Demographic, Epidemiologic, and Economic Analysis for the 100M lives project

The R code implements a Shiny-based dashboard for analyzing the demographic, epidemiologic, and economic impacts of various interventions related to hypertension and sodium intake policies. Here’s a breakdown of how these analyses are performed mathematically and conceptually:

Contents

[1. Project Overview 1](#_Toc189390097)

[2. Methodology 1](#_Toc189390098)

[Demographic Analysis 2](#_Toc189390099)

[Epidemiologic Analysis 2](#_Toc189390100)

[Intervention Impact 2](#_Toc189390101)

[Disease burden 3](#_Toc189390102)

[3. Data Sources & Preprocessing 3](#_Toc189390103)

[4. Model Implementation 4](#_Toc189390104)

[5. Reproducibility & Version Control 4](#_Toc189390105)

## Project Overview

Title: A clear and concise project title.

Abstract: Brief summary of the model’s purpose and expected outcomes.

Objectives: Goals and research questions the model aims to address.

Scope: Definition of the population, time horizon, and health policies under evaluation.

## Methodology

Model Description

- Type of model (e.g., Markov model, agent-based model, system dynamics).

- Key assumptions and justifications.

- Diagram of states and transitions (if applicable).

Mathematical Formulation

- Equations governing state transitions.

- Probability distributions used.

- Calibration methods.

Intervention Scenarios

- Description of health policies being simulated.

- Comparator (baseline) vs. intervention scenarios.

## Demographic Analysis

The code loads demographic data through CSV files (e.g., global\_pop\_0122.csv), which contain population data for various age groups, locations, and other demographic characteristics.

## Epidemiologic Analysis

Epidemiologic modeling involves calculating probabilities of individuals being in certain health states (e.g., hypertension categories) based on their blood pressure, sex, and age. The core function get.bp.prob adjusts probabilities based on various interventions like sodium reduction or antihypertensive therapy.

The model evaluates how interventions affect blood pressure, then computes related probabilities using the Gamma distribution functions (pgamma).

Blood Pressure Distribution Modeling: The shape and scale of blood pressure distributions are calculated based on mean and standard deviation parameters.

Risk Adjustments: Functions like addRR are used to adjust the relative risk (RR) of various cardiovascular diseases (CVD) based on blood pressure categories.

Mathematically, for each year and demographic group, the probability distribution of blood pressure is described by the Gamma distribution:

P(bp ≤ Threshold) = pgamma(Threshold, Shape, Scale)

Intervention impacts are introduced by adjusting mean blood pressure and coverage values in a stepwise manner across years (using linear interpolation).

## Intervention Impact

The interventions (sodium policies and hypertension treatment) are simulated over time (e.g., between 2020 and 2040), scaling their impact linearly based on user input. The health impact of these interventions is calculated by adjusting disease incidence and mortality rates via relative risk reductions and integrating these effects over the population.

State transitions are modeled as:

Sick\_t = Sick\_(t-1) × (1 - Mortality Rate) + Well\_(t-1) × New Incidence Rate

Deaths\_t = Sick\_(t-1) × Case Fatality Rate

## Disease burden

The tool computes per capita costs, cost-effectiveness, and projects results across different scenarios (baseline vs. intervention) for each year from 2020 to 2040. This is displayed in tables and interactive graphs.

1. Compute DALYs averted:
2. Apply a monetary valuation:

## Data Sources & Preprocessing

Data Sources

- Epidemiological data: Disease incidence, prevalence, mortality.

- Demographic data: Population structure, life expectancy.

- Healthcare data: Utilization, treatment effectiveness.

- Economic data: Costs, productivity losses.

- Model parameters: Transition probabilities, effectiveness of interventions.

Data Extraction & Cleaning

- Raw data formats (CSV, JSON, databases).

- Handling missing values.

- Data validation (checking for inconsistencies).

Data Transformations

- Aggregation and disaggregation.

- Converting categorical variables to numeric form.

- Rescaling and normalizing.

- Unit conversions (e.g., per 1,000 population).

Data Adaptation for Model

- Matching model input format requirements.

- Creating age-cohort distributions.

- Estimating transition probabilities from real-world data.

- Mapping different datasets to a unified structure.

- Time series interpolation (if needed).

Final Dataset Preparation

- Summary of final input data.

- Storage format and organization.

## Model Implementation

Software & Tools

- Programming language (Python, R, etc.).

- Libraries and packages used.

Code Structure

- Explanation of scripts/modules.

- Input-output data pipeline.

Computational Requirements

- Hardware and software dependencies.

- Computational complexity and runtimes.

## Reproducibility & Version Control

Instructions for Running the Model

- Step-by-step guide.

- Example input datasets.

Code Repository

- GitHub/GitLab repository structure.

- Versioning and changelog.

Data Management

- Storage and access to datasets.

- Ethical considerations and anonymization.