Mortality
and Longevity

MIM-2021-v3 Data Analysis Tool
User Guide

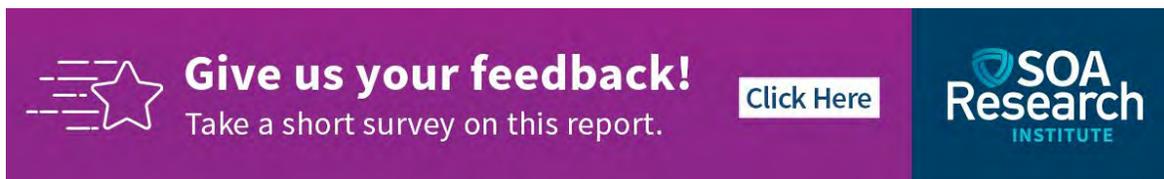OCTOBER | 2022

SOA
Research
INSTITUTE

MIM-2021-v3 Data Analysis Tool User Guide

AUTHOR	MIM Advisory Group	**SPONSOR**	Mortality and Longevity Strategic Research Program Steering Committee

CONTENTS

MIM-2021-v3 Data Analysis Tool
User Guide

Introduction

As described in the MIM-2021-v2 Report[1], the selection of an appropriately chosen historical mortality data set is an essential requirement of the MIM-2021-v3 projection methodology. This document is a step-by-step guide for the MIM-2021-v3 Data Analysis Tool that can help users compare and analyze the various historical mortality data sets available in the MIM-2021-v3, as well as compare those historical rates to previously published SOA mortality tables.

It should be noted the MIM-2021-v3 Data Analysis Tool deals only with historical mortality rates. It does not compare or analyze any of the future mortality rates developed under the MIM-2021-v3. This is the third version of the Data Analysis Tool. There are no changes from the MIM-2021-v2 Data Analysis Tool released in October 2021. The name of the Data Analysis Tool was changed to MIM-2021-v3 to be consistent with the current iteration of the MIM-2021-v3 Application Tool[2].

Throughout this User Guide, the term "NCHS rates" without any further designation refers to all of the following historical data sets:

- Full population mortality rates developed from National Center for Health Statistics (NCHS) death counts starting in 1982 and U.S. census data going back to 1980.
- A variety of subsets of the NCHS full population data set, split into ranges (quintiles and deciles) based on estimated socioeconomic categories; (site Mortality by Socioeconomic Category in the United States). These subsets can be selected individually or blended in combination with other subsets to more closely approximate the anticipated mortality experience of the covered population to which the resulting mortality improvement rates are to be applied.
- An additional group of "static" NCHS decile data sets. These deciles have been based on socioeconomic categories fixed as of 2000[3], as contrasted to the subsets described in the immediately preceding bullet, which reflect changes in socioeconomic deciles as they might have varied over time.

Whenever it becomes necessary to distinguish among any of these NCHS data sets, some additional description will be supplied.

---

[1] Society of Actuaries. 2021. "Developing A Consistent Framework For Mortality Improvement". https://www.soa.org/resources/research-reports/2021/developing-consistency-framework/.

[2] The MIM-2021-v3 Application Tool was released in October 2022 and includes enhancements from the prior version. https://www.soa.org/resources/research-reports/2022/mortality-improvement-model/.

[3] Although the NCHS mortality data begins in 1982, the decennial census as of 2000 was used to establish socioeconomic categories for this group of static NCHS data sets. Given the decennial census occurs only once every 10 years (e.g., 1970, 1980, 1990, etc.), the 2000 census is positioned closest in time to the mid-point year of the NCHS mortality data period (1982-2019).

Tabs 1, 2 and 3: Historical Mortality Datasets

Tabs 1, 2 and 3 include all of the historical SSA and NCHS mortality rates (smoothed[4] and unsmoothed, respectively) required to produce the various graphs, heat maps, and tables generated in the remaining tabs of the spreadsheet. The rates are shown for every calendar year starting in 1982 through 2019.

Two-dimensional Whittaker-Henderson smoothing was performed on the unsmoothed mortality rates starting at age 15 through age 97. Therefore, Tab 2 (using order-2 Whittaker-Henderson) and Tab 3 (using order-3 Whittaker Henderson) include smoothed gender-specific SSA and NCHS mortality rates between those two ages only, while Tab 1 includes the unsmoothed gender-specific SSA[5] and NCHS mortality rates from ages 0 through 110.

In all three -- Tab 1, Tab 2 and Tab 3 -- specific mortality rates can be readily identified by "filtering" on the down arrows in cells B6 through E6:

- Column B, "Data": Choose from the desired historical SSA or NCHS data set(s) from the following options:
    - "d": NCHS decile data sets, with deciles that vary in accordance with each county's socioeconomic categorization over the period 1982 through 2019.
    - "d-2000": NCHS decile data sets, with socioeconomic categories fixed as of 2000
    - "q": NCHS non-static quintile data sets over the period 1982-2019.
    - "q-2000": NCHS decile data sets, with socioeconomic categories fixed as of 2000
    - "ssa": SSA data set
- Column C, "Decile or Quintile": For the NCHS rates, choose from the desired decile(s), quintile(s), or "total"[6]:
    - Numeric value from 1 to 10: The value "1" represents the lowest socioeconomic scored population group for both decile and quintile data sets. The data sets representing the highest socioeconomic scored population groups are "5" for quintiles and "10" for deciles.
    - "total": The full-population data set
- Column D, "Year": Choose the desired calendar year(s); 1982 through 2019
- Column E, "Gender": Choose the desired gender(s); "f" or "m"

Tabs 1, 2, and 3 are "protected" sheets, meaning they cannot be altered. This is necessary because if the structure of the data were inadvertently modified, then various worksheet tabs that use the data could malfunction.

---

[4] See Appendix A of the MIM-2021-v2 report for a description of the methodology used to smooth historical mortality rates. Society of Actuaries. 2021. "Developing A Consistent Framework For Mortality Improvement". https://www.soa.org/resources/research-reports/2021/developing-consistency-framework/.
[5] The SSA mortality rates included in Tab 2 and Tab 3 are those that were published in connection with the annual Trustees' Reports, and subsequently reflect the SSA's standard practice of smoothing their mortality rates within calendar years, but not between calendar years.
[6] For the SSA data set, "total" is the only available option in column C.

Tab 4: Graphical Comparison of National Level SSA and NCHS Mortality Information

As discussed in Section 3 of the MIM-2021-v2 Report, the SSA and NCHS data sets were derived from different sources for ages 65 and over. Tab 4 allows users to visualize differences in the underlining mortality rates and assess the impact of those differences on various related metrics.

Figure 1 displays an example of the inputs for Tab 4.
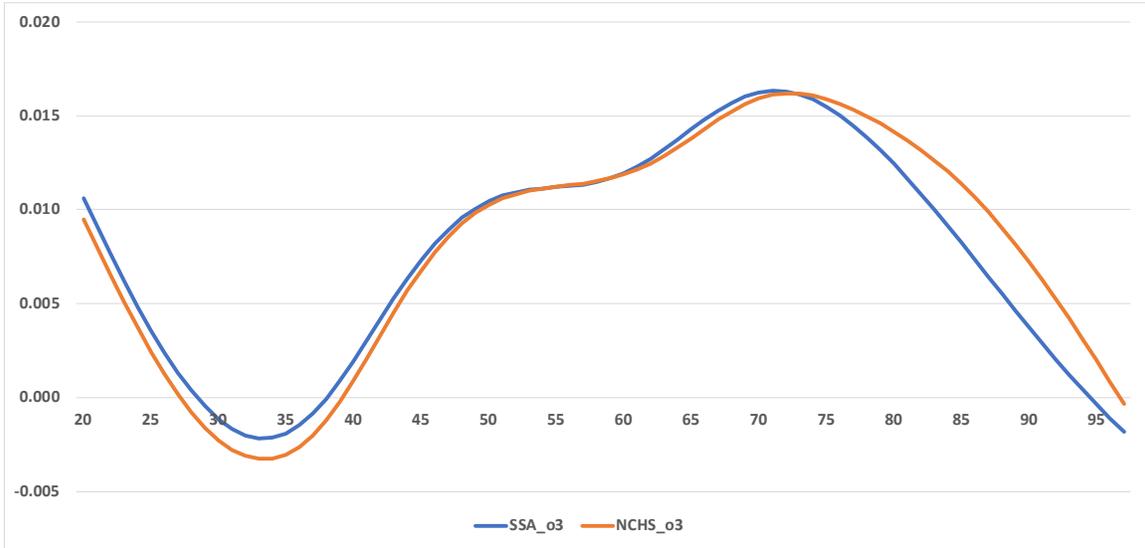
Figure 1
TAB 4 INPUTS

Select Parameters

10	Metric (see list below)
m	Sex
1982	Earliest year
2019	Final year
20	Min age to plot
97	Max age to plot

Select Datasets to Display

No	SSA unsmoothed (*)
No	SSA with order 2 smoothing
Yes	SSA with order 3 smoothing
No	NCHS unsmoothed
No	NCHS with order 2 smoothing
Yes	NCHS with order 3 smoothing

List of Metrics

1. qx in earliest year
2. qx in final year
3. log qx in earliest year
4. log qx in final year
5. qx in final year as % of qx in earliest year
6. qx in final year minus qx in earliest year
7. expected lifetime in earliest year
8. expected lifetime in final year
9. metric 8 minus metric 7
10. annualized geometric rate of improvement from earliest year to final year

The graph produced with this set of inputs (Figure 2) compares the annualized geometric rate of mortality improvement from 1982 through 2019 for males aged 20 through 97, based on the smoothed SSA and NCHS data sets.

Figure 2
TAB 4 GRAPH SHOWING ANNUALIZED GEOMETRIC RATE OF MORTALITY IMPROVEMENT

National-Level Mortality Data: SSA versus NCHS

Tabs 5 and 6: Graphical Comparisons of NCHS Socioeconomic Quintile and Decile Mortality Information

Users can compare a range of key mortality metrics generated by each of the NCHS quintile (Tab 5) and decile (Tab 6) data sets. Figure 3 displays an example of the inputs for Tab 5.

Figure 3
TAB 5 INPUTS

Select Parameters

|	
1	Data: 1 = dynamic quintiles; 2 = static quintiles
3	Smoothing: 1 = unsmoothed, 2 = smoothed o2; 3 = smoothed o3
Yes	Graph national-level data, in addition to quintiles?
f	Sex
1982	Earliest year
2019	Final year

Select Metric from the List Below

14

1. annualized geometric rate of improvement from earliest year to final year
2. like metric 1, but expressed relative to the national average rate of improvement
3. qx in final year as % of qx in earliest year
4. qx in final year minus qx in earliest year
5. qx in earliest year
6. qx in final year
7. log qx in earliest year
8. log qx in final year
9. qx in earliest year as % of national-level value (quintile qx / national level qx)
10. qx in final year as % of national-level value (quintile qx / national level qx)
11. metric 10 minus metric 9
12. expected lifetime in earliest year
13. expected lifetime in final year
14. metric 13 minus metric 12

The bar graph produced with this set of inputs (Figure 4) shows for each NCHS quintile (and for the national-level NCHS data set) a comparison of the change in period life expectancies between 1982 and 2019 for females at eight ages, based on smoothed mortality data using order-3 Whittaker-Henderson.

Figure 4

TAB 5 BAR GRAPH COMPARING THE CHANGE IN PERIOD LIFE EXPECTANCIES
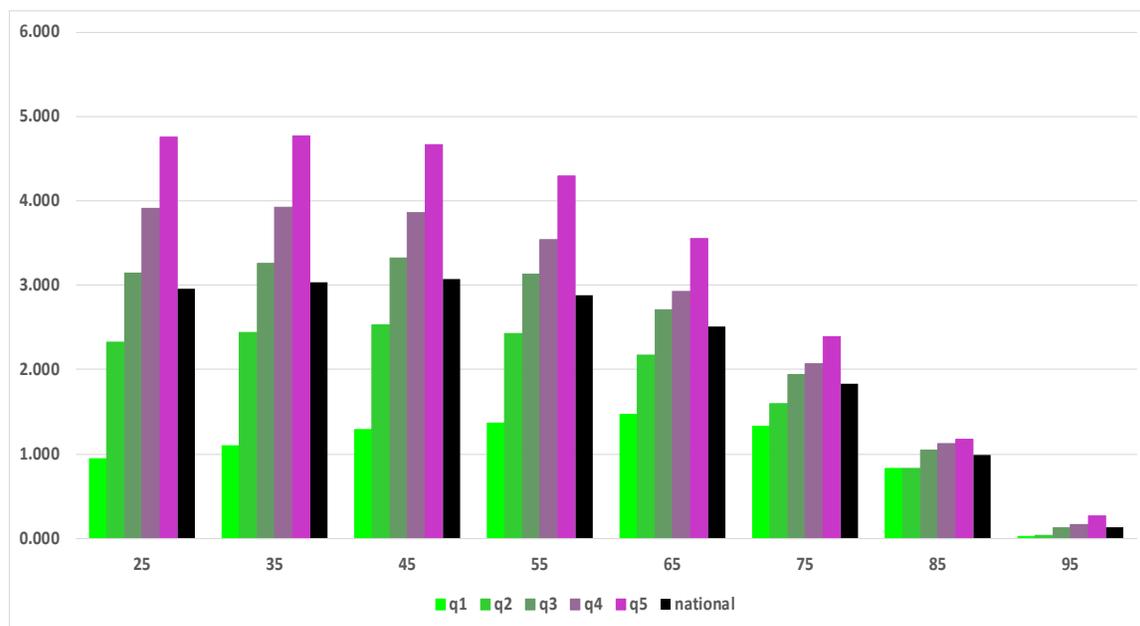
Except for an additional option to select the "static" decile data set (d-2000) for the comparisons, the inputs and analysis in Tab 6 are the identical to those in Tab 5.

Tabs 7 and 8: Graphs of Mortality Rates Across Years and Ages

Tabs 7 and 8 provide the user with an opportunity to compare graphs of mortality rates across a range of years and ages. Users first select in cell C7 of both Tabs the type of output desired:

- Mortality rates
- Natural logarithm of mortality rates
- The ratio of the selected mortality rate to the corresponding mortality rate in 1982.

On Tab 7, users can select up to ten combinations of data set, age and gender. Running the "Redraw Graph" macro produces a graph of the corresponding mortality rates across all years from 1982 through 2019.

Similarly, on Tab 8, after selecting up to ten combinations of data set, calendar year and gender, a graph comparing mortality rates across all ages 20 through 90 is produced.

Tabs 9 and 10: Inputs for Heat Maps

Tabs 9 and 10 offer the user the ability to select mortality data sets to plot on a heat map(s). After selecting the desired data set(s) and pressing RUN, a user is transported to Tab 11 which displays the requested heat maps.

Tabs 9 and 10 are similar in nature, but they each serve a slightly different purpose. Tab 9 is used to select a series of heat maps to produce in sequence, with each heat map reflecting a single mortality data set. For example, one could produce heat maps for each of the five socioeconomic quintiles. Tab 10, in contrast, produces just one pair of heat maps (female and male), but permits a user to blend data sets by entering weights across quintiles or deciles. For example, a user could produce female and male heat maps that reflect a 50%/50% blend of deciles 9 and 10.

Both Tabs 9 and 10 permit a user to specify both the age range (cells B8 and B9) and the year range (cells B10 and B11) to display on the heat maps. Either smoothed (order-2 or order-3) or unsmoothed data can be selected (cell B14). Any of three output types can be selected with cell B15: (1) mortality improvement rates, (2) mortality rates and (3) mortality rates divided by the corresponding rate in a user-specified year (cell B16).

On Tab 9 after choosing the age/year range and the type of output, range B19:B82 is used to select data sets. Figure 5 illustrates a request to produce heat maps for six data sets: a "total" data set includes all quintiles, and then each of the five separate quintiles. The numbering in the example below – 1 through 6 – indicates the order in which the heat maps will output on Tab 10. Simply leave a cell blank if the user does not wish to output a heat map for that data set.

Figure 5
TAB 9 EXAMPLE OF DATA SET SELECTION FOR HEAT MAPS

	SSA-F
	SSA-M
1	NCHS-total-F
	NCHS-total-M
2	q-1-F
	q-1-M
3	q-2-F
	q-2-M
4	q-3-F
	q-3-M
5	q-4-F
	q-4-M
6	q-5-F

Using input range B19:B82, the user may select as many (or as few) data sets as the user wishes, i.e., there is no maximum or minimum limit on the number of selected data sets.

Cell B85 contains an optional input enabling the selected data to be expressed relative to the user-specified data set in cell B85. For example, suppose a user enters "NCHS-total-F" in cell B85, and suppose, with cell B15, a user selects annual mortality improvement rates as the desired output. In this case, the heat maps will show mortality improvement rates expressed relative to those from data set "NCHS-total-F" (the national-level data). Continuing this example, suppose the improvement rate for females aged 55 in 2010 is

1% using national level data, and 1.5% using data for the upper quintile. Then relative improvement rate for the upper quintile is 1.5% minus 1.0%, or 0.5%, and this result would be plotted on the heat map.

Like Tab 9, Tab 10 has inputs for selecting the age and year range to plot on the heat maps. Beginning in row 19, however, the inputs on Tab 10 differ from those on Tab 9. Use cell B19 to select one of 6 data sets to plot. Data sets 1 and 2 are national-level data sets, while data sets 3 through 6 involve quintiles or deciles. For these data sets, the user must enter quintile/decile weightings on row 29. The weights must sum to 100%.

After setting the parameters, press RUN on Tab 9 or 10 to produce the heat maps.

Tabs 11 and 12: Heat Maps

Tab 11 presents the heat map(s) requested on Tab 9 or Tab 10. The default color legend appears in range C6:R6. Alternatively, a user can enter their own set of colors in range C11:R11, and then set cell I13 to "User-Specified". To enter a set of colors on row 11, simply format each cell's background color to match the desired color.

Rows 7 through 10 contains the numerical values corresponding with each color. A user may change any of these values, taking care the sequence of values increases monotonically from left to right.

Tab 12 shows the data plotted on the heat maps on Tab 11. To see all the data, set cell B5 and B6 to "1", and then re-run Tab 9 or Tab 10. For a compact presentation of the data, use cells B5 and B6 to show only every "Nth" calendar year and every "Nth" age. After setting cells B5 and B6, re-run Tab 9 or 10 to produce the desire output.

Tabs 13 and 14: Comparison of SSA and NCHS Mortality Rates to Other Mortality Tables

Tab 13 facilitates a comparison of the SSA and NCHS mortality rates against commonly used mortality tables. Use cells B8, B9 and C19:L19 to select an SSA or NCHS data set. This data set can be compared against any of the data sets shown on rows 33 through 226. Select any (or all) of these data sets by entering a "1" in the appropriate cell in range J33:J226. If the user wishes to exclude a data set, simply leave the corresponding row blank in column J.

Cell B23 provides two options for comparing the SSA or NCHS data against the comparison mortality table. Under option 1, the outputs will be SSA/NCHS mortality rates minus the corresponding rates for the comparison table(s). Under option 2, the outputs will be SSA/NCHS mortality rates divided by the comparison table(s).

After setting the various parameters, press "Fetch and Compare the Selected Mortality Tables." The outputs will appear in columns P through AA. Each row shows the results for one comparison table. Consider the following example:

Figure 6
TAB 13 EXAMPLE

Table Name	Year	Sex	National-Level NCHS Mortality Rates Divided by the Corresponding Mortality Rate of the Comparison Table							
			20	30	40	50	60	70	80	90
RP2006_WhiteCol_HealthyAnnuit	2006	Female				150%	158%	133%	126%	116%

This example compares NCHS national-level female mortality rates against the RP-2006 female white-collar employee mortality table. No results appear for ages 20, 30 and 40 because the RP-2006 table begins at age 50. Consider the "133%" value at age 70. This indicates the NCHS mortality rate at age 70 is 33% greater than the corresponding mortality rate from the RP-2006 table.

Note Tab 14 contains the comparison mortality tables. Keep in mind some of these tables do not cover the entire age. For example, many of the annuitant tables begin at age 50.

Acknowledgments

The Society of Actuaries would like to thank the MIM Advisory Group for their support, guidance, direction and feedback throughout the project.

MIM Advisory Group Members:

Larry N. Stern, FSA, MAAA (Chairperson)

Jean-Marc Fix, FSA, MAAA

Sam Gutterman, FSA, MAAA, CERA, FCAS, FCA, HonFIA

Thomas A. Jones, ASA. MAAA

Allen M. Klein, FSA, MAAA

Laurence Pinzur, Ph.D.

At the Society of Actuaries Research Institute:

R. Dale Hall, FSA, MAAA, CERA, CFA, Managing Director of Research

Patrick D. Nolan, FSA, MAAA, Senior Experience Studies Actuary

Jan Schuh, Senior Research Administrator

Ronora Stryker, ASA, MAAA, Senior Practice Research Actuary

Patrick Wiese, ASA, Lead Modeling Researcher

About The Society of Actuaries Research Institute

Serving as the research arm of the Society of Actuaries (SOA), the SOA Research Institute provides objective, data-driven research bringing together tried and true practices and future-focused approaches to address societal challenges and your business needs. The Institute provides trusted knowledge, extensive experience and new technologies to help effectively identify, predict and manage risks.

Representing the thousands of actuaries who help conduct critical research, the SOA Research Institute provides clarity and solutions on risks and societal challenges. The Institute connects actuaries, academics, employers, the insurance industry, regulators, research partners, foundations and research institutions, sponsors and non-governmental organizations, building an effective network which provides support, knowledge and expertise regarding the management of risk to benefit the industry and the public.

Managed by experienced actuaries and research experts from a broad range of industries, the SOA Research Institute creates, funds, develops and distributes research to elevate actuaries as leaders in measuring and managing risk. These efforts include studies, essay collections, webcasts, research papers, survey reports, and original research on topics impacting society.

Harnessing its peer-reviewed research, leading-edge technologies, new data tools and innovative practices, the Institute seeks to understand the underlying causes of risk and the possible outcomes. The Institute develops objective research spanning a variety of topics with its strategic research programs: aging and retirement; actuarial innovation and technology; mortality and longevity; diversity, equity and inclusion; health care cost trends; and catastrophe and climate risk. The Institute has a large volume of topical research available, including an expanding collection of international and market-specific research, experience studies, models and timely research.

Society of Actuaries Research Institute
475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
www.SOA.org