

Data and application for age mortality trajectories during pre-COVID, COVID and post-COVID years 2018–2023 in 113 Countries

Scientific Modelling and Research

Vol.11, No. 1, 1–11, 2026

e-ISSN: 2523-952X



 **Josef Dolejs**

Department of Informatics and Quantitative Methods, University of Hradec Králové, Rokitsanského 62, 500 03 Hradec Králové, Czech Republic.

Email: josef.dolejs@uhk.cz

ABSTRACT

Age trajectories of mortality (ATM) are important indicators of population health. In recent years, major global events have substantially affected ATM. Previous data and applications covered calendar years only up to 2018. However, the period from 2020 to 2023—marked by excess mortality—offers new opportunities to expand the ATM database and its applications. Constructing ATM within individual calendar years from 2018 to 2023 represents a major methodological extension. The years 2018 and 2019 serve as reference periods, 2020 and 2021 correspond to the COVID-19 pandemic, and 2022 and 2023 reflect a period of relaxation. The analysis was implemented using R software and Shiny technology. A new application was created in R using Shiny technology. The application and expanded data enable a more detailed analysis of disease-specific mortality patterns across the calendar years 2018–2023. The data system and application were developed primarily from the perspective of age-dependent mortality. The presented data and application are intended primarily as exploratory tools. Their greatest value lies in enabling more detailed follow-up analyses using filtered, country-specific data. It is likely that the presented application will primarily serve as an initial exploratory tool, with its greatest utility being in enabling more detailed follow-up analyses on filtered data from specific countries.

Keywords: *Age, Application, COVID-19 pandemic, Mortality rate, WHO database, Epidemiology.*

DOI: 10.55284/smr.v11i1.1985

Citation | Dolejs, J. (2026). Data and application for age mortality trajectories during pre-COVID, COVID and post-COVID years 2018–2023 in 113 Countries. *Scientific Modelling and Research*, 11(1), 1–11.

Copyright: © 2026 by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Funding: This study received no specific financial support.

Institutional Review Board Statement: Not applicable

Transparency: The author confirms that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Competing Interests: The author declares that there are no conflicts of interests regarding the publication of this paper.

History: Received: 16 May 2026/ Revised: 21 May 2026/ Accepted: 15 September 2026/ Published: 13 October 2026

Publisher: Online Science Publishing

Highlights of this paper

- Age trajectories of mortality (ATM) data were processed for calendar years 2018–2023.
- A new application was developed in R using Shiny technology.
- The data system and application focus primarily on age-dependent mortality.

1. INTRODUCTION

Between 2020 and 2022, most populations were affected by events that significantly impacted the health, social, and economic conditions of countries worldwide. The World Health Organization (WHO) did not publish global data for this period (up to 2023) until February 2025 [1]. Previous application and data updated 2019 processing related to Age Trajectories of Mortality (ATM) offer a potential framework for evaluating the global events of 2020–2022 [2].

Age-specific mortality trajectories (ATMs) represent the main outcome of the application. Experience with the previous version and the availability of new data covering the critical period 2020–2023 served as the impetus for developing new software, which enables users to compare individual calendar years from 2018 to 2023 [2]. The primary explanatory variable remains age, allowing for the construction of ATMs for specific disease groups or for all causes of death combined (age trajectory of total mortality - ATTM) [3, 4]. If users wish to explore other directions—such as comparing patterns within larger age groups—data files for both deceased and living individuals are available for download within the application.

The primary data processed include the number of deaths by age and cause of death, categorized using the International Classification of Diseases, 10th Revision [5]. Among the three variables—number of deaths, age, and cause of death—the first two are considered highly reliable. However, the cause of death may be subject to further discussion and detailed analysis.

The presented data and application aim to provide a user-friendly platform for a wide range of users, including physicians, epidemiologists, and the general public, to explore ATM patterns based on classifications of causes of death. Users can visualize data directly in the application or download CSV files for further analysis.

For instance, a user may select a country, choose a main ICD-10 chapter, and then select specific causes of death within that category. The application will then display the ATM for selected group of diseases and the calendar years 2018–2023, with the option to save the data as a CSV file.

Many additional options are available. For example, users can download the main data file and extract information for a specific country using a simple code presented in the appendix. Such data may be analysed using widely available tools such as Microsoft Office, SPSS, OpenOffice, or R.

2. MATERIALS AND METHODS

Mortality data were obtained from the WHO Mortality Database, which provides country-level information on deaths classified by age, calendar year, and cause of death [1]. Causes of death are coded using successive revisions of the International Classification of Diseases (ICD), with the most recent revision, ICD-10, serving as the primary classification system in the present analysis [5].

To ensure methodological consistency across countries and calendar years, only records classified under ICD-10 were retained. Mortality data were further harmonized according to age classification. During the first year of life, the WHO database distinguishes four age intervals (0–24 hours, 1–7 days, 7–28 days, and 28–365 days), whereas subsequent ages are grouped into broader categories. Because not all reporting countries apply identical infant age classifications, only those records containing all four infant categories were included. This criterion enables uniform construction of age trajectories of mortality ATMs across populations.

Above infancy, mortality data were available in the following structure: a single category for ages 1–5 years and five-year age intervals from age 5 up to 90–94 years. Application of the ICD-10 and age-structure criteria reduced the number of usable calendar years in some countries but ensured comparability of ATM across all included populations.

2.1. Population at Risk

Calculation of ATM requires information on both the number of deaths and the corresponding population at risk. While the WHO database contains population counts, its coverage does not fully correspond to the mortality data used in this study. Therefore, population data were obtained from the International Data Base of the United States Census Bureau [6].

The dataset “idsingleyear.txt” provides population counts by single-year age groups, calendar year, country, and sex. For each country and year included in the mortality dataset, the corresponding population counts were extracted and aggregated to match the age categories used in the WHO mortality data. This procedure enabled automated computation of ATM across countries and calendar years.

Due to differences in country definitions between databases, some territorial entities were harmonized. England and Wales, Northern Ireland, and Scotland were treated as a single entity corresponding to the United Kingdom. Several entities present in the WHO database but absent from the Census Bureau dataset were excluded. After harmonization, a total of 113 countries were retained for analysis.

2.2. Aggregation of Calendar Years

When mortality trajectories are constructed for narrowly defined causes of death, zero counts may occur in specific age categories within a single calendar year. To mitigate this limitation, aggregation across multiple calendar years was employed. This approach increases the stability of mortality estimates while preserving the age-dependent structure of the data.

The principle of aggregating mortality data across time was first introduced by Edmond Halley in his classical work on life tables [3]. By combining observations across years, stochastic variation unrelated to age is reduced, allowing the underlying age-related pattern of mortality to emerge more clearly.

Mortality can be analyzed either across age categories within a single calendar year (cross-sectional perspective) or longitudinally across cohorts. In the present application, calendar-year aggregation serves as a methodological tool to suppress non-age-related fluctuations while retaining age as the dominant explanatory variable. This strategy enables construction of robust ATM in both individual countries and large aggregated populations.

2.3. Aggregated Populations

In addition to country-specific analyses, ten aggregated populations were constructed to further stabilize ATM estimates and reduce the impact of sparse data. These aggregated populations group countries by geographical or analytical criteria and follow previously established definitions [2].

The largest aggregated population, P113, includes all 113 countries in the dataset and represents approximately one-third of the global population during the study period. Other aggregated populations include regional groupings such as European, South American, Central American, and Asian countries, as well as combinations used in earlier studies. Selection of an aggregated population automatically displays the list of included countries within the application interface.

2.4. Basic Description of the Application

A new panel, "ATMs from Diseases Defined by User in Individual Years 2018–2023" was developed to enable detailed analysis of specific cause-of-death groups selected by the user for individual calendar years from 2018 to 2023. The second new panel "ATTMs in Individual Years 2018–2023 (Total Mortality)" enables to compare age trajectories of total mortality in these years. Additionally, the "ATTMs in Adults, Data Grouped in All Calendar Years: The Gompertz Law" panel was introduced to illustrate age-specific mortality trajectories in adulthood [7-17]. It extends the analysis of data grouped in all calendar years according to Halley method [2, 3]. The full methodology and theoretical possible explanations are in the following panel "Notes 2025 and Theory for ATTMs After the Birth". All other panels from correspond to the previous application. They are grouped in the menu "Menu for All Other Panels with Software from the Previous Application (2019) with Data Updated in 2025". The eighth panels were updated with the most recent data from the WHO database (the previous data were based on the update from January 25, 2019, while the current version reflects the update from February 27, 2025) [2].

Furthermore, 10 large, aggregated populations that included countries from different regions of the world were created [2]. The largest aggregated population, "P113," contained all 113 countries. The number of people alive per year in "P113" was approximately 2.367 billion, representing about one-third of the world's population during the study period. The list of 113 countries may be displayed in the web application if an aggregated population is selected in the panels: "ATTMs in Adults, Data Grouped in All Calendar Years: The Gompertz Law", "ATTMs After the Birth", "CACNS After the Birth" and "Main Chapters".

Calendar years used by a specific country are automatically shown in each panel where country may be selected. More details are in the files "CountriesDnxx2025.txt" (for the presented application) and "CountriesDnxx.txt" (for the previous data and application). They may be downloaded in the panel "Notes 2025 and Theory for ATTMs After the Birth". Figures 1 and 2 provide a diagrammatic representation of the whole process.

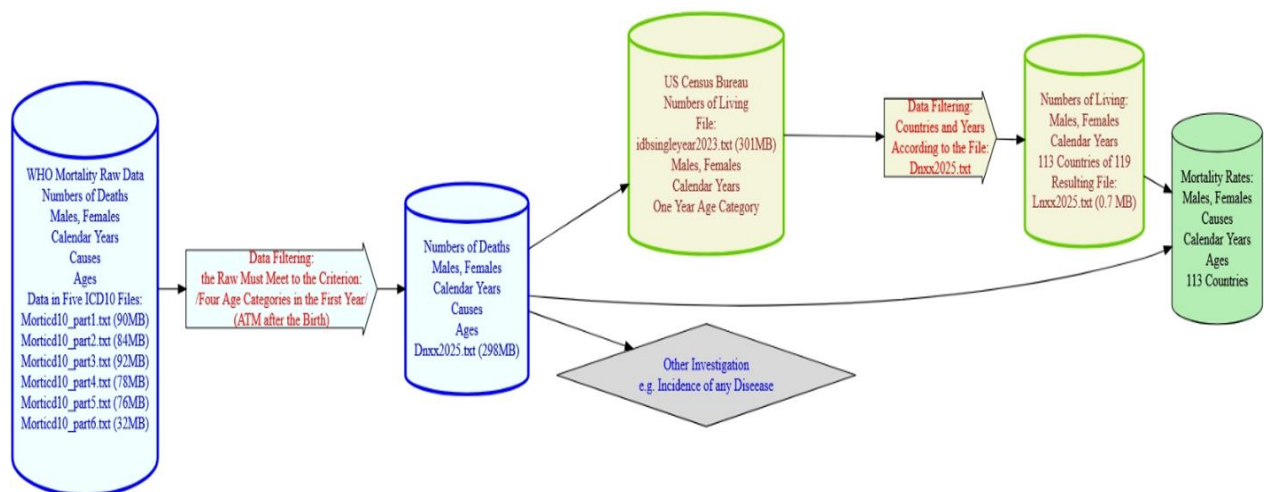


Figure 1. Diagram of the first process.

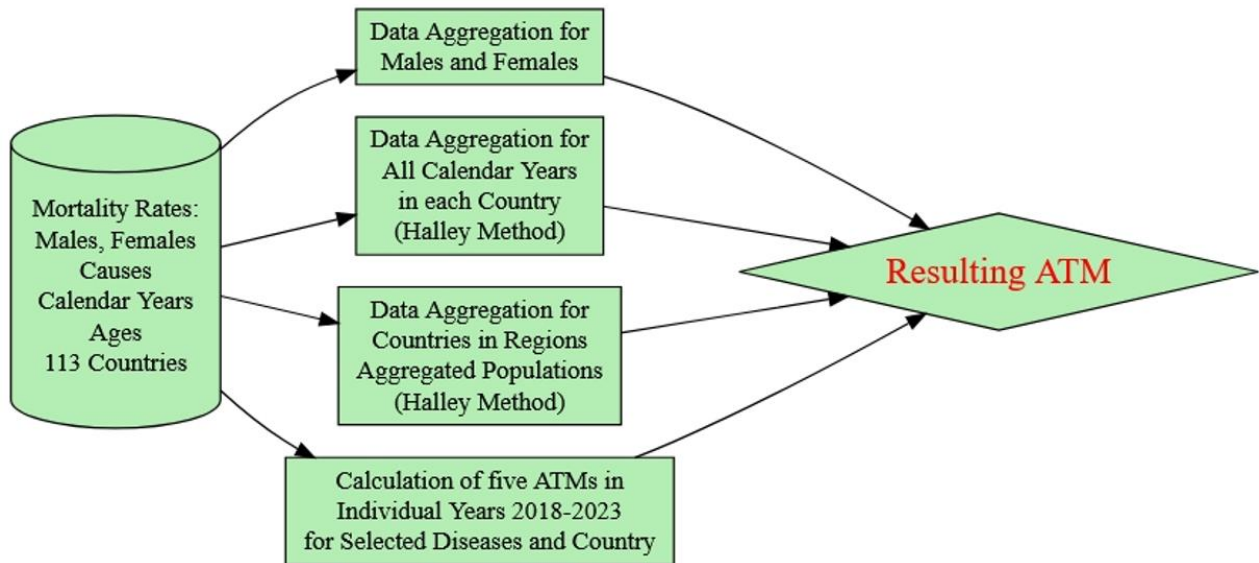


Figure 2. Diagram of the calculation of ATM.

3. RESULTS

3.1. Practical Use

The first panel “*ATMs from Diseases Defined by User in Individual Years 2018–2023*” enables the display of ATMs for the selected group of diseases across up to six separate calendar years: 2018, 2019, 2020, 2021, 2022 and 2023 (aggregated populations are not in the first two panels and user may select individual country). Data for 2023 is already available for 18 countries and for 2022 in 43 countries, while data for 2021 is available in 65 countries. For instance, this panel allows users to examine the situation before the COVID-19 epidemic (2018 and 2019), during the epidemic (2020 and 2021), and in the post-epidemic years 2022 and 2023. If main chapter of ICD10 is selected in the panel then all possible diseases are shown for selected country. User may select own group of disease from the displayed set. Two illustrative examples of a resulting plot drawn for Australia and for the Czech Republic are shown in Figures 3–4 for special ICD10 codes U07.1 and U07.2 (codes for COVID-19). Calculations in the first panel may take more than ten seconds because ATMs are calculated directly from the large file “Dnxx2025.txt”. Resulting plot may be copied similarly as in a standard website and set of six ATMs may be downloaded by button in the panel as a CSV file. Overall view of the whole panel is shown in Figure 5.

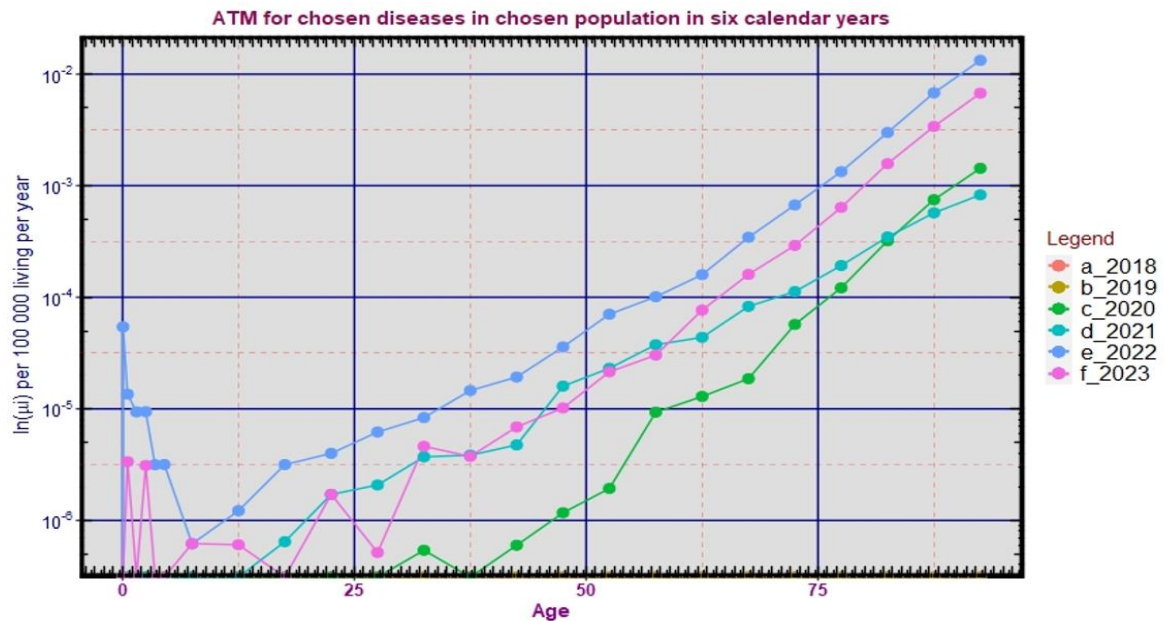


Figure 3. Screenshot of the first panel displaying data for Australia.

Note: Plot of ATMs calculated for Australia using the two ICD-10 codes U07.1 and U07.2 (deaths with a positive COVID-19 test are coded with U07.1, and those suspected of COVID-19 but without a test are coded with U07.2). No ATMs are present for the calendar years 2018 and 2019, as these predate the COVID-19 pandemic. The linear relationship observed corresponds to the exponential law described by Gompertz.

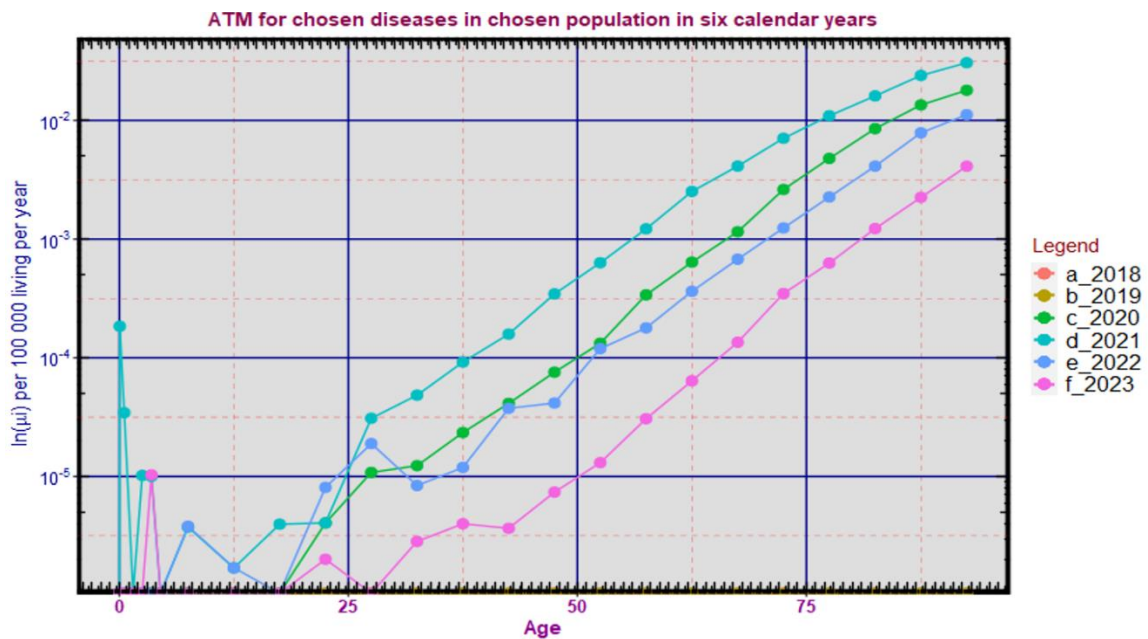


Figure 4. Screenshot of the first panel displaying data for the Czech Republic.

Note: Plot of ATMs calculated for Czech Republic using the two ICD-10 codes U07.1 and U07.2 (deaths with a positive COVID-19 test are coded with U07.1, and those suspected of COVID-19 but without a test are coded with U07.2). No ATMs are present for the calendar years 2018 and 2019, as these predate the COVID-19 pandemic. The linear relationship observed corresponds to the exponential law described by Gompertz.

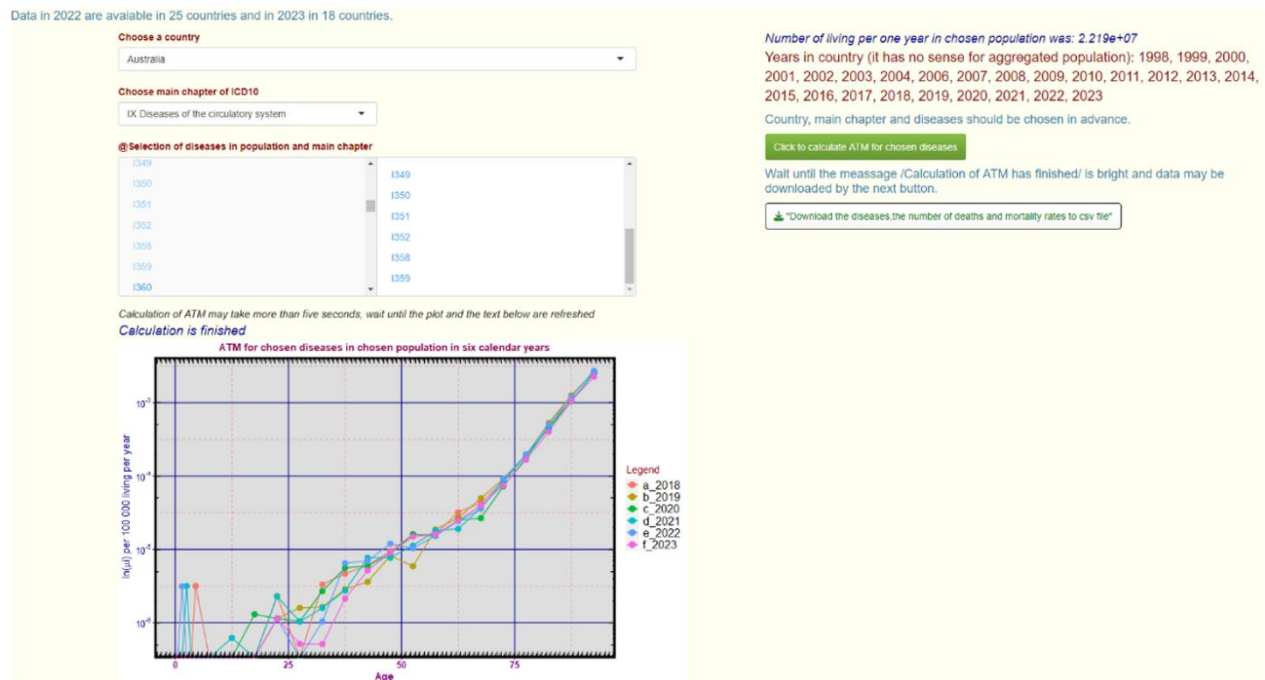


Figure 5. Overall view of the first panel.

Note: The calculation in the image took about one minute. Here Australia and the main chapter of ICD10 "IX Diseases of the circulatory system" was selected. Set of possible calendar years in selected country was shown and the following subset was selected by user: I33 Acute and subacute endocarditis, I34 Nonrheumatic mitral valve disorders, I35 Nonrheumatic aortic valve disorders (specifically I330, I339, I340, I341, I342, I348, I349, I350, I351, I352, I358, I359 in the six years). The specific selection is only illustration and does not represent more detailed analysis. Data may be downloaded as CSV file. The linear relationship observed corresponds to the exponential law described by Gompertz.

The second panel "*ATTMs in Individual Years 2018–2023 (Total Mortality)*" complements the first panel. Total mortality in the six years may be compared visually and downloaded as CSV file. The situation in the United States is shown in Figure 6 where calendar year 2023 is not available.

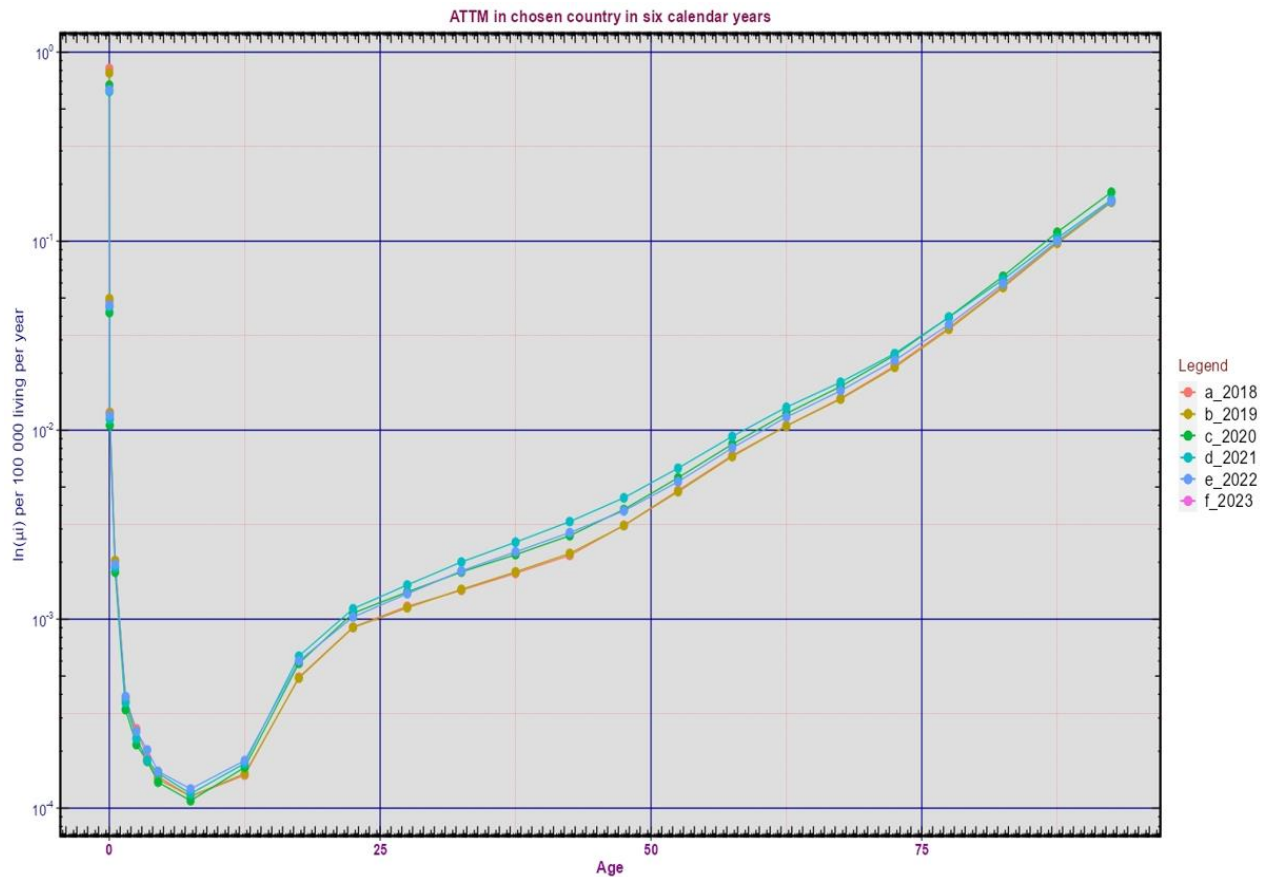


Figure 6. Screenshot of the second panel displaying ATTM in the United States.

Note: ATTM in the six calendar years may be shown (data in 2023 was not available).

The third panel “*ATTMs in Adults, Data Grouped in All Calendar Years: The Gompertz Law*” displays and analyzes the Gompertz relationship on a semilogarithmic scale for ages 40–95 within the age trajectory of total mortality (ATTM) calculated for all possible calendar years combine [2, 4, 7–13, 16, 18, 19]. User may select a specific country or one of ten aggregated populations, and all calculations and plots are generated automatically. If a specific country is selected, the set of available calendar years is also displayed.

The resulting ATTM is presented in a plot, along with the adjusted coefficient of determination ($\text{adj}R^2$) calculated using the Gompertz model in the age interval of 40–95 years. The calendar years are group in the panel and the values of $\text{adj}R^2$ are higher than 0.99 in many populations [2]. Additionally, two empirical mortality percentiles of 113 countries are displayed for the 2.5th and 97.5th percentiles, and a standard confidence band is also shown for the selected ATTM.

Figures 7 and 8 illustrate example plots for the largest aggregated population from all 113 countries and for Egypt, respectively.

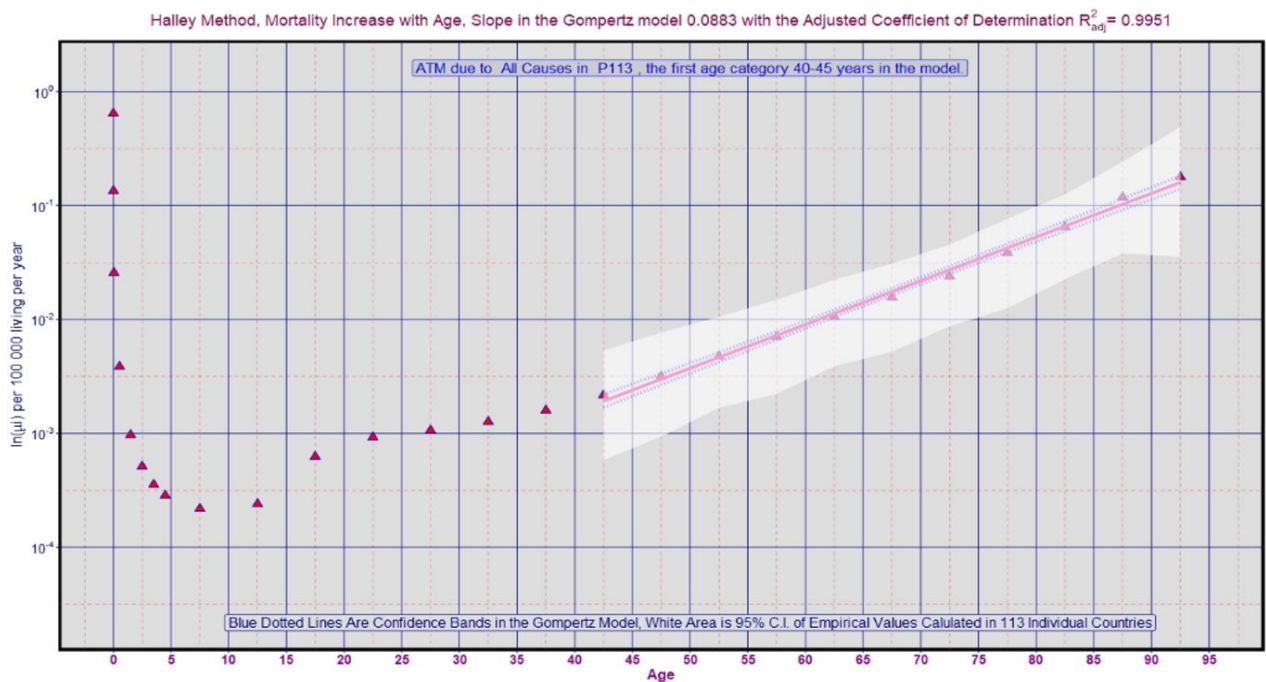


Figure 7. Screenshot of the third panel displaying data for P113.

Note: ATTM in the largest population aggregated from all possible 113 countries calculated using Halley method (all calendar years and all geographical regions were aggregated). The linear relationship observed corresponds to the exponential law described by Gompertz. Slope (0.0883) and coefficient of determination (0.9951) are automatically shown.

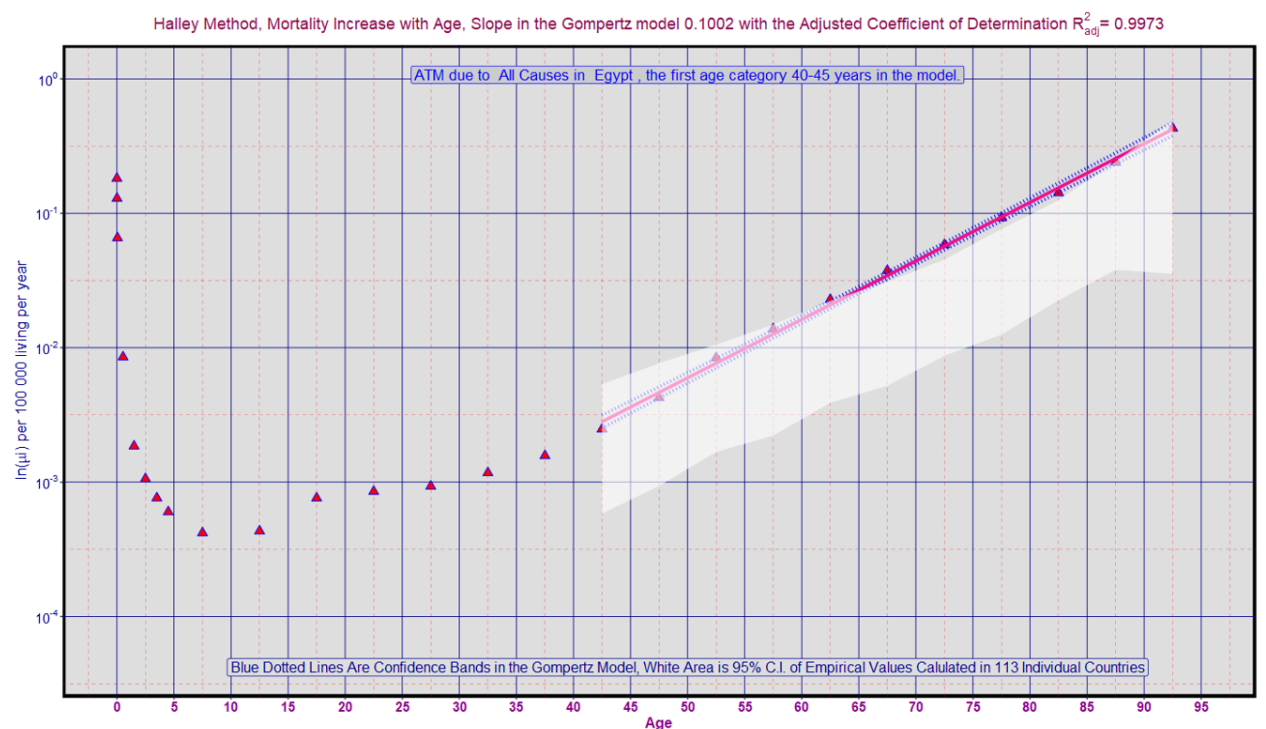


Figure 8. Screenshot of the third panel displaying data for Egypt.

Note: ATTM in Egypt calculated using Halley method (Calendar years were aggregated). The linear relationship observed corresponds to the exponential law described by Gompertz. Slope (0.1002) and coefficient of determination (0.9973) are automatically shown. ATTM is outside of the white region which corresponds to empirical 95 percentile rank in all countries while blue dotted lines are confident bands.

3.2. Data Sharing and Three Main Data Files

The three main data files: "Dnxx2025.txt", "Lnxx2025.txt" and "CountriesDnxx2025.txt" can be downloaded directly from the application in the first panel and in "Notes 2025 and Theory for ATTM After the Birth" panel in

CSV format. The file "Dnxx2025.txt" contains deaths cases in all countries (Rows) in all age categories (Columns) and in all calendar years (Rows). The file "Lnxx2025.txt" contains number of living in countries, calendar years and age categories. The file "CountriesDnxx2025.txt" contains list of countries with information about available calendar years. These files may be used to other epidemiology or demography analysis. For instance, "Dnxx2025.txt" and "Lnxx2025.txt" may be copy to directory in personal device, upload to program R and specific country data may be selected by filter (The simple code is in attachment). These CSV files for specific country may be simply opened also in Excel or other software and furthermore examine.

4. DISCUSSION

The ordering of countries follows the WHO system based on numerical country codes [1]. While this ordering reflects continental groupings, it may not always be ideal for users, as it is not alphabetical. Ten specially created aggregated populations are consistently listed at the end.

The entire data system and application were developed primarily from the perspective of mortality dependence on age. This focus led to a slight reduction in the dataset of deceased individuals. A second reduction was due to the availability of population data for specific countries and calendar years in the U.S. Census Bureau database [6]. However, this reduction does not pose a significant limitation; for example, it only led to the United Kingdom being assessed as a single unit, without distinguishing Scotland, Wales, and England separately.

It is likely that the presented application will primarily serve as an initial exploratory tool, with its greatest utility being in enabling more detailed follow-up analyses on filtered data from specific countries.

The potential outputs and conclusions that can be drawn using the application and shared data are beyond the scope of this study. Given the nature of the data and application, there are numerous such possibilities. In particular, analyses of the data from 2018 to 2023 may yield novel and unexpected findings.

REFERENCES

- [1] World Health Organization, "WHO mortality database (ICD-10). Available at: WHO mortality database," World Health Organization, 2025.
- [2] J. Dolejs, "Database of age trajectories of mortality in 110 countries and web application: Data report," *Frontiers in Public Health*, vol. 10, p. 911589, 2022. <https://doi.org/10.3389/fpubh.2022.911589>
- [3] E. Halley, "An estimate of the degrees of mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslaw, with an attempt to ascertain the price of annuities on lives," *Philos Trans*, vol. 17, pp. 596-610, 1693.
- [4] J. Dolejs and H. Homolková, "Why does child mortality decrease with age? Modeling the age-associated decrease in mortality rate using WHO metadata from 14 European countries," *Frontiers in Pediatrics*, vol. 8, p. 527811, 2020. <https://doi.org/10.3389/fped.2020.527811>
- [5] World Health Organization, "The international classification of diseases," 1997. [Online]. Available: <https://icd.who.int/browse10/2016/en#/>. [Online]. Available: <https://icd.who.int/browse10/2016/en#/>
- [6] Bureau of the Census, "Bureau of the census," 2023. [Online]. Available: <https://www.census.gov/programs-surveys/international-programs/about/idb.html>. [Online]. Available: <https://www.census.gov/programs-surveys/international-programs/about/idb.html>
- [7] B. Gompertz, "XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. FRS &c," *Philosophical Transactions of the Royal Society of London*, no. 115, pp. 513-583, 1825. <https://doi.org/10.1098/rstl.1825.0026>

- [8] W. M. Makeham, "On the law of mortality and the construction of annuity tables," *The Assurance Magazine, and Journal of the Institute of Actuaries*, vol. 8, no. 6, pp. 301-310, 1860.
- [9] B. Strehler and A. Mildvan, "General theory of mortality and aging," *Science*, vol. 132, pp. 14-21, 1960.
- [10] L. Hayflick and P. S. Moorhead, "The serial cultivation of human diploid cell strains," *Experimental Cell Research*, vol. 25, no. 3, pp. 585-621, 1961. [https://doi.org/10.1016/0014-4827\(61\)90192-6](https://doi.org/10.1016/0014-4827(61)90192-6)
- [11] L. Hayflick, "The limited in vitro lifetime of human diploid cell strains," *Experimental Cell Research*, vol. 37, no. 3, pp. 614-636, 1965. [https://doi.org/10.1016/0014-4827\(65\)90211-9](https://doi.org/10.1016/0014-4827(65)90211-9)
- [12] L. Heligman and J. H. Pollard, "The age pattern of mortality," *Journal of the Institute of Actuaries*, vol. 107, no. 1, pp. 49-80, 1980. <https://doi.org/10.1017/S0020268100040257>
- [13] J. E. Riggs, "Longitudinal gompertzian analysis of adult mortality in the US, 1900-1986," *Mechanisms of Ageing and Development*, vol. 54, no. 3, pp. 235-247, 1990.
- [14] J. Dolejs, "The extension of Gompertz law's validity," *Mechanisms of ageing and Development*, vol. 99, no. 3, pp. 233-244, 1997. [https://doi.org/10.1016/S0047-6374\(97\)00104-8](https://doi.org/10.1016/S0047-6374(97)00104-8)
- [15] W. J. Willemse and H. Koppelaar, "Knowledge elicitation of Gompertz'law of mortality," *Scandinavian Actuarial Journal*, vol. 2000, no. 2, pp. 168-179, 2000. <https://doi.org/10.1080/034612300750066845>
- [16] X. S. Lin and X. Liu, "Markov aging process and phase-type law of mortality," *North American Actuarial Journal*, vol. 11, no. 4, pp. 92-109, 2007. <https://doi.org/10.1080/10920277.2007.10597486>
- [17] J. Dolejs, "Age trajectories of mortality from all diseases in the six most populated countries of the South America during the last decades," *Bulletin of Mathematical Biology*, vol. 76, no. 9, pp. 2144-2174, 2014. <https://doi.org/10.1007/s11538-014-0005-0>
- [18] J. Dolejs and P. Marešová, "Onset of mortality increase with age and age trajectories of mortality from all diseases in the four Nordic countries," *Clinical Interventions in Aging*, vol. 12, pp. 161-173, 2017. <https://doi.org/10.2147/CIA.S119327>
- [19] J. Dolejs, "Modeling human mortality from all diseases in the five most populated countries of the European Union," *Bulletin of Mathematical Biology*, vol. 79, no. 11, pp. 2558-2598, 2017. <https://doi.org/10.1007/s11538-017-0341-y>

Online Science Publishing is not responsible or answerable for any loss, damage or liability, etc. caused in relation to/arising out of the use of the content. Any queries should be directed to the corresponding author of the article.