**Additional Details of Methods**

This appendix presents the details of the methods used to estimate intercensal population numbers and excess mortality. The description of the population estimation process also includes some sensitivity analyses comparing different scenarios of the rates of decline of the non-AN population between 1916 and 1918. Five scenarios were considered; results are presented for three of these five scenarios. The same three scenarios were used in the estimation of excess mortality to assess the sensitivity of the results to the population estimates used.

***I. Estimation of intercensal population numbers***

Because of the significant differences in population history between the AN and non-AN populations (described in the main text), different models had to be used to estimate the underlying population numbers needed to analyze excess mortality. The differences in population history also required that the two subpopulations be kept separate.

Historical data suggest that the major declines in the non-AN population occurred between 1916 and 1918 (1). All intercensal population numbers were estimated using Microsoft Excel 365. The process for estimating these numbers for the non-AN group was as follows:

1. An exponential growth rate for the decade between 1910 and 1920 was calculated from the reported population figures in those two years (2,3).
2. Historical evidence indicates that about half the population loss during the decade occurred between 1916 and 1918. Thus, the model used to interpolate the non-AN population initially assumed that 50% of the total change between 1910 and 1920 occurred between those years, with a more gradual decline outside of them.
3. A second exponential model with a growth rate determined from the census populations in 1920 and 1930 was used to estimate the gradual population increase observed between 1920 and 1930 (3,4) (Web Appendix Figure 1).



**Web Appendix Figure 1.** The model used to estimate the non-Alaska Native intercensal populations, 1910-1930.

1. Sensitivity analyses (discussed below) were performed that varied the proportion of the overall 1910-20 population decline that occurred between 1916 and 1918. These proportions ranged from no difference between 1917/1918 and the rest of the decade to 5.0 times faster rate of decline between 1916 and 1918 than in the rest of the decade. In all scenarios the 1920-30 model was one of exponential growth, with the growth rate calculated from the reported 1920 and 1930 census populations. Web Appendix Table 1 gives the growth rates used in each scenario and the ratios between the 1917-18 growth rate and the remaining years.
2. Models with the same basic shape were then used to estimate intercensal populations for all age groups. Estimations were based on exponential growth rates calculated from census data for each age group in 1910, 1920, and 1930.

**Web Appendix Table 1. Scenarios used in the sensitivity analyses of the non-AN population estimation model.**

|  |  |  |  |
| --- | --- | --- | --- |
| Scenario | Annual growth rate 1910-16, 1919-20 | Annual growth rate 1917-18 | Ratio |
| 1 | -0.0315 | -0.0315 | 1.0 |
| 2 | -0.0269 | -0.0500 | 1.9 |
| 3 | -0.0244 | -0.0600 | 2.5 |
| 4 | -0.0226 | -0.0670 | 3.0 |
| 5 | -0.0175 | -0.0875 | 5.0 |

Because historical evidence was used to determine the shape of the population decline throughout the decade from 1910 to 1920, scenarios varying that basic shape were not tested. As shown in Web Appendix Table 2 (below), however, the population estimates are relatively insensitive to assumptions about the rate of decline between 1916 and 1918.

The model used to interpolate the AN population was also an exponential model, with adjustments made to take into account the extreme depopulation that occurred during the influenza pandemic (Web Appendix Figure 2). The procedure used to estimate the intercensal AN populations was as follows:

1. The reported numbers of P&I deaths in 1918 and 1919 were added to the reported 1920 census population to estimate of what the population might have been if the influenza pandemic had not occurred.
2. An exponential growth rate for the decade from 1910 to 1920 was calculated using the reported 1910 census population and the adjusted 1920 population.
3. This growth rate was used to estimate the intercensal populations between 1910 and 1918 under an assumption of exponential growth.
4. Using the exponential growth model, the 1919 population was estimated by subtracting the reported number of P&I deaths from the model-derived 1919 population. The 1920 population was the reported 1920 census population.
5. A new growth rate for the decade from 1920 to 1930 was calculated using the reported 1920 and 1930 populations and the intercensal populations were estimated using this growth rate in an exponential model. The total number of P&I deaths in 1920 was only 143, distributed across all ages, so these deaths were not subtracted from the exponential estimates of age-specific population (see Table 2 in the main text for the population estimates).



**Web Appendix Figure 2.** The model used to estimate the Alaska Native intercensal populations, 1910-1930.

The Alaska Native population estimation model was based on exponential growth combined with mortality data derived from the death certificates and did not depend on any other assumptions. Thus, no sensitivity analyses were conducted for these estimates.

*Sensitivity analyses of the model used to estimate the non-AN population*

Information from the historical record suggests that about half of the population decline between 1910 and 1920 occurred between 1916 and 1918 (1). To derive the initial model used in estimating intercensal population numbers, the total population decline over the decade was calculated from the 1910 and 1920 census figures. It was then assumed that half of this decline occurred in 1917 and 1918, which resulted in the model seen in Web Appendix Figure 1. A sensitivity analysis was performed to assess how sensitive the resulting population estimates were to the assumed 1917-18 decline. As shown above in Web Appendix Table 1, five scenarios were considered. Web Appendix Table 2 shows the estimated subgroup population figures for the years 1910-1930 for all five scenarios. Population estimates for the AN population are included in the table; these values were the same across all scenarios.

**Web Appendix Table 2. Total population estimates for the AN population and the five non-AN scenarios used in the sensitivity analyses.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Year** | **AN pop.** | **Non-AN 1** | **Non-AN 2** | **Non-AN 3** | **Non-AN 4** | **Non-AN 5** |
| **1910a** | 25331 | 39025 | 39025 | 39025 | 39025 | 39025 |
| **1911** | 25553 | 37815 | 37990 | 38085 | 38152 | 38348 |
| **1912** | 25778 | 36642 | 36983 | 37168 | 37298 | 37683 |
| **1913** | 26004 | 35506 | 36002 | 36273 | 36464 | 37029 |
| **1914** | 26232 | 34405 | 35047 | 35400 | 35648 | 36387 |
| **1915** | 26462 | 33338 | 34118 | 34547 | 34851 | 35755 |
| **1916** | 26694 | 32304 | 32454 | 32535 | 32592 | 32760 |
| **1917** | 26929 | 31303 | 30871 | 30641 | 30480 | 30015 |
| **1918** | 27165 | 30332 | 30053 | 29903 | 29798 | 29494 |
| **1919** | 26473 | 29391 | 29256 | 29183 | 29132 | 28983 |
| **1920a** | 26558 | 28480 | 28480 | 28480 | 28480 | 28480 |
| **1921** | 26882 | 28561 | 28561 | 28561 | 28561 | 28561 |
| **1922** | 27210 | 28642 | 28642 | 28642 | 28642 | 28642 |
| **1923** | 27542 | 28723 | 28723 | 28723 | 28723 | 28723 |
| **1924** | 27878 | 28804 | 28804 | 28804 | 28804 | 28804 |
| **1925** | 28219 | 28886 | 28886 | 28886 | 28886 | 28886 |
| **1926** | 28563 | 28967 | 28967 | 28967 | 28967 | 28967 |
| **1927** | 28912 | 29050 | 29050 | 29050 | 29050 | 29050 |
| **1928** | 29264 | 29132 | 29132 | 29132 | 29132 | 29132 |
| **1929** | 29622 | 29214 | 29214 | 29214 | 29214 | 29214 |
| **1930a** | 29983 | 29297 | 29297 | 29297 | 29297 | 29297 |

**AN – Alaska Native; non-AN – non-Alaska Native**

**a Reported in data from the U.S. Census Bureau (2-4)**

Although there are clearly differences in the estimated populations across the five scenarios, in all cases the difference between the highest and lowest values is well under 10%, and in over half of the years these differences are less than 3%.

Intercensal populations by 5-year age class and year were estimated for scenarios 1, 4, and 5 by calculating the overall growth rate between 1910 and 1920 for each age class (under the assumption of exponential growth) and setting the proportion of the growth rate per year in such a way that the ratios presented in Web Appendix Table 2 were maintained. In other words, the shape of the population curve for each age class was the same as in scenarios 1, 4, and 5, but the estimated growth rate for each age class was calculated from the reported census numbers for that age. The resulting age-specific population numbers can be found at the project github site, <https://github.com/anthlsMU/AKNL-AJE23>.

To further test the sensitivity of model assumptions, the excess mortality analyses were run for scenarios 1, 4, and 5. Results were similar for all scenarios (Web Appendix Table 3).

**Web Appendix Table 3. Excess mortality per 10,000 persons a for scenarios 1, 4, and 5.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Group** | **Scenario 1** | **Scenario 4** | **Scenario 5** |
| Total population | 55.0 (40.0, 83.0) | 55.4 (40.5, 83.4) | 55.6 (40.7, 83.7) |
| 0-4 | 0.6 (≤0, 91.6) | 0.5 (≤0, 94.9) | 0.4 (≤0, 97.0) |
| 5-14 | 0.7 (≤0, 23.4) | 0.5 (≤0, 23.7) | 0.3 (≤0, 24.0) |
| 15-19 | 25.4 (4.1, 81.6) | 25.6 (4.4, 81.2) | 25.7 (4.5, 81.0) |
| 20-24 | 51.5 (33.4, 93.3) | 54.8 (36.5, 95.6) | 56.7 (38.2, 97.4) |
| 25-29 | 395.1 (272.2, 647.4) | 122.3 (90.6, 174.7) | 126.4 (95.7, 176.6) |
| 30-34 | 110.3 (82.6, 160.4) | 113.2 (86.6, 161.4) | 114.9 (88.8, 162.3) |
| 35-39 | 63.9 (39.2, 110.9) | 65.6 (41.1, 111.7) | 66.5 (42.3, 112.2) |
| 40-44 | 34.9 (15.2, 63.0) | 35.7 (15.6, 63.1) | 36.1 (16.2, 63.2) |
| 45-49 | 31.3 (8.4, 71.6) | 31.6 (8.9, 71.2) | 31.7 (9.2, 71.1) |
| 50-54 | 50.0 (25.0, 137.1) | 50.0 (24.9, 137.2) | 50.0 (24.8, 137.3) |
| 55-59 | 70.6 (25.8, 168.7) | 69.8 (24.0, 171.7) | 69.3 (22.8, 173.7) |
| 60-64 | 33.8 (≤0, 164.3) | 33.1 (≤0, 170.5) | 32.7 (≤0, 174.7) |
| 65-69 | 51.7 (≤0, 259.9) | 52.3 (≤0, 269.8) | 52.7 (≤0, 277.5) |
| 70-74 | 257.4 (87.6, 563.4) | 249.8 (82.9, 578.7) | 246.5 (79.9, 591.4) |
| 75+ | 279.7 (≤0, 1009.6) | 259.7 (≤0, 1063.2) | 249.4 (≤0, 1106.5) |

a 95% confidence intervals provided in parentheses

***II. Estimation of excess mortality***

We used SAS 9.4 to perform the excess mortality analyses. This analysis first involved determining a seasonal baseline model for the total population against which the actual mortality could be compared. Serfling [5] and Serfling-like regression have been widely used in previous works to estimate excess mortality (e.g., 6-10). We use a similar approach here.

For the baseline mortality rate for total population, we tested three models:

Model 1:

Model 2:

Model 3:

is the estimated expected death rate at time t,

is the constant term,

is the secular trend term,

is the coefficient of the quadratic term,

and are the terms associated with annual seasonality, and are the terms associated with semi-annual seasonality and and are the terms associated with quarterly seasonality.

is the error term at time t.

We assessed the coefficient of determination (R2) for each of these models; R2 for model 1, model 2 and model 3 were 39.9%, 41.9% and 46.6%, respectively. We also assessed each model’s Akaike Information Criteria (AIC) values. Model 1 had an AIC of 92.3%, Model 2 had an AIC of 94.1% and Model 3 had an AIC of 92.6% (Web Appendix Table 4). Therefore, based on the R2 and AIC value evaluation, we decided to use model 3 as the final model. Web Appendix Table 5 provides the parameter estimates for this model.

**Web Appendix Table 4. Evaluation metrics for three models.**

|  |  |  |
| --- | --- | --- |
| Model | R2 | AIC |
| 1 | 39.9 | 92.3 |
| 2 | 41.9 | 94.1 |
| 3 | 46.6 | 92.6 |

**Web Appendix Table 5. Parameter estimates for the final baseline model for the total population.**

A table of numbers and symbols

Description automatically generated with medium confidence

We then assessed the diagnostics for Model 3. We provide the standard diagnostic plots for the final baseline model for the total population used to estimate the total excess mortality rate. The population used for this plot includes values for the AN population estimated using an exponential model and scenario 5 population values (as described above) for the non-AN population (Web Appendix Figure 1). The “residuals vs predicted value” and “scale location” plots justify assumptions of linearity and homoscedasticity, respectively. The residual vs quantile plot and the distribution of residuals exhibit a normal distribution (Web Appendix Figure 3).

A screenshot of a graph

Description automatically generated

**Web Appendix Figure 3.** Diagnostics for the final baseline model for the total population. The residuals vs predicted value and scale location (Studentized residual vs Predicted value) plots justify assumptions of linearity and homoscedasticity, respectively. The residual vs quantile plot and the distribution of residuals exhibit a normal distribution.

We used the predictors in Model 3 to estimate the baseline mortality rate for subpopulations (AN, non-AN, and different age categories in both these groups). We also provide the standard diagnostics for the AN and non-AN populations in Web Appendix Figures 4 and 5, respectively. The parameter estimates for the baseline model for the AN and non-AN populations are provided in Web Appendix Tables 6 and 7. The standard diagnostics and parameter estimates for some of the AN and non-AN age groups are also provided in Web Appendix Figures 6-11 and Web Appendix Tables 8-13.

All-cause excess mortality rates per 10,000 population were estimated separately among both the AN and non-AN populations of Alaska across 15 different age groups (0-4, 5-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, and 75 years and above. The mortality rate above a seasonal baseline of expected mortality rates in the absence of influenza activity was computed as in previous studies (6-10). The procedure is as follows:

1. Determine the most likely periods of pandemic influenza activity from the time series of monthly all-cause deaths. Months associated with these periods are excluded from modeling the baseline mortality rate.
2. Fit Model 3 described above to all-cause deaths for the total population by ethnic group in non-influenza pandemic months. The resulting models provide estimates of the expected influenza mortality (the baseline) at any time within the study interval. Confidence intervals surrounding the estimates are also calculated.
3. Define periods of pandemic influenza circulation as the months in 1918 through 1920 where the observed total all-cause mortality rate per 10,000 people exceeds the baseline’s upper 95% confidence interval (CI).
4. Use these defined pandemic periods to estimate excess mortality rates according to indigenous status and age group. Excess mortality is the difference between the observed and model-adjusted baseline mortality rates for each month constituting a pandemic outbreak.
5. Sum the excess death rates across all pandemic months from 1918-1920 to get the overall pandemic excess mortality for the AN and non-AN groups and these two groups by age. Negative excess mortality estimates are replaced by zero in our analyses.

A screenshot of a graph

Description automatically generated

**Web Appendix Figure 4.** Diagnostics for final baseline model for the AN population. The residuals vs predicted value and scale location (Studentized residual vs Predicted value) plots justify assumptions of linearity and homoscedasticity, respectively. The residual vs quantile plot and the distribution of residuals exhibit a normal distribution.

**Web Appendix Table 6. Parameter estimates for the final baseline model for the AN population.**

A table with numbers and symbols

Description automatically generated

A screenshot of a graph

Description automatically generated

**Web Appendix Figure 5.** Diagnostics for the final baseline model for excess mortality among the non-AN population. The residuals vs predicted value and scale location (Studentized residual vs Predicted value) plots justify assumptions of linearity and homoscedasticity, respectively. The residual vs quantile plot and the distribution of residuals exhibit a near normal distribution.

**Web Appendix Table 7. Parameter estimates for the final baseline model for the non-AN population.**

A table of numbers and symbols

Description automatically generated with medium confidence

A screenshot of a graph

Description automatically generated

**Web Appendix Figure 6.** Diagnostics for final baseline model for excess mortality among 25-29 year old AN population. The residuals vs predicted value and scale location (Studentized residual vs Predicted value) plots provide reasonable justification for assumptions of linearity and homoscedasticity, respectively. The residual vs quantile plot and the distribution of residuals exhibit a near normal distribution.

**Web Appendix Table 8. Parameter estimates for the final baseline model for 25-29 year old AN population.**

A table with numbers and symbols

Description automatically generated

A chart of data analysis

Description automatically generated with medium confidence

**Web Appendix Figure 7.** Diagnostics for final baseline model for excess mortality among the 0-4 year old non-AN population. The residuals vs predicted value and scale location (Studentized residual vs Predicted value) plots justify assumptions of linearity and homoscedasticity, respectively. The residual vs quantile plot and the distribution of residuals exhibit a near normal distribution.

**Web Appendix Table 9. Parameter estimates for the final baseline model for the 0-4 year old non-AN population.**

A table with numbers and a number of error

Description automatically generated with medium confidence

A screenshot of a graph

Description automatically generated

**Web Appendix Figure 8.** Diagnostics for the final baseline model for excess mortality among the 0-4 year old AN population. The residuals vs predicted value and scale location (Studentized residual vs Predicted value) plots justify assumptions of linearity and homoscedasticity, respectively. The residual vs quantile plot and the distribution of residuals exhibit a near normal distribution.

**Web Appendix Table 10. Parameter estimates for the final baseline model for the 0-4 year old AN population.**

A table with numbers and a number of objects

Description automatically generated with medium confidence

A screenshot of a computer screen

Description automatically generated

**Web Appendix Figure 9.** Diagnostics for the final baseline model for excess mortality among the 40-44 year old non-AN population. The residuals vs predicted value and scale location (Studentized residual vs Predicted value) plots justify assumptions of linearity and homoscedasticity, respectively. The residual vs quantile plot and the distribution of residuals exhibit a near normal distribution.

**Web Appendix Table 11. Parameter estimates for the final baseline model for the 40-44 year old non-AN population.**

A table with numbers and symbols

Description automatically generated

A screenshot of a graph

Description automatically generated

**Web Appendix Figure 10.** Diagnostics for the final baseline model for excess mortality among the 65-69 year old AN population. The residuals vs predicted value and scale location (Studentized residual vs Predicted value) plots justify assumptions of linearity and homoscedasticity, respectively. The residual vs quantile plot and the distribution of residuals exhibit a near normal distribution.

**Web Appendix Table 12. Parameter estimates for the final baseline model for the 65-69 year old AN population.**

A table with numbers and symbols

Description automatically generated

A screenshot of a diagram

Description automatically generated

**Web Appendix Figure 11.** Diagnostics for the final baseline model for excess mortality among the 65-69 year old non-AN population. The residuals vs predicted value and scale location (Studentized residual vs Predicted value) plots justify assumptions of linearity and homoscedasticity, respectively. The residual vs quantile plot and the distribution of residuals exhibit a near normal distribution.

**Web Appendix Table 13. Parameter estimates for the final baseline model for the 65-69 year old non-AN population.**

A table with numbers and letters

Description automatically generated

**REFERENCES CITED**

1. Huskey L. Alaska's economy: The First World War, frontier fragility, and Jack London. *Northern Rev*. 2017;44:327-346
2. U.S. Census Bureau. *Statistics for Alaska. Thirteenth Census of the United States Taken in the Year 1910.* Government Printing Office, Washington, D.C., 1913.
3. U.S. Census Bureau. *Population: Alaska. Fourteenth Census of the United States: 1920.* Government Printing Office, Washington, D.C., 1921.
4. U.S. Census Bureau. *Population—Alaska. Fifteenth Census of the United States: 1930*. Government Printing Office, Washington, D.C., 1933.
5. Serfling RE. Methods for current statistical analysis of excess pneumonia-influenza deaths. *Public Health Rep* *(1896)*, 1963;78(6):494-506.
6. Chowell G, Simonsen L, Flores J, et al. Death patterns during the 1918 influenza pandemic in Chile. *Emerg Infect Dis*. 2014;20(11):1803-1811.
7. Viboud C, Eisenstein J, Reid AH, et al. Age- and sex-specific mortality associated with the 1918-1919 influenza pandemic in Kentucky. *J Infect Dis*. 2013;207(5):721-729
8. Modig K, Lambe M, Ahlbom A, et al. Excess mortality for men and women above age 70 according to level of care during the first wave of COVID-19 pandemic in Sweden: A population-based study. *Lancet Reg Health Eur.* 2021;4:100072.
9. Dahal S, Banda JM, Bento AI, et al. Characterizing all-cause excess mortality patterns during COVID-19 pandemic in Mexico. *BMC Infect Dis*. 2021;21:E432.
10. Dahal S, Jenner M, Dinh L, et al. Excess mortality patterns during 1918-1921 influenza pandemic in the state of Arizona, USA. *Ann Epidemiol*. 2019;28(5):273-280.