankey, and Miller, 1996). During this period, similar decreases have been observed in UK (Quinn and Allen, 1995) and Canada (Chu, Tarone, Kessler, Ries, Hankey, and Miller, 1996). This may be due to increased use of adjuvant therapy, particularly tamoxifen therapy, and the use of early detection (Briele, Walker, Wild, Wood, Greager, Schneebaum, Silva-Lopez, Han, Gunter, Tapas, and Gupta, 1990; Quinn and Allen, 1995). A large percentage of young white women (age 20 to 59) years were obtaining regular clinical breast cancer examinations (Makuc, Fried, and Klienman, 1989) and that may have contributed to long term decline of breast cancer mortality in this group (Foster, Worden, Costanza, and Solomon, 1992).

There are many reasons for the slower decline in mortality rates for black women including racial differences in mammography use. The reasons for this disparity may be differential access to mammography among black and white women, differential quality in mammography and differential access to the breast cancer treatments (Hirschman, Whitman, and Ansell, 2007). The percentages of women who have ever received a clinical breast examination or mammogram are lower in older black women than in older white women (Caplan, Wells, and Haynes, 1992). Black women are usually less likely than white women to be aware of the use of screening tests and the women with low education level and low cancer knowledge are less likely to be clinical breast exam or mammogram compliant (Harris, Miller, and Davis, 2003).

Other possible explanations for the racial trends in breast cancer mortality include socioeconomic status, stage of disease and tumor biology (Marbella and Layde, 2001). Studies showed that black women usually have lower socioeconomic status (Farley and Flannery, 1989) and are more likely to be diagnosed with advanced stage of breast cancer than white women (Beral, Hermon, Reevas, and Peto, 1995; Eley, Hill, and Chen, 1994; Elledge, Clark, Chamness, and Osbome, 1994). In breast cancer tumor characteristics, such as hormone receptor status, significant racial differences have been found so that black women exhibit more aggressive aspects of disease (Elmore, Meceri, Carter, and Larson, 1998; Gapstur, Dupuis, Gann, Collila, and Winchester, 1996). However, one study showed that despite black women tending to present with higher-stage of breast cancer disease, their tumors do not seem to be more aggressive and equal treatments may yield equal outcomes (Chu, Tarone, and Brawley, 1999). Black women have been shown to be less likely than whites to have any insurance, and are less likely to be covered by a private insurance carrier (Blendon, Aiken, Freeman, and Corey, 1989). This suggests that disparity may be due to different access to health care and more cultural and economic programs are needed to remove barriers from early detection for black women. More breast cancer education programmes are needed in order to attract the attention of black women and to educate them about screening programs (Sadler, Ko, Cohn, White, Weldon, and Wu,

2007). In another study, Newman and Alfonso (1997) found that breast cancer mortality may be high among young black women and biologically, their cancers are more aggressive among the subset of young women. They recommended that young black women should be targeted for aggressive breast cancer screening.

Women aged 40–64 years can benefit from mammography in reducing the risk of death from breast cancer (Norman, Localio, Weber, Coates, Zhou, Bernstein, Malone, Marchbanks, Weiss, Lee, and Nadel, 2007). Mammography screening is more effective in older women (aged 50–64 years) than in younger women (aged 40–49). Screening efficacy is usually greater in post-menopausal women; as the breast density affects mammographic sensitivity (Carney, Miglioretti, Yankaskas, Kerlikowske, Rosenberg, Rutter, Geller, Abraham, Taplin, Dignan, Cutter, and Ballard-Barbash, 2003) and density decreases after menopause (Stone, Gunasekara, Martin, Yaffe, Minkin, and Boyd, 2003).

Among black women, higher mortality rates may be due to lower serum 25-hydroxyvitamin D [25(OH)D] caused by lower vitamin D production rates from solar ultraviolet-B irradiance due to their darker skin (Grant, 2006). Black Americans generally have lower 25(OH)D levels than white Americans (Harkness and Cromer, 2005; Looker, Hughes, Calvo, Gunter, and Sahyoun, 2002). Several cancers (including breast cancer) behave more aggressively in blacks than whites due to this lower rate of 25(OH)D (Studzinski and Moore, 1995); this aggressiveness affects survival rates from cancer (Giovannucci, 2005).

However, some results are slightly different. Jatoi, Anderson, Rao, and Devesa (2005) observed that for both white and black women, the overall mortality rates increased with age, but black women have higher rates at ages younger than 57. They observed a decreasing trend in mortality rates for both black and white women until the late 1970s. Thereafter, a sharp decline was evident in the calendar period slope for whites, whereas blacks experienced an increase in this slope which stabilized around 1995. After that period, mortality rates decreased for black women with a different rate of change than those for whites and nonwhites. However, their analysis is based on age-adjusted (2000 US standard) mortality rates per 100,000 woman-years . The age-related trends in mortality were adjusted for period (and cohort) factors and calendar period trends were adjusted for age (and cohort) factors.

In contrast, we performed the analysis for all age groups of white and black women and found that during the most recent years, the mortality rates were higher for blacks in all age groups than the mortality rates for whites except for the oldest age group (see Figure 1). We found that the breast cancer mortality rates for the young white women (age 45–60 years) decreased very slowly from 1970 until 1990 whereas they increased for older white women (age 65 years and

above) until 1988. After that they started to decline for all ages and more sharply for the older white women.

However for black women, the rates were slightly increasing for younger women (age 45–59 years) during the study period until 1995 and then started to decline. For older black women (age 65 and above), there is a prominent increasing trend until 1992 and then the rates declined but little bit slowly as compared to white. However for women age 80 and above, the rates are too much fluctuated and declined in most recent years.

Marbella and Layde (2001) found that black women had no significant change in breast cancer mortality during the periods 1989–1992 and 1993–1996, whereas white women experienced a significant decrease during this time period. They analyzed the data from 1981 to 1996. In our analysis, we found a decline in breast cancer mortality rates for whites from 1989 and for blacks from 1996.

4.1 Strengths

Here, we are using an alternative approach to the most widely used APC models for forecasting mortality rates. We use a recently developed method for demographic forecasting by Hyndman and Ullah (2007) which makes use of all available historical data rather than only the most recent observations. This new approach has a number of advantages. At first, it allows us to model mortality rates as continuous functions of age to capture the patterns of variation between years. The data are first smoothed before estimating the basis functions and their coefficients to reduce observational error. Our approach forecasts the entire function for future time periods with prediction intervals. Our method also has the potential to incorporate important covariates such as mammographic screening and treatment effects such as hormone replacement therapy. In contrast, the APC models include strong parametric assumptions about the distribution of error and assume exact linear dependency between the three variables (age, period and cohort). Also, there are problems in identification of parameters as these models are sensitive to changes in recent patterns of birth cohorts. The extended APC model proposed by Hanayama (2007) provides a better fit to the data than traditional models, but they analysed the data for all US females. Also, they used 10-year age-groups and periods so detailed age-based discussion is not possible. In their analysis, the residuals showed systematic departures from proposed models and this suggests the model is inadequate.

4.2 Limitations

However, there are some limitations in our new method. Our models only consider calendar period (year of death) effects and do not incorporate cohort (year of birth) effects. Calendar period mortality trends reflect the impact of new medical interventions whereas birth cohort mortality trends can reflect changes in risk factors. There may be some other factors, such as screening and treatment options associated with hormone replacement therapy, that could change the entire shape of mortality curve. We are not considering these options here, but in future we aim to extend our models so cohort effects can also be incorporated.

In summary, this study utilises recent approaches in the statistical analysis of age-specific mortality rates to model future age-specific breast cancer induced mortality rates amongst two groups of American women, namely whites and blacks. These projections provide a clear indication that although overall mortality of breast cancer is set to decline, the current disparity between the two groups will increase, with some variation between age groups, thus highlighting those groups of American women most at risk of death due to breast cancer.

References

American Cancer Society (2007). *Breast Cancer Facts and Figures: 2007–2008*. Atlanta, GA.

Beral, E., C. Hermon, G Reevas, and R Peto (1995). Sudden fall in breast cancer death rates in England and Wales [Letter]. *Lancet* **345**, 1642–3.

Blendon, R. J, L. H. Aiken, H. E Freeman, and C. R. Corey (1989). Access to medical care for black and white Americans: a matter of continuing concern. *Journal of the American Medical Association* **261**, 278–81.

Briele, H., M. J. Walker, L. Wild, D. Wood, J. Greager, L. Schneebaum, E. Silva-Lopez, M. Han, T. Gunter, K. Tapas, and Gupta (1990). Results of treatment of stage I-III breast cancer in black Americans: the COOK County Hospital experience 1973-1987. *Cancer* **65**, 1062–71.

Canto, M., W. Anderson, and O. Brawley (2001). Geographic Variation in Breast Cancer Mortality for White and Black Women. *CA Cancer J Clin* **51**, 367–370.

Caplan, L. S., B. L Wells, and S. Haynes (1992). Breast Cancer Screening among Older Racial/Ethnic Minorities and Whites: Barriers to early detection. *Journal of Gerontology, Health Module* **47**, 101–110.

Carney, P. A., D. L. Miglioretti, B. C. Yankaskas, K. Kerlikowske, R. Rosenberg, C. M. Rutter, B. M. Geller, L. Abraham, S. H. Taplin, M. Dignan, G. Cutter, and R. Ballard-Barbash (2003). Individual and Combined effects of age, breast density and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med* **138**, 168–175.

Chelbowski, R., Z. Chen, G. L. Anderson, T. Rohan, A. Aragasi, D. Lane, N. C. Dolan, E. D. Paskett, A. McTiernan, F. A. Hubbel, L. Adams-Campbell, and R. Prentice (2005). Ethnicity and breast cancer: factors influencing differences in incidence and outcome. *J Natl Cancer Inst* **97**(6), 439–448.

Cheverley, F. and E White (1997). Recent trends in breast cancer mortality among White and Black US women. *Am J Public Health* **87**(5), 775–781.

Chu, K. C, R. E Tarone, and O. W Brawley (1999). Breast cancer trends of black women compared with white women. *Arch Fam Med* **8**, 521–528.

Chu, K. C., R. E Tarone, L. G. Kessler, L. A. Ries, B. F Hankey, and B. A. Miller (1996). Recent trends in US breast cancer incidence, survival and mortality rates. *J Natl Cancer Inst* **88**(6), 1571–91.

Cleveland, W. S. and S. J. Devlin (1988). Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association* **83**, 596–610.

Eley, J. W, H. A Hill, and V. W Chen (1994). Racial differences in survival from breast cancer: Results of the National Cancer institute Black/White Cancer Survival Study. *Journal of the American Medical Association* **272**(12), 947–54.

Elledge, R. M, G. M Clark, G. C Chamness, and C. K Osbome (1994). Tumour biologic factors and breast cancer prognosis among white, Hispanic, and black women in the United States. *J Natl Cancer Inst* **86**, 705–712.

Elmore, J. G, V. M Meceri, D Carter, and E. B. Larson (1998). Breast carcinoma tumor characteristics in Black and White women. *Cancer* **83**(12), 2509–2515.

Erbas, B., R. J. Hyndman, and D. M. Gertig (2007). Forecasting age-specific breast cancer mortality using functional data models. *Statistics in Medicine* **26**, 458–470.

Farley, T. A and J. T Flannery (1989). Late-stage diagnosis of breast cancer in women of lower socioeconomic status: public health implications. *Am J Public Health* **79**(11), 1508–12.

Foster, R. S, J. K Worden, M. C Costanza, and L. J Solomon (1992). Clinical breast examination and breast self-examination, past and present effect on breast cancer survival. *Cancer* **69**, 1992–1998.

Gapstur, S. M, J Dupuis, P Gann, S Collila, and D. P. Winchester (1996). Hormone receptor status of breast tumours in Black, Hispanic, and non-Hispanic White women: an analysis of 13,239 cases. *Cancer* **77**(8), 1465–1471.

Giovannucci, E. (2005). The epidemiology of vitamin D and cancer incidence and mortality: A review (United States). *Cancer Causes and Controls* **16**, 83–95.

Grant, W. B. (2006). Lower vitamin-D production from solar ultraviolet-B irradiance may explain some differences in cancer survival rates. *J Natl Med Assoc* **98**(3), 357–64.

Hanayama, N. (2007). An extended age period cohort model for analysing (age, period)-tabulated data. *Stat Med.* **26**(18), 3459–75.

Harkness, L. S and B. A Cromer (2005). Vitamin D deficiency in adolescent females. *Journal of adolescent health* **37**, 75.e1–75.e5.

Harris, D. M, J. E Miller, and D. M Davis (2003). Racial differences in breast cancer screening, knowledge and compliance. *J Natl Med Assoc* **95**(8), 693–701.

Hirschman, J., S. Whitman, and D. Ansell (2007). The Black:White Disparity in Breast Cancer Mortality: The Example of Chicago. *Cancer Causes Control* **18**(3), 323–333.

Hyndman, R. J and H. L Shang (2010). Rainbow plots, bagplots and boxplots for functional data. *Journal of Computational and Graphical Statistics* **19** (1), 29–45.

Hyndman, R. J and M. S. Ullah (2007). Robust forecasting of mortality and fertility rates: a functional data approach. *Computational Statistics & Data Analysis* **51**, 4942 –4956.

Hyndman, R. J., A. B. Koehler, J. K. Ord, and R. D. Snyder (2008). *Forecasting with exponential smoothing: the state space approach*. Berlin: Springer-Verlag. www.exponentialsmoothing.net.

Jatoi, I., W. F. Anderson, S. R. Rao, and S. S. Devesa (2005). Breast Cancer trends Among Black and White Women in the United States. *Journal of Clinical Oncology* **23**(31), 7836–7841.

Jemal, A, A Thomas, T Murray, and M Thun (2002). Cancer Statistics 2002. *CA Cancer J Clin* **52**, 23–47.

Joslyn, S. A. (2002). Racial differences in treatment and survival from early-stage breast carcinoma. *Cancer* **95**, 1759–1766.

Li, C. L., K. E. Malone, and J. R Daling (2003). Differences in breast cancer stage; Treatment and survival by race and ethnicity. *Arch Intern Med* **163**(1), 49–56.

Looker, A. C, B. D Hughes, M. S. Calvo, E. W. Gunter, and N. R Sahyoun (2002). Serum 25-hydroxyvitamin D status of adolescents and adults in two seasonal subpopulations from NHANES III. *Bone* **30**(5), 771–777.

Makuc, D. M, V. M Fried, and J. C. Klienman (1989). National trends in the use of preventive health care by women. *Am J Public Health* **79**(1), 21–26.

Marbella, A. M. and P. M. Layde (2001). Racial Trends in Age-Specific Breast Cancer Mortality Rates in US Women. *American Journal of Public Health* **91**(1), 118–121.

Newman, L. A and A. E. Alfonso (1997). Age related differences in breast cancer stage at diagnosis between black and white patients in an urban community hospital. *Annals of Surgical Oncology* **4**(8), 655–662.

Norman, S. A, A. R Localio, A. L Weber, R. J Coates, L. Zhou, L. Bernstein, K. E. Malone, P. A. Marchbanks, L. K Weiss, N. C. Lee, and M. Nadel (2007). Protection of mammography screening against death from breast cancer in women aged 40–64 years. *Cancer Causes Control* **18**(9), 909–918.

Quinn, M. and E. Allen (1995). Changes in incidence of and mortality from breast cancer in England and Wales since introduction of screening. *BR Med J* **311**, 1391–5.

Ramsay, J. O and B. W. Silverman (2005). *Functional Data Analysis*. Springer.

Ruppert, D., M. Wand, and R. J. Carroll (2003). *Semiparametric Regression*. New York: Cambridge University Press.

SEER (2007). *Surveillance, Epidemiology, and End Results (SEER) Program . SEER*Stat Database: Mortality – All COD, Public-Use With State, Total US (1969-2004), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch*. released April 2007, Underlying mortality data provided by NCHS (www.cdc.gov/nchs). www.seer.cancer.gov.

Sadler, G. R, C. M Ko, J. A Cohn, M White, R. N. Weldon, and P Wu (2007). Breast cancer knowledge, attitudes, and screening behaviours among African American women: the Black cosmetologists promoting health program. *BMC Public Health* **17**, 7–57.

Smigal, C., A. Jemal, E. Ward, V. Cokkinides, R. Smith, H. L. Howe, and M. Thun (2006). Trends in Breast Cancer by Race and Ethnicity: Update 2006. *CA: A Cancer Journal for Clinicians* **56**, 168–183.

Stone, J, A. Gunasekara, L. J. Martin, M Yaffe, S Minkin, and N. F. Boyd (2003). The detection of change in mammographic density. *Cancer Epidemiol Biomarkers Prev* **12**, 625–630.

Studzinski, G. P. and D. C. Moore (1995). Sunlight—Can it prevent as well as cause cancer? *Cancer Research* **55**, 4014–4022.

Tarone, R. E, K. C. Chu, and L. A. Gaudette (1997). Birth Cohort and Calendar Period trends in Breast Cancer Mortality in the United States and Canada. *Journal of the National Cancer Institute* **89**(3), 251–256.

Tyczynski, J. E., T. D. Hill, and H. J. Berkel (2006). Why do Postmenopausal African-American Women not benefit from overall breast cancer mortality decline? *Ann Apidemiol* **16**(3), 180–190.

Wood, S. N. (1994). Monotonic Regression splines fitted by cross validation. *SIAM Journal of Scientific Computing* **15**(5), 1126–1133.

Appendix

Suppose $y_{k,t}(x_i)$ denotes the log mortality rate at age x_i in year t for the kth group ($k = 1, 2$). Here, white and black women are considered to be group 1 and 2 respectively. We assume that

$$y_{k,t}(x_i) = f_{k,t}(x_i) + \sigma_{k,t}(x)\varepsilon_{k,t,i}, \tag{1}$$

where $\{f_{k,t}(x)\}$ is a smooth function of x, $\sigma_{k,t}(x)$ allows the amount of noise to vary with x, and $\{\varepsilon_{k,t,i}\}$ are considered to be independent and identically distributed random variables with zero mean and unit variance.

The first step is to estimate these smooth functions from the discrete noisy data. This is possible using nonparametric smoothing methods such as penalized regression splines (Ruppert, Wand, and Carroll, 2003) or loess curves (Cleveland and Devlin, 1988).

We then decompose the smoothed curves via a basis function expansion using the following model

$$f_{k,t}(x) = \mu_k(x) + \sum_{j=1}^{J} \beta_{k,j,t}\phi_{j,k}(x) + e_{k,t}(x), \tag{2}$$

where $\mu_k(x)$ is the mean log mortality across years for the kth group and $\{\phi_{j,k}(x)\}$ is a set of orthogonal basis functions for each k. The values $\{\beta_{k,j,1}, \ldots, \beta_{k,j,n}\}$ form a univariate time series for each $k = 1, 2$ and $j = 1, \ldots, J$. The basis functions $\{\phi_{j,k}(x)\}$ are computed using functional principal components (Ramsay and Silverman, 2005) applied to the smooth curves $f_{k,t}(x)$.

The h-step ahead forecast of $\hat{y}_{n+h,k}(x)$ can be obtained as

$$\hat{y}_{n+h,k}(x) = \hat{\mu}_k(x) + \sum_{j=1}^{J} \hat{\beta}_{k,j,n+h}\hat{\phi}_{j,k}(x), \tag{3}$$

where $\hat{\mu}_k(x)$ and $\hat{\phi}_{j,k}(x)$ are the estimates of the mean function and basis functions respectively and $\hat{\beta}_{k,j,n+h}$ denotes the h-step ahead forecast of $\beta_{k,j,n+h}$ computed using a state space exponential smoothing model (Hyndman, Koehler, Ord, and Snyder, 2008).

Hyndman and Ullah (2007) showed that the forecast variance can be obtained by adding the variances of all individual terms as

$$\mathrm{Var}[\hat{y}_{n+h}(x)] = \hat{\sigma}_{\mu}^2(x) + \sum_{j=1}^{J} u_{k,j,n,h}\hat{\phi}_j^2(x) + \nu(x) + \sigma_{n+h}^2(x), \tag{4}$$

where $\hat{\sigma}_\mu^2(x)$ is the variance obtained using the smoothing method, $u_{k,j,n,h} = \text{Var}(\beta_{k,j,n+h} \mid \beta_{k,j,1},\ldots,\beta_{k,j,n})$ is the forecast variance from the time series model, $v(x)$ is obtained using the sum of squared residuals, and $\sigma_{n+h}^2(x)$ can be obtained using a Poisson approximation (see Hyndman and Ullah, 2007, for details). After substituting the values of all these terms in (4), one may obtain prediction intervals for the entire mortality curve, assuming a normal distribution for the forecast errors.