

Original Paper

# Racial Disparities in Mortality Among American Film Celebrities: A Wikipedia-Based Retrospective Cohort Study

Hannah Speaks<sup>1</sup>, MSPH; Alyssa Falise<sup>1</sup>, MSPH; Kaitlin Grosgebauer<sup>1</sup>, MSPH; Dustin Duncan<sup>2</sup>, SCD, MPH; Adam Carrico<sup>1</sup>, PhD

<sup>1</sup>University of Miami, Department of Public Health Sciences, Miami, FL, United States

<sup>2</sup>Columbia University, Mailman School of Public Health, New York City, NY, United States

**Corresponding Author:**

Adam Carrico, PhD

University of Miami

Department of Public Health Sciences

1120 NW 14th St

Office 1005

Miami, FL, 33136

United States

Phone: 1 4155338653

Email: [a.carrico@miami.edu](mailto:a.carrico@miami.edu)

## Abstract

**Background:** In the United States, well-documented racial disparities in health outcomes are frequently attributed to racial bias and socioeconomic inequalities. However, it remains unknown whether racial disparities in mortality persist among those with higher socioeconomic status (SES) and occupational prestige.

**Objective:** As the celebrity population is generally characterized by high levels of SES and occupational prestige, this study aimed to examine survival differences between black and white film celebrities.

**Methods:** Using a Web-based, open-source encyclopedia (ie, Wikipedia), data for 5829 entries of randomly selected American film actors and actresses born between 1900 and 2000 were extracted. A Kaplan-Meier survival curve was conducted using 4356 entries to compare the difference in survival by race. A Cox semiparametric regression analysis examined whether adjusting for year of birth, gender, and cause of death influenced differences in survival by race.

**Results:** Most celebrities were non-Hispanic white (3847/4352, 88.4%), male (3565/4352, 81.9%), and born in the United States (4187/4352, 96.2%). Mean age at death for black celebrities (64.1; 95% CI 60.6-67.5 years) was 6.4 years shorter than that for white celebrities (70.5; 95% CI 69.6-71.4 years;  $P < .001$ ). Black celebrities had a faster all-cause mortality rate using Kaplan-Meier survival function estimates and a log-rank test. However, in a Cox semiparametric regression, there was no longer a significant difference in survival times between black and white celebrities (hazard ratio 1.07; 95% CI 0.87-1.31).

**Conclusions:** There is some evidence that racial disparities in all-cause mortality may persist at higher levels of SES, but this association was no longer significant in adjusted analyses. Further research is needed to examine if racial disparities in mortality are diminished at higher levels of SES among more representative populations.

(*Interact J Med Res* 2019;8(4):e13871) doi: [10.2196/13871](https://doi.org/10.2196/13871)

**KEYWORDS**

continental population groups; internet; mortality; race

## Introduction

The Heckler Report, published in 1985, was the first to report statistical differences in health and mortality among races in the United States after the Civil Rights Act in 1964 [1]. Over 30 years later, there have been significant improvements to minority lifespan and health, but notable variations in mortality

rate, cause of death, and prevalence of disease have persisted [2-4]. Blacks (national born and immigrants) make up a rising 13% of the US population [5]; yet, they experience the highest mortality rates for 8 of the 10 most common causes of death in the United States. In addition, they have maintained the highest mortality rate for 8 of the 10 most common causes of death as

well as a 6-year and 4-year shorter life span than that of white American males and females, respectively [4].

Socioeconomic status (SES), as defined by income, education, occupation, and social status, is a product of policy and culture often studied in context with its effect on health [3,6-8]. Unlike fixed demographic variables such as gender, race, and birth date, SES measures such as education, geographic location, marital status, financial hardship, and social status can fluctuate throughout a lifetime and generally represent a dynamic measure of social influence and affluence in society [9,10]. Economic census data throughout the past century have established significant economic and educational disparities between blacks and whites in the United States. A study of US Panel Income Dynamics over 1997-2007 found that only 12.6% of those identifying as black had completed a college education compared with 32.7% of white counterparts, whereas 22.3% of blacks and 9.3% of whites had not attained a high school diploma [11]. Wealth among non-Hispanic, American whites has been found to be orders of magnitude greater than black Americans since the 1970s [12] and highly correlated to health and mortality outcomes [13-15]. Therefore, disparities in health outcomes and mortality are generally thought to be due to historical and enduring structural economic inequities such as redlining of predominantly black neighborhoods by lenders [3,7,10,11]. Recent studies quantifying racial health and mortality disparities that controlled for SES have found that differences among races are reduced but are still present [15-17]. However, there is a growing body of literature that measures not only income but also wealth and associated variables such as social standing.

Social status and wealth are understudied aspects of SES, especially in the American context. Wealth has been defined as the total sum of income, assets, and debts to quantify personal net worth. Hajat et al found a significant inverse relationship between wealth and mortality [18], which is also consistent with a previous systematic review [9]. Another study found that wealth, rather than income alone, was a more predictive measure of SES differences in international health care utilization [19]. In contrast, social status entails combinations of SES variables and public opinion of occupational groups. These measurements of occupational and social prestige use tools such as Duncan SES index and Siegel Prestige Score for common occupational groups in a population [20]. The combination of these scales indicates a higher risk of death for those in lower social positions, consistent with previous studies, but still did not examine moderation by race and gender within each level of social class as well as occupation [21]. Although these measures address wealth and occupational prestige, both associated with cultural notability or societal influence, they do not account for individuals known as celebrities and the racial disparities that may exist within these population subsets.

Celebrities are a subset of the population that have occupations associated with higher influence, notability, and wealth. Among celebrities, actors and actresses usually maintain societal standing and wealth because of the ongoing opportunities extended once fame is achieved. Owing to their top-tier social standing and relative affluence to the general population, the high societal influence of actors and actresses is hypothesized

to be associated with lower mortality rates and relatively good health. Redelmeier and Singh studied trends in mortality for Oscar-nominated actors and actresses and found that the primary causes of death followed the top-reported causes of death for the general population, with an average lifespan higher than that for the general population (76-79 years) [22]. Racial diversity in American film actors has also grown steadily throughout the years, with Smith et al reporting nearly 13% black actors and 73% white actors in 2014, which closely reflects the US census demographics for the decade [23].

To date, no research has been published that observes racial disparities in mortality among American celebrities. This study sampled Wikipedia to examine racial disparities in a retrospective cohort of actors and actresses born in the 20th century. Choosing actors and actresses born in this time frame includes those that started acting at all stages of life and allows for both an ample sample size and sufficient follow-up time. On the basis of what has been found in previous studies, it is hypothesized that black American film actors and actresses experienced faster mortality rates than their white American counterparts.

## Methods

### Wikipedia Data Source and Tools

To operationalize societal standing and notability as a celebrity, Wikipedia's list of American film actors and actresses was used as a sampling frame. Wikipedia is a nonprofit database that is maintained internationally by the public and moderated by a large group of Wikipedia editors. These editors use a set of criteria for notability, ensuring that data entered are verifiable and each subject page has enough reference and societal influence for inclusion in the online encyclopedia. If entries are under review or do not meet all criteria, they are clearly marked with a banner at the top of the page [24]. Notability is considered for every Wikipedia page and is determined by the quality and quantity of third-party sources rather than the page or article content. This editor gatekeeping method creates a unique, reliable benchmark for determining occupational prestige for actors and actresses. Wikipedia has been verified as a quality informational source by many scholars, with the most notable being *The Wisdom of Crowds* [25], which found Wikipedia to be comparable with the Encyclopedia Britannica. Others such as Vrandečić and Krötzsch have verified the reliability and value of Wikipedia as a collectively maintained database, with the diversity of authors, the number of contributors, and the ease of access continuing to improve over time, providing a sound sampling frame [26,27].

Microsoft Excel was used for data entry, observation randomization, and creation of the final dataset file. SAS 9.4 (SAS Institute 2013) was used for descriptive statistics and both survival curve estimates using LIFETEST and PHREG procedures for the Kaplan-Meier curve and Cox semiparametric logistic regression curve, respectively.

### Creating a Dataset

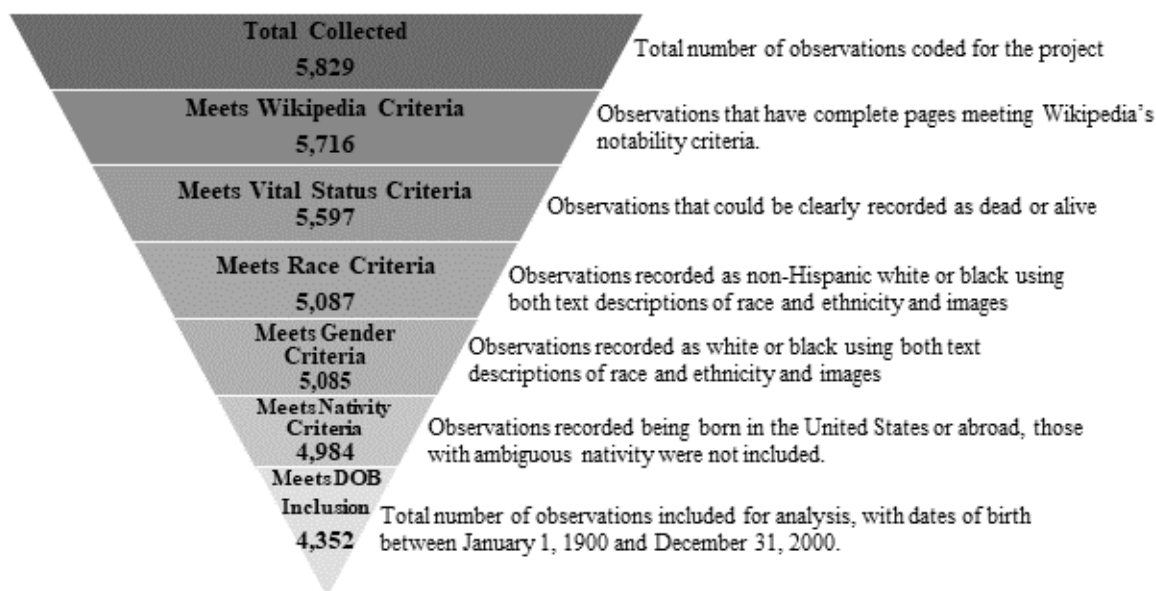
A dataset of URLs for US film actors and actresses was created by importing the HTML page links from the Wikipedia category

pages for American film actors and actresses. These data are open source under the Creative Commons Attribution-Share-Alike License. On May 31, 2017, 12,164 entries with actor names and the Wikipedia page URL were imported from Wikipedia, formatted into a comma-separated values (CSV) dataset using a general text editor program, and imported into Microsoft Excel 2016 for randomization and data entry. The randomized data were then split among 3 people for data entry for each observation, and once there were at least 5500 entries, the Excel datasets were merged into 1 CSV file and imported into SAS for data cleaning and analysis. To check accuracy and inter-rater reliability, 200 random entries from each of the 3 coding individuals were checked for accuracy, particularly for consistent coding of race.

For each film celebrity entry, a total of 6 variables were collected: gender, race, age, US nativity (yes or no), mortality status (dead, alive, or unknown), and cause of death. Wikipedia notability was confirmed by the absence of a banner indicating questionable reliability of information to confirm notability, and only those with confirmed notability were included in the

final dataset. Race was recorded using the Center for Disease Control and Prevention's definition for race based on perception of skin tone and relevant content contained in the Wikipedia entry. Collected categories were white, black, and other. Causes of death were grouped into the following categories: cardiovascular disease (CVD), suicide, accident, drug overdose, cancer, cause unknown, cause missing, and alive (no cause of death). Categories were chosen from the most common causes of death, with the addition of drug overdose and suicide because they were included in other studies focusing on similar musician populations [28,29]. CVD, suicide, and drug overdose were coded as defined by the World Health Organization, and other or unknown cause was coded as *other* [30]. The final inclusion criteria included actors that had notable Wikipedia status, were coded as either non-Hispanic white or black, had known birth and death dates (if applicable), and were born within the 20th century (January 1, 1900, to December 31, 2000). Death data were recorded with reference to May 27, 2017. In total, 5829 observations were made and 4352 (74.66%) met all inclusion criteria for study analysis (Figure 1).

**Figure 1.** Participant selection based on inclusion criteria. DOB: date of birth.



## Statistical Analyses

Microsoft Excel was used for data entry, observation randomization, and creation of the final dataset file. SAS 9.4 (SAS Institute 2013) was used for descriptive statistics and both survival curve estimates using LIFETEST and PHREG procedures for the Kaplan-Meier curve and Cox semiparametric logistic regression curve, respectively. Following data collection, descriptive statistics were run to characterize the distributions and bivariate associations among variables. A student *t* test was used to compare the average age at death by race. Normality and proportionality tests were run for age at death by race and gender to select the most appropriate survival analysis methods. In this analysis, survival curves use death as an event, and the specified period is the lifespan of the celebrity. The Kaplan-Meier survival curve was created through the SAS LIFETEST procedure to plot survival for each race from date

of birth to date of death, and the log-rank test was used to test statistically significant differences between each curve by race ( $P < .05$ ). A Cox semiparametric logistic regression model was used via the SAS PHREG procedure to compare survival rates while adjusting for year of birth, gender, and cause of death. A model was built via stepwise regression to see the effect of each covariate (including race) on survival outcomes. Only covariates showing significant effect on survival rate were retained in the model, and the final regression model was then used to compare races and obtain an adjusted hazard ratio (aHR) for the all-cause mortality rate among black celebrities, with white male celebrities as the reference group.

## Results

Of the 4352 observations in the analytic sample, 1335 (30.88%) actors died during the follow-up period. Of 505 black actors

observed, 405 (80.2%) were alive at the end of the follow-up period, and of 3847 white actors, 3017 (78.42%) were alive at the end of follow-up period. As shown in [Table 1](#), most of the sample (3565/4352, 81.92%) was male, and 4187/4352 (96.21%) had US nativity. Most of the causes of death (621/1335, 46.52%) observed were recorded as dead with unknown cause, with CVD and cancer as the top 2 causes of death among our sample claiming 339/1335 (25.39%) and 251/1335 (18.80%) of those dead, respectively.

Using the Kaplan-Meier function to compare survival rates between each race with celebrities' genders combined ([Figure 2](#)), there was a significant difference between white and black actors when using a log-rank test for equality for each race group ( $P<.001$ ). This was consistent with a two-sample student *t* test comparing age at death for white (70.5; 95% CI 69.6-71.4) and

black (64.1; 95% CI 69.6-67.5) actors and actresses, with a significant difference in mean age at death ( $P<.001$ ; [Figure 3](#)). On average, black actors and actresses died 6.4 years faster than their white counterparts.

Cox semiparametric logistic regression model was created via stepwise regression, with the significance threshold for inclusion set at  $P<.05$ . Date of birth was included as a continuous covariate with an expected hazard ratio (HR) of 1.00. The completed logistic regression model included gender ( $P=.001$ ), cause of death ( $P<.001$ ), and date of birth ( $P<.001$ ) as predictors of age at death. Race was not a significant predictor of age at death using stepwise regression ( $P=.13$ ), and when observing HRs for black actors with reference to their white counterparts, it was not found to be statistically significant (aHR 1.07; 95% CI 0.87-1.31).

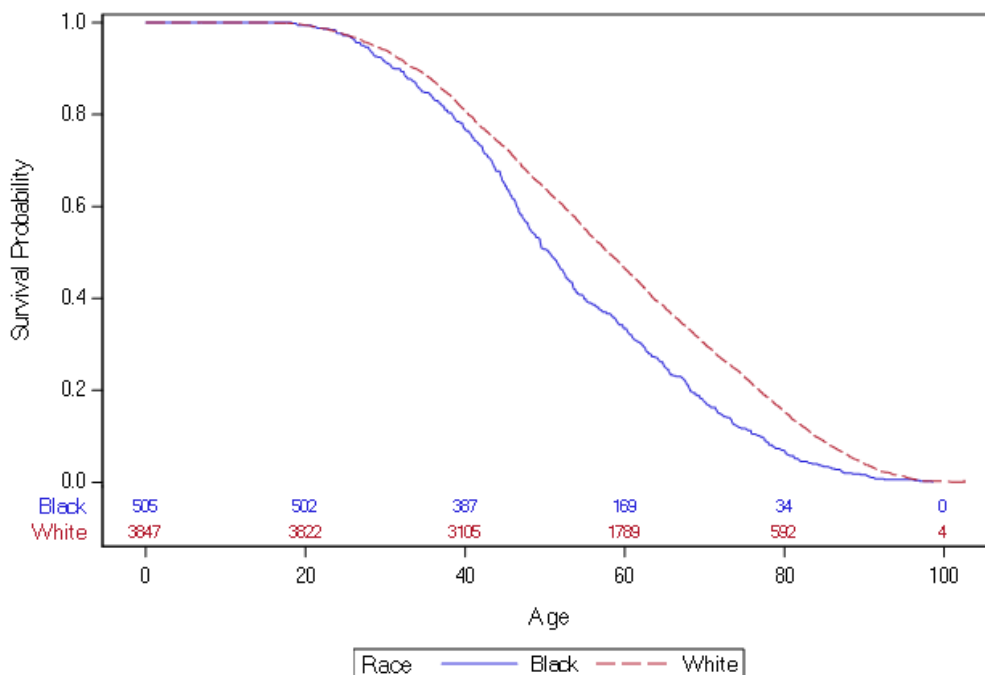
**Table 1.** Results for analysis sample demographics and adjusted hazard ratios.

Study covariate	Total (N=4352), n (%)	White race (n=3847), n (%)	Black race (n=505), n (%)	<i>P</i> value	Adjusted hazard ratio <sup>a</sup> (95% CI)
<b>Gender</b>				<.01	1.32 (1.120-1.56)
Male	3565 (81.92)	3139 (81.60)	426 (84.4)		
Female	787 (18.08)	708 (18.40)	79 (15.6)		
<b>Nativity</b>				.21	1.17 (0.91-1.50)
American	4187 (96.21)	3702 (96.23)	485 (96.0)		
Other	165 (3.79)	145 (3.77)	20 (4.0)		
<b>Vital status</b>				— <sup>b</sup>	—
Dead	1335 (30.68)	1235 (32.10)	100 (19.8)		
Alive	3017 (69.32)	2612 (67.90)	405 (80.2)		
<b>Cause of death</b>				<.01	1.00 (1.00-1.01)
Cardiovascular disease	339 (7.79)	304 (7.90)	35 (6.9)		
Cancer	251 (5.77)	227 (5.90)	24 (4.8)		
Accident	57 (1.31)	54 (1.40)	3 (0.0)		
Suicide	34 (0.78)	33 (0.86)	1 (0.2)		
Homicide	15 (0.34)	12 (0.31)	3 (0.6)		
Drug overdose	18 (0.41)	16 (0.42)	2 (0.4)		
Cause unknown	621 (14.27)	589 (15.31)	32 (6.3)		
Alive	3017 (69.32)	2612 (67.90)	405 (80.2)		
Date of birth				<.01	1.00 (1.00-1.00)

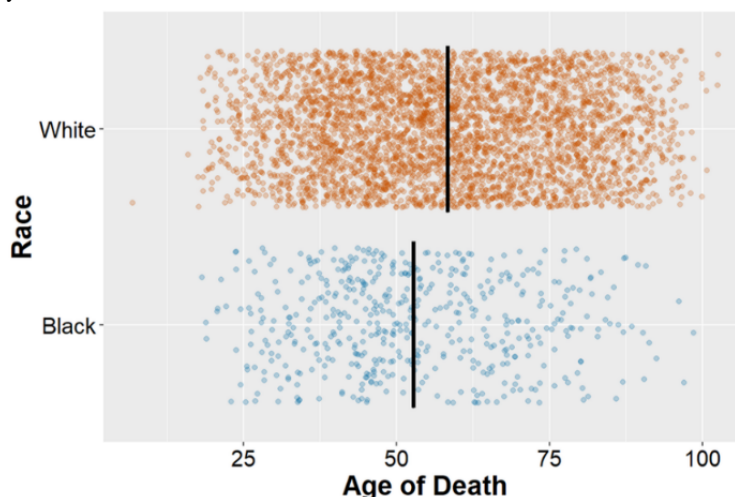
<sup>a</sup>White males were used as the reference group.

<sup>b</sup>Not applicable.

**Figure 2.** Survival probability over time by race (Kaplan-Meier plot).



**Figure 3.** Average age of death by race.



## Discussion

Findings from this Wikipedia study indicated that black film celebrities died 6.4 years faster than their white counterparts. Although hastened mortality among black film celebrities was observed, this association did not persist after adjusting for gender, cause of death, and date of birth. These results provide some indication that racial disparities in mortality may persist even at the highest levels of SES, and this association was no longer significant in adjusted analyses.

The discrepancies in the results between different survival comparison methods (Kaplan-Meier and Cox semiparametric logistic regression model) reflect important social changes in the representation of black actors and actresses in the American film industry during the 20th century. Black actors and actresses have historically experienced racism and other structural barriers to full participation in the American film industry, such that

increased representation began only during the 1960s [31,32]. The proportion of black film celebrities in this sample was approximately 13%, which is consistent with the proportion of black Americans recorded in US census data from 2010 [4]. Nearly all black actors and actresses included in the sample were born during the second half of the 20th century. Therefore, it is likely that there has not been sufficient time to adequately assess racial disparities in mortality after adjusting for date of birth. Future research should continue to examine racial disparities in mortality among American film celebrities, and further research is also needed to elucidate if SES modifies racial disparities in mortality for black Americans.

One noteworthy finding was that only 787 (18.1%) of the random sample of 4352 Wikipedia entries were women. This may reflect the fact that many women began entering workforce in larger numbers during the last 30 years of the 20th century. However, this may also be attributable to enduring concerns related to gender bias in the American film industry [33,34].

The primary objective of this study was to examine racial disparities in a population of film celebrities with high SES. However, these stark differences in the proportional representation of women underscore the questionable generalizability of this unique sample to the broader population of the United States. Our findings clearly demonstrate that white actors appear to be the overwhelming majority in the 20th century American film industry.

Although this study underscores the potential benefits of leveraging open access sources such as Wikipedia, findings should be interpreted in the context of several important limitations. Owing to key social changes in the representation of black actors and actresses in the American film industry, it is likely that there has not been a sufficient amount of time to examine racial disparities when adjusting for date of birth. American film actors and actresses also represent a subset of celebrities from a broader population containing distinct occupational groups such as musicians and athletes. In this study, these groups were not included for several reasons. For example, inclusion of athletes introduces a fitness bias. There

were also important limitations to the retrospective cohort design using Wikipedia data. Wikipedia itself has been found to have a deficit of female editors, fewer notable female Wikipedia page entries than notable male Wikipedia page entries, and biased language on female Wikipedia pages [35-37]. Most notably, some data that may have served as key confounders, such as relationship status, education level, childhood SES, and geographic residence, were not extracted [9,11,13]. Future studies with primary data collection would provide more nuanced information regarding potential confounders and effect modifiers of racial disparities in all-cause mortality.

Despite these limitations, this study provides some of the first estimates of racial disparities in mortality among American film celebrities. Results from this study have observed a 6.4-year higher mortality rate among black film celebrities, which underscores that racial disparities may persist in the United States even at the highest levels of SES. Findings support the need for further research to examine the social and psychological mechanisms that could explain the profound racial disparities in mortality experienced by black Americans.

## Acknowledgments

The authors would like to acknowledge the Wikimedia Foundation as well as the community of Wikipedia editors for their efforts toward publishing and maintaining online databases to support unique projects such this study. The authors would also like to acknowledge the efforts of the University of Miami's Public Health 651 Research Methods class for their thoughts and efforts toward the research concept and methods.

## Conflicts of Interest

None declared.

## References

1. United States Department of Health and Human Services, Task Force on Black and Minority Health. Digital Collections - National Library of Medicine. Report of the Secretary's Task Force on Black & Minority Health URL: <http://health-equity.lib.umd.edu/3005/1/ANDERSON.pdf> [accessed 2019-11-15]
2. Cunningham TJ, Croft JB, Liu Y, Lu H, Eke PI, Giles WH. Vital signs: racial disparities in age-specific mortality among blacks or African Americans - United States, 1999-2015. *MMWR Morb Mortal Wkly Rep* 2017 May 5;66(17):444-456 [FREE Full text] [doi: [10.15585/mmwr.mm6617e1](https://doi.org/10.15585/mmwr.mm6617e1)] [Medline: [28472021](https://pubmed.ncbi.nlm.nih.gov/28472021/)]
3. Center for Medicare and Medicaid Services. 2017. Racial and Ethnic Disparities by Gender in Health Care in Medicare Advantage URL: <https://www.cms.gov/About-CMS/Agency-Information/OMH/Downloads/Health-Disparities-Racial-and-Ethnic-Disparities-by-Gender-National-Report.pdf> [accessed 2019-11-15]
4. Geronimus AT, Bound J, Waidmann TA, Colen CG, Steffick D. Inequality in life expectancy, functional status, and active life expectancy across selected black and white populations in the United States. *Demography* 2001 May;38(2):227-251. [doi: [10.1353/dem.2001.0015](https://doi.org/10.1353/dem.2001.0015)] [Medline: [11392910](https://pubmed.ncbi.nlm.nih.gov/11392910/)]
5. DeNavas-Walt C, Proctor BD, Smith JC. Census Bureau. 2010. Income, Poverty, and Health Insurance Coverage in the United States: 2010 URL: <https://www2.census.gov/library/publications/2011/demo/p60-239/p60-239.pdf> [accessed 2019-11-15]
6. Hudson DL, Puterman E, Bibbins-Domingo K, Matthews KA, Adler NE. Race, life course socioeconomic position, racial discrimination, depressive symptoms and self-rated health. *Soc Sci Med* 2013 Nov;97:7-14. [doi: [10.1016/j.socscimed.2013.07.031](https://doi.org/10.1016/j.socscimed.2013.07.031)] [Medline: [24161083](https://pubmed.ncbi.nlm.nih.gov/24161083/)]
7. Huie SB, Krueger PM, Rogers RG, Hummer RA. Wealth, race, and mortality. *Soc Sci Q* 2003;84(3):667-684. [doi: [10.1111/1540-6237.8403011](https://doi.org/10.1111/1540-6237.8403011)]
8. Berkman LF, Kawachi I, Glymour M. *Social Epidemiology*. Second Edition. Oxford: Oxford University Press; 2000.
9. Pollack CE, Chideya S, Cubbin C, Williams B, Dekker M, Braveman P. Should health studies measure wealth? A systematic review. *Am J Prev Med* 2007 Sep;33(3):250-264. [doi: [10.1016/j.amepre.2007.04.033](https://doi.org/10.1016/j.amepre.2007.04.033)] [Medline: [17826585](https://pubmed.ncbi.nlm.nih.gov/17826585/)]
10. Sequist TD, Schneider EC. Addressing racial and ethnic disparities in health care: using federal data to support local programs to eliminate disparities. *Health Serv Res* 2006 Aug;41(4 Pt 1):1451-1468 [FREE Full text] [doi: [10.1111/j.1475-6773.2006.00549.x](https://doi.org/10.1111/j.1475-6773.2006.00549.x)] [Medline: [16899018](https://pubmed.ncbi.nlm.nih.gov/16899018/)]

11. Do DP, Frank R, Finch BK. Does SES explain more of the black/white health gap than we thought? Revisiting our approach toward understanding racial disparities in health. *Soc Sci Med* 2012 May;74(9):1385-1393. [doi: [10.1016/j.socscimed.2011.12.048](https://doi.org/10.1016/j.socscimed.2011.12.048)] [Medline: [22405688](https://pubmed.ncbi.nlm.nih.gov/22405688/)]
12. Green RD, Oliver ML, Shapiro TM. Black wealth, white wealth: a new perspective on racial inequality. *J Negro Educ* 1995;64(4):477. [doi: [10.2307/2967270](https://doi.org/10.2307/2967270)]
13. Braveman PA, Cubbin C, Egerter S, Chideya S, Marchi KS, Metzler M, et al. Socioeconomic status in health research: one size does not fit all. *J Am Med Assoc* 2005 Dec 14;294(22):2879-2888. [doi: [10.1001/jama.294.22.2879](https://doi.org/10.1001/jama.294.22.2879)] [Medline: [16352796](https://pubmed.ncbi.nlm.nih.gov/16352796/)]
14. Betancourt JR, Green AR, Carrillo JE, Ananeh-Firempong O. Defining cultural competence: a practical framework for addressing racial/ethnic disparities in health and health care. *Public Health Rep* 2003;118(4):293-302 [FREE Full text] [doi: [10.1093/phr/118.4.293](https://doi.org/10.1093/phr/118.4.293)] [Medline: [12815076](https://pubmed.ncbi.nlm.nih.gov/12815076/)]
15. Weinick RM, Zuvekas SH, Cohen JW. Racial and ethnic differences in access to and use of health care services, 1977 to 1996. *Med Care Res Rev* 2000;57(Suppl 1):36-54. [doi: [10.1177/1077558700057001S03](https://doi.org/10.1177/1077558700057001S03)] [Medline: [11092157](https://pubmed.ncbi.nlm.nih.gov/11092157/)]
16. Williams DR, Collins C. Racial residential segregation: a fundamental cause of racial disparities in health. *Public Health Rep* 2001;116(5):404-416 [FREE Full text] [doi: [10.1093/phr/116.5.404](https://doi.org/10.1093/phr/116.5.404)] [Medline: [12042604](https://pubmed.ncbi.nlm.nih.gov/12042604/)]
17. Williams DR, Jackson PB. Social sources of racial disparities in health. *Health Aff (Millwood)* 2005;24(2):325-334. [doi: [10.1377/hlthaff.24.2.325](https://doi.org/10.1377/hlthaff.24.2.325)] [Medline: [15757915](https://pubmed.ncbi.nlm.nih.gov/15757915/)]
18. Hajat A, Kaufman JS, Rose KM, Siddiqi A, Thomas JC. Long-term effects of wealth on mortality and self-rated health status. *Am J Epidemiol* 2011 Jan 15;173(2):192-200 [FREE Full text] [doi: [10.1093/aje/kwq348](https://doi.org/10.1093/aje/kwq348)] [Medline: [21059808](https://pubmed.ncbi.nlm.nih.gov/21059808/)]
19. Allin S, Masseria C, Mossialos E. Measuring socioeconomic differences in use of health care services by wealth versus by income. *Am J Public Health* 2009 Oct;99(10):1849-1855. [doi: [10.2105/AJPH.2008.141499](https://doi.org/10.2105/AJPH.2008.141499)] [Medline: [19150899](https://pubmed.ncbi.nlm.nih.gov/19150899/)]
20. Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annu Rev Public Health* 1997;18:341-378. [doi: [10.1146/annurev.publhealth.18.1.341](https://doi.org/10.1146/annurev.publhealth.18.1.341)] [Medline: [9143723](https://pubmed.ncbi.nlm.nih.gov/9143723/)]
21. Gregorio DI, Walsh SJ, Paturzo D. The effects of occupation-based social position on mortality in a large American cohort. *Am J Public Health* 1997 Sep;87(9):1472-1475. [doi: [10.2105/ajph.87.9.1472](https://doi.org/10.2105/ajph.87.9.1472)] [Medline: [9314799](https://pubmed.ncbi.nlm.nih.gov/9314799/)]
22. Redelmeier DA, Singh SM. Survival in Academy Award-winning actors and actresses. *Ann Intern Med* 2001 May 15;134(10):955-962. [doi: [10.7326/0003-4819-134-10-200105150-00009](https://doi.org/10.7326/0003-4819-134-10-200105150-00009)] [Medline: [11352696](https://pubmed.ncbi.nlm.nih.gov/11352696/)]
23. Smith SL, Choueiti M, Pieper K, Gillig T, Lee C, DeLuca D. USC Annenberg School for Communication and Journalism. 2015. Inequality in 700 Popular Films: Examining Portrayals of Gender, Race & LGBT Status from 2007 to 2014 URL: [http://assets.uscannenberg.org/docs/inequality\\_in\\_700\\_popular\\_films\\_8215\\_final\\_for\\_posting.pdf](http://assets.uscannenberg.org/docs/inequality_in_700_popular_films_8215_final_for_posting.pdf) [accessed 2019-11-15]
24. Wikipedia. Wikipedia: Notability URL: <https://en.wikipedia.org/wiki/Wikipedia:Notability> [accessed 2019-11-15]
25. Giles J. Internet encyclopaedias go head to head. *Nature* 2005 Dec 15;438(7070):900-901. [doi: [10.1038/438900a](https://doi.org/10.1038/438900a)] [Medline: [16355180](https://pubmed.ncbi.nlm.nih.gov/16355180/)]
26. Vrandečić D, Krötzsch M. Wikidata: a free collaborative knowledgebase. *Commun ACM* 2014;57(10):78-85. [doi: [10.1145/2629489](https://doi.org/10.1145/2629489)]
27. Arazy O, Morgan W, Patterson R. Wisdom of the Crowds: Decentralized Knowledge Construction in Wikipedia. *SSRN J* 2006. [doi: [10.2139/ssrn.1025624](https://doi.org/10.2139/ssrn.1025624)]
28. Tippett R. Carolina Demography. 2010. Mortality and Cause of Death, 1900 v 2010 URL: <https://demography.cpc.unc.edu/2014/06/16/mortality-and-cause-of-death-1900-v-2010/> [accessed 2019-11-15]
29. Kenny DT, Asher A. Life expectancy and cause of death in popular musicians: is the popular musician lifestyle the road to ruin? *Med Probl Perform Art* 2016 Mar;31(1):37-44. [doi: [10.21091/mppa.2016.1007](https://doi.org/10.21091/mppa.2016.1007)] [Medline: [26966963](https://pubmed.ncbi.nlm.nih.gov/26966963/)]
30. Porta M, Last J. *A Dictionary of Public Health*. Geneva: World Health Organization; 2018.
31. Quinn E. Closing doors: Hollywood, affirmative action, and the revitalization of conservative racial politics. *J Am Hist* 2012;99(2):466-491. [doi: [10.1093/jahist/jas302](https://doi.org/10.1093/jahist/jas302)]
32. Hesmondhalgh D, Saha A. Race, ethnicity, and cultural production. *Popular Commun* 2013;11(3):179-195. [doi: [10.1080/15405702.2013.810068](https://doi.org/10.1080/15405702.2013.810068)]
33. Bielby D, Bielby W. Women and men in film: gender inequality among writers in a culture industry. *Gend Soc* 1996;10(3):248-270. [doi: [10.1177/089124396010003004](https://doi.org/10.1177/089124396010003004)]
34. de Pater IE, Judge TA, Scott BA. Age, gender, and compensation: a study of Hollywood movie stars. *J Manag Inq* 2014;23(4):407-420. [doi: [10.1177/1056492613519861](https://doi.org/10.1177/1056492613519861)]
35. Lam S, Uduwage A, Dong Z, Sen S, Musicant DR, Terveen L, et al. WP:Clubhouse?: An Exploration of Wikipedia's Gender Imbalance. In: *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*. 2011 Presented at: WikiSym'11; October 3-5, 2011; Mountain View, California p. 1-10. [doi: [10.1145/2038558.2038560](https://doi.org/10.1145/2038558.2038560)]
36. Wagner C, Graells-Garrido E, Garcia D, Menczer F. Women through the glass ceiling: gender asymmetries in Wikipedia. *EPJ Data Sci* 2016;5(1). [doi: [10.1140/epjds/s13688-016-0066-4](https://doi.org/10.1140/epjds/s13688-016-0066-4)]
37. Graells-Garrido E, Lalmas M, Menczer F. First Women, Second Sex: Gender Bias in Wikipedia. In: *Proceedings of the 26th ACM Conference on Hypertext & Social Media*. 2015 Presented at: HT'15; September 1 - 4, 2015; Guzelyurt, Northern Cyprus p. 165-174. [doi: [10.1145/2700171.2791036](https://doi.org/10.1145/2700171.2791036)]

---

**Abbreviations**

**aHR:** adjusted hazard ratio  
**CSV:** comma-separated values  
**CVD:** cardiovascular disease  
**HR:** hazard ratio  
**SES:** socioeconomic status

---

*Edited by G Eysenbach; submitted 01.03.19; peer-reviewed by A Menking, L Rasberry; comments to author 03.08.19; revised version received 17.08.19; accepted 31.08.19; published 10.12.19*

*Please cite as:*

*Speaks H, Falise A, Grosgebauer K, Duncan D, Carrico A*

*Racial Disparities in Mortality Among American Film Celebrities: A Wikipedia-Based Retrospective Cohort Study*

*Interact J Med Res 2019;8(4):e13871*

URL: <http://www.i-jmr.org/2019/4/e13871/>

doi: [10.2196/13871](https://doi.org/10.2196/13871)

PMID: [31821148](https://pubmed.ncbi.nlm.nih.gov/31821148/)

©Hannah Speaks, Alyssa Falise, Kaitlin Grosgebauer, Dustin Duncan, Adam Carrico. Originally published in the Interactive Journal of Medical Research (<http://www.i-jmr.org/>), 10.12.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.i-jmr.org/>, as well as this copyright and license information must be included.