

UC Davis

UC Davis Electronic Theses and Dissertations

Title

Cancer Incidence and Survival among the Armenian Population in California: Construction of the Armenian Surname List (ASL) and its Application in the California Cancer Registry (CCR)

Permalink

<https://escholarship.org/uc/item/18m4x88n>

Author

Movsisyan, Ani Smbatovna

Publication Date

2024

Peer reviewed|Thesis/dissertation

Cancer Incidence and Survival among the Armenian Population in California:
Construction of the Armenian Surname List (ASL) and its Application
in the California Cancer Registry (CCR)

By

ANI SMBATOVNA MOVSISYAN

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Public Health Sciences

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Theresa Keegan, Chair

Laura Fejerman

Jeffrey Hoch

Committee in Charge

2024

Dedicated to my parents:
Gayane Hovhannisyan and Smbat Movsisyan

Acknowledgements

I am incredibly honored to have had Dr. Theresa Keegan as my dissertation chair and major professor. Dr. Keegan's guidance throughout this dissertation has been invaluable, as she consistently provided thorough reviews, clear and focused feedback, and support for exploring innovative ideas. Dr. Keegan's passion for research challenged me and fueled my growth and development as an academic researcher. I am grateful to my other dissertation committee members, Dr. Laura Fejerman and Dr. Jeffrey Hoch, for all of their expertise, support, guidance, and encouragement throughout this project. Additionally, as members of my qualifying exam committee, Dr. Michelle Ko, Dr. Keith Watenpaugh, Dr. Lorena Garcia, and Dr. Rosemary Cress provided constructive feedback and thoughtful questions that helped me shape the chapters of this dissertation. I am grateful to Dr. Kiumarss Nasserri for his expertise in creating the Middle Eastern Surname list and guidance for developing the Armenian Surname List. I am also grateful to Dr. Susan Stewart for her guidance on statistical analyses for the second and third chapters, and for the support of my colleagues, Dr. Cyllene Morris, Brenda Hofer, and Dr. Fran Maguire, for their years of experience and expertise in using California's cancer registry data, and for sharing their advice and encouragement in various parts of this project.

This dissertation would not have been possible without the support of my family. As survivors of the Armenian Genocide, my grandparents' and great-grandparents' will to live is a daily reminder to pursue all opportunities wholeheartedly and never to take this gift of life for granted. Without the sacrifices of my parents, who moved our family to the United States in 2000, I would not have had the opportunity to pursue this level of education. From a very young age, my mother taught me to value and love education, and to complete any and all tasks with wholehearted devotion. My father, a survivor of the 1988 Spitak Earthquake who faced heartbreaking loss, inspires me with his courage, contagious passion, and love for life. Without my parents' sacrifices, perseverance, and hope for the future, I would not be who I am. My siblings, Araksya and Gagik, have not only been the best sister and brother I could have asked for, but they have also been incredible friends throughout the various stages of our lives. From building mud houses in our village in Armenia, to moving to the United States and learning to speak a new language, to doing hours of homework together on our dining room table, applying to colleges and graduating, I simply cannot imagine life without

them. Throughout this dissertation project, they have continually encouraged me, expressed their belief in my abilities to continue, and uplifted me with their incredible humor in moments of discouragement. During this dissertation project, I met my amazing husband, John, who has consistently helped me overcome challenges in this project by patiently listening, encouraging me when I was overwhelmed, making limitless cups of coffee, and reminding me of the reasons behind this dissertation. His faith in my abilities gave me the strength to keep moving forward when I faced roadblocks. I am beyond grateful for his unwavering support and love, all while wedding planning, getting married, and moving to a new city to start our life together.

I am thankful for my cohort in the Public Health Sciences program for walking this journey together, as they understand the unexpected challenges of completing a doctoral program during the height of the COVID-19 pandemic. My classmate and friend, Pauline deLange Martinez, has been a constant source of support, encouragement, and understanding as we moved through this program together. Our routine walks at the arboretum were the highlight of many difficult weeks during the COVID-19 pandemic and the 2020 Artsakh War.

Lastly, I am grateful for our neighbors who became family in Sarakap and Gyumri, where I spent my childhood, as our tight-knit community instilled in me values of faith, community, hard work, and generosity. In the diaspora, I am thankful for and inspired by Armenian communities in California and all around the world for their resilient spirit and passion for life after overcoming unimaginable hardships in our history. Many members of Armenian communities worldwide have reached out to me with their support for this dissertation and expressed the need to represent the Armenian community in research studies. Only with the support of my committee, community, colleagues, family, and friends was I able to persevere and complete this dissertation. I am eternally grateful and could not have done this alone.

Table of Contents

Acknowledgements	iii
Abstract	vii
Introduction	1
Chapter 1.....	6
Abstract	7
Introduction.....	8
Methods.....	9
Data Sources	9
Analyses.....	10
Results	12
Construction of the Armenian Surname List	12
Evaluation of the Armenian Surname List	13
Discussion	13
Figure 1.1 Flowchart of the Armenian Surname List (ASL) construction process.....	17
Figure 1.2 Counts of overlapping surnames between the Armenian Surname List (ASL), California Cancer Registry (CCR) Birthplace in Armenia List, and Middle Eastern Surname List (MESL) from the CCR	18
Figure 1.3 Percentage of surnames in the Armenian Surname List (ASL) and California Cancer Registry (CCR) Birthplace in Armenia List ranked by most likely country of origin and second-most likely country of origin ^a	19
Table 1.1 Number and percentage of surnames identified by the Middle Eastern Surname List (MESL) ^a via probabilistic linkage with California Public Use Death File (CPUDF), 1905-2020.....	20
Supplemental Table 1.1 Manual review categorization rules, California Public Use Death File (CPUDF) 1905-2020 and Middle Eastern Surname List (MESL) Armenian Surname probabilistic linkage ^b results	20
Chapter 2.....	21
Abstract	22
Introduction.....	23
Methods.....	24
Data Sources	24
Study Population	25
Analyses.....	25
Results	26
Discussion	28
Table 2.1 Characteristics of Armenians diagnosed with cancer in California, California Cancer Registry, 1988-2019.....	33
Table 2.2 Age-adjusted proportional incidence ratios for the 10 most common cancers in Armenian patients compared to NHW patients, California Cancer Registry, 1988-2019.....	34
Table 2.3 Age-Adjusted incidence rates and rate ratios and proportional incidence ratios of the top 10 cancers among Armenian patients compared with NHW patients, California Cancer Registry, 2006-2015	35

Supplemental Table 2.1 Armenian population demographics in California, American Community Survey, 2006-2015	36
Supplemental Table 2.2 Characteristics of Armenian and NHW patients diagnosed with cancer in California, California Cancer Registry, 1988-2019	37
Chapter 3.....	39
Abstract	40
Introduction.....	41
Methods.....	42
Data Sources	42
Study Population	42
Sociodemographic and Clinical Variables	43
Analyses.....	43
Results	44
Discussion	46
Table 3.1 Characteristics of Armenian and NHW patients diagnosed with stomach, lung, colorectal, and bladder cancer in California, 1988-2019	50
Table 3.2 Odds ratios for regional/remote stage (versus localized) cancer diagnosis among Armenian patients in California, 1988-2019.....	52
Table 3.3 Cancer-specific survival of Armenian patients compared to NHW patients in California by cancer type, 1988-2019	53
Figure 3.1 Frequency distribution of stage at diagnosis by cancer type among Armenian and NHW patients in California, 1988-2019	54
Figure 3.2 Odds ratios for regional/remote stage (versus localized) cancer diagnosis among Armenian patients compared to NHW patients in California by cancer type, 1988-2019.....	55
Supplemental Table 3.1 Selected tumor sites based on International Classification of Diseases for Oncology, 3rd edition/World Health Organization (ICD-O-3/WHO) 2008 site codes	56
Supplemental Table 3.2 Cancer-specific survival of Armenian patients compared to NHW patients in California by cancer type considering effect of lost to follow-up (LTFU), 1988-2019	57
Conclusion	58
References.....	61

Abstract

Armenians, a growing ethnic minority population in the United States, have historically faced war, economic instability, and persecution including the Armenian Genocide in the early 20th century. In search of refuge, many Armenians relocated to the United States and over time, California has become the state with the largest population of Armenians. However, research on the population is limited due to the lack of accurate Armenian population estimates, a standardized collection of Armenian ethnicity or an Armenian surname list. Consequently, many Armenians are included in the 'White' or 'Some Other Race' race/ethnicity categories in research studies despite possible differences in genetic and sociodemographic factors affecting health patterns, including cancer incidence and survival. As cancer patterns specifically among Armenians in California have remained unknown, the overall purpose of this dissertation was to fill this research gap by constructing an Armenian Surname List (ASL) and using the ASL to identify and compare cancer incidence, stage at diagnosis and survival between Armenian and non-Hispanic White (NHW) patients in California.

In the first chapter, we extracted Armenian surnames from the Middle Eastern Surname List and constructed the ASL using California death records from 1905 to 2020. We used NamSor, an onomastic classification tool, to evaluate the ASL by ranking surname country of origin. In the second chapter, we used the ASL and CCR birthplace information to identify Armenians with cancer and American Community Survey data to identify the Armenian population in California. We then selected 10 most common cancers among Armenian males and females and calculated age-adjusted and sex stratified proportional incidence ratios (PIR) and incidence rate ratios (IRR) compared with NHW patients. In the third and final chapter, we used the ASL and birthplace information in the CCR to compare late-stage at diagnosis and cause-specific cancer survival for four tobacco-related cancers (stomach, lung, colorectal, and bladder) between Armenian and NHW patients. We used Cox proportional hazards models to calculate cancer-specific survival and logistic regression models to calculate odds of late-stage diagnoses.

In the first chapter, we constructed and evaluated the ASL and found that 81% of the 3,428 unique surnames in the ASL had Armenia as the most or second-most likely country of origin. In the second chapter, we identified Armenians (n=27,212) diagnosed with common cancers in California and found

significant differences in cancer incidence compared to NHW patients. Among Armenian males, the 10 most frequent cancers were prostate, lung, colorectal, bladder, non-Hodgkin lymphoma (NHL), stomach, leukemia, kidney, pancreas, and liver and intrahepatic bile duct (IBD). Among Armenian females, the 10 most frequent cancers were breast, colorectal, lung, uterine, NHL, thyroid, ovary, stomach, pancreas, and leukemia. Among the 10 most frequent cancers, Armenian males had higher proportions of seven cancers including stomach (PIR=2.39, 95% CI=2.19-2.60), bladder (PIR=1.53, 95% CI=1.45-1.61), colorectal (PIR=1.29, 95% CI=1.23-1.35), lung (PIR=1.16, 95% CI=1.11-1.21), leukemia (PIR=1.16, 95% CI=1.06-1.26), liver and IBD (PIR=1.19, 95% CI=1.05-1.34), and kidney (PIR=1.11, 95% CI=1.01-1.21) cancers. Armenian females also had higher proportions of seven cancers including stomach (PIR=3.24, 95% CI=2.93-3.57), thyroid (PIR=1.47, 95% CI=1.36-1.60), colorectal (PIR=1.29, 95% CI=1.23-1.35), pancreatic (PIR=1.20, 95% CI=1.09-1.33), leukemia (PIR=1.20, 95% CI=1.08-1.32), NHL (PIR=1.15, 95% CI=1.05-1.24), and ovarian (PIR=1.14, 95% CI=1.04-1.24) cancers. Exploratory IRR analyses showed higher stomach (IRR=1.78, 95% CI=1.55-2.06), bladder (IRR=1.13, 95% CI=1.05-1.23), and colorectal (IRR=1.12, 95% CI=1.03-1.20) cancers among Armenian males and higher stomach (IRR=2.54, 95% CI=2.15-2.99) cancer among Armenian females. In the third chapter, we found higher odds of late-stage colorectal (OR=1.12, 95% CI=1.03-1.22), lung (OR=1.26, 95% CI=1.12-1.42), and stomach (OR=1.43, 95% CI=1.17-1.74) cancers diagnosed among Armenian patients compared to NHW patients, and a moderate survival advantage of Armenian patients for stomach (HR=0.85, 95% CI=0.76-0.94), lung (HR=0.86, 95% CI=0.82-0.91), colorectal (HR=0.82, 95% CI=0.77-0.88), and bladder (HR=0.87, 95% CI=0.76-0.99) cancers. Among Armenian patients, we found that lower neighborhood socioeconomic status was associated with late-stage stomach and bladder cancer diagnoses, and public health insurance was associated with late-stage lung, colorectal, and bladder cancer diagnoses.

This dissertation project is the first to create an Armenian surname list, the ASL, that can be used to identify Armenians in databases for epidemiological research and in other research fields concerning Armenian populations. With the use of the ASL and CCR, we found that Armenian patients had higher incidence of several common cancers, which may relate to risk factors including tobacco-use, obesity, and exposure to harmful environmental agents. Our findings of a higher risk for late-stage diagnoses of numerous tobacco-related cancers compared to NHW patients, and that Armenian patients with lower

socioeconomic status and public health insurance had higher likelihood of late-stage diagnoses, reveal a need for increased access to reliable health care and cancer screening. Our findings also show the need for further research to address risk factors associated with specific cancers among Armenians and to understand factors associated with the modest survival advantage observed among the Armenian patients compared to NHW patients. Overall, this dissertation provides a new identification tool and data to inform evidence-based cancer control efforts for the growing Armenian population in California.

Introduction

Understanding factors associated with adverse health outcomes among minority racial/ethnic populations in the United States is important for informing health policies and distribution of resources to communities.¹ Across the country, many race/ethnic minority groups have a history of trauma, institutional racism, political instability, oppression, and other adversities that have a long-term impact on physical and mental health.² Among immigrant populations, stressors often continue after resettlement in the form of social exclusion, discrimination, and disconnection from their communities.² Long-term exposures to trauma and chronic stress are linked to adverse mental and physical health outcomes, including depression, anxiety, feelings of alienation, hopelessness, identity issues, diabetes, heart disease, and weight gain leading to various obesity-related illnesses including cancer.³

For public health research to have a significant impact on the Armenian community outside of Armenia, it is critical to understand the historical, cultural, and social context surrounding Armenian communities in the United States. Many Armenians, a growing ethnic minority population in the United States, have faced long-term trauma including persecution, torture, violence, and war prior to immigration.^{4,5} While the first Armenians in America arrived in the early 1600s to pursue educational and vocational opportunities, mass emigration from Armenia is historically believed to have begun in the late 1800s following rising violence and persecution by the Ottoman Empire.⁴ Following the Hamidian Massacres of 1895-1897 and the Armenian Genocide of 1915 that led to the extermination of an estimated 1.5 million Armenians, many survivors forcibly fled their native land and settled as refugees around the world, primarily in countries including Syria, Lebanon, Iran, Iraq, Turkey, Egypt, Russia, France, and the United States.^{4,6-9} Subsequent waves of emigration occurred after the Immigration and Nationality Act of 1965, the Jackson-Vanik amendment (1974), the Lebanese Civil War (1975-1990), the Iranian Revolution (1979), the Iran-Iraq War (1980-1988), the Baku and Sumgait Massacres (1988-1990), the collapse of the Soviet Union (1991), and the subsequent economic crash of the Republic of Armenia beginning in 1992.^{2,7,10,11} Additional destabilization in the country followed after the 1988 Spitak Earthquake that killed 25,000, injured 100,000, and displaced half a million people, and the 1988-1994 and 2020 Wars in Artsakh, also known as Nagorno-

Karabagh, with neighboring Azerbaijan.¹² Most recently, Azerbaijan's nine-month blockade of the Lachin Corridor, the sole route of transport of medicine, food, fuel, and other necessary supplies from Armenia to Artsakh, led to the United Nations declaration of humanitarian emergency in the region in August 2023 as the population of an estimated 120,000 ethnic Armenians in Artsakh faced severe shortages of food and medication.¹³ Despite international outcry for protection of Artsakh residents' human rights, Azerbaijan followed the blockade with a 24-hour military offensive on September 19, 2023, leading to the displacement and mass exodus of the entirety of the Artsakh population from their native land of Artsakh to Armenia in less than a month, and resulting in yet another humanitarian crisis as Armenia struggles to provide shelter and permanent housing for Artsakh refugees.¹⁴⁻¹⁷

As a result of these continued extreme hardships, thousands of Armenians relocated to the United States due to involuntary immigration and many are refugees who face added stressors as they go through the process of resettling in a new country.² The largest population of Armenians gradually shifted from the eastern states of New York, Massachusetts, and Pennsylvania to settle in California and over time, the majority of the population in California moved from residing primarily in Fresno to Los Angeles county, creating large ethnic enclaves in the cities of Glendale, Pasadena, and Hollywood.⁷ Today, approximately 40% of the population of Glendale is of Armenian descent, making the Armenian community a vital and integral part of the city's culture, economy, and government.^{7,18}

Armenian communities in the diaspora have shown incredible resiliency through the process of immigration and integration and have made notable contributions to their respective societies meanwhile enriching and preserving their culture, heritage, language, and the Armenian Church.¹⁹ This resiliency has shown through the experience of continued discrimination due to their Armenian identity in numerous forms in host countries.⁷ In Fresno, California, Armenians were regularly excluded from social activities and country clubs, and faced housing discrimination as restrictive laws and discriminatory advertisements promoted the protection of neighborhoods from 'undesirables', referring to racial/ethnic minority populations, including Armenians.²⁰ Discrimination against Armenians continues, with recent examples including a major California bank's discrimination against applicants of Armenian descent from 2015-2021, a hate crime outside an Armenian church in Watertown, Massachusetts, in September 2023, and most

recently, a targeted violent attack using lethal weapons on Armenian bishops, priests, and deacons in Jerusalem, Israel, on December 28, 2023.^{21–23}

In Armenia, waves of emigration, a poorly funded healthcare system, and lack of a national cancer registry have presented several challenges for epidemiologic surveillance of the population, including cancer surveillance.^{9,24,25} However, considerable progress has been made with the initiation of the first population-based cancer registry in 2019 and recent public health efforts to reduce tobacco use.^{26,27} With a population of nearly three million people, the cancer burden in Armenia is a significant cause of death and has received considerable attention and resources in recent years by the Ministry of Health.²⁸ Cancer is the second leading cause of death in Armenia, and Armenia has the 17th highest cancer mortality rate in the world.^{28–32} Lung, colorectal, breast, prostate, and stomach are the most common cancers, and lung cancer has the highest mortality rate.^{33–35} Further, previous studies have found lower cancer survival in Armenia compared to high-income countries, including the United States.³⁶ Differences in survival may relate to stage at diagnosis, as nearly half (49%) of all cancers in Armenia, and a majority of several tobacco-related cancers including stomach (70%), lung (82%), and colorectal (56%), are diagnosed at late-stage.^{36–38} Tobacco-use, a leading carcinogen, is widespread with 28% of the population, 52% of men and 2% of women, being current smokers.²⁷ While cancer research and treatment have improved, many Armenians struggle to receive timely cancer screening and are at risk for late-stage diagnoses.²⁸

Despite Armenia's geographic proximity to Middle-Eastern countries, and overlap of some cultural and social traditions, similar adverse cancer patterns are not observed among its closest neighboring countries, including Iran, Georgia, and Azerbaijan.²⁸ Cancer risk factors, including tobacco-use, and incidence and mortality rates for several cancers are higher in Armenia compared to surrounding countries.^{32,39,40} Age-adjusted cancer incidence and mortality rates per 100,000 in Armenia are 201.6 and 116.7, respectively, compared to Georgia (197.5 and 111.7), Azerbaijan (152.3 and 93.2), Iran (149.0 and 96.5), and Iraq (136.6 and 84.4).⁴¹ Further, while Armenians have lived in the United States for hundreds of years, public health and epidemiological research on the population is limited. Existing research studying the Armenian population in the United States focuses on the prevalence and predictors of diabetes, cardiovascular disease, hypertension, post-traumatic stress disorder, effects of relocation, and longitudinal cohort studies following survivors of the 1988 Spitak Earthquake.^{15,42}

Cancer patterns among Armenians after immigration are unknown as Armenians have remained unrepresented in public health research studies in the United States.⁴³ According to the US Census, there are approximately 470,000 Armenians living in the United States, of which more than half (259,430) reside in California.⁴⁴ This is likely an underestimate as many Armenians are likely included in the 'White' or 'Some Other Race' race/ethnicity categories.⁴⁵ For example, in the California Cancer Registry (CCR), Armenians are among 42 other race/ethnicities represented in the non-Hispanic White (NHW) category in the race variable.⁴⁶ However, genetic, environmental, and cultural factors may differentiate cancer incidence, stage at diagnosis, and survival patterns from White and other race/ethnic populations.²⁸ Providing cancer patterns among Armenians in California has numerous public health implications, including providing educational and awareness materials in the Armenian language (Eastern and Western dialects) and understanding barriers to and increasing access to screening and healthcare to improve cancer outcomes.

The challenge of identifying Armenians in the United States has precluded the representation of this population in public health and epidemiological studies.^{47,48} Previously, surname lists have been used to study Middle-Eastern populations in California.⁴⁸⁻⁵¹ Studies on cancer among Middle-Eastern populations in California included the Armenian population, but solely due to Armenia's geographic location bordering Middle-Eastern countries, and showed notably different cancer patterns between the Middle-Eastern and NHW populations.⁴⁸ This prior work in Middle-Eastern populations highlighted the need for future research to refine procedures pertaining to ethnic overlaps between Armenians, Turks, Iranians, and Coptic and Christian Arabs in the surname list as surnames used by first generation immigrants from Turkey, Lebanon, Iran, and the Caucasian Republic of the former Soviet Union were dominated with Armenian surnames.⁴⁸ Furthermore, when the Middle Eastern Surname List (MESL) was linked with the CCR, analyses by place of birth and Arab ethnicity showed that over 90% of patients were born in Afghanistan, Armenia, Iran, Jordan, Lebanon, and Syria.⁴⁸ While the inclusion of Armenian surnames in the MESL was certainly an important step toward better representation of Armenians, the lack of studies specific to the Armenian population still presented a significant gap in public health research in California. As California is home to the largest population of Armenians in the United States, creating a surname identification tool that can be linked to the CCR provides the opportunity to examine cancer patterns specifically among Armenians. Therefore, in the first chapter of this dissertation, we constructed an Armenian Surname List (ASL) for public

health research using the California public death records from 1905 to 2020 and an extract of Armenian surnames from the MESL, and evaluated the ASL by ranking surname country of origin using NamSor, an onomastic classification tool.^{52,53} In the second chapter, we linked the ASL with the CCR to estimate proportional cancer incidence ratios and cancer incidence rates using population denominators from the American Community Survey in Armenian patients compared with NHW patients for the 10 most common cancers among Armenian males and females. In the third and final chapter, using the ASL linked to the CCR, we compared late-stage of cancer diagnosis and cancer survival between Armenian and NHW cancer patients for four tobacco-related cancers.

Chapter 1

Title: Construction of the Armenian Surname List (ASL) for Public Health Research

Authors:

Ani S. Movsisyan^{1,2}, Kiumarss Nasser³, and Theresa H. Keegan^{1,2}

1. Department of Public Health Sciences, University of California Davis, Davis, CA
2. UC Davis Comprehensive Cancer Center, University of California Davis Medical Center, Sacramento, CA
3. Emeritus Professor of Epidemiology, International Health and Epidemiology Research Center

Abstract

Introduction: According to the US Census American Community Survey, there are an estimated 470,000 Armenians in the United States, and more than half live in California. Armenians are generally represented within the 'White' or 'Some Other Race' race/ethnicity categories in population-based research studies. While Armenians have been included in studies focused on Middle-Eastern populations, there are no studies focused exclusively on Armenians due to a lack of standardized collection of Armenian ethnicity in the United States or an Armenian surname list. To fill this research gap, we sought to construct and evaluate an Armenian Surname List (ASL) for use as an identification tool in public health and epidemiological research studies focused on Armenian populations.

Methods: Data sources for the ASL included the California Public Use Death Files (CPUDF) and the Middle Eastern Surname List (MESL). For evaluation of the ASL, the California Cancer Registry (CCR) database was queried for surnames with birthplace in Armenia and identified by the MESL.

Results: There are a total of 3,428 surnames in the ASL. Nearly half (1,678) of surnames in the ASL were not identified by the MESL. The ASL captured 310 additional Armenian surnames in the CCR than the MESL.

Conclusions: The ASL is the first surname list for identifying Armenians in major databases for epidemiological research.

Introduction

US Census race/ethnicity categories provide critical information allowing for health disparities to be recognized and addressed.⁴⁵ These categories have lacked details to capture existing health disparities affecting specific communities, including Armenians.⁴⁵ According to the US Census American Community Survey, the premier source for population information in the United States, there are an estimated 468,342 people of Armenian ancestry living in the United States, 44% (204,549) are foreign-born, and more than half (259,430) reside in California.^{44,54,55} This is likely an underestimate as the US Census does not have a standardized data collection of Armenian ethnicity, and the Embassy of Armenia to the United States of America estimates that the Armenian population in the United States is anywhere between 500,000 to 2,000,000 million.⁵⁶ Armenians are officially included within the 'White' or 'Some Other Race' race/ethnicity categories in population-based research studies,⁵⁷ while genetic and cultural factors are likely to underlie different health patterns among Armenians compared with other race/ethnic groups.^{28,58} For example, while health-related research among Armenians is limited in the United States, a prospective cohort study of Armenian and non-Armenian patients in a county hospital based in Los Angeles, California, showed that genetic, cultural, and dietary factors may contribute to Armenian ethnicity being associated with a higher risk of cardiovascular disease.⁴²

Using a common surname list is often employed to identify and study various ethnic groups.^{48,51,59} In particular, surname lists have been used to study cancer patterns among Arab-American and Middle-Eastern populations in California.^{48–51} Studies using the Middle Eastern Surname List (MESL), which included Armenians, showed notably different cancer patterns between the Middle-Eastern and non-Hispanic, non-Middle-Eastern White populations.⁴⁸ While Armenia is geographically proximal to Middle-Eastern countries and some cultural and social traditions are shared, similar adverse cancer patterns are not observed among its closest neighboring countries, including Iran, Georgia, and Azerbaijan.²⁸ Other studies on cancer among Middle-Eastern populations in California included Armenians within a broad Middle-Eastern race/ethnic group, but the challenge of identifying Armenians as a separate group has precluded studies specific to the Armenian population.^{48–50,60} A surname list specific to Armenians will allow

for the study of cancer patterns among this population, and can serve as a valuable tool for studying the Armenian population in various epidemiological subfields.⁴³

While the inclusion of Armenians in the Middle-Eastern studies is certainly an important step towards better representation of Armenians in research, studies specifically on Armenians are necessary to better understand health patterns among Armenians around the world.^{48–50,60} We sought to bridge this gap by creating and evaluating an Armenian Surname List (ASL) to identify Armenians in large population-based research databases. We created the ASL by linking an extract of Armenian surnames from the MESL to the California Public Use Death Files (CPUDF) and we used birthplace in Armenia to identify additional Armenian surnames. We evaluated the ASL by comparing with a population-based cancer registry database containing Armenian surnames and birthplace and using a name checking technology to rank the country of origin of the surnames in the ASL.

Methods

Data Sources

California Public Use Death Files

The CPUDF is available through the Center for Health Statistics and Informatics (CHSI) branch of the California Department of Public Health (CDPH) and contains information about in-state California deaths.⁶¹ An annual file is released approximately four months after the end of the calendar year. The CPUDF from 1905-2020 contains the following information: first name, middle name, last name, sex, date of birth, place of birth, place of death, date of death, and father's last name. Father's last name was available from 1940-2020. Using SAS version 9.4 (SAS Institute, Inc.), individual year death files from 1905-2020 were consolidated into one large file. The consolidated CPUDF from 1905-2020 contained a total of 16,949,541 records and 1,121,710 unique surnames.

Middle Eastern Surname List (MESL) Armenian Surname Extract

The MESL was created in 2007 and was extracted from a NUMIDENT extract file, a data source provided by the US Social Security Administration that began collecting place of birth in 1979.⁴⁸ The Arab surname list was also used as a source of additional Middle-Eastern surnames.⁵¹ The MESL contained 49,610 surnames.⁴⁸ Based on birthplace in Armenia, 1,332 surnames were extracted from the MESL and used in

this project. Surnames in the MESL are truncated at 10 digits. As such, the extract file contains some truncated Armenian surnames.⁴⁸

California Cancer Registry

The California Cancer Registry (CCR) is a comprehensive state-wide cancer surveillance program and meets SEER standards for quality and completeness of data.⁶² The CCR is also Gold Certified by the North American Association of Central Cancer Registries (NAACCR) and is one of the largest cancer registries in the world.⁶² The CCR research file provides patient demographics, country of birth, primary tumor site, tumor morphology and stage at diagnosis, first course of treatment, and follow-up for vital status on all cancers except for non-melanoma skin cancers.⁶⁰ We used the CCR research file with cancer diagnoses from 1988 to 2021. We used patient last name, maiden name, and father's surname when available to create the list of CCR surnames. Because an estimated 98% of people born in Armenia are of Armenian ancestry, birthplace in Armenia is a relatively reliable measure for Armenian ancestry.⁶³ Therefore, we used the country of birth variable in the CCR to retrieve all surnames with birth in Armenia. In the CCR, there is a variable indicating whether a patient's last name is on the MESL.

Analyses

Construction of the Armenian Surname List

Step 1: California Public Use Death Files (CPUDF) and Armenian Surnames from Middle Eastern Surname List (MESL) Probabilistic Linkage

We used Match*Pro to link surnames in the CPUDF with Armenian surnames from the MESL. Match*Pro is a probabilistic linkage software developed by IMS (Information Management Services, Inc.) under contract with the National Cancer Institute (NCI). The Match*Pro configuration page includes blocking and matching methods, adjusting for blocking sensitivity, and manual review tools to filter categorized matches. Last Name was the comparator of choice and the blocking strategy selected was *Soundex*, which is a phonetic algorithm for indexing names based on sound rather than spelling.⁶⁴ The *Soundex* feature allowed for truncated Armenian last names in the MESL to be included in record comparisons. It also allowed flexibility in the matching process for surnames with special characters including spaces, hyphens, and

capitalizations. For example, the surname ‘Ter-Minasyan’ was scored as a perfect match with ‘Ter Minasyan’.

All linked records were manually reviewed using the Match*Pro Linkage Results interface. In this step, records were reviewed to sort matches at the surname level. Linkage results were manually separated into two categories of match or non-match described in Supplemental Table 1.1. This linkage was not limited by birthplace, allowing for the identification of Armenian surnames with birth outside of Armenia.

Step 2: Selection of Surnames by Birthplace in Armenia

Surnames from the CPUDF that were born in Armenia, but did not link with existing Armenian surnames in the MESL (Step 1), were manually reviewed and surnames meeting the following two criteria were deleted from the list: 1) had a length of less than five characters or 2) did not have a common Armenian surname suffix (*ian, yan, ians, yans, iants, yants*), a common Russian patronymic suffix (*ov, ova*),^{65,66} and were less than 12 characters long. These deletions were necessary to remove common non-Armenian surnames such as ‘Smith’, ‘Abad’, ‘Ryan’, and ‘Weatherford’, and to retain Armenian surnames with less common suffixes and at least 12 characters long such as ‘Hambartsumanz’ and ‘Ter-Prakhourany’. Because many Armenians changed the suffix of their last names during the Union of Soviet Socialist Republics (USSR) period to follow Russian patronymic suffixes, we included Armenian last names with a common Russian suffix. All of the 1,332 unique Armenian surnames from the MESL that were not identified by the CPUDF/MESL Armenian surnames linkage nor by the birthplace in Armenia queries in the CPUDF were retained in the ASL.

Evaluation of the Armenian Surname List

Step 1: Compare Exact Surnames between the ASL and California Cancer Registry (CCR) List of Surnames with Known Birthplace in Armenia

Since birthplace in Armenia is a reliable measure for Armenian ancestry,⁶³ we compared the ASL with the CCR Birthplace in Armenia surname list to evaluate the performance of the ASL in identifying Armenian surnames in large research databases. To compare the ASL with the CCR Birthplace in Armenia List, we calculated the number of surnames that were identified in both versus only one of the lists. We presented results in a Venn diagram, with counts representing unique surnames. We calculated the proportion of

surnames in the CCR with birthplace in Armenia that were in the MESL. We also compared the proportion of surnames in the CCR Birthplace in Armenia List that were identified by the MESL and the ASL.

Step 2: NamSor Country of Origin Ranking of ASL and CCR Birthplace in Armenia List

We also sought to evaluate the country of origin of surnames in the ASL compared to the CCR Birthplace in Armenia List because surnames with birthplace in Armenia are likely to have originated in Armenia and therefore represent Armenian ethnic identity.^{63,67} We used an independent onomastic classification tool, NamSor, to compare the country of origin rank of surnames in each list. NamSor is a name checking technology that uses applied onomastics to classify names by gender, country of origin, ethnicity, and diaspora.^{52,68,69} Naïve Bayes Classifiers are a class of algorithms used by NamSor for ranking and classification purposes.⁶⁹ The NamSor Country of Origin feature returns a list of the top 10 countries of origin, ranked from most likely to least likely. We compared the country of origin of surnames in the ASL to the CCR Birthplace in Armenia List.

Results

Construction of the Armenian Surname List

After manual review of linked surnames between the existing Armenian surnames in the MESL and the CPUDF, we selected 1,290 unique surnames to include in the ASL (Supplemental Table 1.1, Figure 1.1).

The Soundex blocking strategy allowed for truncated surnames in the MESL to match with full Armenian surnames in the CPUDF. For example, the surnames 'Haroutunia' and 'Martirosya' from the MESL matched with the surnames 'Haroutunian' and 'Martirosyan', respectively, from the CPUDF. Soundex also allowed for flexibility in the spelling of last names and surnames with different spellings were included in the ASL as unique surnames. For example, the surname 'Mahtesyan' from the MESL matched with the surname 'Mahtesian' from the CPUDF and both surnames were included in the ASL. The Soundex blocking strategy allowed for the identification of 232 (17.9%) unique surnames from this linkage that would not have been identified by an exact matching strategy.

There were an additional 2,018 unique surnames in the CPUDF with a birthplace of Armenia that were not identified by the Armenian surnames in the MESL (Table 1.1). After manually reviewing these

surnames, 158 were removed, and the remaining 1,860 surnames were included in the ASL (Figure 1.1). From the existing list of 1,332 Armenian surnames in the MESL, 278 surnames did not appear in the CPUDF and were retained in the ASL, including 'Ter-Galoustian', 'Ambartsumi', 'Mikhailov', 'Nerses', and 'Oganesov'. Because we retained all Armenian surnames from the MESL, only those surnames from the MESL were truncated at 10 digits in the ASL. The final ASL has a total of 3,428 unique surnames.

Evaluation of the Armenian Surname List

There were 1,698 unique surnames identified in the CCR with birthplace in Armenia. Of the surnames in the CCR Birthplace in Armenia Surname List, 921 (54.2%) were also identified by the ASL (Figure 1.2). There were 18,944 surnames in the CCR file with a known birthplace country that had MESL surnames. Out of 921 surnames, the ASL identified 310 (33.7%) additional surnames in the CCR Birthplace in Armenia List than the MESL. Of the 3,428 surnames in the ASL, 1,678 (48.9%) were not identified in the CCR Birthplace in Armenia List nor the MESL.

Of the 3,428 surnames in the ASL, 2,349 (68.5%) had Armenia ranked first as the most likely country of origin and 431 (12.6%) as the second-most likely country of origin from NamSor (Figure 1.3). There were 257 (7.5%) with Iran listed first, 114 (3.3%) with Georgia Republic listed first, 63 (1.8%) with Russia listed first, and 59 (1.7%) with Romania listed first. Of the 1,698 surnames in the CCR Birthplace in Armenia List, 1,159 (68.3%) had Armenia ranked first as the most likely country of origin from NamSor (Figure 1.3). Iran followed with 138 (8.1%) of surnames, then Georgia Republic with 43 surnames (2.5%), Russia with 33 (1.9%) surnames, and Lebanon with 27 (1.6%) surnames.

Discussion

To our knowledge, this is the first study to create and evaluate a specific Armenian surname list for research purposes. Prior work in Middle-Eastern populations highlighted the need for future research to refine procedures pertaining to ethnic overlaps in the MESL.⁴⁸ The ASL is a significant step towards increasing clarity and understanding of public health patterns in Armenians. The ASL builds upon Armenian surnames in the MESL by additionally including surnames from over a hundred years of death records in California.

Access to death records in California to create the ASL is particularly advantageous because Armenian immigration to the state dates back to the 1800s and California remains the state with the largest population of Armenians in the United States.⁷ The utilization of the Soundex blocking strategy in the linkage between the MESL Armenian surnames and the CPUDF allowed for the identification of nearly 18% of the surnames added to the ASL. As such, when the ASL is used in probabilistic linkages with large databases, we recommend selecting Soundex as a blocking strategy during the linkage configuration process to maximize the number of Armenians identified.⁶⁴

During the process of creating the ASL, we considered historical and political events that may have impacted data collection and representation of countries of birth. Specifically, Armenia gained its independence from the USSR in 1991, raising the question of whether surnames of those born in Armenia during the USSR (1920-1991) had Armenia listed as their country of birth.⁷⁰ Our analyses showed that approximately 95% of records in the California death files with birthplace in Armenia were born between 1920-1990, suggesting that surnames of people born in Armenia during the USSR were successfully identified and included in the ASL. A common challenge faced by researchers during the creation of surname lists is the changes of surnames over time. Because Russian was the official language of the USSR, individual countries within the USSR developed onomastic systems that followed Russian naming conventions.⁷¹ Since Armenia was under USSR rule for the majority of the 20th century, we included Armenian surnames with common Russian suffixes -ov and -ova in the ASL, consistent with the MESL.^{66,72} Out of the 3,428 unique surnames in the ASL, 73 (2%) have a suffix of -ov or -ova and 70 (96%) of these surnames had a birthplace of Armenia in the CPUDF. Given the unique characteristics of Armenian surnames and the small percentage of surnames in the ASL with -ov or -ova suffixes, we do not expect a remarkable number of non-Armenian Russians to be identified by the ASL.^{65,66}

Another challenge in creating surnames lists is changing surnames due to marriage or to follow naming conventions of a new culture and society after immigration. To address this issue, we included all available father's last names from the California death records for both males and females. For example, father's last name of 'Danielian' was identified with the same record where the last name followed a different naming convention of 'Danielson'. A common limitation we faced is the lack of mother's last name, which

may have precluded the identification of children of Armenian mothers with an Armenian last name and fathers with non-Armenian last names.

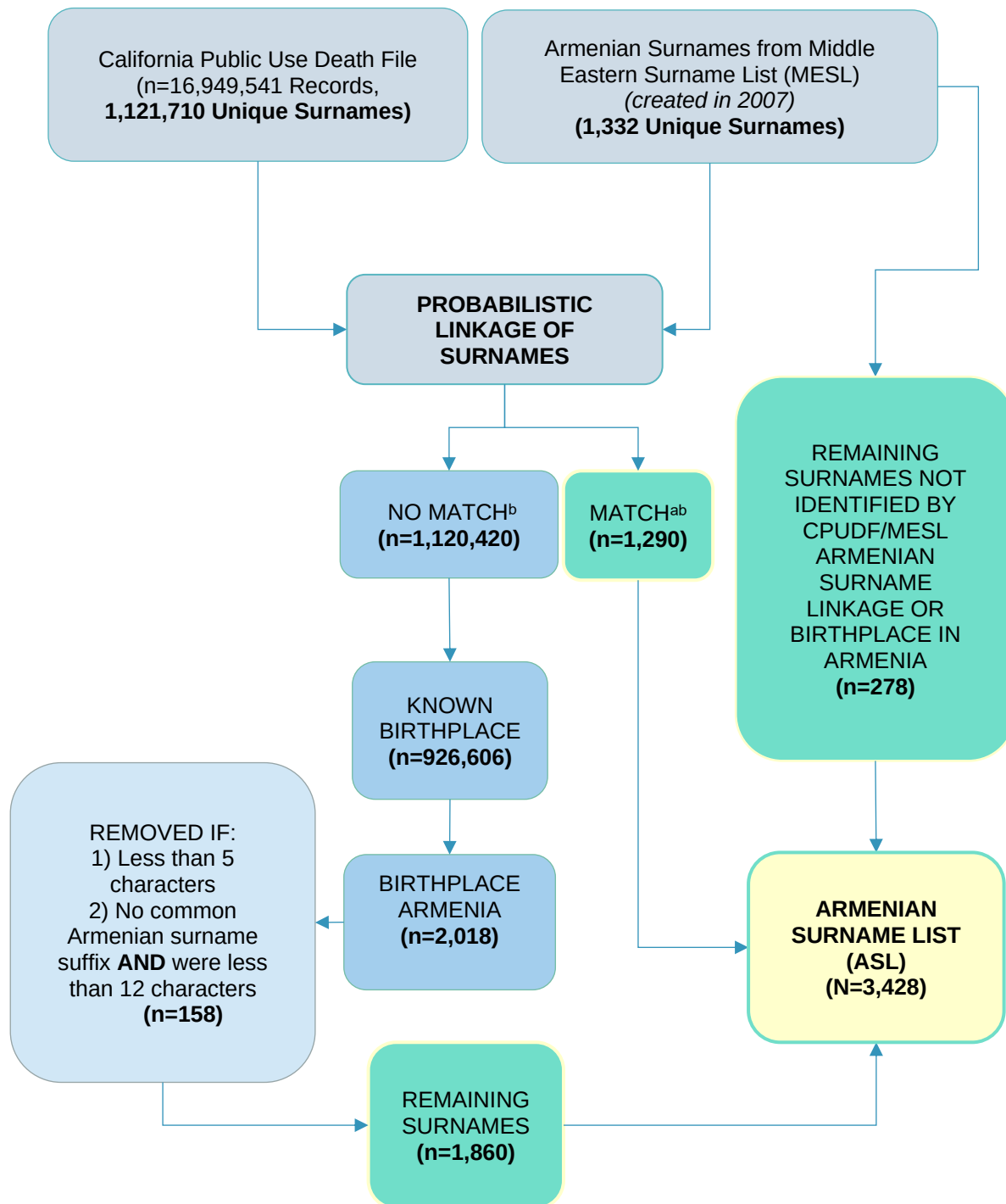
Using full surnames, as we did when creating the ASL, is particularly advantageous in large database linkages for several reasons. In probabilistic linkages, full surnames will have a higher similarity score when they match a full versus truncated version of the surname. Also, deterministic linkages can be utilized with non-truncated surnames. Lastly, name-checking technologies, such as NamSor, will return more accurate rankings of surname country of origin with full surnames, as many naming conventions, including Armenian, rely on the suffix to identify ethnic origin. Due to the novelty of this surname list, we did not have a 'gold standard' or contact people to confirm their ethnic origin to validate the ASL. While previous researchers who have created surname lists, such as the Arab surname list, have used self-reported ethnicity from telephone surveys as a 'gold standard' when validating new surname lists,⁷³ we did not have self-reported ethnicity for the purposes of this study. Therefore, we compared overlap of surnames between the CCR and the ASL.

Comparisons of the ASL with a list of MESL surnames from the CCR were suggestive that the ASL identifies a considerable percentage (34%) of Armenian surnames not identified by the MESL. This may relate to a few characteristics of the list of Armenian surnames extracted from the MESL. In addition, the Armenian surnames were extracted from NUMIDENT using a filter of birth in Armenia and may exclude surnames of people who were born in other countries and who had different surname conventions than those who were born in Armenia. For example, in the CPUDF and MESL Armenian surname linkage results, we observed that 83% of surnames with birthplace in Russia followed a surname convention of the suffix beginning with the letter *I* (*ian, ians, iants*), while 22% of surnames with birthplace in Armenia had surname suffixes of *ian, ians, iants* and 77% had surname suffixes beginning with the letter *Y* (*yan, yans, yants*). We attempted to address this within the ASL, as the surnames added to the ASL were not limited to birth in Armenia. A potential limitation of the ASL is that a small percentage of people born in Armenia may have a non-Armenian last name or may not have Armenian origin because we retained Armenian surnames from the MESL. Further research application of the ASL will clarify whether any enhancements to the ASL are warranted.

Country of origin checks using NamSor for the ASL surnames and the CCR Birthplace in Armenia surnames retrieved similar results, with Armenia ranked as the most likely or second-most likely country of origin for more than three-fourths of surnames, suggesting that the ability of the ASL to identify Armenian surnames is comparable to the ability of a list containing surnames identified solely by birthplace in Armenia. Because the two lists had only about 30% overlap between surnames, it may be worthwhile to consider the combination of the two lists to optimize the identification of Armenian surnames in large population-based databases.

The use of the ASL can increase the representation of Armenians in public health research. For example, the ASL can be used to identify Armenians in population-based cancer registry databases and observe cancer patterns by demographic and clinical factors. This will result in a clearer representation and understanding of cancer patterns among racial/ethnic categories where Armenians are commonly included, such as Middle-Eastern and White. Such data can inform health education and awareness campaigns to increase access to preventive healthcare services, such as screening, in the Armenian population. The ASL allows for the identification and representation of Armenians in a myriad of research endeavors and lays the groundwork as a valuable resource for research studies in public health, epidemiology, and other fields concerning Armenian communities.

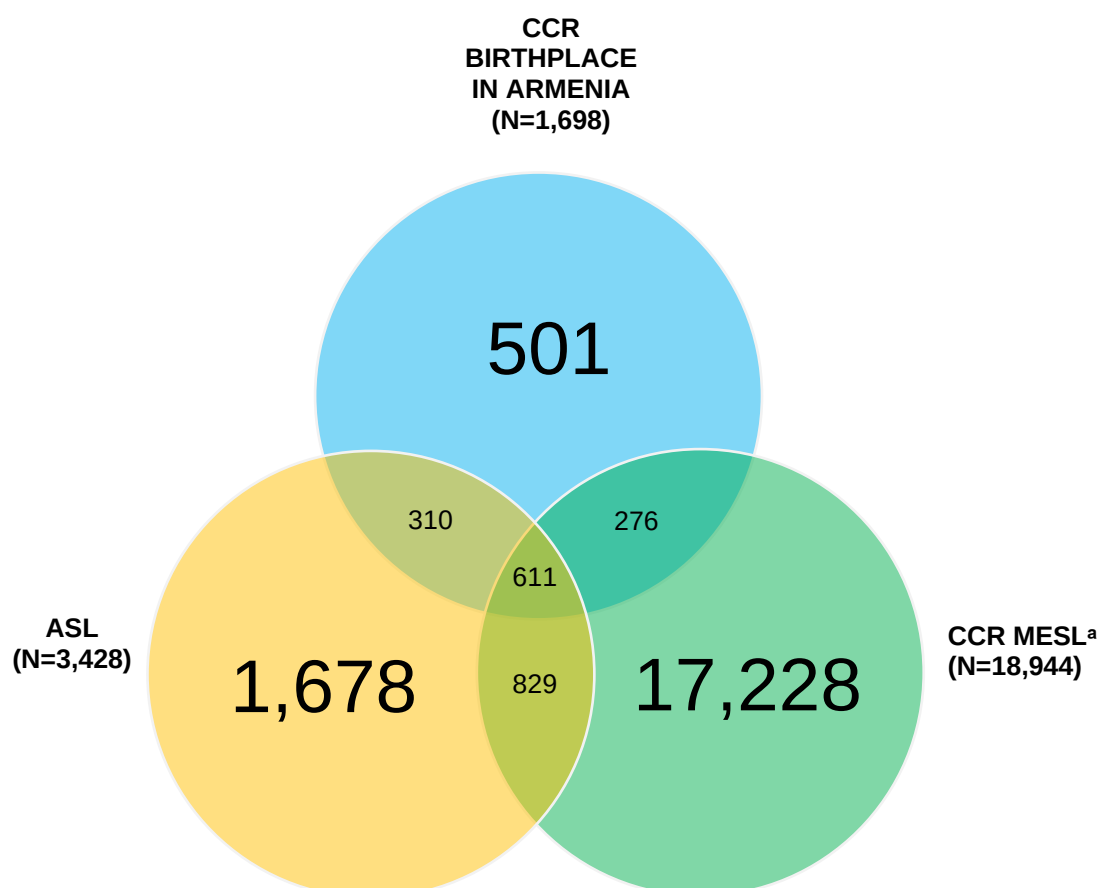
Figure 1.1 Flowchart of the Armenian Surname List (ASL) construction process



^aSoundex Blocking Strategy was used in Match*Pro Linkage Software, allowed for identification and addition of surnames with different spellings of the 1,332 MESL Armenian surnames.⁶⁴

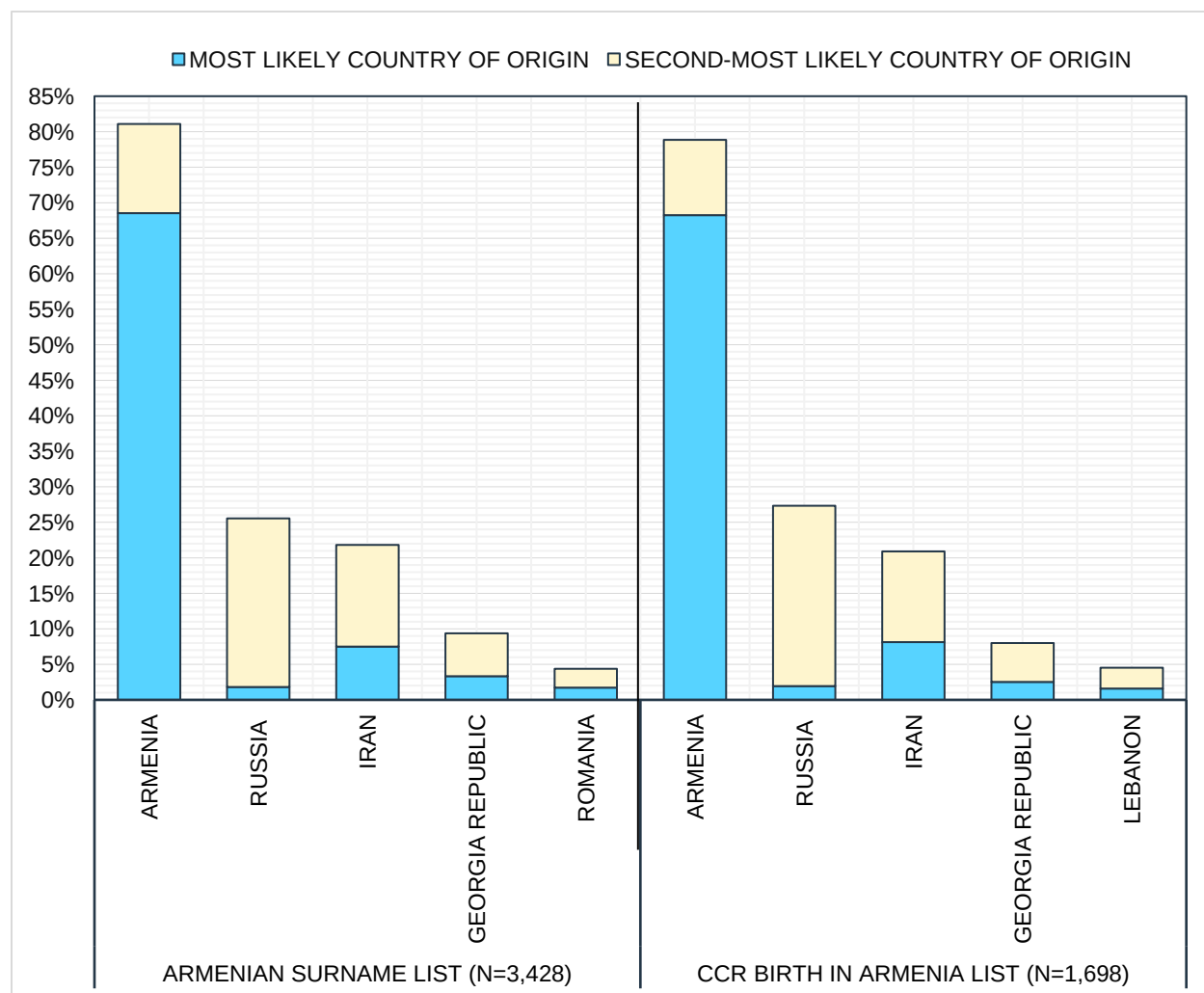
^bRefer to Supplemental Table 1.1 for Manual Review Categorization Rules.

Figure 1.2 Counts of overlapping surnames between the Armenian Surname List (ASL), California Cancer Registry (CCR) Birthplace in Armenia List, and Middle Eastern Surname List (MESL) from the CCR



^aIncludes all surnames with known country of birth in the California Cancer Registry that indicate surname is on the MESL created in 2007.

Figure 1.3 Percentage of surnames in the Armenian Surname List (ASL) and California Cancer Registry (CCR) Birthplace in Armenia List ranked by most likely country of origin and second-most likely country of origin^a



^aThe Country of Origin feature in NamSor was used to rank country of origin for each surname and returned a list of the top 10 countries of origin ordered from most likely to least likely.^{52,68} Only the top five countries of origin ordered by combined percentage of most likely and second most-likely country of origin are presented for each list.

Table 1.1 Number and percentage of surnames identified by the Middle Eastern Surname List (MESL)^a via probabilistic linkage with California Public Use Death File (CPUDF), 1905-2020

Identified by MESL Armenian Surnames ^a	Birthplace in Armenia in the California Public Use Death File (CPUDF)			
	Yes	No	Unknown	Total
Yes	720 (26.30)	485 (0.05)	85 (0.04)	1,290 (0.11)
No	2,018 (73.70)	924,679 (99.95)	193,819 (99.96)	1,120,516 (99.89)
Total	2,738 (100)	925,164 (100)	193,904 (100)	1,121,806 (100)

^aList of Armenian Surnames from the MESL created in 2007.

Supplemental Table 1.1 Manual review categorization rules, California Public Use Death File (CPUDF) 1905-2020 and Middle Eastern Surname List (MESL) Armenian Surname probabilistic linkage^b results

Rule	N Records	%	Similarity ^a
Match			
1 Exact Match	31,176	92.72	1.00
2 Spelling variation or truncation AND 'yan', 'ian', 'ians', 'yans', 'iants', 'yants', ending convention	2,135	6.35	0.87-0.98
3 Spelling variation or truncation AND ≥10 digits	30	0.09	0.88-0.98
Non-Match			
4 Did not meet any match categorization rule	282	0.84	0.87-0.98
Total	33,623	100	

^aBased on matching and blocking threshold settings selected on the linkage configuration page, the Match*Pro results file provides a similarity measure from 0-1 for each linked pair.

^bMatch*Pro Probabilistic Linkage Configuration:

- 1) On the Matching tab, the last name and father's last name variables from the CPUDF were selected as the File 1 Field, and the last name from the list of Armenian surnames from the MESL was selected as the File 2 Field.
- 2) On the Blocking tab, the only blocking field selected was last name.

Chapter 2

Title: Cancer incidence among Armenians in California

Authors: Ani S. Movsisyan, MS^{1,2}, Jeffrey S. Hoch, PhD^{1,3}, Laura Fejerman, PhD, MSc^{1,2}, and Theresa H. Keegan, PhD, MS^{1,2}

1. Department of Public Health Sciences, University of California Davis, Davis, CA
2. UC Davis Comprehensive Cancer Center, University of California Davis Medical Center, Sacramento, CA
3. Center for Healthcare Policy and Research, University of California, Davis, CA

Abstract

Introduction: California is home to the largest population of Armenians in the United States. The historical categorization of Armenians as 'White' or 'Some Other Race' in population databases has likely masked cancer incidence patterns in this population. This is the first study considering cancer incidence among Armenians in California.

Methods: We used the Armenian Surname List and birthplace information in the California Cancer Registry to identify Armenians with cancer diagnosed during 1988-2019. We calculated proportional incidence ratios (PIR) among Armenians compared with non-Hispanic Whites (NHWs). As an exploratory analysis, we calculated incidence rate ratios (IRR) during 2006-2015 using Armenian population denominators from the American Community Survey (ACS). We selected PIR as our primary method given uncertainty regarding the use of ACS population estimates for rate calculations.

Results: There were 27,212 cancer diagnoses among Armenians in California, 13,754 among males and 13,458 among females. Armenian males had notably higher proportions of stomach (PIR=2.39), thyroid (PIR=1.45), and tobacco-related cancers including bladder (PIR=1.53), colorectal (PIR=1.29), and lung (PIR=1.16) cancers. Higher proportional incidence of cancers including stomach (PIR=3.24), thyroid (PIR=1.47), and colorectal (PIR=1.29) were observed among Armenian females. Exploratory IRR analyses showed higher stomach (IRR=1.78), bladder (IRR=1.13), and colorectal (IRR=1.12) cancers among Armenian males and higher stomach (IRR=2.54) cancer among Armenian females.

Conclusion: We observed higher stomach, colorectal and thyroid cancer incidence among males and females and tobacco-related cancers among males. Further research is needed to refine Armenian population estimates and understand and address risk factors associated with specific cancers among Armenians in California.

Introduction

California is the largest and most diverse state in the United States,^{74,75} and home to the largest population of Armenians in the United States, with immigration to the state dating back to the 1870s.^{7,76} Given California's racial and ethnic diversity, numerous recent research studies have focused on cancer occurrence patterns in specific ethnic groups, including Hispanic, Asian, Middle-Eastern, and Arab populations in California.^{77–81} However, Armenians remain unrepresented in public health and epidemiological research due to their categorization as 'White' or 'Some Other Race' in population databases.^{7,43,82}

Prior studies on cancer occurrence among Middle-Eastern populations in California included the Armenian population and found proportionally higher stomach and thyroid cancers and proportionally lower lung and cervical cancers among Middle-Eastern females compared with non-Hispanic White (NHW) females.^{48,81} In addition, proportionally higher thyroid, bladder, and stomach cancers and lower lung cancer were observed among Middle-Eastern males compared with NHW males.^{50,81} While Armenians were included in the broad Middle-Eastern group, cancer risk factors, including tobacco-use, are higher in Armenia compared to surrounding countries and cancer mortality rates in Armenia are notably higher than neighboring Middle-Eastern countries, such as Iran and Iraq.^{32,39} Additionally, a hospital in Los Angeles County, where the majority of Armenians in California reside, found that a majority of their patients with stomach cancer were Armenian.⁸³ However, state-level analyses of cancer incidence by cancer type and sex among the Armenian population in California and in the United States have remained unknown, and it is unclear whether cancer patterns among Armenians differ from the NHW population with which they have been historically categorized in population-based research studies.^{47,82}

No prior studies, to our knowledge, have studied cancer incidence specifically among the Armenian population in California.⁴⁴ Therefore, we utilized the recently developed Armenian Surname List (ASL)⁴⁷ and birthplace data in the population-based California Cancer Registry to identify Armenians with cancer. Research on this growing minority group in California can reveal the cancers disproportionately impacting Armenians.

Methods

Data Sources

California Cancer Registry

The California Cancer Registry (CCR) is a statewide population-based cancer surveillance system and has collected cancer diagnoses in California since 1988, with reporting guidelines similar to those of the Surveillance, Epidemiology, and End Results (SEER) database.^{62,84} Consistently meeting data quality standards, the CCR is Gold Certified by the North American Association of Central Cancer Registries (NAACCR) and is one of the largest cancer registries in the world.^{62,85} As mandated by NAACCR, data reported to the CCR meet standards of completeness, accuracy, timeliness, and quality control measures.^{85,86} The CCR provides patient demographics, tumor characteristics including primary site, behavior, histology, morphology, and stage at diagnosis, first course of treatment, country of birth, county of diagnosis, and follow-up for vital status for all cancers, except for non-melanoma skin cancers.⁶⁰ For the presented analyses, we used variables including age at diagnosis, year of diagnosis, race/ethnicity, sex, tumor site, and behavior. We additionally presented demographic variables including county of diagnosis and country of birth.

Population Estimates

The American Community Survey (ACS) is a nationwide survey provided by the US Census Bureau and was used to obtain Armenian population estimates.⁸⁷ US Census respondents were asked “What is your ancestry or ethnic origin?”, and the ancestry variable is comprised of responses to this question.^{88–90} We used this ancestry variable to retrieve population detailed tables for the Armenian population in California, and among this population, we obtained additional demographic variables, including age, sex, place of birth, and county of residence.^{44,87,91} Armenian population estimates in California were retrieved from the ACS 2010 and 2015 5-year selected population detailed tables and included sex, age, and county variables, with data available for 19 of the 58 counties in California.^{44,87} Population estimates for the NHW population were obtained from the CCR, provided by the California Department of Finance and the Centers for Disease Control and Prevention National Center for Health Statistics branch.^{92,93}

Study Population

To identify Armenian cancer diagnoses, we used Match*Pro probabilistic linkage software to link the ASL with the CCR December 2021 incidence research file extract that includes primary cancers considered complete for diagnosis years 1988 through 2019.⁹⁴ We developed the ASL using data from the California Public Use Death Files from years 1905-2020, and an extract of Armenian surnames from the Middle Eastern Surname List that was developed in 2007.⁴⁷ The ASL contains 3,428 unique surnames that can be linked with last names in research databases to identify Armenians.⁴⁷ We also selected surnames from the CCR with birthplace in Armenia that were not in the ASL, and we linked this list of surnames to the CCR incidence research file extract. After manual review of both linkages, we combined the resulting matched records into one file. In the incidence file, we included all malignant cancer diagnoses among males and females using the International Classification of Diseases for Oncology, 3rd edition/World Health Organization (ICD-O-3/WHO) 2008 site codes.^{95,96} We extracted Armenians from the NHW, non-Hispanic Black, Hispanic, Asian/Pacific Islander, American Indian, Other/Unknown CCR race and ethnicity categories to create a separate Armenian group. We included Armenian and NHW patients of all ages who were male and female due to the availability of population estimates by sex. We compared Armenians to the NHW group because 96.2% of patients we identified as Armenian were previously identified as NHW in the CCR and this historical categorization of Armenians as NHW has likely masked true cancer incidence among the Armenian population in California.

Analyses

Demographic characteristics and cancer characteristics were stratified by sex and presented for the Armenian and NHW population. We calculated proportional incidence ratios (PIR), incidence rates (IR), and rate ratios (IRR) for the most frequent cancers among Armenian males and females compared with NHWs. Armenian population denominators were obtained from the American Community Survey (ACS) for IR calculations. The PIR is the observed number of Armenian cancer cases divided by the number of Armenian cancer cases expected if the Armenian population has the same proportion of cancer as the NHW population.^{81,97,98} Age-adjusted PIRs were calculated for cancers diagnosed from 1988-2019 and from 2006-2015 to compare directly with IRRs available only for years 2006-2015. Due to concerns raised

previously regarding the accuracy of ACS data denominators for small populations, we calculated the PIR as the primary method in this study.⁹⁹ The following PIR formula was used.

$$PIR = \frac{\sum A_i}{\sum B_i \left(\frac{C_i}{D_i} \right)}$$

where

A_i = number of Armenian site-specific cases for age group i

B_i = total number of Armenian cases for age group i

C_i = number of NHW site-specific cases for age group i

D_i = total number of NHW cases for age group i

ACS Armenian population denominator estimates from 2006-2015 were used to obtain IRs per 100,000 people age-adjusted to the US 2000 standard population in 18 age categories. We presented IRs and IRRs for the 10 most common cancers among Armenian males and females. We then compared the IRs of these top 10 cancers observed among Armenian males and females with the NHW racial and ethnic group as the reference population with IRRs and 95% confidence intervals during 2006-2015. SEER*Stat software (version 8.4.0) was used to retrieve age-adjusted IRs, while SAS version 9.4 software (SAS institute Inc., Cary, NC, USA) was used to obtain descriptive demographic data, PIRs, and IRRs.

Results

Relevant characteristics of Armenians diagnosed with cancer in California between years 1988-2019 are presented in Table 2.1 and Supplemental Table 2.2 is provided for comparison of characteristics with NHW patients. A total of 27,212 malignant cancer diagnoses among Armenians in California met inclusion criteria with nearly an equal frequency of cancer cases among males (50.5%) and females (49.5%). About 13% of diagnoses occurred among people less than 50 years old, 43% were between 50-69 years old, and 44% were aged 70 or greater. Most (76.3%) of the cancer cases were diagnosed in Los Angeles County, followed by Fresno (5.7%) and Orange (3.3%) counties. Most (55.2%) of the study population was foreign-born, primarily in Armenia (20.1%), Iran (12.7%), Lebanon (4.3%), Syria (3.6%), and Russia (3.5%). Overall, 12.8% were United States-born (US-born) and 32.0% had unknown place of birth.

Based on the Armenian population demographics in California as reported by the ACS, there were an estimated 241,323 Armenians in California from 2006-2010 and 259,430 from 2011-2015, representing a 7.5% population increase between the two time periods (Supplemental Table 2.1). Based on the 2015 5-year estimate, 61% of the population was <50 years old, 26% was 50-69 years old, and 13% was ≥70. A majority (75.6%) of the population resided in Los Angeles County, followed by Fresno (3.9%) and Orange (3.5%) counties. A majority (60.7%) of the population was foreign-born, primarily in Armenia (25.4%), Iran (18.2%), Lebanon (4.7%), Syria (2.7%), and Iraq (1.7%). More than a third (39.3%) of the population was born in the United States.

Among 13,754 Armenian males diagnosed with cancer between 1988-2019, the top 10 most frequent cancers in descending order were prostate (22.8%), lung (15.2%), colorectal (12.3%), bladder (11.1%), non-Hodgkin lymphoma (NHL) (4.3%), stomach (3.9%), leukemia (3.8%), kidney (3.6%), pancreas (2.5%), and liver and intrahepatic bile duct (IBD) (1.9%). Among 13,458 Armenian females diagnosed with cancer, the top 10 most frequent cancers in descending order were breast (32.6%), colorectal (12.1%), lung (6.2%), uterine (6.1%), NHL (4.3%), thyroid (4.2%), ovary (3.7%), stomach (3.0%), pancreas (2.9%), and leukemia (2.8%) (Table 2.2). For cancers diagnosed from 1988-2019, Armenian males had a significantly higher proportion of stomach (PIR=2.39, 95% CI=2.19-2.60), bladder (PIR=1.53, 95% CI=1.45-1.61), colorectal (PIR=1.29, 95% CI=1.23-1.35), lung (PIR=1.16, 95% CI=1.11-1.21), leukemia (PIR=1.16, 95% CI=1.06-1.26), liver and IBD (PIR=1.19, 95% CI=1.05-1.34), and kidney (PIR=1.11, 95% CI=1.01-1.21) cancers, and a significantly lower proportion of prostate (PIR=0.84, 95% CI=0.81-0.87) cancer compared to NHW males. In addition, while thyroid cancer was not among the top 10 most frequent for Armenian males, we observed a higher proportion of thyroid (PIR=1.45, 95% CI=1.25-1.67) cancer among Armenian males compared with NHW males. From 1988-2019, Armenian females had a significantly higher proportion of stomach (PIR=3.24, 95% CI=2.93-3.57), thyroid (PIR=1.47, 95% CI=1.36-1.60), colorectal (PIR=1.29, 95% CI=1.23-1.35), pancreatic (PIR=1.20, 95% CI=1.09-1.33), leukemia (PIR=1.20, 95% CI=1.08-1.32), NHL (PIR=1.15, 95% CI=1.05-1.24), and ovarian (PIR=1.14, 95% CI=1.04-1.24) cancers, and a significantly lower proportion of lung (PIR=0.49, 95% CI=0.46-0.53) cancer compared to NHW females.

In Table 2.3, we show the IRR and PIR for the top 10 most common cancers among Armenian males and females diagnosed from 2006 to 2015. Our results show significantly higher IRRs of stomach (IRR=1.78, 95% CI=1.55-2.06), bladder (IRR=1.13, 95% CI=1.05-1.23), and colorectal (IRR=1.12, 95% CI=1.03-1.20) cancers, and significantly lower prostate (IRR=0.64, 95% CI=0.60-0.68), NHL (IRR=0.72, 95% CI=0.63-0.81), leukemia (IRR=0.84, 95% CI=0.73-0.96), kidney (IRR=0.79, 95% CI=0.69-0.89), liver and IBD (IRR=0.72, 95% CI=0.59-0.87), and pancreatic (IRR=0.72, 95% CI=0.62-0.86) cancers among Armenian males compared to NHW males. Armenian males had an insignificantly lower IRR for thyroid (IRR=0.89, 95% CI=0.72-1.12) cancer. PIRs for males followed a similar pattern as described in Table 2.2, except that PIRs for liver and IBD (PIR=0.92, 95% CI=0.75-1.11), leukemia (PIR=1.10, 95% CI=0.95-1.26), kidney (PIR=1.05, 95% CI=0.92-1.19), and thyroid (PIR=1.19, 95% CI=0.95-1.47) cancers were not statistically significant. Among Armenian females, we observed significantly higher IRRs of stomach cancer (IRR=2.54, 95% CI=2.15-2.99) and significantly lower IRRs of breast (IRR=0.74, 95% CI=0.70-0.77), lung (IRR=0.36, 95% CI=0.33-0.41), and uterine (IRR=0.76, 95% CI=0.68-0.85) cancers. PIRs for females followed a similar pattern as in Table 2.2.

Discussion

To our knowledge, this is the first study showing cancer incidence specifically among the Armenian population in California. While other studies have previously included Armenians in a broad Middle-Eastern group in California,^{48,60} we were able to use our recently developed surname tool, the ASL,⁴⁷ and birthplace data to identify Armenians diagnosed with cancer in California. Compared with individuals in the NHW category, where Armenians have most commonly been included, we found significant differences in proportional incidence and incidence rates of many cancers. Notable findings include higher proportional incidence of stomach, colorectal, and thyroid cancer among Armenian males and females and tobacco-related cancers among Armenian males. Our findings can guide healthcare professionals and researchers toward targeted cancer screening interventions and etiologic studies for these cancers in Armenians.

Among the 10 most frequent cancers, Armenian males had higher proportions of seven cancers and higher incidence rates of three cancers, and Armenian females had higher proportions of seven cancers and a higher incidence rate of one cancer compared to NHWs. The correct interpretation of the

two methods is crucial for a thorough understanding of our study results; PIR analyses show which cancers are proportionally higher among Armenians diagnosed with cancer in California compared to NHWs, while IRR analyses show risk of cancer incidence compared to NHWs. For instance, a higher PIR for a particular cancer among Armenians does not mean that the risk of the respective cancer is higher than NHWs, but rather that the proportion of that cancer is higher than it is among NHWs.⁹⁷ The two methods convey different but equally important findings, as proportional analyses can guide use of cancer-related healthcare resources within the Armenian population, and incidence rate analyses can be used to discover and address risk factors for higher risk of certain cancer types among Armenians compared to NHWs. In the context of our study, we think our incidence rate calculations should be interpreted with caution due to uncertainty regarding accuracy of ACS population estimates.⁹⁹ While we encourage future studies to further refine Armenian population estimates, we believe the incidence rate data can provide complementary information to our proportional incidence analyses.

Our observation of a substantially higher proportion of stomach cancer among Armenians suggests a need for further exploration into the etiology of this cancer among this population. Risk factors for stomach cancer include a diet high in salt, smoked, processed, grilled, or barbecued meat, and low in fruits, vegetables, and fiber.^{100,101} Obesity, a known risk factor for stomach, colorectal, and multiple other cancers, is a known public health issue in Armenia, with over half the population being overweight (22%) or obese (29%).¹⁰² While the prevalence of obesity in the Armenian population of California or the United States is unknown, a previous study in a Los Angeles County hospital showed that Armenian patients had a higher proportion of hyperlipidemia compared to non-Armenian patients that may be related to either a diet high in saturated fats, red meat, or alcohol, or a possible genetic predisposition for hyperlipidemia among Armenians.⁴² In addition to diet and obesity, tobacco and alcohol use, low physical activity, family history of stomach cancer, and environmental exposure to infections such as *Helicobacter pylori* (*H. pylori*) and hepatitis B are all known risk factors.^{100,101} Some studies also suggest an association between *H. pylori* exposure and colorectal cancer, which was also proportionally higher among Armenians in our study.¹⁰³ The highest stomach cancer incidence rates in the world have been observed in Eastern Asia followed by Central and Eastern Europe,¹⁰⁰ and evidence suggests that first-generation immigrants to the United States may have had higher exposure to these harmful environmental agents in comparison to second-generation

immigrants.⁸¹ Further, a study conducted at a hospital in Los Angeles County with a large patient population of Armenian immigrants found that Armenians made up 61% of all stomach cancer cases diagnosed at the hospital.⁸³ Additionally, the study showed that Armenian immigrants from Armenia had lower rates of stomach cancer compared to immigrants from Middle-Eastern and other countries, providing supporting evidence of environmental effects on this observed disparity, as opposed to a genetic predisposition for stomach cancer.⁸³ Our findings of high stomach and colorectal cancer proportions suggest the need for further research considering environmental and generational factors, to better understand and address cancer risk factors among the Armenian population in California.

We also observed higher proportions of several tobacco-related cancers particularly among Armenian males, including lung, bladder, leukemia, kidney, liver and IBD.¹⁰⁴ Tobacco-use, mainly in the form of cigarettes, has been an ongoing public health issue among Armenian men.²⁷ Armenian men are estimated to have the 12th highest smoking rate in the world, and 28% of the Armenian population, 51.8% of men and 1.6% of women, are current smokers.^{27,105} While the percentage of women in Armenia who smoke is much lower than men, over half of women in Armenia are exposed to second-hand smoke at home.¹⁰⁶ In addition, among 30 low-middle income countries studied, pregnant women in Armenia have the highest percentage (70%) of daily second-hand smoke exposure.^{106,107} As of 2022, lung cancer had the highest incidence in Armenia among all cancers studied.²⁴ It was the most frequent cancer among men, and the fifth-most frequent among women, accounting for 22.5% of new cases among men and 5% among women.²⁴ In our study, lung cancer was the second-most common among men and third-most common among women. We also found that men had a significantly higher proportion, while women had a significantly lower proportion and incidence rate of lung cancer compared to NHWs, and a possible explanation for this finding may be partly attributed to the higher percentage of smoking among NHW women in California (7%), compared to Armenian women in Armenia (1.6%).¹⁰⁵ Further studies regarding tobacco-use behaviors among Armenians in California are needed to understand whether the high smoking prevalence observed in Armenia remains a pattern post-immigration. Our results, along with the high percentages of tobacco-use, lung cancer, and multiple tobacco-related cancers among men in Armenia,¹⁰⁸ suggest that tobacco-use behaviors among Armenian men in California warrant attention, including the implementation of culturally competent tobacco-cessation programs.

Our findings of higher proportional incidence of thyroid cancer among Armenians are consistent with previous studies among the Middle-Eastern group in California that included Armenians compared to NHWs.^{50,81} Although the primary known risk factor for thyroid cancer is exposure to radiation, thyroid cancer is often preceded by benign thyroid disease.¹⁰⁹ Iodine deficiency is a known risk factor for thyroid disease and may therefore be indirectly linked with eventual development of thyroid cancer.¹⁰⁹ Sources of radiation exposure over time in Armenia include the Chernobyl accident of 1986, the Armenian Nuclear Power Plant, tests of nuclear weapons in the 1950s and 1960s, mining, cosmic rays in mountainous regions, and natural radioactivity of urban soil.¹¹⁰ Research from the 1970s showed Armenia to be an iodine-deficient country after discovering high prevalence of goiter among children and adults.^{111,112} In 1995, a survey found that half of pregnant women and 40% of 6-12 year old children had goiter.¹¹² In the 1990s, even after a major salt producer in Armenia iodized table salt, goiter prevalence remained high in the next 10 years.¹¹¹ As such, in 2004, the Government of Armenia initiated a salt iodization strategy, requiring the fortification of food and salt with iodine.¹¹¹ While this national strategy successfully reduced levels of iodine deficiency, recent studies have still observed possible high risk of thyroid disorders in Armenia.¹¹³ Our findings highlight the need to better understand the etiology of and risk factors for thyroid cancer among different generations of Armenians in California.

Our study has limitations. One limitation is that we used ACS data to calculate age-adjusted incidence rates among Armenians and previous studies have shown inconsistent incidence rates that were calculated using ACS population estimates compared with SEER population estimates among small populations.⁹⁹ Specifically, our population estimates would not capture Armenians not identified by the ancestry variable in ACS, potentially underestimating the numbers of Armenians in California. Given this limitation, we advocate for continued efforts in collecting accurate population estimates to further assess cancer incidence among Armenians to confirm our findings. Another limitation is the high percentage of missing country of birth in the CCR (32%), impacting our ability to compare birthplace between the CCR and ACS data. However, prior studies observed that those with missing country of birth in the CCR are more likely to be US-born rather than foreign-born,^{114,115} suggesting that a significant proportion of the patients with unknown country of birth are actually US-born. This is further supported by the only somewhat lower proportions of foreign-born Armenians identified in the CCR (55.2%) than the ACS (60.7-61.6%).

While we used available father's last names and maiden names, in addition to last names, to identify Armenians with cancer, a common limitation we faced was the lack of mother's last name in the CCR, which may have prevented the identification of cancer patients who have Armenian mothers with an Armenian last name and fathers with non-Armenian last names. In addition, Armenians with non-Armenian last names, either due to marriage or other changes, may have not been identified particularly in cases of missing maiden names and father's last names. Additionally, the data we used to calculate incidence rates included Armenians from different generations, and different generations of Armenians likely have differing cancer incidence due to the effects of acculturation on cancer risk.⁸¹ While we did not have a 'gold standard' list or contact people to confirm their ethnic origin for validation of the ASL, we evaluated the ASL previously, including using NamSor, an onomastic classification tool, and found that 81% of surnames had Armenia listed as the most or second-most likely country of origin.⁴⁷ Despite these limitations, a significant strength in our study was that this was the first research application of the newly developed ASL as a tool for identifying Armenians in large research databases that can be utilized by other cancer registries.

Our study findings highlight numerous cancer patterns among the Armenian population, including disproportionately higher stomach, colorectal, and thyroid cancers among men and women and tobacco-related cancers among men. Our findings of significantly different incidence rates of several cancers among Armenians diagnosed between 2006 to 2015 have potential to guide healthcare professionals serving the Armenian community in California toward targeted cancer screening interventions. Most notably, the significantly higher incidence rate ratios and proportional incidence ratios of stomach cancer observed among Armenian males and females provide further confirmation that public health interventions to understand stomach cancer etiology are necessary. In addition, our findings suggest the need for further research considering nativity and generational factors, to better understand and address cancer etiology among the Armenian population in California. We encourage researchers and healthcare professionals, particularly in areas of California with the largest Armenian populations, such as Los Angeles and Fresno counties, to use these data as a basis to examine cancers disproportionately impacting Armenian populations and for developing culturally competent interventions to address risk factors such as tobacco-use, obesity, and exposure to harmful environmental agents.

Table 2.1 Characteristics of Armenians diagnosed with cancer in California, California Cancer Registry, 1988-2019

Characteristics	Male (n=13,754)		Female (n=13,458)		Total (N=27,212)	
	n	%	n	%	N	%
Age at diagnosis						
0 to 9	69	(0.5)	51	(0.4)	120	(0.4)
10 to 19	66	(0.5)	65	(0.5)	131	(0.5)
20 to 29	157	(1.1)	173	(1.3)	330	(1.2)
30 to 39	312	(2.3)	574	(4.3)	886	(3.3)
40 to 49	717	(5.2)	1,380	(10.3)	2,097	(7.7)
50 to 59	2,048	(14.9)	2,474	(18.4)	4,522	(16.6)
60 to 69	4,002	(29.1)	3,197	(23.8)	7,199	(26.5)
70+	6,383	(46.4)	5,544	(41.2)	11,927	(43.8)
County						
Los Angeles	10,722	(78.0)	10,054	(74.7)	20,776	(76.3)
Fresno	724	(5.3)	840	(6.2)	1,564	(5.7)
Orange	441	(3.2)	465	(3.5)	906	(3.3)
San Diego	230	(1.7)	264	(2.0)	494	(1.8)
Sacramento	159	(1.2)	207	(1.5)	366	(1.3)
Santa Clara	162	(1.2)	202	(1.5)	364	(1.3)
San Francisco	145	(1.1)	179	(1.3)	324	(1.2)
Other	1,171	(8.5)	1,247	(9.3)	2,418	(8.9)
Country of Birth						
Armenia ^a	2,775	(20.2)	2,687	(20.0)	5,462	(20.1)
United States	1,589	(11.6)	1,898	(14.1)	3,487	(12.8)
Iran	1,803	(13.1)	1,652	(12.3)	3,455	(12.7)
Lebanon	604	(4.4)	555	(4.1)	1,159	(4.3)
Syria	537	(3.9)	436	(3.2)	973	(3.6)
Russia	458	(3.3)	508	(3.8)	966	(3.5)
Turkey	463	(3.4)	346	(2.6)	809	(3.0)
Greece	244	(1.8)	211	(1.6)	455	(1.7)
Iraq	176	(1.3)	179	(1.3)	355	(1.3)
Egypt	116	(0.8)	115	(0.9)	231	(0.8)
Other	545	(4.0)	599	(4.5)	1,144	(4.2)
Unknown	4,444	(32.3)	4,272	(31.7)	8,716	(32.0)

^aIncludes Asian Republics of the USSR.

Table 2.2 Age-adjusted proportional incidence ratios for the 10 most common cancers in Armenian patients compared to NHW patients, California Cancer Registry, 1988-2019

Cancer Site	Observed	Expected	PIR	95% CI
Male				
Prostate	3,132	3,741.82	0.84*	(0.81-0.87)
Lung	2,088	1,799.24	1.16*	(1.11-1.21)
Colorectal	1,695	1,318.04	1.29*	(1.23-1.35)
Bladder	1,531	1,002.81	1.53*	(1.45-1.61)
NHL	587	622.39	0.94	(0.87-1.02)
Stomach	530	222.15	2.39*	(2.19-2.60)
Leukemia	519	448.59	1.16*	(1.06-1.26)
Kidney	501	452.73	1.11*	(1.01-1.21)
Pancreas	348	338.80	1.03	(0.92-1.14)
Liver and IBD	257	216.01	1.19*	(1.05-1.34)
Female				
Breast	4,380	4,349.37	1.01	(0.98-1.04)
Colorectal	1,634	1,269.37	1.29*	(1.23-1.35)
Lung	839	1,702.06	0.49*	(0.46-0.53)
Uterine	818	813.01	1.01	(0.94-1.08)
NHL	573	499.97	1.15*	(1.05-1.24)
Thyroid	564	382.58	1.47*	(1.36-1.60)
Ovary	498	437.15	1.14*	(1.04-1.24)
Stomach	400	123.56	3.24*	(2.93-3.57)
Pancreas	393	327.10	1.20*	(1.09-1.33)
Leukemia	385	321.59	1.20*	(1.08-1.32)

PIR: Proportional Incidence Ratio

NHL: non-Hodgkin lymphoma

IBD: intrahepatic bile duct

*Statistically significant.

Table 2.3 Age-Adjusted incidence rates and rate ratios and proportional incidence ratios of the top 10 cancers among Armenian patients compared with NHW patients, California Cancer Registry, 2006-2015

	Armenian			NHW								
Cancer Site	Count	IR	(95% CI)	Count	IR	(95% CI)	IRR	(95% CI)	O	E	PIR	(95% CI)
Male												
Prostate	1,169	80.0	(75.4-84.6)	131,406	125.3	(124.6-126.0)	0.64*	(0.60-0.68)	1,169	1,409.16	0.83*	(0.78-0.88)
Lung	832	57.7	(53.7-61.6)	59,198	59.9	(59.4-60.4)	0.96	(0.89-1.03)	832	643.42	1.29*	(1.21-1.38)
Colorectal	711	50.3	(46.5-54.0)	44,391	45.1	(44.6-45.5)	1.12*	(1.03-1.20)	711	479.65	1.48*	(1.38-1.60)
Bladder	649	45.1	(41.7-48.8)	38,850	39.8	(39.4-40.2)	1.13*	(1.05-1.23)	649	421.97	1.54*	(1.42-1.66)
NHL	253	18.2	(15.9-20.2)	24,407	25.4	(25.1-25.7)	0.72*	(0.63-0.81)	253	265.26	0.95	(0.84-1.08)
Stomach	198	13.9	(12.0-16.0)	7,692	7.8	(7.6-8.0)	1.78*	(1.55-2.06)	198	83.19	2.38*	(2.06-2.74)
Leukemia	211	15.7	(13.6-18.0)	17,606	18.8	(18.5-19.1)	0.84*	(0.73-0.96)	211	192.27	1.10	(0.95-1.26)
Kidney	228	15.9	(13.9-18.2)	20,158	20.3	(20.0-20.6)	0.79*	(0.69-0.89)	228	217.24	1.05	(0.92-1.19)
Pancreas	145	10.1	(8.5-11.9)	13,810	13.8	(13.6-14.1)	0.72*	(0.62-0.86)	145	149.35	0.97	(0.82-1.14)
Liver and IBD	108	7.4	(6.1-9.0)	10,988	10.3	(10.1-10.5)	0.72*	(0.59-0.87)	108	117.78	0.92	(0.75-1.11)
Female												
Breast	1,664	102.9	(97.8-108.0)	150,964	139.7	(138.9-140.4)	0.74*	(0.70-0.77)	1,664	1,716.53	0.97	(0.92-1.02)
Colorectal	661	36.8	(33.9-39.7)	41,490	35.3	(34.9-35.6)	1.04	(0.96-1.13)	661	459.94	1.44*	(1.33-1.55)
Lung	343	18.7	(16.7-20.7)	60,259	50.7	(50.3-51.2)	0.36*	(0.33-0.41)	343	661.20	0.52*	(0.47-0.58)
Uterine	330	19.4	(17.3-21.5)	29,217	25.6	(25.3-25.9)	0.76*	(0.68-0.85)	330	327.29	1.01	(0.90-1.12)
NHL	257	14.9	(13.1-17.0)	18,780	16.6	(16.3-16.8)	0.90	(0.79-1.02)	257	210.09	1.22*	(1.08-1.38)
Thyroid	254	17.3	(15.1-19.5)	16,646	19.1	(18.8-19.4)	0.91	(0.79-1.03)	254	204.43	1.24*	(1.09-1.41)
Ovary	200	11.8	(10.2-13.7)	14,382	13.1	(12.9-13.4)	0.90	(0.78-1.04)	200	163.44	1.22*	(1.06-1.41)
Stomach	152	8.7	(7.3-10.3)	4,041	3.4	(3.3-3.5)	2.54*	(2.15-2.99)	152	44.63	3.41*	(2.89-3.99)
Pancreas	168	9.2	(7.9-10.8)	12,790	10.5	(10.3-10.7)	0.88	(0.75-1.02)	168	139.64	1.20*	(1.03-1.40)
Leukemia	164	9.7	(8.2-11.4)	12,108	11.2	(10.9-11.4)	0.87	(0.74-1.02)	164	136.33	1.20*	(1.03-1.40)

IR: Age-adjusted incidence rate per 100,000 people, IRR: Rate Ratio, PIR: Age-adjusted Proportional Incidence Ratio, CI: Confidence Interval, O: Observed, E: Expected, NHL: non-Hodgkin lymphoma, IBD: intrahepatic bile duct.

*Statistically significant at $p < 0.05$.

Supplemental Table 2.1 Armenian population demographics in California, American Community Survey, 2006-2015

Characteristics	2006-2010		2011-2015	
	N	%	N	%
Population				
California population	36,637,290		38,421,464	
Armenian population	241,323	(0.7)	259,430	(0.7)
Sex				
Male	117,867	(48.8)	125,075	(48.2)
Female	123,456	(51.2)	134,355	(51.8)
Age, y				
0 to 9	24,222	(10.0)	25,183	(9.7)
10 to 19	30,333	(12.6)	29,099	(11.2)
20 to 29	36,357	(15.1)	38,276	(14.8)
30 to 39	30,122	(12.5)	34,539	(13.3)
40 to 49	35,819	(14.8)	32,731	(12.6)
50 to 59	34,985	(14.5)	38,764	(14.9)
60 to 69	21,885	(9.1)	27,920	(10.8)
70+	27,600	(11.4)	32,918	(12.7)
County				
Los Angeles	179,279	(74.3)	196,075	(75.6)
Fresno	10,602	(4.4)	10,052	(3.9)
Orange	10,226	(4.2)	9,136	(3.5)
San Diego	5,503	(2.3)	5,292	(2.0)
Santa Clara	3,349	(1.4)	3,098	(1.2)
Sacramento	3,318	(1.4)	5,440	(2.1)
San Mateo	2,791	(1.2)	2,379	(0.9)
Other	18,856	(7.8)	21,605	(8.3)
Unknown	7,399	(3.1)	6,353	(2.4)
Nativity				
Foreign-Born	148,638	(61.6)	157,566	(60.7)
United States-Born	92,685	(38.4)	101,864	(39.3)
Country of Birth				
United States	92,685	(38.4)	101,864	(39.3)
Armenia	61,300	(25.4)	65,776	(25.4)
Iran	39,817	(16.5)	47,197	(18.2)
Lebanon	12,606	(5.2)	12,267	(4.7)
Syria	7,733	(3.2)	6,947	(2.7)
Iraq	4,484	(1.9)	4,369	(1.7)
Turkey	3,156	(1.3)	2,749	(1.1)
Russia	2,910	(1.2)	3,023	(1.2)
Other	16,632	(6.9)	15,238	(5.9)

Source: American Community Survey 2010 and 2015 5-Year Selected Population Profile.⁵⁴

Supplemental Table 2.2 Characteristics of Armenian and NHW patients diagnosed with cancer in California, California Cancer Registry, 1988-2019

Characteristics	Armenian (n=27,212)		NHW (n=3,094,510)	
	n	%	n	%
Year of Diagnosis				
1988-1998	5,572	(20.5)	1,025,176	(33.1)
1999-2009	9,658	(35.5)	1,081,555	(35.0)
2010-2019	11,982	(44.0)	987,779	(31.9)
Age at Diagnosis				
0 to 9	120	(0.4)	9,705	(0.3)
10 to 19	131	(0.5)	10,663	(0.3)
20 to 29	330	(1.2)	34,552	(1.1)
30 to 39	886	(3.3)	92,192	(3.0)
40 to 49	2,097	(7.7)	223,596	(7.2)
50 to 59	4,522	(16.6)	484,491	(15.7)
60 to 69	7,199	(26.5)	801,715	(25.9)
70+	11,927	(43.8)	1,437,596	(46.5)
Sex				
Male	13,754	(50.5)	1,600,579	(51.7)
Female	13,458	(49.5)	1,493,931	(48.3)
Cancer Type				
Breast	4,422	(16.3)	478,260	(15.5)
Colorectal	3,329	(12.2)	300,226	(9.7)
Prostate	3,132	(11.5)	434,593	(14.0)
Lung	2,927	(10.8)	405,193	(13.1)
Bladder	1,893	(7.0)	152,827	(4.9)
NHL	1,160	(4.3)	129,014	(4.2)
Stomach	930	(3.4)	40,212	(1.3)
Leukemia	904	(3.3)	86,913	(2.8)
Uterine	818	(3.0)	89,563	(2.9)
Kidney	764	(2.8)	81,330	(2.6)
Other	6,933	(25.5)	896,379	(29.0)
County				
Los Angeles	20,776	(76.3)	607,854	(19.6)
Fresno	1,564	(5.7)	60,976	(2.0)
Orange	906	(3.3)	274,888	(8.9)
San Diego	494	(1.8)	295,728	(9.6)
Sacramento	366	(1.3)	132,726	(4.3)
Santa Clara	364	(1.3)	133,973	(4.3)
San Francisco	324	(1.2)	68,389	(2.2)

Other	2,418	(8.9)	1,519,976	(49.1)
Nativity				
United States-Born	3,487	(12.8)	1,575,466	(50.9)
Foreign-born	15,009	(55.2)	177,886	(5.7)
Unknown	8,716	(32.0)	1,341,158	(43.3)

NHL: non-Hodgkin lymphoma

IBD: intrahepatic bile duct

Chapter 3

Title: Stage at diagnosis and cancer-specific survival for stomach, lung, colorectal, and bladder cancers among Armenians in California

Authors: Ani S. Movsisyan, MS^{1,2}, Jeffrey S. Hoch, PhD^{1,3}, Laura Fejerman, PhD, MSc^{1,2}, and Theresa H. Keegan, PhD, MS^{1,2}

1. Department of Public Health Sciences, University of California Davis, Davis, CA
2. UC Davis Comprehensive Cancer Center, University of California Davis Medical Center, Sacramento, CA
3. Center for Healthcare Policy and Research, University of California, Davis, CA

Abstract

Introduction: While California is home to the largest population of Armenians in the United States, the historical categorization of Armenians as 'White' or 'Some Other Race' in population databases has limited our knowledge of stage diagnosis and cancer-specific survival for tobacco-related cancers among Armenians.

Methods: We used the Armenian Surname List and birthplace information in the California Cancer Registry to identify Armenians with cancer diagnosed during 1988-2019. We used Cox proportional hazards models to calculate cancer-specific survival and logistic regression models to calculate odds of late-stage diagnoses among Armenian and non-Hispanic White (NHW) patients for stomach, lung, colorectal, and bladder cancers.

Results: There were 639,224 stomach, lung, colorectal, and bladder cancer diagnoses identified from 1988 to 2019, 6,642 among Armenian patients and 632,582 among NHW patients. Higher proportions of Armenians resided in the lowest socioeconomic status (SES) neighborhoods (31.1%), had public insurance (56.4%), and were foreign-born (57.9%) compared to NHW individuals (23.6%, 24.7%, and 5.8%, respectively). Armenian patients were significantly more likely to be diagnosed with late-stage colorectal (OR=1.12, 95% CI=1.03-1.22), lung (OR=1.26, 95% CI=1.12-1.42), and stomach (OR=1.43, 95% CI=1.17-1.74) cancers, and low neighborhood SES and public insurance were significantly associated with late-stage diagnoses. In multivariable models adjusted for sociodemographic and clinical factors, Armenian patients had better survival than NHWs for stomach (HR=0.85, 95% CI=0.76-0.94), lung (HR=0.86, 95% CI=0.82-0.91), colorectal (HR=0.82, 95% CI=0.77-0.88), and bladder (HR=0.87, 95% CI=0.76-0.99) cancers.

Conclusion: While Armenian patients were at greater risk of late-stage diagnoses of colorectal, lung, and stomach cancers, we observed better survival among Armenian patients compared to NHW patients. Further research is needed to understand factors impacting survival in Armenians, including genetic, behavioral, and social factors. Our findings of lower SES and public health insurance associated with late-stage diagnoses suggest a need for increased access to care and cancer screening among Armenians.

Introduction

Cancer continues to be a leading cause of death in California, with nearly 60,000 deaths per year.¹¹⁶ While previous population-based research studies in the large and racially/ethnically diverse state of California have observed differences in cancer survival among individuals in the non-Hispanic Black, Hispanic, Asian/Pacific Islander, and American Indian categories compared to those in the non-Hispanic White (NHW) category, cancer survival has not been studied among many growing race/ethnic sub-categories, including Armenians.^{77,117–122} California is home to the largest population of Armenians in the United States.⁷ Out of the 259,430 Armenians in California as estimated by the American Community Survey, a majority (60.7%) are foreign-born.⁵⁴ However, due to a lack of a direct and consistent identification method, Armenians in the United States have historically been included in the ‘White’ or ‘Some Other Race’ race/ethnicity categories in population-based research studies despite this population having unique genetic, behavioral, environmental, and social factors that may impact survival.^{42,43,123,124} The historical grouping of Armenians within a broad NHW category has likely masked cancer survival patterns among the Armenian population.⁴⁷

Tobacco-use, a significant risk factor for the development of several cancers, remains a major public health issue in Armenia.²⁷ An estimated 52% of men and 2% of women are current smokers, and over half of tobacco-related annual deaths occur among individuals younger than 70 years old.²⁷ In addition to causing cancer, tobacco-use negatively impacts effectiveness of cancer treatment and is associated with lower survival after diagnosis.^{125,126} Our prior work using the recently developed Armenian Surname List (ASL) as well as birthplace and surname data from the California Cancer Registry (CCR) showed significantly higher incidence of several cancers that are tobacco-related including stomach, lung, colorectal, and bladder cancers among Armenians compared to NHWs in California.¹²⁷ However, no studies, to our knowledge, have assessed cancer survival data among Armenians compared to NHWs in California. Cancer survival may differ, as cancer mortality rates are notably higher in Armenia (10.1 per 100,000) than in NHW individuals in California (2.1 per 100,000), and mortality rates in Armenia have been increasing in recent years.^{37,128,129} Further, previous studies have also provided evidence for lower cancer survival in Armenia compared to high-income countries, including the United States.³⁶ Differences in

survival may also relate to stage at diagnosis, as nearly half (49%) of all cancers, and a majority of several tobacco-related cancers including stomach (70%), lung (82%), and colorectal (56%) cancers in Armenia, are diagnosed at late-stage.^{36–38}

Understanding differences in stage at diagnosis and survival is a critical step to identify and reduce any disparities in cancer outcomes among Armenians. Therefore, we examined cancer survival for four common tobacco-related cancers observed among Armenians.⁴⁷ Specifically, we compared odds of late-stage diagnosis and cancer-specific survival for stomach, lung, colorectal, and bladder cancers between Armenian and NHW patients. Additionally, we examined characteristics associated with late-stage diagnoses among the Armenian patient population.

Methods

Data Sources

The California Cancer Registry (CCR) is a statewide population-based cancer surveillance system with similar reporting guidelines as the Surveillance, Epidemiology, and End Results (SEER) database and has collected cancer diagnoses in California since 1988.¹³⁰ The CCR is one of the largest cancer registries in the world and is Gold Certified by the North American Association of Central Cancer Registries (NAACCR).¹³⁰ Information available in the CCR includes patient demographics, primary tumor site, tumor morphology, stage at diagnosis, first course of treatment, country of birth, and county of diagnosis. The CCR obtains vital status and cause of death information through routine linkages with state and national databases, providing follow-up information for vital status for all cancers except non-melanoma skin cancers.

Study Population

To identify Armenians with cancer diagnoses, we used Match*Pro probabilistic linkage software to link the ASL and surnames from the CCR with birthplace in Armenia that were not in the ASL with the CCR December 2021 incidence file extract containing primary cancers complete for years of diagnosis 1988 through 2019.^{47,94,127} We extracted Armenians from the NHW and other race and ethnicity categories to create a separate, mutually exclusive Armenian group. We selected data from the CCR on Armenian and

NHW patients diagnosed with first-primary, malignant stomach, lung, colorectal, and bladder cancers from 1988 to 2019. We excluded autopsy and death certificate diagnoses and records with missing survival time (n=13,112). We also excluded unknown/other sex categories (n=72). Tumor sites were selected based on International Classification of Diseases for Oncology, 3rd edition/World Health Organization (ICD-O-3/WHO) 2008 site codes and are shown in Supplemental Table 3.1.^{95,96}

Sociodemographic and Clinical Variables

Sex was categorized as male and female. We used SEER Summary Stage to define stage at diagnosis (localized, regional, remote, and unknown), with late-stage defined as regional or remote.¹³¹ Health insurance types were organized into five categories: 1) not insured; 2) private included private insurance not otherwise specified (NOS), fee for service, managed care health maintenance organization (HMO), or preferred provider organization (PPO), Medicare with supplement, NOS, Medicare with private supplement, and TRICARE; 3) public included Medicaid, Medicaid administered through managed care plan, Medicare without supplement, Medicare, NOS, Medicare administered through a managed care plan, and Medicare with Medicaid eligibility; 4) other public included Military, Veterans Affairs, Indian/Public Health Service, and county funded, NOS; and 5) unknown, as done previously.¹³² For logistic regression analyses, we combined unknown with not insured, and public with other public due to small numbers. Neighborhood SES (nSES) was categorized into lowest, middle, and highest groups.¹³³ We also considered marital status (married/in partnership, previously married, single/never married, and unknown), nativity (US-born, foreign-born, and unknown), and number of comorbidities (none, one or two, three or more, and unknown).^{122,134} We categorized known cigarette use (not including electronic cigarettes) as of diagnosis date into current use and never/former use for patients diagnosed during 2014-2019 when this data was more routinely collected.

Analyses

Demographic characteristics and cancer characteristics were presented for the Armenian and NHW population. For descriptive analyses of cigarette use during 2014-2019, we excluded unknown values of known cigarette use at the time of diagnosis (n=32,634). To examine odds of late-stage diagnoses among Armenian patients compared to NHW patients, and to assess sociodemographic factors associated with late-stage diagnoses among Armenian patients, we used multivariable logistic regression models adjusted

for year of diagnosis, age at diagnosis, sex, nSES, insurance type, marital status, comorbidity, and reporting regional cancer registry. Given the general Medicare eligibility age of 65 years,¹³⁵ we additionally examined the relationship between insurance type and late-stage diagnoses within separate age groups of <65 and 65+ years old.

We used multivariable Cox proportional hazards regression models to calculate hazard ratios (HRs) and associated 95% confidence intervals (Cis). Survival time was calculated using date of cancer diagnosis and date of last contact (follow-up) or death. Patients who died of other or unknown causes, or patients who were alive at the study cut-off date of December 31, 2019, were censored at that time. We compared stomach, lung, colorectal, and bladder cancer survival among Armenians and NHWs, adjusting for variables consistent with previous cancer survival studies, including year of diagnosis, reporting regional registry, marital status, age, tumor site and histology, nSES, sex, insurance type, first course of treatment, comorbidity, and stage at diagnosis.⁷⁷ For each model, we assessed the proportional hazards assumption by visually inspecting survival curves (log(-log) of the survival probability) by log (survival months of active follow-up).¹³⁶ Variables that violated the proportional hazards assumption were then stratified in the final model. After observing a significant interaction between race/ethnicity and stage at diagnosis for lung cancer ($p=0.017$), we assessed survival by each level of cancer stage. To estimate patients lost to follow-up (LTFU), we calculated the time interval in years between last follow-up date and study cut-off date among those who were alive. Alive patients with last follow-up date more than two years prior to the study cut-off date were defined as LTFU. Sensitivity analyses assuming all patients LTFU and only foreign-born patients who were LTFU died at the date of last contact were conducted to determine the impact on survival differences between Armenians and NHWs. All analyses were conducted using SAS version 9.4 software (SAS Institute Inc., Cary, NC, USA).

Results

A total of 639,224 cancer cases were identified from the CCR from 1988 to 2019, 6,642 among Armenians and 632,582 among NHW individuals (Table 3.1). Among Armenians, a higher percentage (43.4%) of diagnoses occurred in the most recent years of diagnosis compared to NHW patients (26.1%). More diagnoses were observed among Armenian males (62.4%) and NHW males (54.9%) compared to females

in each group. Higher proportions of Armenians resided in the lowest nSES neighborhoods (31.1%), were married (64.3%), had public insurance (56%), and were foreign-born (57.9%) compared to NHW individuals (23.6%, 53.8%, 24.7%, and 5.8%, respectively). Analyses of known cigarette use during 2014-2019 (68.9% of Armenians and 66.8% of NHWs) showed that 29.1% of Armenian patients (83.0% male) compared to 22.4% of NHW cancer patients (56.2% male) were current cigarette users.

We observed somewhat higher proportions of late-stage diagnoses for stomach (69.5% vs. 66.1%), lung (79.8% vs. 73.4%), and colorectal (60.7% vs. 57.9%) cancers among Armenians compared to NHWs (Figure 3.1). We observed significantly higher odds of late-stage colorectal (OR=1.12, 95% CI=1.03-1.22), lung (OR=1.26, 95% CI=1.12-1.42), and stomach (OR=1.43, 95% CI=1.17-1.74) cancers among Armenian patients compared to NHW patients (Figure 3.2). In multivariable models by cancer type in Armenians, lowest (OR=1.73, 95% CI=1.03-2.92) and middle (OR=1.79, 95% CI=1.10-2.90) nSES status was associated with late-stage stomach cancer (Table 3.2). Male sex (OR=1.70, 95% CI=1.28-2.25) and public health insurance (OR=1.57, 95% CI=1.18-2.09) was associated with late-stage lung cancer. For colorectal cancer, public health insurance among patients younger than 65 years (OR=1.54, 95% CI=1.12-2.13) was associated with late-stage disease. Patients residing in the lowest nSES (OR=1.99, 95% CI=1.22-3.26) and with public health insurance (OR=1.99, 95% CI=1.22-3.26) were significantly more likely to be diagnosed with late-stage bladder cancer.

In multivariable models adjusting for sociodemographic and clinical factors, including stage at diagnosis, for each of the four cancers studied, we observed lower hazard of death among Armenians compared to NHWs, showing a modest survival advantage among Armenians for stomach (HR=0.85, 95% CI=0.76-0.94), lung (HR=0.86, 95% CI=0.82-0.91), colorectal (HR=0.82, 95% CI=0.77-0.88), and bladder (HR=0.87, 95% CI=0.76-0.99) cancers (Table 3.3). Additionally, survival estimates of Armenians relative to NHW individuals by stage at diagnosis for lung cancer showed no significant differences for localized and regional stage, and better survival for remote stage (HR=0.84, 95% CI=0.79-0.89). Among 111,811 patients alive at last follow-up, n=1,918 Armenian and n=109,893 NHW, 7.7% of Armenian and 3.4% of NHW patients were LTFU. We also observed a greater percentage of LTFU among foreign-born patients among Armenian (9.04%) and NHW (9.25%) compared to patients with United States/unknown place of birth (4.8%, 2.8%, respectively). In two sensitivity analyses, when we assumed that all patients LTFU died at the

date of last contact or all foreign-born LTFU patients died at the date of last contact, we observed similar results to our findings in Table 3.3, except for bladder cancer where there was similar survival between Armenian and NHW patients (Supplemental Table 3.2).

Discussion

To our knowledge, this is the first population-based study to consider stage at diagnosis and cancer-specific survival among the Armenian compared to the NHW population in California. We observed a higher likelihood of late-stage diagnoses for colorectal, lung, and stomach cancers among Armenian compared to NHW patients, suggesting potential barriers to timely cancer screening for lung and colorectal cancer in Armenians. After consideration of stage at diagnosis and other sociodemographic and clinical factors, we additionally found a moderate survival advantage among Armenian compared to NHW patients for all four cancers studied. Our findings highlight the need to understand factors impacting the survival advantage in Armenians, including the role of genetics, behavioral and social factors, or biases in ascertainment of survival data.

The higher likelihood of late-stage diagnoses for three of the four tobacco-related cancers among Armenians compared to NHW patients may indicate barriers to reliable healthcare among the Armenian population in California. Lung and colorectal cancers are both screen-detectable, and the disparity in late-stage diagnoses among the Armenian population may signal potential barriers to care, including routine annual examinations and timely cancer screening for screen-detectable cancers. Consistent with previous studies,^{137–139} public health insurance was associated with late-stage lung and colorectal cancers among younger Armenians (<65 years old), suggesting that increased access to reliable health insurance may be needed for early cancer detection among the Armenian community. Further, Armenians residing in lower SES neighborhoods were more likely to be diagnosed with late-stage stomach and bladder cancers, signaling a need for additional research to understand specific barriers among the Armenian population experiencing economic disadvantage. Previous studies have identified that socioeconomic disadvantage presents numerous risks for late-stage diagnoses, including unreliable access to health care and lack of job security and transportation.¹³⁹ The higher proportion of foreign-born Armenian patients may also suggest additional difficulties receiving adequate healthcare, as supported by studies showing that

immigrant populations face increased barriers to access, including linguistic barriers, and are less likely to use preventive health services including cancer screening.^{140–142}

The modest survival advantage among Armenian patients may be related to a variety of factors, such as reliable community and family support after cancer diagnosis, the salmon bias, or better overall health among Armenians.⁷ Previous research has shown a significant connection between ethnic identity, cultural beliefs and values, and self-care of cancer survivors, and this is relevant to our study as a majority of Armenian patients (57.9%) in our study were foreign-born and may closely adhere to traditional Armenian cultural norms and behaviors.¹⁴³ Further, the better survival observed among Armenians may be related to the salmon bias where patients with cancer may have left California to return to their native country in later years of life or after illness.¹⁴⁴ The salmon bias has previously been used to explain lower mortality among foreign-born Hispanic cancer patients in California, and better survival of Middle-Eastern women (including Armenian women) with advanced stage breast cancer compared to NHW women.^{49,145} The higher LTFU among Armenian compared to NHW patients raises the question of whether these patients returned to Armenia or other native countries post-diagnosis. However, the similar survival patterns observed in our sensitivity analyses suggest that LTFU does not explain the survival advantage observed among Armenian patients in our study. Our findings may also relate to the healthy immigrant effect, which has been used to explain the better health of immigrants in the United States compared to US-born populations.¹⁴⁶ However, additional information on the health among Armenian cancer patients would be required to confirm its role in cancer survival statistics.

To understand the survival advantage of Armenian cancer patients more fully, genetic and epigenetic studies are also needed. Previous epigenetic studies have examined the Hispanic Paradox in the United States to better understand the lower mortality observed among Hispanic patients compared to NHW patients despite lower SES, higher percentage uninsured, higher cardio-metabolic disease, higher prevalence of obesity, and type 2 diabetes.^{145,147} Among other findings, researchers found that Hispanic populations of Mexican ancestry in California have lower rates of intrinsic aging in their blood and suggest this finding has a more plausible connection to the lower overall mortality risk as opposed to the healthy immigrant effect.¹⁴⁷ Genetic and epigenetic studies to understand cancer survival among Armenians are necessary given the genetic continuity maintained among the Armenian population since the Bronze Age,¹⁴⁸

and a previously discovered link between Armenian ethnicity and increased cardiovascular disease risk in California.⁴²

Given the continued public health burden of tobacco-use in Armenia, and the majority foreign-born population of Armenians in California, it is important to understand tobacco-use among Armenians after immigration.²⁸ While tobacco-use data among the Armenian population in California is currently unavailable, we observed a higher percentage of cigarette use among Armenian cancer patients, and a clear gender disparity in cigarette use among Armenian male and female patients. Our findings point to the need for educational interventions pertaining to the negative effects of tobacco-use on cancer treatment and survival.¹²⁵ While cigarette use patterns in California appear to follow those observed in Armenia, our findings should be interpreted with caution due to the more recent availability and high proportion of missing cigarette use data in the CCR, and uncertainty as to whether these variables were missing at random between the two race/ethnic groups and among males and females.

While the use of the ASL allowed the study of cancer survival among Armenians for the first time, the ASL is subject to common limitations arising with the use of surname lists in general.⁴⁷ A common limitation is the possibility of missing Armenian patients with non-Armenian last names who may have changed their last names due to marriage or other reasons. While we included maiden name and father's surname where available to minimize this limitation, it is still possible we may have missed identifying Armenian patients with non-Armenian last names and missing maiden and father's surnames. Additionally, we may have missed Armenian patients with non-Armenian fathers and Armenian mothers who may have had a non-Armenian last name, maiden name, and father's surname.⁴⁷

Another limitation is the possibility of missed deaths during linkage to death records, which has previously been shown to bias survival estimates in survival studies using population-based cancer registry data.¹⁴⁹ Additionally, death linkages rely on variables including valid social security numbers, and populations with large foreign-born patients, including Armenians in the present study, may have precluded the identification of foreign-born Armenians who have died, and these patients would have been considered alive in this study.¹⁴⁹ However, complete follow-up with at least five years of vital status information is one way to mitigate the effects of missing deaths, and the 31 years of follow-up information used in the present study may have helped reduce the potential survival bias related to missing deaths.¹⁴⁹

While Armenian patients were at greater risk of late-stage diagnoses of colorectal, lung, and stomach cancers, we observed better survival among Armenian patients compared to NHW patients. Our findings of higher proportions of late-stage diagnoses and continued tobacco-use among Armenian cancer patients suggest the need for interventions to increase access to care, cancer screening, and culturally competent tobacco cessation programs. Given the collectivist and traditional Armenian culture and a majority foreign-born demography,⁷ public health interventions in the Armenian language (Eastern and Western dialects), including partnerships among public health professionals, clinicians, and trusted community leaders in the Armenian community, are needed for successful community outreach and engagement regarding cancer risk factors and access to resources including screening for screen-detectable cancers.^{142,150} Further research is needed to understand the observed survival advantage among Armenians, including genetic, behavioral, and social factors or biases in ascertainment of survival data.

Table 3.1 Characteristics of Armenian and NHW patients diagnosed with stomach, lung, colorectal, and bladder cancer in California, 1988-2019

Characteristic	Armenian (N=6,642)		Non-Hispanic White (N=632,582)	
	n	(%)	n	(%)
Year of Diagnosis				
1988-1998	1,406	21.2	248,476	39.3
1999-2009	2,355	35.5	218,788	34.6
2010-2019	2,881	43.4	165,318	26.1
Sex				
Male	4,145	62.4	347,410	54.9
Female	2,497	37.6	285,172	45.1
Age at diagnosis (years)				
0-29	25	0.4	1,662	0.3
30-49	455	6.9	35,527	5.6
50-69	2,873	43.3	259,232	41.0
70+	3,289	49.5	336,161	53.1
Cancer Type				
Stomach	774	11.7	31,428	5.0
Lung	2,348	35.4	304,962	48.2
Colorectal	2,750	41.4	234,080	37.0
Bladder	770	11.6	62,112	9.8
Neighborhood Socioeconomic Status				
Lowest	2,065	31.1	149,149	23.6
Middle	2,756	41.5	243,785	38.5
Highest	1,821	27.4	239,648	37.9
Marital Status				
Married/in partnership	4,274	64.3	340,495	53.8
Previously married	1,473	22.2	198,639	31.4
Single, never married	688	10.4	75,058	11.9
Unknown	207	3.1	18,390	2.9
Insurance Type				
Not insured	102	1.5	5,969	0.9
Private	1,890	28.5	299,732	47.4
Public	3,660	55.1	135,599	21.4
Other public	61	0.9	13,207	2.1
Unknown	929	14.0	178,075	28.2
Stage at diagnosis				
Localized	1,923	29.0	189,144	29.9
Regional	1,931	29.1	169,880	26.9
Remote	2,319	34.9	223,694	35.4
Unknown	469	7.1	49,864	7.9
Nativity				
United States-Born	635	9.6	329,852	52.1
Foreign-born	3,847	57.9	36,920	5.8
Unknown	2,160	32.5	265,810	42.0
Comorbidity				
3+	819	12.3	57,134	9.0
1-2	2,184	32.9	193,345	30.6
0	2,491	37.5	217,014	34.3
Unknown	1,148	17.3	165,089	26.1
Vital Status				
Dead	4,724	71.1	522,689	82.6
Alive	1,918	28.9	109,893	17.4

Known Cigarette Use^a, 2014-2019	N=1,234		N=64,413	
Current User	359	29.1	14,456	22.4
Male	298	83.0	8,128	56.2
Female	61	17.0	6,328	43.8
Never/former User	875	70.9	49,957	77.6
Male	489	55.9	26,568	53.2
Female	386	44.1	23,389	46.8

^aKnown cigarette use at date of diagnosis in the California Cancer Registry; analyses were limited to years of diagnosis 2014-2019 when this data was more routinely collected and excluded those with unknown values (n=558 Armenians; n=32,076 Non-Hispanic Whites). NHW: Non-Hispanic White.

Table 3.2 Odds ratios for regional/remote stage (versus localized) cancer diagnosis among Armenian patients in California, 1988-2019

Characteristic	Stomach OR (95% CI)	Lung OR (95% CI)	Colorectal OR (95% CI)	Bladder OR (95% CI)
Year of Diagnosis				
1988-1998	1.46 (0.76-2.73)	0.66 (0.45-0.96)	1.35 (1.02-1.78)	0.70 (0.35-1.37)
1999-2009	1.34 (0.86-2.08)	0.97 (0.73-1.29)	0.94 (0.78-1.13)	0.85 (0.56-1.28)
2010-2019	reference	reference	reference	reference
Age at diagnosis				
0 to 49	reference	reference	reference	reference
50 to 69	0.57 (0.25-1.29)	0.93 (0.53-1.61)	0.78 (0.57-1.07)	0.76 (0.35-1.65)
70+	0.29 (0.13-0.66)	0.94 (0.54-1.65)	0.78 (0.56-1.08)	0.51 (0.23-1.14)
Sex				
Male	1.11 (0.73-1.68)	1.70 (1.28-2.25)	1.05 (0.88-1.25)	0.75 (0.47-1.18)
Female	reference	reference	reference	reference
Neighborhood SES				
Lowest	1.73 (1.03-2.92)	0.99 (0.71-1.39)	1.06 (0.85-1.33)	1.99 (1.22-3.26)
Middle	1.79 (1.10-2.90)	1.09 (0.80-1.49)	0.94 (0.77-1.16)	1.29 (0.81-2.06)
Highest	reference	reference	reference	reference
Insurance Type				
All ages (years)				
Private	reference	reference	reference	reference
Public	1.09 (0.68-1.74)	1.57 (1.18-2.09)	1.23 (1.01-1.49)	1.59 (1.02-2.49)
Unknown	1.13 (0.52-2.48)	1.52 (0.95-2.42)	1.12 (0.81-1.55)	1.46 (0.68-3.16)
Ages <65 (years)				
Private	reference	reference	reference	reference
Public	0.53 (0.21-1.31)	2.05 (1.22-3.44)	1.54 (1.12-2.13)	1.48 (0.71-3.10)
Unknown	1.98 (0.39-9.94)	1.65 (0.79-3.45)	1.37 (0.83-2.26)	1.17 (0.35-3.91)
Ages 65+ (years)				
Private	reference	reference	reference	reference
Public	1.38 (0.77-2.45)	1.44 (1.02-2.04)	1.04 (0.81-1.35)	1.62 (0.90-2.91)
Unknown	0.84 (0.33-2.13)	1.54 (0.84-2.83)	0.89 (0.58-1.39)	1.44 (0.50-4.19)
Marital Status				
Married/in partnership	reference	reference	reference	reference
Previously married	1.23 (0.74-2.06)	1.05 (0.75-1.45)	1.03 (0.83-1.28)	0.95 (0.58-1.56)
Single, never married	0.98 (0.46-2.08)	0.82 (0.55-1.21)	0.94 (0.73-1.22)	0.81 (0.43-1.51)
Unknown	0.63 (0.22-1.78)	2.82 (0.86-9.21)	0.48 (0.30-0.79)	0.39 (0.08-1.81)
Comorbidity				
3+	0.68 (0.36-1.31)	0.72 (0.48-1.07)	0.80 (0.59-1.09)	1.51 (0.85-2.68)
1-2	0.99 (0.63-1.59)	0.79 (0.58-1.07)	0.95 (0.78-1.16)	1.31 (0.84-2.03)
0	reference	reference	reference	reference
Unknown	0.68 (0.37-1.28)	1.19 (0.78-1.82)	0.75 (0.59-0.96)	0.61 (0.36-1.04)

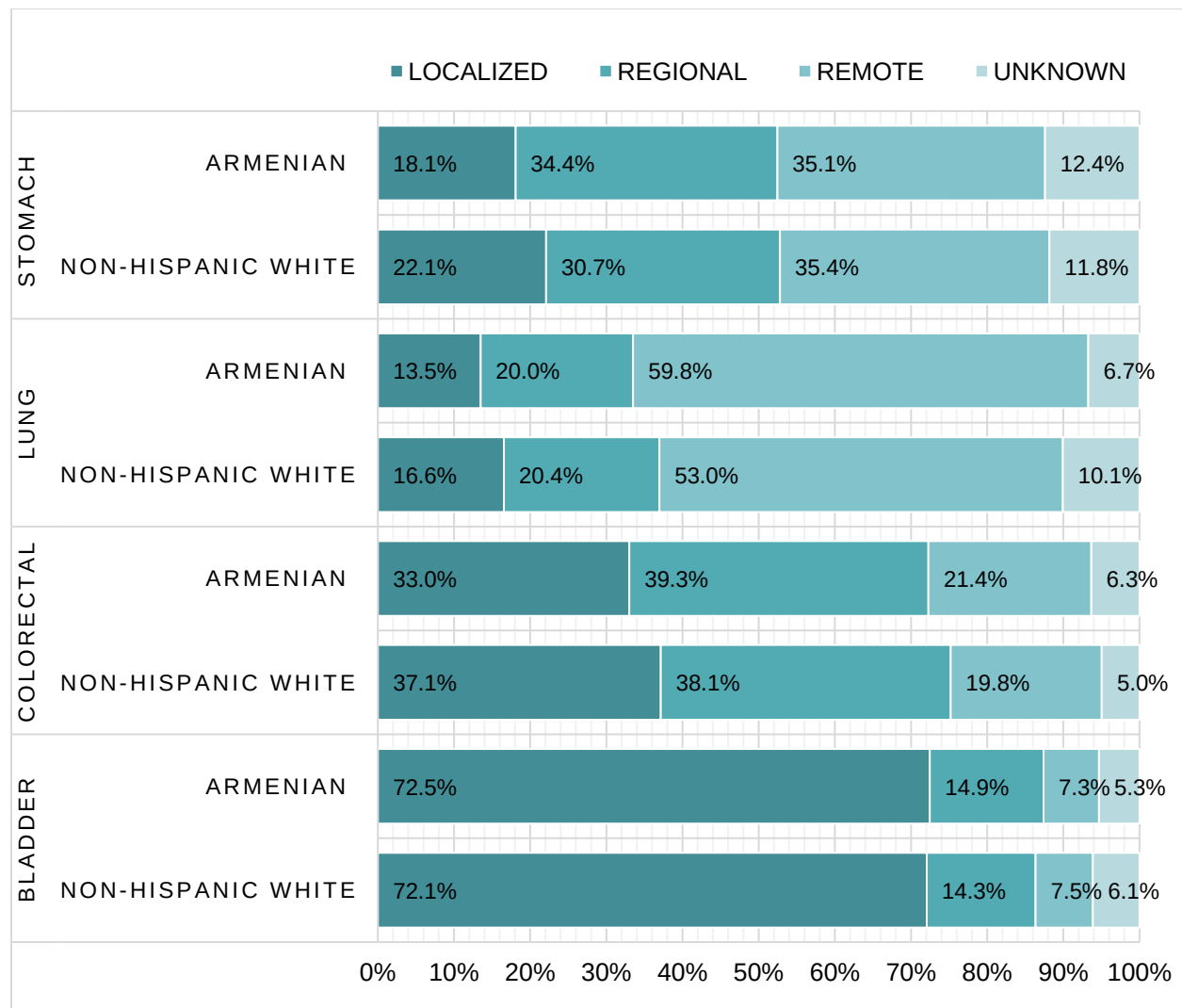
Models were adjusted for variables in the table and reporting regional cancer registry. Bold: Statistically significant at p<0.05.

Table 3.3 Cancer-specific survival of Armenian patients compared to NHW patients in California by cancer type, 1988-2019

Cancer Type	HR ^a	95% CI	p-value
Stomach	0.85*	(0.76-0.94)	0.0015
Lung ^b	0.86*	(0.82-0.91)	<0.0001
<i>Localized</i>	0.99	(0.82-1.21)	0.9678
<i>Regional</i>	0.98	(0.87-1.11)	0.7723
<i>Remote</i>	0.84*	(0.79-0.89)	<0.0001
Colorectal	0.82*	(0.77-0.88)	<0.0001
Bladder	0.87*	(0.76-0.99)	0.0308

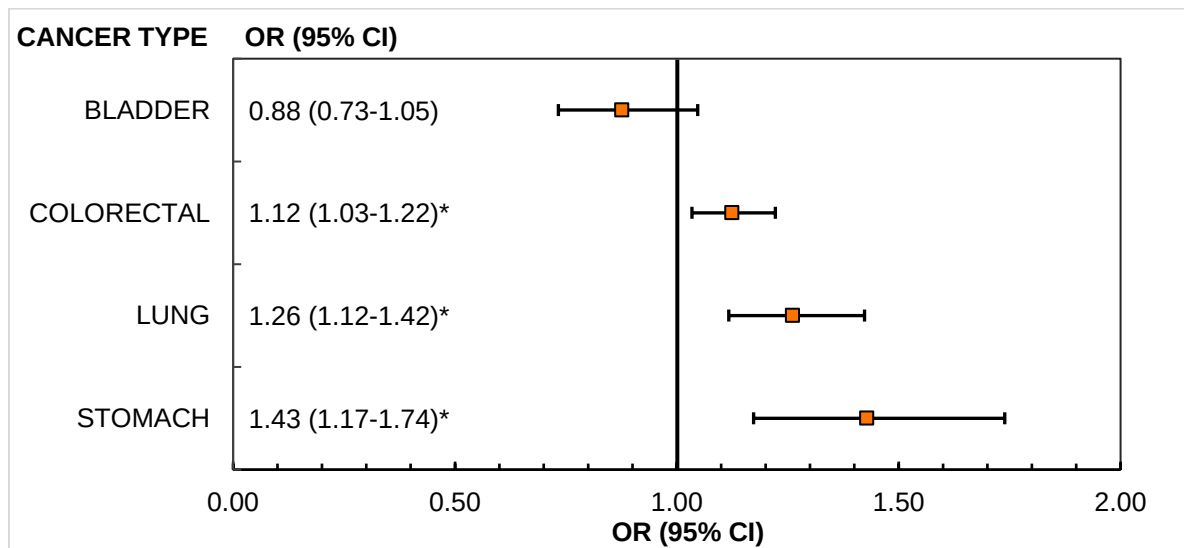
^aModels adjusted for age, year of diagnosis, sex, marital status, insurance type, comorbidity, reporting regional cancer registry, socioeconomic status, first course of treatment, histology, cancer site, and stage at diagnosis. HR: Hazard Ratio, CI: Confidence interval. ^bHR shown by stage at diagnosis for lung cancer due to observed interaction between race/ethnicity and stage at diagnosis. *Statistically significant at p<0.05. NHW: Non-Hispanic White.

Figure 3.1 Frequency distribution of stage at diagnosis by cancer type among Armenian and NHW patients in California, 1988-2019



NHW: Non-Hispanic White.

Figure 3.2 Odds ratios for regional/remote stage (versus localized) cancer diagnosis among Armenian patients compared to NHW patients in California by cancer type, 1988-2019



Models adjusted for age, year of diagnosis, sex, marital status, insurance type, comorbidity, reporting regional cancer registry, and socioeconomic status. OR: Odds Ratio, CI: 95% Confidence interval.

*Statistically significant at $p < 0.05$. NHW: Non-Hispanic White.

Supplemental Table 3.1 Selected tumor sites based on International Classification of Diseases for Oncology, 3rd edition/World Health Organization (ICD-O-3/WHO) 2008 site codes

Tumor Site	ICD-O-3 Code
Stomach	
Cardia	C160
Non-Cardia	C161,C162,C163,C164,C165,C166
Overlapping, NOS	C168, C169
Lung	
Upper lobe	C341
Middle lobe	C342
Lower lobe	C343
Main bronchus	C340
Overlapping	C348, C349
Colorectal	
Colon	C180, C182-C189, C260
Rectum	C199, C209
Appendix	C181
Bladder	
Trigone of bladder	C670
Dome of bladder	C671
Lateral wall of bladder	C672
Anterior wall of bladder	C673
Posterior wall of bladder	C674
Bladder neck	C675
Ureteric orifice	C676
Urachus	C677
Overlapping lesion of bladder	C678
Bladder, NOS	C679

Supplemental Table 3.2 Cancer-specific survival of Armenian patients compared to NHW patients in California by cancer type considering effect of lost to follow-up (LTFU), 1988-2019

Cancer Type	HR ^a	95% CI	p-value
Assumption 1: All patients LTFU died at the date of last contact			
Stomach	0.86*	(0.78-0.95)	0.0040
Lung	0.87*	(0.83-0.92)	<0.0001
Colorectal	0.86*	(0.80-0.91)	<0.0001
Bladder	0.94	(0.83-1.07)	0.3686
Assumption 2: All foreign-born patients LTFU died at the date of last contact			
Stomach	0.85*	(0.77-0.94)	0.0019
Lung	0.87*	(0.83-0.92)	<0.0001
Colorectal	0.85*	(0.80-0.91)	<0.0001
Bladder	0.94	(0.83-1.07)	0.3318

^aModels adjusted for age, year of diagnosis, sex, marital status, insurance type, comorbidity, reporting regional cancer registry, socioeconomic status, first course of treatment, histology, cancer site, and stage at diagnosis. HR: Hazard Ratio, CI: Confidence interval. *Statistically significant at p<0.05. NHW: Non-Hispanic White.

Conclusion

The genetic continuity in the Armenian population, resulting from thousands of years of cultural and geographic isolation, makes Armenians an ideal population for public health and epidemiological research to understand the health of the population after emigration to other countries.^{123,124,151} Estimates of Armenians in the United States range from 450,000 to 2,000,000, with the large range due to broad race/ethnic categories and an insufficient identification method of minority race/ethnic populations in the United States.^{56,68,152} According to the US Census, there are approximately 470,000 Armenians living in the United States, of which more than half (259,430) live in California.⁴⁴ However, these are likely underestimates as the Embassy of Armenia to the United States of America estimates that anywhere between 500,000 to 2,000,000 Armenians live in the United States.⁵⁶ The challenge of identifying Armenians in the United States has hindered the representation of this population in research, including in public health and epidemiological research studies.^{47,48} While California is home to the largest population of Armenians in the United States, cancer patterns among the Armenian population have not been previously studied due to the lack of accurate Armenian population estimates, a standardized collection of Armenian ethnicity, or an Armenian surname list. This dissertation is the first to identify the Armenian population in a large, population-based cancer registry in the country and to provide an analysis of cancer incidence and survival patterns.¹⁵³

In the first chapter of this dissertation, we created the Armenian Surname List (ASL) using an extract of Armenian surnames from the Middle Eastern Surname List and California public death records from 1905 to 2020.^{47,48} The ASL not only serves as a tool to identify Armenians in cancer registries, but it also lays the groundwork as a valuable resource for research studies in broad epidemiologic, public health, and other fields concerning Armenians in California and the United States.⁴⁷ This is especially true for studies based on linkages that require surnames as a matching variable.¹⁵⁴ In addition, since many Armenian names follow ethnically unique conventions, many matching algorithms will produce a higher match score resulting in more accurate study results.^{43,48,154}

In the second chapter, the ASL allowed for the identification of Armenian cancer patients in the California Cancer Registry (CCR).⁴⁷ We used the ASL and birthplace information in the CCR to identify

Armenians with cancer diagnosed during 1988-2019 and calculated proportional incidence ratios (PIR) among Armenians compared with non-Hispanic Whites (NHWs). We also calculated incidence rate ratios (IRR) during 2006-2015 using Armenian population denominators from the American Community Survey (ACS), as an exploratory analysis. Among Armenian males, we found higher PIRs of stomach, bladder, colorectal, lung, leukemia, liver and IBD, and kidney cancers. While thyroid cancer was not among the top 10 most frequent for Armenian males, we also observed a higher proportion of thyroid cancer among Armenian males compared with NHW males. Among Armenian females, we found higher PIRs of stomach, thyroid, colorectal, pancreatic, leukemia, NHL, and ovarian cancers. Exploratory IRR analyses showed higher stomach, bladder, and colorectal cancers among Armenian males and higher stomach cancer among Armenian females. The significantly higher PIRs and IRRs of stomach cancer among Armenian patients suggest a need for public health interventions to understand stomach cancer etiology, including obesity, dietary behaviors, and exposure to infections such as *Helicobacter pylori*.^{100,101} While tobacco-use data among Armenians in California is currently unavailable, the high smoking rate among Armenian men in Armenia, and the higher PIRs of several tobacco-related cancers among Armenians, including stomach, lung, colorectal, and bladder, present a need for additional research investigating the relationship between tobacco-use and cancer risk among Armenians in California.^{27,105} Our findings suggest a need for additional research addressing risk factors including tobacco-use, obesity, iodine deficiency, and exposure to harmful environmental agents associated with specific cancers among Armenians in California.

In the third and final chapter, we used the ASL and birthplace information to identify Armenians in the CCR diagnosed with cancer during 1988-2019 and analyzed cancer survival and stage at diagnosis for four tobacco-related cancers: stomach, lung, colorectal, and bladder. Notable findings in this chapter include higher odds of late-stage colorectal, lung, and stomach cancers diagnosed among Armenian patients compared to NHW patients. Among Armenian patients, we found that lower neighborhood socioeconomic status was associated with late-stage stomach and bladder cancer diagnoses, and public health insurance was associated with late-stage lung, colorectal, and bladder cancer diagnoses. Armenian patients had a moderate survival advantage for stomach, lung, colorectal, and bladder cancers compared to NHW patients. This survival advantage may be attributed to cultural customs and norms including reliable community and family support, the salmon bias, the healthy immigrant effect, or overall better health among

Armenians.^{7,143,144} However, the survival advantage remained even after considering lost-to-follow-up assumptions in sensitivity analyses. Our findings that Armenian patients with lower neighborhood socioeconomic status and public health insurance had higher likelihood of late-stage cancer diagnoses have significant public health implications including the need for increased access to reliable health care, cancer screening, and culturally competent tobacco cessation programs for the Armenian population in California. Further research including genetic and epigenetic factors can better help understand the underlying factors for the better survival observed among Armenians compared to NHW cancer patients.^{42,145,147,148}

This dissertation project is the first to create a surname list to be used for the identification of Armenians for public health research and to identify and study cancer successfully among the Armenian population in California. The ASL not only provided a database required for our study, but it serves as a resource for a myriad of research projects concerning Armenian communities in California and the United States. With the use of the ASL, further research on this growing ethnic minority group has potential for revealing health disparities and addressing barriers to timely cancer screening and care.¹⁵⁵ Through the examination of cancer incidence, stage at diagnosis, and survival patterns, this dissertation provides novel data to inform evidence-based cancer control efforts for the growing Armenian population in California.

References

1. Schillinger D. The Intersections Between Social Determinants of Health, Health Literacy, and Health Disparities. *Student Heal Technol Informatics*. 2020;269:22-41. doi:10.3233/SHTI200020
2. Awad GH, Kia-Keating M, Amer MM. A model of cumulative racial-ethnic trauma among Americans of Middle Eastern and North African (MENA) Descent. *Am Psychol*. 2019;74(1):76-87. doi:10.1037/amp0000344
3. Chronic stress puts your health at risk - Mayo Clinic. Accessed May 23, 2020. <https://www.mayoclinic.org/healthy-lifestyle/stress-management/in-depth/stress/art-20046037>
4. Malcom MV. *The Armenians in America*. The Pilgrim Press; 1919.
5. Kalayjian A, Weisberg MM. Generational Impact of Mass Trauma: The Post-Ottoman Turkish Genocide of the Armenians. In: Piven, JS, Boyd, C, Lawton H, ed. *Jihad and Sacred Vengeance*. Writers Club Press; 2002:254-279. Accessed December 13, 2021. <https://www.humiliationstudies.org/documents/KalayjianGenerationalTransmissionChapter.pdf>
6. Watenpaugh KD. The League of Nations' Rescue of Armenian Genocide Survivors and the Making of Modern Humanitarianism, 1920-1927. *Am Hist Rev*. Published online 2010. Accessed April 26, 2021. <https://academic.oup.com/ahr/article/115/5/1315/41289>
7. Fittante D. But Why Glendale? A History of Armenian Immigration to Southern California. *Calif Hist*. 2017;94(3):1-31. Accessed April 14, 2021. <https://escholarship.org/uc/item/6v04w06x>
8. Mekdjian S, Nanterre PX. *The Armenian Community in Glendale (2000)*. Accessed April 14, 2021. www.census.gov
9. Barseganyan VS, Nersesyan AK. Cancer incidence in Armenia (1970 - 1998). *Arch Oncol*. 2000;8(4):187.
10. Sharafyan N, Martirosyan A. Organization, Management, Economic Consequences, and Legal Qualifications of the Massacres of Armenians in January 1990 in Baku. *Alternative*. Published online 2022. doi:http://dx.doi.org/10.55528/18292828-2022.1-171
11. Arushanyan Y. *Sumgait: The Silence*. American University of Armenia; 2011. Accessed December 27, 2023. https://dspace.aua.am/xmlui/bitstream/handle/123456789/539/Yevgenia_Arushanyan.pdf?sequence=1&isAllowed=y
12. Demirchyan A, Petrosyan D, Armenian HK. Rate and predictors of postpartum depression in a 22-year follow-up of a cohort of earthquake survivors in Armenia. *Arch Womens Ment Health*. 2014;17(3):229-237. doi:10.1007/S00737-013-0404-5
13. UN experts urge Azerbaijan to lift Lachin corridor blockade and end humanitarian crisis in Nagorno-Karabakh | OHCHR. Accessed December 27, 2023. <https://www.ohchr.org/en/press-releases/2023/08/un-experts-urge-azerbaijan-lift-lachin-corridor-blockade-and-end>
14. Campbell RD, Griffin MJ, Bragagnoio LJ, Yanev PI. *The December 7,1988 Armenia Earthquake Effects on Selected Power, Industrial and Commercial Facilities.*; 1991.
15. Goenjian AK, Walling D, Steinberg AM, Karayan I, Najarian LM, Pynoos R. A prospective study of posttraumatic stress and depressive reactions among treated and untreated adolescents 5 years after a catastrophic disaster. *Am J Psychiatry*. 2005;162(12):2302-2308. doi:10.1176/appi.ajp.162.12.2302
16. Mills C. What is happening in Nagorno-Karabakh? Accessed December 27, 2023. <https://commonslibrary.parliament.uk/research-briefings/cbp-9862/>
17. Otte J. 'This is a forced migration': the ethnic Armenians fleeing Nagorno-Karabakh | Nagorno-Karabakh | The Guardian. Published 2023. Accessed December 27, 2023. <https://www.theguardian.com/world/2023/oct/03/this-is-a-forced-migration-the-ethnic-armenians->

fleeing-nagorno-karabakh

18. City Council | City of Glendale, CA. Accessed December 28, 2023. <https://www.glendaleca.gov/government/city-council>
19. Armenian Institute. *Armenian Diaspora Survey*.; 2020. <https://www.armeniandiasporasurvey.com/2019-survey>
20. Minasian, Don A. *Settlement Patterns of Armenians in Fresno, California*. San Fernando Valley State College; 1972. Accessed December 28, 2023. <http://dspace.calstate.edu/handle/10211.2/4756>
21. CFPB Orders Citi to Pay \$25.9 Million for Intentional, Illegal Discrimination Against Armenian Americans | Consumer Financial Protection Bureau. Accessed December 28, 2023. <https://www.consumerfinance.gov/about-us/newsroom/cfpb-orders-citi-to-pay-25-9-million-for-intentional-illegal-discrimination-against-armenian-americans/>
22. Armenian Christians attacked in Jerusalem, some in serious condition - Crime in Israel - The Jerusalem Post. Accessed December 28, 2023. <https://www.jpost.com/israel-news/crime-in-israel/article-779881>
23. Sign posted outside an Armenian church in Watertown is being investigated as a hate crime – Boston 25 News. Accessed December 28, 2023. <https://www.boston25news.com/news/local/sign-posted-outside-an-armenian-church-watertown-is-being-investigated-hate-crime/NAQQVUE2FZF3JCQGYGNZEKWWDE/>
24. Zohrabyan D, Karapetyan N, Danielyan S, et al. Lung Cancer in Armenia. *J Thorac Oncol*. 2023;18(4):402-409. doi:10.1016/j.jtho.2022.12.015
25. Chekijian S, Truzyan N, Stepanyan T, Bazarchyan A. Healthcare in transition in the Republic of Armenia: the evolution of emergency medical systems and directions forward. *Int J Emerg Med*. 2021;14(5):1-11. doi:10.1186/s12245-020-00328-3
26. Cancer Surveillance | The Cancer Atlas. Accessed December 28, 2023. <https://canceratlas.cancer.org/taking-action/cancer-surveillance/>
27. Ministry of Health Armenia, RTI International, United Nations Development Programme, Secretariat of the WHO Framework Convention on Tobacco Control, World Health Organization. *The Case for Scaling up WHO FCTC Implementation Investment Case for Tobacco Control in Armenia*.; 2021. Accessed September 11, 2023. <https://www.undp.org/sites/g/files/zskgke326/files/migration/am/ARMENIA-FCTC-INVESTMENT-CASE-ENGLISH.pdf>
28. Berg CJ, Harutyunyan A, Paichadze N, Hyder AA, Petrosyan V. Addressing cancer prevention and control in Armenia: tobacco control and mHealth as key strategies. *Int J Equity Health*. 2021;20:4. doi:10.1186/s12939-020-01344-8
29. Armenia | Institute for Health Metrics and Evaluation. Accessed May 2, 2021. <http://www.healthdata.org/armenia>
30. All Cancers Death Rate by Country. Accessed May 2, 2021. <https://www.worldlifeexpectancy.com/cause-of-death/all-cancers/by-country/>
31. Jemal A, Torre L, Street W, Bray F. *The Cancer Atlas*.; 2019. www.cancer.org/canceratlas
32. International Agency for Research on Cancer WHO. Cancer Today. Published 2020. Accessed February 7, 2023. https://gco.iarc.fr/today/online-analysis-map?v=2020&mode=population&mode_population=continents&population=900&populations=900&key=asr&sex=0&cancer=39&type=1&statistic=5&prevalence=0&population_group=0&ages_group%5B%5D=0&ages_group%5B%5D=17&nb_items=10&gr
33. World Health Organization. *Cancer Country Profile 2020: Armenia*.; 2020. Accessed January 8, 2024. <https://www.who.int/publications/m/item/cancer-arm-2020>
34. Armenia. Accessed May 17, 2021. <https://www.who.int/data/gho/data/countries/country->

details/GHO/armenia?countryProfileId=b36fc47a-744e-4e14-82de-85b7a3a1bdde

35. World Health Organization. *Statistics at a Glance, Armenia.*; 2022. Accessed February 26, 2024. <https://gco.iarc.who.int/media/globocan/factsheets/populations/51-armenia-fact-sheet.pdf>
36. Bardakhchyan S, Mkhitarian S, Zohrabyan D, et al. Treatment and Outcomes of Colorectal Cancer in Armenia: A Real-World Experience From a Developing Country. *JCO Glob Oncol.* 2020;(6):1286-1297. doi:10.1200/go.20.00251
37. Bedirian K, Aghabekyan T, Mesrobian A, et al. Overview of Cancer Control in Armenia and Policy Implications. *Front Oncol.* 2022;11:1-12. doi:10.3389/fonc.2021.782581
38. World Health Organization. *Global Country Profiles on Burden of Cancer.*; 2020.
39. Smoking - Our World in Data. Accessed October 25, 2022. <https://ourworldindata.org/smoking>
40. Who smokes more, men or women? - Our World in Data. Accessed February 26, 2024. <https://ourworldindata.org/who-smokes-more-men-or-women>
41. Cancer Today. Accessed February 26, 2024. <https://gco.iarc.fr/today/en/fact-sheets-populations#countries>
42. Rostomian AH, Soverow J, Sanchez DR. Exploring Armenian Ethnicity as an Independent Risk Factor for Cardiovascular Disease: Findings from a Prospective Cohort of Patients in a County Hospital. *JRSM Cardiovasc Dis.* 2020;9:1-7. doi:10.1177/2048004020956853
43. Aintablian H, Melkonian C, Galoustian N, et al. Why we should study the Armenian population: a goldmine of public health information. *MOJ Public Heal.* 2018;7(1):11-12. doi:10.15406/mojph.2018.07.00196
44. B01001: Census Bureau Table. Accessed October 24, 2022. https://data.census.gov/cedsci/table?q=armenian&g=0100000US_0400000US06,06%240500000&tid=ACSDT5YSPT2015.B01001
45. About Race. Accessed April 24, 2021. <https://www.census.gov/topics/population/race/about.html>
46. California Cancer Registry. *Cancer Reporting in California Appendix O: Race and Nationality.*; 2023. Accessed January 8, 2024. https://docs.google.com/viewer?url=https%3A%2F%2Fwww.ccrca.org%2Fwp-admin%2Fadmin-ajax.php%3Fjuwpfisadmin%3Dfalse%26action%3Dwpfd%26task%3Dfile.download%26wpfd_category_id%3D303%26wpfd_file_id%3D9088%26token%3D%26preview%3D1&embedded=true
47. Movsisyan AS, Nasseri K, Keegan TH. Construction of the Armenian Surname List (ASL) for public health research. *BMC Med Res Methodol.* 2023;23:29. doi:10.1186/s12874-023-01848-1
48. Nasseri K. Construction and validation of a list of common Middle Eastern surnames for epidemiological research. *Cancer Detect Prev.* 2007;31(5):424-429. doi:10.1016/j.cdp.2007.10.006
49. Nasseri K. Breast cancer in the Middle Eastern Population of California, 1988-2004. *Breast J.* 2009;15(2):182-188. doi:10.1111/j.1524-4741.2009.00694.x
50. Nasseri K. Thyroid cancer in the Middle Eastern population of California. *Cancer Causes Control.* 2008;19(10):1183-1191. doi:10.1007/s10552-008-9185-y
51. Schwartz K, Beebani G, Sedki M, et al. Enhancement and Validation of an Arab Surname Database. *J Registry Manag.* 2013;40(4):176-179.
52. Namsor: name checker for gender, origin and ethnicity classification. Accessed April 2, 2022. <https://namsor.app/>
53. Sebo P. How well does NamSor perform in predicting the country of origin and ethnicity of individuals based on their first and last names? *PLoS One.* 2023;18(11). doi:10.1371/JOURNAL.PONE.0294562
54. B05006: Place of Birth for the Foreign-Born Population in the United States. United States Census

- Bureau. Accessed February 7, 2023.
[https://data.census.gov/table?q=armenian&g=0100000US_0400000US06,06\\$0500000&tid=ACSDT5YSPT2015.B05006](https://data.census.gov/table?q=armenian&g=0100000US_0400000US06,06$0500000&tid=ACSDT5YSPT2015.B05006)
55. B05012: Nativity in the United States. United States Census Bureau. Accessed February 7, 2023.
[https://data.census.gov/table?q=armenian&g=0100000US_0400000US06,06\\$0500000&tid=ACSDT5YSPT2015.B05012](https://data.census.gov/table?q=armenian&g=0100000US_0400000US06,06$0500000&tid=ACSDT5YSPT2015.B05012)
 56. About Community - Embassy of Armenia to the United States of America. Accessed February 13, 2020. <http://usa.mfa.am/en/community-overview/>
 57. Movsisyan AS, Hofer BM, Maguire FB, et al. *Cancer in California, 1988-2016.*; 2019.
 58. Nikoghosyan M, Hakobyan S, Hovhannisyan A, Loeffler-Wirth H, Binder H, Arakelyan A. Population Levels Assessment of the Distribution of Disease-Associated Variants With Emphasis on Armenians - A Machine Learning Approach. *Front Genet.* 2019;10(APR). doi:10.3389/FGENE.2019.00394
 59. Singh-Carlson S, Wong F, Oshan G, Lail H. Recognition to Identify Patients of South Asian Ethnicity within the Cancer Registry. *J Oncol Nurs.* 2016;3(1):86-92. doi:10.4103/2347-5625.170224
 60. Nasser K, Mills PK, Allan M. Cancer Incidence in the Middle Eastern Population of California, 1988-2004. *Asian Pacific J Cancer Prev.* 2007;8:405-411. doi:10.1111/j.1524-4741.2009.00694.x
 61. VSB Data Applications. Accessed May 19, 2022.
<https://www.cdph.ca.gov/Programs/CHSI/Pages/Data-Applications.aspx>
 62. About California Cancer Registry - California Cancer Registry. Accessed December 9, 2019.
<https://www.ccrca.org/learn-about-ccr/about-cancer-registries/>
 63. About Armenia - Demographics - The Government of the Republic of Armenia. Accessed May 19, 2022. <https://www.gov.am/en/demographics/>
 64. Soundex System | National Archives. Accessed April 14, 2022.
<https://www.archives.gov/research/census/soundex>
 65. Yumaguzin V, Vinnik M. Surnames in modern Russia. *Ann Hum Biol.* 2019;46(6):475-490. doi:10.1080/03014460.2019.1685130/FORMAT/EPUB
 66. Clair-sobell JST, Carlsen I. The Structure of Russian Surnames. *Can Slavon Pap / Rev Can des Slavistes.* 1959;4:42-60. <https://www.jstor.org/stable/40866066%0AJSTOR>
 67. Bayar Y. Navigating the nationalist landscape of exclusion: Armenian citizens of Turkey and the politics of naming. *Br J Sociol.* 2021;72(4):1127-1140. doi:10.1111/1468-4446.12846
 68. Savchick L. Skills Mapping Through Big Data a Case Study of Armenian Diaspora in the United States of America and France. *Int Organ Migr.* Published online 2019:1-66. Accessed May 17, 2021. www.iom.int/countries/Armenia
 69. NamSor. *Inferring Gender from Names in Any Region, Language, or Alphabet.*; 2019.
 70. Armenia - Countries - Office of the Historian. Accessed May 19, 2022.
<https://history.state.gov/countries/armenia>
 71. Azhniuk B, Azhniuk L. *Translating Personal Names in the USSR Successor States: Cross-Cultural and Sociolinguistic Perspectives.* doi:10.2436/15.8040.01.30
 72. Understanding Russian-Language Naming Conventions to Enhance Investigations - Sayari Learn. Accessed May 24, 2022. <https://learn.sayari.com/understanding-russian-language-naming-conventions-to-enhance-investigations/>
 73. Quan H, Wang F, Schopflocher D, et al. Development and validation of a surname list to define Chinese ethnicity. *Med Care.* 2006;44(4):328-333. doi:10.1097/01.mlr.0000204010.81331.a9
 74. Race/Ethnicity and the 2020 Census — Census 20/20. Accessed April 13, 2021.
<https://www.census2020now.org/faces-blog/same-sex-households-2020-census-r3976>

75. Bureau UC. State Population Totals and Components of Change: 2020-2021. Accessed October 12, 2022. <https://www.census.gov/data/tables/time-series/demo/popest/2020s-state-total.html>
76. Yeretdzian AS. *A History of Armenian Immigration to America with Special Reference to Conditions in Los Angeles*. University of Southern California; 1923.
77. Klapheke AK, Carvajal-Carmona LG, Cress RD. Racial/ethnic differences in survival among gastric cancer patients in California. *Cancer Causes Control*. 2019;30:687-696. doi:10.1007/s10552-019-01184-0
78. Bergmans R, Soliman AS, Ruterbusch J, et al. Cancer incidence among Arab Americans in California, Detroit, and New Jersey SEER Registries. *Am J Public Health*. 2014;104(6):83-91. doi:10.2105/AJPH.2014.301954
79. Ellis L, Abrahao R, McKinley M, et al. Colorectal Cancer Incidence Trends by Age, Stage, and Racial/Ethnic Group in California, 1990-2014. *Cancer Epidemiol Biomarkers Prev*. Published online 2018;9. doi:10.1158/1055-9965.EPI-18-0030
80. Ziadeh C, Ziogas A, Jiang L, Anton-Culver H. Breast cancer characteristics in Middle Eastern Women Immigrants Compared with Non-Hispanic White Women in California. *JNCI Cancer Spectr*. 2018;2(2). doi:10.1093/jncics/pky014
81. Ziadeh C, Ziogas A, Anton-Culver H. Cancer risk in different generations of Middle Eastern immigrants to California, 1988-2013. *Int J Cancer*. 2017;141:2260-2269. doi:10.1002/ijc.30928
82. Phelan J, Jones NA, Konya S, et al. 2015 National Content Test Race and Ethnicity Analysis Report. Published online 2017.
83. Glendale Adventist Medical Center. *2014 Cancer Services Annual Report*.; 2014.
84. California Cancer Registry, Chronic Disease Surveillance and Research Branch. Policies and Procedures for Access to and Disclosure of Confidential Data From the California Cancer Registry January 2014. 2014;2(January). http://www.ccrca.org/Cancer_Reporting/Reporting_Legislation_and_Regulation.shtml
85. Hofferkamp J. Standards for Cancer Registries Volume III. Published online 2008. Accessed February 12, 2024. <http://www.naaccr.org>.
86. Tucker TC, Durbin EB, McDowell JK, Huang B. Unlocking the potential of population-based cancer registries. *Cancer*. 2019;125(21):3729-3737. doi:10.1002/cncr.32355
87. *American Community Survey: Data Suppression*.; 2016. Accessed February 12, 2024. <https://www.census.gov/programs-surveys/acs/technical-documentation/data-suppression.html>
88. Table 1. First, Second, and Total Responses to the Ancestry Question by Detailed Ancestry Code: 2000. Accessed January 19, 2023. www.census.gov/prod/cen2000/doc/sf3.pdf
89. Frequently Asked Questions (FAQs) About Ancestry. Accessed February 9, 2023. <https://www.census.gov/topics/population/ancestry/about/faq.html>
90. Ancestry | American Community Survey | U.S. Census Bureau. Accessed February 9, 2023. <https://www.census.gov/acs/www/about/why-we-ask-each-question/ancestry/>
91. Census Bureau U. American Community Survey and Puerto Rico Community Survey 2021 Code List.
92. U.S. Census Populations With Bridged Race Categories. Accessed August 29, 2023. https://www.cdc.gov/nchs/nvss/bridged_race.htm
93. Race / Ethnic Population with Age and Sex Detail, 1970–1989 | Department of Finance. Accessed August 29, 2023. <https://dof.ca.gov/forecasting/demographics/estimates/race-ethnic-population-with-age-and-sex-detail-1970-1989/>
94. Download Match*Pro Software. Accessed April 13, 2021. <https://surveillance.cancer.gov/matchpro/download>

95. Site Recode ICD-O-3/WHO 2008 - SEER Data Reporting Tools. Accessed November 8, 2020. https://seer.cancer.gov/siterecode/icdo3_dwho/home/index.html
96. Fritz A, Percy C, Jack A, et al. ICD-O International Classification of Diseases for Oncology First Revision. Published online 2013. Accessed September 1, 2023. www.who.int
97. Boyle P, Parkin DM. Cancer registration: principles and methods. *Statistical methods for registries. IARC Sci Publ.* 1991;(95):126-158.
98. Schwartz KL, Kulwicki A, Weiss LK, et al. Cancer among Arab Americans in the Metropolitan Detroit Area. *Ethn Dis.* 2004;14:141-146. Accessed July 9, 2020. http://digitalcommons.wayne.edu/fmhs_pubs/1
99. Mantey J, Ruterbusch J, Meza R, Schwartz K. Cancer incidence trends using American Community Survey estimates are not consistent with SEER for small populations. *Cancer Epidemiol.* 2016;43:87-91. doi:10.1016/j.canep.2016.06.014
100. Ilic M, Ilic I. Epidemiology of stomach cancer. *World J Gastroenterol.* 2022;28(12):1187-1203. doi:10.3748/wjg.v28.i12.1187
101. Ma K, Baloch Z, He TTB, Xia X. Alcohol Consumption and Gastric Cancer Risk: A Meta-Analysis. *Med Sci Monit.* 2017;23:238-246. doi:10.12659/MSM.899423
102. Andreasyan D. The problem of being overweight among the Armenian population. *Public Heal Panor.* 2017;3(4):764-771.
103. Zuo Y, Jing Z, Bie M, Xu C, Hao X, Wang B. Association between Helicobacter pylori infection and the risk of colorectal cancer A systematic review and meta-analysis. *Medicine (Baltimore).* 2020;99(37):1-8. doi:10.1097/MD.00000000000021832
104. Maguire FB, Movsisyan AS, Morris CR, Parikh-Patel A, Keegan THM, Tong EK. Evaluation of Cancer Deaths Attributable to Tobacco in California, 2014-2019. *JAMA Netw Open.* 2022;5(12):e2246651-e2246651. doi:10.1001/JAMANETWORKOPEN.2022.46651
105. Smoking Rates by Country 2023. Accessed September 12, 2023. <https://worldpopulationreview.com/country-rankings/smoking-rates-by-country>
106. Farrington J, Kontsevaya A, Fediaev D, et al. *Prevention and Control of Noncommunicable Diseases in Armenia: The Case for Investment.*; 2019. Accessed September 13, 2023. <http://www.euro.who.int/pubrequest>
107. Reece S, Morgan C, Parascandola M, Siddiqi K. Secondhand smoke exposure during pregnancy: a cross-sectional analysis of data from Demographic and Health Survey from 30 low-income and middle-income countries. *Tob Control.* 2019;28(4):420-426. doi:10.1136/TOBACCOCONTROL-2018-054288
108. Arshakyan M, Rocchi M, Arshakyan G, Chiarantini L. Disparities in Cancer Incidence in Armenia, 2013-2016: Setting Priorities for Preventive Interventions. *Int J Healthc Sci.* 6:158-165. Accessed September 20, 2023. www.researchpublish.com
109. Kitahara CM, Schneider AB. Epidemiology of Thyroid Cancer. *Cancer Epidemiol Biomarkers Prev.* 2022;31(7):1284-1297. doi:10.1158/1055-9965.EPI-21-1440
110. Arakelyan VB, Khachatryan GE, Nalbandyan-Schwarz AG, Mothersill CE, Seymour CB, Korogodina VL. Main radiation pathways in the landscape of Armenia. *Int J Radiat Biol.* 2023;99(8):1178-1187. doi:10.1080/09553002.2023.2172623
111. Hutchings N, Aghajanova E, Baghdasaryan S, et al. A stratified cross-sectional cluster model survey of iodine nutrition in Armenia after a decade of universal salt iodization. *Endocr Pract.* 2019;25(10):987. doi:10.4158/EP-2018-0634
112. Rossi L, Branca F. Salt iodisation and public health campaigns to eradicate iodine deficiency disorders in Armenia. *Public Health Nutr.* 2003;6(5):463-469. doi:10.1079/PHN2003461

113. Ghukasyan NN. Management of pregnancy and the possibility of normal delivery in pregnant women with thyroid cancer in Armenia. *World J Adv Res Rev.* 2022;2022(03):634-637. doi:10.30574/wjarr.2022.16.3.1365
114. Gomez SL, Glaser SL, Kelsey JL, Lee MM. Bias in completeness of birthplace data for Asian groups in a population-based cancer registry (United States). *Cancer Causes Control.* 2004;15(3):243-253. doi:10.1023/B:CACO.0000024244.91775.64/METRICS
115. Gomez SL, Glaser SL. Quality of cancer registry birthplace data for Hispanics living in the United States. *Cancer Causes Control.* 2005;16(6):713-723. doi:10.1007/S10552-005-0694-7/METRICS
116. Stats of the States - Cancer Mortality. Accessed November 30, 2023. https://www.cdc.gov/nchs/pressroom/sosmap/cancer_mortality/cancer.htm
117. Ellis L, Canchola AJ, Spiegel D, Ladabaum U, Haile R, Gomez SL. Racial and ethnic disparities in cancer survival: The contribution of tumor, sociodemographic, institutional, and neighborhood characteristics. *J Clin Oncol.* 2018;36(1):25-33. doi:10.1200/JCO.2017.74.2049
118. Thornton RLJ, Glover CM, Cené CW, Glik DC, Henderson JA, Williams DR. Evaluating Strategies For Reducing Health Disparities By Addressing The Social Determinants Of Health. *Health Aff.* 2016;35(8):1416-1423. doi:10.1377/hlthaff.2015.1357
119. Most & Least Diverse States in America (2023). Accessed November 8, 2023. <https://wallethub.com/edu/most-least-diverse-states-in-america/38262>
120. Maguire FB, Movsisyan AS, Morris CR, Parikh-Patel A, Keegan THM, Wun T. Assessing the Burden of Cancer among Californians, 1988-2018. *Calif Cancer Report Epidemiol Surveill Program, Univ Calif Davis Compr Cancer Center, Univ Calif Davis.* Published online 2021:53. Accessed February 6, 2023. <https://health.ucdavis.edu/cancer/about/cancer-registry-data.html>
121. Morris CR, Movsisyan AS, Hofer BM, Parikh-Patel A, Kizer KW. Cancer Burden among Native Americans in California. *Calif Cancer Report Epidemiol Surveill Program, Inst Popul Heal Improv Univ Calif Davis.* Published online 2019. Accessed February 6, 2023. <https://health.ucdavis.edu/cancer/about/cancer-registry-data.html>
122. Morris CR, Movsisyan AS, Villazana RM, et al. *The Burden of Cancer among Hispanic/Latinos in California.*; 2022. <https://health.ucdavis.edu/cancer/pdfs/cancer-burden-hispanic-latinos-in-california.pdf>
123. Yepiskoposyan L, Hovhannisyan A, Khachatryan Z. Genetic Structure of the Armenian Population. *Arch Immunol Ther Exp (Warsz).* 2016;64:113-116. doi:10.1007/s00005-016-0431-9
124. Aintablian H. Direct ancestry to a genocide survivor has transgenerational effects on mental health; a case of the Armenian population. *MOJ Public Heal.* 2018;7(4):233-239. doi:10.15406/mojph.2018.07.00235
125. Choi SH, Terrell JE, Bradford CR, et al. Does Quitting Smoking Make a Difference Among Newly Diagnosed Head and Neck Cancer Patients? *Nicotine Tob Res.* Published online 2016:2216-2224. doi:10.1093/ntr/ntw189
126. Kvaavik E, Weemes Grøtting M, Halkjelsvik T, et al. The effect of a smoking cessation program for patients in cancer treatment: a quasi-experimental intervention study. *Acta Oncol (Madr).* Published online 2023. doi:10.1080/0284186X.2023.2277883
127. Movsisyan Vernon AS, Hoch JS, Fejerman L, Keegan TH. Cancer incidence among Armenians in California. *Cancer Med.* 2024;13(5):e7100. doi:10.1002/CAM4.7100
128. American Cancer Society. *Cancer Facts and Figures 2023.*; 2023.
129. Maguire FB, Movsisyan AS, Morris CR, Parikh-Patel A, Keegan THM, Wun T. *Cancer in California: Assessing the Burden of Cancer among Californians 1988-2019.*; 2022. Accessed April 6, 2023. <https://health.ucdavis.edu/cancer/about/cancer-registry-data.html>
130. Learn About CCR - California Cancer Registry. Accessed April 22, 2021.

<https://www.ccrca.org/learn-about-ccr/>

131. Summary Staging | SEER Training. Accessed January 30, 2024. <https://training.seer.cancer.gov/staging/systems/summary/>
132. Ellis L, Canchola AJ, Spiegel D, Ladabaum U, Haile R, Gomez SL. Trends in cancer survival by health insurance status in California from 1997 to 2014. *JAMA Oncol.* 2018;4(3):317-323. doi:10.1001/jamaoncol.2017.3846
133. Patel MI, McKinley M, Cheng I, Haile R, Wakelee H, Gomez SL. Lung cancer incidence trends in California by race/ethnicity, histology, sex, and neighborhood socioeconomic status: An analysis spanning 28 years. *Lung Cancer.* 2017;108:140-149. doi:10.1016/J.LUNGCAN.2017.03.014
134. Morris CR, Movsisyan AS, Hofer BM, et al. *The Burden of Cancer among Black/African Americans in California.*; 2021. Accessed April 6, 2023. <https://www.ccrca.org/learn-about->
135. Who's eligible for Medicare? | HHS.gov. Accessed January 30, 2024. <https://www.hhs.gov/answers/medicare-and-medicaid/who-is-eligible-for-medicare/index.html>
136. Kuitunen I, Ponkilainen VT, Uimonen MM, Eskelinen A, Reito A. Testing the proportional hazards assumption in cox regression and dealing with possible non-proportionality in total joint arthroplasty research: methodological perspectives and review. *BMC Musculoskelet Disord.* 2021;22(1):1-7. doi:10.1186/S12891-021-04379-2/TABLES/2
137. Keegan THM, Parsons HM, Chen Y, et al. Impact of Health Insurance on Stage at Cancer Diagnosis Among Adolescents and Young Adults. *J Natl Cancer Inst.* 2019;111(11):1152-1160. doi:10.1093/jnci/djz039
138. Primm KM, Huepenbecker SP, Zhao H, et al. Impacts of Medicaid Expansion on Stage at Cancer Diagnosis by Patient Insurance Type. *Am J Prev Med.* 2022;63(6):915-925. doi:10.1016/j.amepre.2022.06.001
139. Alcaraz KI, Wiedt TL, Daniels EC, Yabroff KR, Guerra CE, Wender RC. Understanding and addressing social determinants to advance cancer health equity in the United States: A blueprint for practice, research, and policy. *CA Cancer J Clin.* 2020;70(1):31-46. doi:10.3322/caac.21586
140. Dee EC, Gomez SL. Cancer among Immigrants: Diverse Histories, Diverse Disparities, Diverse Opportunities to Promote Equity. *Cancer Epidemiol Biomarkers Prev.* 2022;31(7):1251-1253. doi:10.1158/1055-9965.EPI-22-0337
141. Breen N, Rao SR, Meissner HI. Immigration, health care access, and recent cancer tests among Mexican-Americans in California. *J Immigr Minor Heal.* 2010;12(4):433-444. doi:10.1007/s10903-008-9198-3
142. Siddiq H, Pavlish C, Alemi Q, Menten J, Lee E. Beyond Resettlement: Sociocultural Factors Influencing Breast and Colorectal Cancer Screening Among Afghan Refugee Women. *J Cancer Educ.* Published online 2020:10. doi:10.1007/s13187-020-01822-1
143. Yeom JW, Yeom IS, Park HY, Lim SH. Cultural factors affecting the self-care of cancer survivors: An integrative review. *Eur J Oncol Nurs.* 2022;59:102165. doi:10.1016/J.EJON.2022.102165
144. Turra CM, Elo IT. The Impact of Salmon Bias on the Hispanic Mortality Advantage: New Evidence from Social Security Data. *Popul Res Policy Rev.* 2008;27:515-530. doi:10.1007/s11113-008-9087-4
145. Miller KD, Ortiz AP, Pinheiro PS, et al. Cancer statistics for the US Hispanic/Latino population, 2021. *CA Cancer J Clin.* 2021;71(6):466-487. doi:10.3322/CAAC.21695
146. Markides KS, Rote S. The Healthy Immigrant Effect and Aging in the United States and Other Western Countries. *Gerontologist.* 2019;59(2):205-214. doi:10.1093/GERONT/GNY136
147. Horvath S, Gurven M, Levine ME, et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. *Genome Biol.* 2016;17(1):1-23. doi:10.1186/S13059-016-1030-0/FIGURES/7

148. Movsesian AA, Mkrtchyan RA, Simonyan HG. The Bronze and Iron Age populations of the Armenian Highland in the genetic history of Armenians. *Am J Phys Anthropol.* 2020;173(1):158-167. doi:10.1002/AJPA.24060
149. Pinheiro PS, Morris CR, Liu L, Bungum TJ, Altekruze SF. The Impact of Follow-up Type and Missed Deaths on Population-Based Cancer Survival Studies for Hispanics and Asians. *JNCI Monogr.* 2014;2014(49):210-217. doi:10.1093/JNCIMONOGRAPHS/LGU016
150. Ka'opua LSI, Park SH, Ward ME, Braun KL. Testing the feasibility of a culturally tailored breast cancer screening intervention with Native Hawaiian women in rural churches. *Health Soc Work.* 2011;36(1):55-65. doi:10.1093/hsw/36.1.55
151. Haber M, Mezzavilla M, Xue Y, et al. Genetic evidence for an origin of the Armenians from Bronze Age mixing of multiple populations. *Eur J Hum Genet.* 2016;24:931-936. doi:10.1038/ejhg.2015.206
152. US Armenians Need Better Demographic Studies: Changes in the US Census Will Help - The Armenian Mirror-Spectator. Accessed May 5, 2020. <https://mirrorspectator.com/2015/09/10/us-armenians-need-better-demographic-studies-changes-in-the-us-census-will-help/>
153. American colonies | Facts, History, and Definition | Britannica. Accessed January 2, 2024. <https://www.britannica.com/topic/American-colonies>
154. Doidge JC. Demystifying probabilistic linkage: Common myths and misconceptions. *Int J Popul Data Sci J Website www.ijpds.org.* 2018;3(1). doi:10.23889/ijpds.v3i1.410
155. MDAT. Accessed April 20, 2021. <https://data.census.gov/mdat/#/search?ds=ACSPUMS5Y2019&rv=ucgid,ANC1P&wt=PWGTP&g=0400000US06>