

## Article

# Risk-Adjusted Estimation and Graduation of Transition Intensities for Disability and Long-Term Care Insurance: A Multi-State Model Approach

Beatriz A. Curioso <sup>1,2,†</sup> , Gracinda R. Guerreiro <sup>1,2,\*,†</sup>  and Manuel L. Esquível <sup>1,2,†</sup> 

<sup>1</sup> NOVA School of Science and Technology, Universidade Nova de Lisboa, Campus de Caparica, 2829-516 Caparica, Portugal; b.curioso@campus.fct.unl.pt (B.A.C.); mle@fct.unl.pt (M.L.E.)

<sup>2</sup> Center for Mathematics and Applications (NOVA Math), NOVA School of Science and Technology, Universidade Nova de Lisboa, Campus de Caparica, 2829-516 Caparica, Portugal

\* Correspondence: grg@fct.unl.pt

† These authors contributed equally to this work.

## Abstract

This paper introduces a methodology for estimating transition intensities in a multi-state model for disability and long-term care insurance. We propose a novel framework that integrates observable risk factors, such as demographic (age and sex), lifestyle (smoking and exercise habits) and health-related variables (body mass index), into the estimation and graduation of transition intensities, using a parametric approach based on the Gompertz–Makeham law and generalised linear models. The model features four states—autonomous, dead, and two intermediate states representing varying disability levels—providing a detailed view of disability/lack of autonomy progression. To illustrate the proposed framework, we simulate a dataset with individual risk profiles and model trajectories, mirroring Portugal’s demographic composition. This allows us to derive a functional form (as a function of age) for the transition intensities, stratified by relevant risk factors, thus enabling precise risk differentiation. The results offer a robust basis for developing tailored pricing structures in the Portuguese market, with broader applications in actuarial science and insurance. By combining granular disability modelling with risk factor integration, our approach enhances accuracy in pricing structure and risk assessment.

**Keywords:** multi-state models; long-term care; disability insurance; transition intensity approach; graduation; data simulation



Academic Editor: Annamaria Olivieri

Received: 30 April 2025

Revised: 13 June 2025

Accepted: 16 June 2025

Published: 27 June 2025

**Citation:** Curioso, Beatriz A., Gracinda R. Guerreiro, and Manuel L. Esquível. 2025. Risk-Adjusted Estimation and Graduation of Transition Intensities for Disability and Long-Term Care Insurance: A Multi-State Model Approach. *Risks* 13: 124. <https://doi.org/10.3390/risks13070124>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. 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/>).

## 1. Introduction

From a socioeconomic perspective, disability insurance and, in particular, long-term care (LTC) insurance have recently become areas of interest in developed countries as a result of the observed demographic trends. The accelerated ageing of the population, caused by the increase in life expectancy and the decline in fertility rates, is expected to intensify in the coming years, as is discussed in OECD (2023).

With the increase in the elderly—and consequently dependent—population, several countries, including Germany, the United Kingdom, France, and Switzerland, have developed public long-term care networks and also offer LTC products through private insurance companies; see Haberman and Pitacco (2018). Recently, Lee et al. (2023) conducted a detailed comparative analysis of LTC systems across 26 OECD countries, focusing on the

source of financing (namely tax-based systems, health insurance and LTC insurance) while providing a characterisation of each financing mechanism.

Regardless of the financing mechanism, LTC systems are inherently challenging to implement, as explored in [Barreira et al. \(2023\)](#). LTC insurance products, in particular, are complex to design due to their broad scope of coverage (both in terms of benefits offered and the extended duration of the policies), which often results in high premiums that are not affordable to the general population. Moreover, since this field and its associated products are relatively new, there is a significant lack of data concerning disability events, which is essential for developing accurate models for premium and reserve calculations.

Population ageing is also an observable trend in Portugal, see [INE \(2023\)](#). To address this phenomenon, the country has established the National Network of Integrated Continuous Care (RNCCI), a public network that provides healthcare and social support to individuals in situations of dependence. Several studies have been conducted using RNCCI data, mainly focusing on characterising key features of the Portuguese population with dependency needs. Notable works include those by [Lopes et al. \(2020, 2021\)](#) and the other references therein, which explore various aspects of the system in place, such as the demographic profile of individuals receiving care, factors influencing the duration of care, and the effectiveness of integrated continuous care systems in addressing dependency challenges. These studies provide valuable insights for improving care strategies and policies targeting the dependent population in Portugal. Portugal has been identified as one of the EU-27 member states with the highest rates of care provided by informal carers and with the highest share of direct out-of-pocket funding for LTC, see [European Commission \(2021\)](#).

Although there are no private companies in Portugal providing LTC insurance products, mostly due to the high risk involved and the lack of data, see [Esquível et al. \(2021\)](#), some research work has been conducted in this area. In [Oliveira et al. \(2017\)](#), using 2015 data from the RNCCI database and a clustering-based approach, the authors estimate the discrete-time transition probabilities and [Oliveira \(2017\)](#) develops the simulation and calibration of a multi-state model, in discrete time, for the calculation of the associated costs and premiums for an LTC product in Portugal. In [Esquível et al. \(2021, 2024\)](#), using the same dataset and relaxing the assumption of a discrete model, the authors present a method to estimate and calibrate transition intensities for non-homogeneous continuous-time Markov chains, and they present a general multi-state model using official Portuguese data. In this paper, we aim to develop further stochastic and statistical methods for the formulation of disability and LTC insurance products in Portugal by incorporating the demographic characteristics of the population into the framework.

From an actuarial perspective, the main concern when dealing with disability and LTC insurance (as with other types of insurance) is the modelling of the underlying risks. In this case, we are presented with a higher complexity, not only due to the wide range of possible statuses of an individual's health but also because of the variety of benefits that can be provided by this type of insurance. Therefore, robust models and accurate estimates are required for the accurate risk modelling of the different events and underlying benefits, as well as for the respective premium and reserve calculations.

Multi-state models are the most common actuarial models used for the study of disability and LTC insurance. A well-established mathematical study of the theory of Markov chains and their relation with multi-state models is presented, for instance, in [Ross \(1996\)](#), and [Wolthuis \(2003\)](#). The application of multi-state models in disability and LTC insurance has also been thoroughly studied, see [Pitacco \(2014\)](#), [Haberman and Pitacco \(2018\)](#), and [Dickson et al. \(2019\)](#). Regarding the methodology for the formulation of these models, the mentioned studies follow the transition intensity approach (or TI-approach) proposed in [Waters \(1984\)](#), in which the transition intensities are considered

the fundamental quantities (or parameters) of multi-state models. For the estimation and graduation of the transition intensities, methods that derive from the TI-approach have been developed—see [Waters \(1984\)](#), [Dickson et al. \(2019\)](#) and [Haberman and Pitacco \(2018\)](#)—including parametric approaches that apply generalised linear models (GLMs) as graduation techniques, as described in [Rensham and Haberman \(1995\)](#).

Several studies, employing diverse methodologies involving Markov multi-state models, predominantly focus on age and gender as the key risk factors in the analysis of long-term care dynamics. [Cui et al. \(2022\)](#) estimates the transition probabilities to project LTC costs, explicitly accounting for age and gender differences. Similarly, [Pritchard \(2013\)](#) evaluates the transition intensities and associated costs under LTC insurance, distinguishing between males and females. [Fong et al. \(2015\)](#) develops age- and sex-specific functional status transition rates using GLMs, particularly for older individuals. [Baione and Levantesi \(2014\)](#) addresses data limitations by estimating transition intensities based on aggregated mortality and morbidity information, stratified by age and gender. [Leung \(2006\)](#) applies graduation by mathematical formulae to obtain graduated transition intensities and then compute LTC premiums, with separate considerations for males and females. Collectively, these studies highlight the significance of age and gender in assessing LTC risks and costs.

The primary goal of the work presented in this paper is to develop a parametric framework for the estimation and graduation of transition intensities within a multi-state model, specifically designed for disability and LTC insurance applications. Recent studies on the analysis of long-term transition dynamics employ methodological approaches that closely resemble ours, similarly using age and gender as baseline risk factors, though differing in the selection of additional covariates. [Czado and Rudolph \(2002\)](#) employs a semiparametric hazard model that, while centred on diagnoses as key predictors, also includes age, sex and nursing home status as explanatory variables, which aligns with our interest in multidimensional risk factors. Similarly, [Fuino and Wagner \(2018\)](#) models transition probabilities by age and gender but extends the analysis to account for duration in dependency. [Kabuche et al. \(2024\)](#) maintains a demographic focus (age and gender) but introduces additional stratification by functional disability and chronic illness, mirroring our emphasis on layered risk assessment. [Guo \(2024\)](#), [Arrighi et al. \(2017\)](#) and [Casasnovas and Nicodemo \(2016\)](#) further expand the scope by incorporating socioeconomic variables (education, economic status, or residence), reinforcing our methodological commitment to capture diverse influences on LTC transitions. While these studies share important conceptual similarities with our approach (particularly in incorporating additional risk covariates and addressing data limitations in LTC insurance), our methodology introduces novelty in the considered risk factors and through the application of GLMs. In contrast with the mentioned authors, which predominantly employ semiparametric hazard models, in our GLM-based framework, we take advantage of the flexibility and simplicity offered by this type of models while maintaining robust modelling capabilities.

In the absence of real portfolio data for the application of the graduation techniques, we resort to data simulation to generate a dataset of individuals with demographic and lifestyle characteristics (such as age, sex, smoking habits, body mass index, exercise habits and region of residence) representative of the Portuguese population.

The ultimate objective of this work is to present functional forms for the transition intensities as a function of age, where the parameters are dependent on the demographic characteristics of the population. In other words, the proposed methodology differentiates individuals based on their risk profiles when estimating their transition intensities. This approach properly adjusts the probabilities used in the calculation of premiums and reserves for disability and long-term care insurance, ensuring that premiums are appropriately aligned with individual risk profiles.

Personalised risk premiums in LTC insurance can significantly benefit insurers by improving risk assessment accuracy, enhancing profitability and promoting market sustainability. By tailoring premiums to individual risk factors, insurers can better align pricing with actual expected costs, reducing adverse selection and financial instability. Advanced modelling techniques, including transition probability analysis and multi-state frameworks, allow for more precise premium differentiation, thus ensuring that high-risk policyholders pay appropriate premiums while lower-risk individuals are not overcharged. This tailored pricing improves underwriting efficiency, optimises reserve allocations, and may expand insurability by making coverage more accessible to lower-risk groups. Accurate personalised premiums in LTC insurance enhance risk management by reducing financial volatility through precise claim forecasting, improving underwriting efficiency via data-driven models, and strengthening competitiveness by attracting lower-risk policyholders with fair pricing. This approach also supports regulatory compliance by ensuring adequate reserves and solvency margins, while enabling dynamic pricing adjustments for long-term portfolio stability and profitability. Overall, personalised LTC premiums strengthen insurers' financial resilience while incentivising data-driven product innovation.

In this work, we present a model that is both generalizable and applicable in a real-world context. Section 2 outlines our proposed framework step by step, ensuring straightforward adaptation to other model formulations and/or real-world datasets. Section 2.1 introduces a multi-state model designed for disability and long-term care insurance applications, which serves as the foundation for illustrating the proposed approach, which relies on the estimation and graduation of the transition intensities considering multiple risk factors. To address the critical lack of reliable data for disability and LTC insurance in Portugal, we developed a sophisticated synthetic portfolio in Section 3.1 that accurately mirrors the following: (a) Portugal's demographic structure; (b) lifestyle risk factors representative of the population; and (c) policyholder behaviour patterns. This simulated dataset enables Portuguese insurers to (i) test our methodology without proprietary data constraints; (ii) validate models against Portugal-specific risk profiles; and (iii) develop products tailored to the country's ageing population challenges. The simulation framework, which may be extended for other risk factors, provides an essential tool for insurers to overcome data limitations while maintaining actuarial rigour, making it a crucial advancement for this developing insurance market. The proposed methodology directly addresses the need for sophisticated yet practical tools to support the growing disability and LTC insurance sector.

The framework developed in this paper offers a flexible foundation for multi-state modelling across diverse populations and/or insurance portfolios and, while broadly applicable, the results obtained in this paper may be used to propose a risk-adjusted pricing structure tailored to the Portuguese population's profile.

## 2. Risk-Adjusted Estimation and Graduation of Transition Intensities

This section presents the proposed methodology for the estimation and risk-adjusted graduation of transition intensities within continuous-time Markov multi-state models in a structured sequence, with each step corresponding to the fully detailed technical development in subsequent subsections.

The framework proposed in this paper can be summarised into three key steps, briefly described as follows:

### 1. Multi-state model specification

- Design of a continuous-time Markov model with the definition of a finite state space, with states representing the underlying condition or status of individuals.
- Identification of possible transitions between states and the characterisation of the corresponding transition intensity matrix.

## 2. Estimation of transition intensities

Estimation of piecewise constant transition intensities, stratified by age intervals, following a maximum likelihood approach.

## 3. Risk-adjusted graduation of transition intensities

### (a) Data preparation

Collection of transition counts between states and state-specific sojourn times, incorporating relevant risk factor data.

### (b) Modelling framework

Implementation of a Poisson GLM (log-link function) with the formal inclusion of risk factors through the covariate selection and estimation of smooth, risk-adjusted intensity estimates.

### (c) Functional form specification and parameter estimation

Definition of a functional form for risk-adjusted transition intensities (e.g., Gompertz–Makeham), using non-linear regression models for parameters estimation.

The technical details of these steps are presented comprehensively in Sections 2.1–2.3.

### 2.1. Multi-State Model Specification

We begin by defining the specific multi-state model employed throughout this paper to illustrate the results, while emphasizing that the proposed methods and findings are generalizable to a broader class of multi-state models.

For a general continuous-time multi-state model, with a finite state space  $S = \{1, \dots, n\}$ ,  $n \in \mathbb{N}$ , each state represents a distinct condition of an individual at a given time. Instantaneous transitions are possible between selected pairs of states, and the model describes the random transitions of the individual over time as they move between the possible states.

For each  $x \geq 0$ , the random variable  $X(x)$  takes one of the values  $\{1, \dots, n\}$ , which represents the state of the individual at a given age  $x$ . The set  $\{X(x); x \geq 0\}$  is a continuous-time stochastic process and, for states  $i, j \in S$  and times  $x, t \geq 0$ , we define the transition probabilities,

$${}_t p_x^{ij} \equiv P_{ij}(x, t) = P[X(x+t) = j | X(x) = i],$$

and the transition intensities of the multi-state model,

$$\mu_x^{ij} \equiv \mu_{ij}(x) = \lim_{h \rightarrow 0^+} \frac{{}_h p_x^{ij}}{h}, \quad i \neq j.$$

Furthermore, assuming that the probability of future transitions depends solely on the current state of the process (i.e., the transition probability  ${}_t p_x^{ij}$  is independent of any past values of the process prior to time  $x$ ), we conclude that the process  $\{X(x); x \geq 0\}$  is a continuous-time Markov chain.

Multi-state models can be designed for disability and LTC insurance purposes by considering several states to represent different degrees of disability/lack of autonomy.

Disability insurance (DI) and long-term care insurance (LTCI) both address disability risk but differ fundamentally in scope and purpose: DI provides income replacement when policyholders cannot work due to illness/injury, typically paying monthly benefits tied to pre-disability earnings and requiring occupational incapacity proof; meanwhile, LTCI covers actual care costs (such as nursing homes or home care) triggered by functional impairments in Activities of Daily Living (ADLs) or cognitive decline, reimbursing specific expenses regardless of employment status. DI primarily serves working-age populations

with time-limited benefits, whereas LTCI targets older adults or those with chronic conditions, offering lifelong coverage for care needs. Although distinct, both products mitigate financial risks arising from disability as follows: DI preserves income streams, while LTCI protects against catastrophic care expenditures. The methodology proposed in this paper serves both types of insurance, with premiums defined according to the respective coverages and benefits.

To illustrate the framework proposed in this paper, we consider a general four-state continuous-time Markov model, with two disability states (State 1 ≡ Autonomous, State 2 ≡ Mildly Disabled, State 3 ≡ Severely Disabled, and State 4 ≡ Dead), where a direct recovery from a severe disability is not possible, as illustrated in Figure 1.

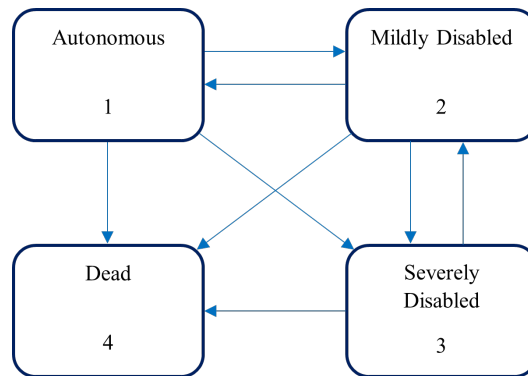


Figure 1. Disability/LTC model with two disability states.

For  $x \geq 0$ , the corresponding transition intensity matrix of the model is given by

$$Q(x) = \begin{bmatrix} -(\mu_x^{12} + \mu_x^{13} + \mu_x^{14}) & \mu_x^{12} & \mu_x^{13} & \mu_x^{14} \\ \mu_x^{21} & -(\mu_x^{21} + \mu_x^{23} + \mu_x^{24}) & \mu_x^{23} & \mu_x^{24} \\ 0 & \mu_x^{32} & -(\mu_x^{32} + \mu_x^{34}) & \mu_x^{34} \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{1}$$

and, for  $x, t \geq 0$ , the transition probability matrix is given by

$$P(x, t) = \begin{bmatrix} {}_t p_x^{11} & {}_t p_x^{12} & {}_t p_x^{13} & {}_t p_x^{14} \\ {}_t p_x^{21} & {}_t p_x^{22} & {}_t p_x^{23} & {}_t p_x^{24} \\ {}_t p_x^{31} & {}_t p_x^{32} & {}_t p_x^{33} & {}_t p_x^{34} \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

with  ${}_t p_x^{ij}$  easily derived from the Kolmogorov forward equations given the transition intensities, see (Haberman and Pitacco 2018, p. 18).

While the model in Figure 1 considers only two intermediate disability/LTC states, we emphasise that both the methodology and results presented in this paper can be easily extended to more complex multi-state disability/LTC frameworks.

### 2.2. Estimation of Transition Intensities

In this paper, we adopt the formulation of multi-state models based on the transition intensity approach (TI-approach), in which the transition intensities are treated as the models’ fundamental quantities (or parameters). For the estimation of the transition intensities, we follow a maximum likelihood approach to obtain crude estimates, assuming that the transition intensities are piecewise constant functions of age, i.e., they are constant between consecutive ages. This technique was introduced in Waters (1984) and also studied in Haberman and Pitacco (2018) and Dickson et al. (2019).

In the following, let us consider a population of individuals under observation during a certain period of consecutive years and, for illustration purposes, the disability multi-state model presented in Figure 1. For each individual, we collect data referring to the transitions between states during the observation period, including the sojourn times in each state, as well as the number of transitions of each type. Hence, each age  $x$  individual symbolises an independent realisation of the underlying continuous-time Markov process  $\{X(x); x \geq 0\}$ .

Without loss of generality, let us select an interval of consecutive ages, say  $[x, x + 1)$ , over which we can reasonably assume that the transition intensities are constant, i.e.,

$$\mu_{x+s}^{ij} = \mu_x^{ij}, \forall s \in [0, 1), i, j \in S. \tag{2}$$

Let us also assume that there are  $N$  individuals under observation with age within the considered range  $[x, x + 1)$ .

### Transition Intensity Estimation Approach

The proposed methodology, based on the model depicted in Figure 1, but easily adapted for different multi-state models, can be summarised by the following steps:

#### 1. Individual likelihood function

Determine the likelihood function for each individual under observation as a function of the transition intensities.

For a single individual, see (Haberman and Pitacco 2018, p. 147), each stay of duration  $t \in [0, 1)$  in a given state, followed by a transition to a different state, contributes to the likelihood of the sample path with a factor of the form

$${}_t p_x^{\bar{i}i} \mu_x^{ij}.$$

Following (Dickson et al. 2019, p. 296) and considering assumption (2), for  $t \in [0, 1)$ , each sojourn probability can be rewritten as

$${}_t p_x^{\bar{i}i} = \exp \left\{ - \int_0^t \sum_{\substack{k \in S \\ k \neq j}} \mu_{x+s}^{ik} ds \right\} = \exp \left\{ - \int_0^t \sum_{\substack{k \in S \\ k \neq j}} \mu_x^{ik} ds \right\} = \exp \left\{ - \left( \sum_{\substack{k \in S \\ k \neq j}} \mu_x^{ik} \right) t \right\}.$$

Thus, for  $i, j \in S$ , and defining

$T_k^i$ —total sojourn time of individual  $k$  in state  $i$ ;

$N_k^{ij}$ —total number of transitions of individual  $k$  from state  $i$  to state  $j$ ;

the likelihood function for individual  $k$ , as a function of the transition intensities  $\boldsymbol{\mu} = (\mu_x^{12}, \mu_x^{13}, \mu_x^{14}, \mu_x^{21}, \mu_x^{23}, \mu_x^{24}, \mu_x^{32}, \mu_x^{34})$ , is given by

$$\mathcal{L}_k(\boldsymbol{\mu}) = e^{-T_k^1(\mu_x^{12} + \mu_x^{13} + \mu_x^{14})} e^{-T_k^2(\mu_x^{21} + \mu_x^{23} + \mu_x^{24})} e^{-T_k^3(\mu_x^{32} + \mu_x^{34})} \times (\mu_x^{12})^{N_k^{12}} (\mu_x^{13})^{N_k^{13}} (\mu_x^{14})^{N_k^{14}} (\mu_x^{21})^{N_k^{21}} (\mu_x^{23})^{N_k^{23}} (\mu_x^{24})^{N_k^{24}} (\mu_x^{32})^{N_k^{32}} (\mu_x^{34})^{N_k^{34}}.$$

#### 2. Population likelihood function

Obtain the likelihood function for the whole population under observation.

Assuming the mutual independence of the individuals' sample paths, the likelihood function for the observed population is given by

$$\mathcal{L}(\boldsymbol{\mu}) = e^{-T^1 \mu_x^{12}} (\mu_x^{12})^{N^{12}} e^{-T^1 \mu_x^{13}} (\mu_x^{13})^{N^{13}} e^{-T^1 \mu_x^{14}} (\mu_x^{14})^{N^{14}} e^{-T^2 \mu_x^{21}} (\mu_x^{21})^{N^{21}} \times e^{T^2 \mu_x^{23}} (\mu_x^{23})^{N^{23}} e^{-T^2 \mu_x^{24}} (\mu_x^{24})^{N^{24}} e^{-T^3 \mu_x^{32}} (\mu_x^{32})^{N^{32}} e^{-T^3 \mu_x^{34}} (\mu_x^{34})^{N^{34}}.$$

with  $T^j = \sum_{k=1}^N T_k^j$  and  $N^{ij} = \sum_{k=1}^N N_k^{ij}$ ,  $i, j \in \mathcal{S}$ .

### 3. Estimation of transition intensities

Obtain the maximum likelihood estimators for the transition intensities.

Factorizing the likelihood function with respect to each parameter to be estimated, we obtain terms of the form

$$\mathcal{L}(\mu_x^{ij}) = e^{-T^i \mu_x^{ij}} (\mu_x^{ij})^{N^{ij}}, \quad (3)$$

resulting in the log-likelihood function

$$\ell(\mu_x^{ij}) = \log \mathcal{L}(\mu_x^{ij}) = -T^i \mu_x^{ij} + N^{ij} \log(\mu_x^{ij}).$$

The maximum likelihood estimators for the transition intensities, for a given age  $x$ , are given by

$$\hat{\mu}_x^{ij} = \frac{N^{ij}}{T^i}, \quad i, j \in \mathcal{S}. \quad (4)$$

Following the steps outlined above, for each consecutive age interval within the observation period, crude estimates of the transition intensities for each year of age are obtained, resulting in a piecewise constant function between ages (as initially assumed).

This methodology and assumption may compromise certain desired properties, such as the smoothness of the transition intensity function across ages, see (Haberman and Pitacco 2018, p. 145). Therefore, graduation of these raw estimates is usually employed to maintain such desired characteristics.

#### 2.3. Risk-Adjusted Graduation of Transition Intensities

In this work, a parametric approach using generalised linear models (GLMs), specifically Poisson regression models, with predictors of the Gompertz–Makeham form is adopted. For completeness, we begin with a brief introduction of the necessary concepts.

The Gompertz–Makeham mortality law, see Gompertz (1825) and Makeham (1860), is a fundamental model for age-dependent death rates, and it can be expressed as

$$\mu_x = A + Bc^x, \quad A \geq 0, B > 0, c > 1.$$

While the Gompertz–Makeham law is primarily used to model mortality rates, it can also be applied to describe other age-dependent transitions in multi-state models, such as long-term care and disability insurance, see (Haberman and Pitacco 2018, p. 100).

Moreover, it is among the first expressions employed for the graduation of mortality rate estimates, see Forfar et al. (1988). This approach is also applicable to the graduation of other transition intensities in multi-state models and, notably, it can be integrated into graduation processes using GLMs, see Renshaw (1991), as demonstrated in the graduation process outlined in this paper.

For the risk-adjusted graduation of transition intensities, consider a vector  $\mathbf{y} = (y_i)$ ,  $i = 1, 2, \dots, n$ , of observations (or response variables), assumed to be the realisation of  $\mathbf{Y} = (Y_i)$ , a vector of independent random variables. Additionally, let  $\mathbf{x}_j = (x_{ij})$ ,  $j = 1, 2, \dots, p$ , represent the vectors of covariates (or explanatory variables), each containing  $n$  known observations, which model the response variables.

In order to estimate the expected value of  $\mathbf{Y}$  given the observations  $\mathbf{X}$ , i.e.,  $\mathbf{m} = E(\mathbf{Y}|\mathbf{X})$ , GLMs assume that there exists a non-linear injective and differentiable transformation  $g$ , known as the link function, such that

$$g(\mathbf{m}) = \boldsymbol{\eta} \Leftrightarrow \mathbf{m} = g^{-1}(\boldsymbol{\eta}) = g^{-1}(X\boldsymbol{\beta})$$

where  $\boldsymbol{\eta} = X\boldsymbol{\beta}$  is a linear predictor and  $\beta_j, j = 1, 2, \dots, p$ , are unknown parameters to be estimated by maximum likelihood, based on observations  $\mathbf{y}$ .

To graduate the raw estimates of the transition intensities obtained from the methodology described in Section 2.2, we consider a Poisson regression, i.e., a GLM with a Poisson-distributed response variable, with a logarithmic link function.

This parametric graduation method based on GLMs was first proposed in Renshaw (1991) for the graduation of the force of mortality. In Rensham and Haberman (1995), it was further applied to the graduation of the transition intensities in the standard sickness–death model, noting that it can be generalised to other multi-state models, provided the required data are available. While Rensham and Haberman (1995) considers as the explanatory variables the age at the onset of sickness and, for the transitions from the sick state, the sickness duration, this paper incorporates additional observable risk factors, which is one of the key contributions of this work.

Regarding the modelling of mortality indicators, several methods have been proposed to account for population heterogeneity and the impact of various observable variables on mortality. In Rensham (1994), a graduation of the force of mortality is performed, incorporating multiple explanatory variables, such as gender, smoking habits, medical status, and age.

In this paper, we generalise the approach presented in Rensham and Haberman (1995) to the multi-state model introduced in Section 2.1, expanding the explanatory variables to include additional demographic risk factors (such as age, sex, smoking habits, body mass index, exercise habits, and region of residence), in line with the methodology of Rensham (1994).

The proposed risk-adjusted graduation process consists of the following steps:

### 1. Data collection

Collect and process the data regarding the transitions between states.

For the graduation of each transition intensity, organise the data in a grid of units or cells denoted by  $u \equiv (x_1, x_2, \dots)$ , where  $x_1, x_2, \dots$  correspond to the risk factors. The observed data consist of a set of ordered pairs  $(i_u, e_u)$ , where  $i_u$  is the number of transitions accruing from the central exposures (or sojourn times)  $e_u$ , for each unit  $u$ . The covariates included in the definition of units  $u$  will be different for each intensity, as well as the definition of  $i_u$  and  $e_u$ , as follows:

- For  $\mu_u^{12}, \mu_u^{13}$ , and  $\mu_u^{14}$ ,  $i_u$  represents the observed number of transitions from the Autonomous state to the Mildly Disabled, Severely Disabled, and Dead states, respectively, accruing from corresponding exposures  $e_u$ , which, in this case, is the total sojourn time in the Autonomous state.
- For  $\mu_u^{21}, \mu_u^{23}$ , and  $\mu_u^{24}$ ,  $i_u$  represents the observed number of transitions from the Mildly Disabled state to the Autonomous, Severely Disabled, and Dead states, respectively, accruing from corresponding exposures  $e_u$ , which, in this case, is the total sojourn time in the Mildly Disabled state.
- For  $\mu_u^{32}$  and  $\mu_u^{34}$ ,  $i_u$  represents the observed number of transitions from the Severely Disabled state to the Mildly Disabled and Dead states, respectively, accruing from corresponding exposures  $e_u$ , which, in this case, is the total sojourn time in the Severely Disabled state.

### 2. GLM framework

Without loss of generality, let  $\mu$  be one of the transition intensities available in the disability model with two disability states, presented in Section 2.1. We assume that

there exists a differentiable and injective (and thus invertible) function  $g$  that links the intensities  $\mu_u$  to a linear predictor

$$\eta_u = \sum_j h_j(u) \beta_j, \quad (5)$$

where functions  $h_j$  define the known covariate structure of the current model and coefficients  $\beta_j$  are the unknown regression parameters. In other words,

$$g(\mu_u) = \eta_u \Leftrightarrow \mu_u = g^{-1}(\eta_u) = g^{-1}\left(\sum_j h_j(u) \beta_j\right). \quad (6)$$

### 3. Parameter estimation

Estimate the  $\beta_j$  parameters using a maximum likelihood approach.

Assuming constant transition intensities between consecutive ages, we showed that the likelihood function for the transition intensities is expressed by Equation (3), which is equivalent to assuming that the number of transitions is Poisson-distributed. Thus, considering  $i_u$  the response variables of the GLM, we assume they are independent and Poisson-distributed, with the expected number of transitions given by the product of the exposures to risk and the respective transition intensity.

However, since the same transition between two states may be observed multiple times for each individual (with the exception of transitions to the Dead state), the variance of the observed number of transitions may exceed the mean, indicating the presence of overdispersion.

To account for the possibility of overdispersion, we assume that  $i_u$  are independent overdispersed Poisson variables,

$$i_u \sim \text{Poi}(e_u \mu_u), \text{ independently for all } u,$$

with mean and variance, respectively, given by

$$m_u = E(i_u) = e_u \mu_u \text{ and } \text{Var}(i_u) = \phi m_u,$$

where  $\phi > 1$  is the scale or dispersion parameter. Furthermore, from (5) and (6), the expected value is expressed as

$$m_u = e_u \mu_u = e_u g^{-1}(\eta_u) = e_u g^{-1}\left(\sum_j h_j(u) \beta_j\right), \quad (7)$$

which establishes a relationship between  $m_u$  and the unknown parameters  $\beta_j$ .

Considering  $\mathbf{m} = (m_u)$  and  $\mathbf{i} = (i_u)$ , the log-likelihood function associated with the Poisson-distributed response variables is given by

$$\ell(\mathbf{i}; \mathbf{m}) = \log \mathcal{L}(\mathbf{i}; \mathbf{m}) = \sum_u (-m_u + i_u \log(m_u) - \log(i_u!)), \quad (8)$$

which can be expressed in terms of the  $\beta_j$  using Equation (7). The estimates  $\hat{\beta}_j$  are obtained by maximizing the log-likelihood function (8).

### 4. Link function

Choose the function  $g(\cdot)$  that links the transition intensities to the linear predictor containing the estimated parameters  $\beta_j$ . We employ the log-link function, as it serves as the canonical link for Poisson-distributed response variables.

Using the logarithm in Equation (6), we have

$$\log(\mu_u) = \eta_u \Leftrightarrow \mu_u = \exp\left(\sum_j h_j(u)\hat{\beta}_j\right),$$

replacing the linear predictor by its expression in Equation (5). By the properties of the logarithmic function,

$$\log(m_u) = \log(e_u\mu_u) = \log(e_u) + \sum_j h_j(u)\hat{\beta}_j,$$

and we observe that the terms  $\log(e_u)$  act in the linear model as an offset. Therefore, the graduated transition intensities are given by

$$m_u = e_u\hat{\mu}_u \Leftrightarrow \hat{\mu}_u = \frac{m_u}{e_u} \Leftrightarrow \hat{\mu}_u = \exp\left(\sum_j h_j(u)\hat{\beta}_j\right). \quad (9)$$

It is important to highlight that graduation methods based on GLMs are particularly valuable, not only due to their broad range of actuarial applications but also for their versatility. These methods can be applied to the transition intensities in multi-state models, extending beyond their traditional use in modelling mortality indicators.

### 3. Empirical Approach

In this section, we present a practical application of the previously developed methods for estimating and graduating the transition intensities of the multi-state model illustrated in Figure 1.

Using publicly available data for Portugal—for Health statistics, see [INE \(2020a, 2024\)](#), and for Census data, see [INE \(2021a, 2021b, 2021c\)](#)—we employ simulation techniques to generate a dataset/portfolio that reflects the characteristics of Portuguese individuals, as well as their trajectories through the states of the adopted multi-state model over a given observation period. By applying the previously introduced GLM framework, we obtain graduated estimates of the transition intensities, which depend not only on the individuals' age but also on other relevant risk factors.

The simulation procedures, model fitting and additional computations have been implemented using R (version 4.2.3) software.

#### 3.1. Dataset/Portfolio Simulation

In Portugal, there is currently a lack of data related to LTC insurance, as referred to, for instance, in [Esquível et al. \(2021\)](#). Therefore, it is challenging to obtain the necessary data to estimate and graduate transition intensities following the TI-approach (namely, number of transitions of each type and the corresponding central exposures per age).

For the purpose of this study, in the absence of a real dataset, we start by simulating a dataset/portfolio reflecting the demographic characteristics of the Portuguese population based on the most recent health statistics, see [INE \(2020a, 2024\)](#), and Census data, see [INE \(2021a, 2021b, 2021c\)](#). Employing the calibrated transition intensities derived in [Esquível et al. \(2021\)](#) for a similar multi-state model and from real data of the Portuguese population, we use the transition intensity matrix of the model in Figure 1 to simulate trajectories for the individuals within the portfolio over a given observation period.

Current demographic and health-related trends observed in Portugal (many of which are shared with other nations), such as the rapidly ageing population—projected to reach 3 million elderly by 2050, according to [INE \(2020b\)](#)—the high prevalence of chronic

conditions among older adults, and regional disparities in healthcare access, highlight the need for LTC insurance products. By integrating Portugal-specific characteristics—such as age and sex distribution, smoking and exercise habits, body mass index (BMI), and region of residence—as risk factors into the modelling of transition intensities, this study provides a foundation for developing tailored premium models. These models align with market needs and public health priorities, while addressing Portugal’s critical LTC coverage gap through more accurate, equitable, and sustainable risk-based pricing.

The 2021 Population and Housing Census, conducted by the Portuguese National Institute of Statistics (INE), provides data on the distribution of the resident population in Portugal by age group, sex, and regional population size, see INE (2021a, 2021b, 2021c). Additionally, the 2019 National Health Inquiry, also conducted by the INE, provides data on the distribution of Portugal’s resident population by age group and sex for each of the remaining covariates (BMI, smoking habits, and exercise habits), see INE (2020a). Note that BMI serves as a high-level indicator of an individual’s health status based on their height and weight.

Based on the mentioned publicly available data for the Portuguese population, see INE (2020a, 2021a, 2021b, 2021c, 2024), the risk factors and the corresponding levels considered to model the transition intensities are described in Table 1. Note that the variable region is categorised in terms of population size (in number of individuals) and the exercise habits are measured by the number of hours spent performing physical activity in a normal week.

**Table 1.** Risk factors and levels considered for the simulated population.

Risk Factors	Levels
Age	Five-year age groups, from 0 to 100 years old
Sex	Male   Female
Region	0–1999   2000–4999   5000–9999   10,000–19,999   20,000–99,999   100,000 or more
BMI	Underweight (BMI < 18.5)   Normal weight ( $18.5 \leq \text{BMI} < 25$ )   Overweight I ( $25 \leq \text{BMI} < 27$ )   Overweight II ( $27 \leq \text{BMI} < 30$ )   Obese (BMI $\geq 30$ )
Smoking	Smoker   Non-smoker
Exercise	Less than 2 h   2 to less than 3 h   3 to less than 5 h   5 h or more

A dataset of 100,000 mutually independent individuals was simulated, in which the explanatory variables for each individual are sampled according to the available statistical data for the Portuguese population.

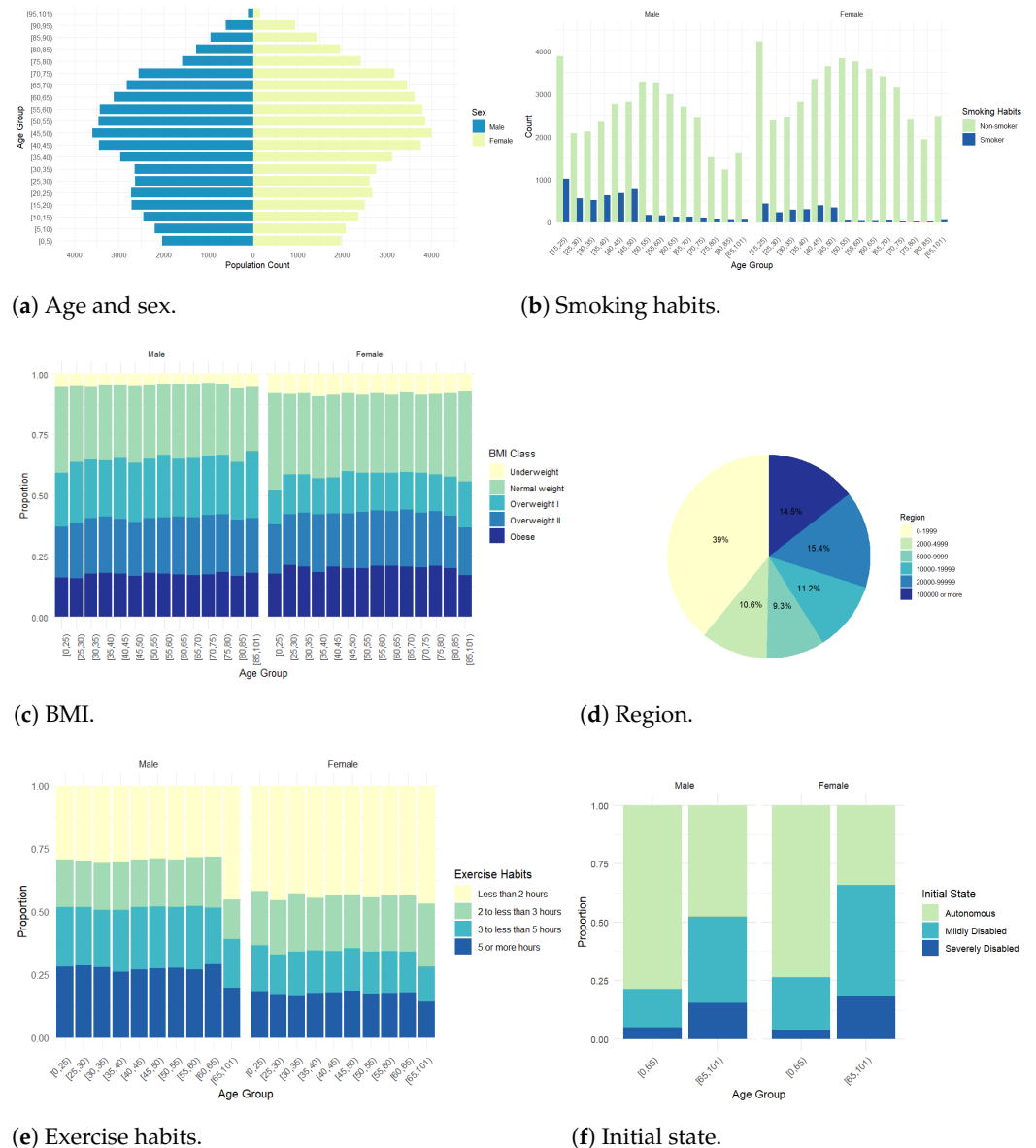
The simulation of the individual trajectories through the states of the multi-state model is based on each individual’s characteristics or, in other words, their risk profile. The following elements were simulated for each individual:

- Initial state: the condition or status of the individual at the beginning of the observation period, serving as the starting point for their trajectory.
- Exposure: the length of time during which the individual is observed (or at risk) within the study period.
- Transition intensities: the intensities of the transitions performed by each individual, according to the transition intensity matrix (1) associated with the adopted multi-state model.

For the purpose of the data simulation, we assume the initial state of the trajectories may correspond to any state of the model except the Dead state. Initial states are then

sampled for each individual in the population using the information available on the Health Statistics published by INE (2024) for the year 2022, which provides the distribution of the resident population in Portugal by sex and age groups, categorised according to limitations in performing daily activities (“Not limited” corresponding to the Autonomous state, “Limited but not severely” to the Mildly Disabled state, and “Severely limited” to the Severely Disabled state). Note that when dealing with real portfolio data regarding a disability or LTC insurance, the Autonomous state would be the initial state for all individuals (which would impact the transition intensities estimates).

Figure 2 illustrates the distribution of the generated population by each risk factor and initial state.



**Figure 2.** Distribution of the simulated population by variable. Source: Authors preparation.

For the simulation of the exposure (or exposure to risk) of each individual, we considered that the observation period is uniformly distributed between 6 and 12 months, and that the exposure is independent from their demographic characteristics.

For the transition intensities included in the transition intensity matrix (1), following Esquível et al. (2021) and Haberman and Pitacco (2018), firstly we assume intensities of the Gompertz–Makeham type, i.e., of the form

$$\mu_x^{ij} = \gamma_{ij} + 10^{\alpha_{ij}x + \beta_{ij}}, \alpha_{ij}, \beta_{ij}, \gamma_{ij} \in \mathbb{R}, \tag{10}$$

in which the transition intensities exponentially increase with the age of the individual. This is a reasonable assumption that is commonly used in other countries for disability insurance, see Haberman and Pitacco (2018), for instance.

As a starting point, we consider the calibrated transition intensity parameters obtained in Esquível et al. (2021) for a multi-state model with three dependence states, using data from 2015 of the Portuguese RNCCI. Therefore, the initial transition intensities for the individuals of the dataset simulated in this paper are of the form of Equation (10), with the parameters for each transition presented in Table 2.

**Table 2.** Initial parameters for the Gompertz–Makeham-type transition intensities. Source: Esquível et al. (2021).

State	<i>i</i>	<i>j</i>	$\gamma_{ij}$	$\alpha_{ij}$	$\beta_{ij}$
Autonomous	1	2	0.00040	0.060	−5.46
	1	3	0.00044	0.052	−5.46
	1	4	0.00050	0.038	−4.12
Mildly Disabled	2	1	0.00040	0.060	−5.46
	2	3	0.00043	0.054	−5.46
	2	4	0.00050	0.037	−4.12
Severely Disabled	3	2	0.00043	0.054	−5.46
	3	4	0.00042	0.054	−4.12

By calibrating the Gompertz–Makeham parameters, we aim to model transition intensities that reflect individual risk profiles. Table 3 illustrates the most frequent characteristics (i.e., most common risk profile) in the simulated population/portfolio.

**Table 3.** Most frequent risk profile of the simulated portfolio.

Sex	Region	BMI	Smoking	Exercise
Female	0–1999	Normal weight	Non-smoker	Less than 2 h

For these individuals, transition intensities are defined by the parameters specified in Table 2. To account for different risk profiles, we modify these parameters using multiplicative adjustment factors. Specifically, for a transition intensity of the form (10), the adjusted parameters are expressed as

$$\alpha_{ij}^* = a \cdot \alpha_{ij}, \beta_{ij}^* = b \cdot \beta_{ij}, \gamma_{ij}^* = c \cdot \gamma_{ij}, \quad a, b, c \in \mathbb{R}.$$

Note that, for individuals whose characteristics match those in Table 3, all multiplicative adjustment factors are equal to 1. Moreover, the factors are cumulative, i.e., for individuals with multiple characteristics differing from the reference profile in Table 3, the final adjustment factor equals the product of all applicable individual adjustment factors.

For the intensities from the Autonomous state to the Mildly Disabled state,  $\mu_x^{12}$ , the adjustment factors according to each explanatory variable are presented in Table 4. To identify the optimal risk differentiation, and using published health statistics from WHO (2009, 2024), multiple combinations of adjustment factors were tested, and those that

best captured meaningful variations in covariate effects, while ensuring model coherence, were selected. The final chosen factors were designed to reflect both clinically plausible risk gradients and actuarially significant distinctions between covariate levels, maintaining an appropriate balance between sensitivity and stability across the pricing framework.

**Table 4.** Adjustment factors according to each variable for the transition intensity  $\mu_x^{12}$ .

Variable	Level	a	b	c
Sex	Male	1.05	0.9998	1.3
Region	2000–4999	1.0005	0.9999	1.01
	5000–9999	1.0010	0.9998	1.02
	10,000–19,999	1.0015	0.9997	1.03
	20,000–99,999	1.0020	0.9996	1.04
	100,000 or more	1.0040	0.9995	1.05
BMI	Underweight	1.01	0.9999	1.2
	Overweight I	1.01	0.9999	1.2
	Overweight II	1.03	0.9995	1.25
	Obese	1.05	0.9990	1.3
Smoking	Smoker	1.055	0.997	1.3
Exercise	2 to less than 3 h	0.995	0.9997	0.87
	3 to less than 5 h	0.985	0.9995	0.83
	5 or more hours	0.965	0.9990	0.77

Similar procedures for the remaining transition intensities were followed, considering appropriate adjustments to account for the effects of the selected risk factors, thereby ensuring the risk differentiation is consistently reflected across all transitions.

For simulation purposes, and given the assumption of Gompertz–Makeham transition intensities, we consider a normally distributed error term with 0 mean and standard deviation proportional to the transition intensity, i.e.,

$$\mu_{x, \text{final}}^{ij} = \mu_x^{ij} + \varepsilon_x^{ij},$$

with  $\varepsilon_x^{ij} \sim \mathcal{N}(0, 0.05 \cdot \mu_x^{ij})$ .

Finally, the trajectory of each individual in the population or, in other words, the sequence of states visited and the corresponding durations spent in each state throughout the observation period was simulated. Under the assumption of piecewise constant intensities between consecutive ages, the sojourn times and state transitions can be derived from the transition intensities, as described in (Ross 1996, p. 165).

Let us consider an individual, with simulated initial state and exposure, as well as transition intensity matrix (1), with the transition intensities depending on their risk profile. Starting from the initial state, the individual will remain in that state for a certain amount of time before they transition to the next state. The sojourn time in each state  $i$  is exponentially distributed, with rate equal to the total intensity of leaving state  $i$ , which we shall denote by  $\mu_x^i$ , with  $\mu_x^i = \sum_{j \neq i} \mu_x^{ij}$ .

If the generated sojourn time is greater than the observation period, the trajectory ends with the member having made no transitions. Otherwise, the next state  $j$  ( $j \neq i$ ) is randomly chosen, with probability  $\mu_x^{ij} / \mu_x^i$ . The trajectory ends when the absorbing state (Dead) is reached or when the aggregated sojourn times reach the end of the observation period.

Having simulated the trajectories for the 100,000 individuals, with observation periods between 6 months and 12 months, the number of transitions of each type and the times spent in each state observed for each individual were determined, which allows to proceed

with the estimation and graduation of transition intensities according to the risk profile of the individual, following the procedures described in Sections 2.2 and 2.3.

### 3.2. Results and Discussion

Following the GLM framework from Section 2.3 and incorporating the simulated risk factors, risk-adjusted graduated transition intensities, for the disability/LTC multi-state model defined in Section 2.1, were obtained. The ultimate objective is to derive parametric forms for these transition intensities that capture population risk heterogeneity.

First, we organise the available data into a grid of units  $u$ , where the categorical variables represent the risk factors and levels outlined in Table 1, where individuals under the age of 45 were grouped into a single age group.

The most frequent levels in the grid of units  $u$  for each risk factor are presented in Table 5. This set of characteristics constitutes the intercept (base level) of the fitted models.

**Table 5.** Base level of the fitted models.

Age Group	Sex	Region	BMI	Smoking	Exercise
[0, 45)	Male	0–1999	Normal weight	Non-smoker	Less than 2 h

#### 3.2.1. Transition Intensities from Autonomous to Mildly Disabled

In this section, for illustration purposes, we perform a detailed presentation of the estimation of risk-adjusted transition intensities from the Autonomous state to the Mildly Disabled state,  $\mu_u^{12}$ .

For each cell  $u$  of the grid, the required data consist of the observed number of transitions  $i_u$  from the Autonomous state to the Mildly Disabled state, accruing from corresponding exposure  $e_u$  (total sojourn time in the Autonomous state).

By assuming that  $i_u$  are independent overdispersed Poisson variables, we fit a quasi-Poisson regression to the data, using a log-link function and with  $\log(e_u)$  offsetting the model.

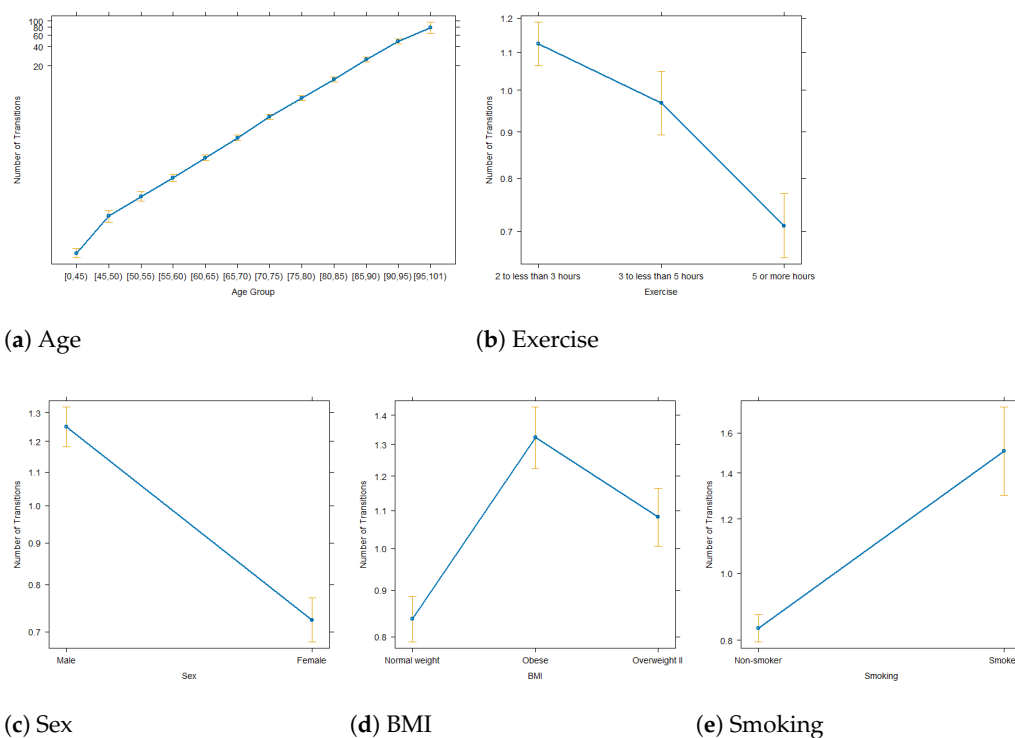
Table 6 presents the final GLM model for the transition intensities  $\mu_u^{12}$ , which includes the maximum likelihood estimates of the model’s coefficients, along with the standard error and statistical significance. Furthermore, we estimate the dispersion parameter of the quasi-Poisson model ( $\hat{\phi} = 2.082165$ ) and compute the AIC (10,627.89), as well as the residual deviance (3840.9) and the degrees of freedom (3533) of the final model. Model adequacy was further assessed through deviance residual analysis, which indicated a satisfactory goodness of fit. It is worth noting that all levels of the explanatory variables are found to be statistically significant.

Figure 3 illustrates the estimated effects of each covariate on the response variable in the final model. Each plot displays the predicted number of transitions by level of the explanatory variable (dots), the fitted value trend across levels (blue lines), and the associated 95% confidence intervals (orange bars).

Consistent with prior expectations, the estimated transition intensity from the Autonomous state (State 1) to the Mildly Disabled state (State 2) increases substantially with age. Furthermore, statistically significant differences are observed between smokers and non-smokers and between males and females, with higher intensities associated with the former in both comparisons. A negative association is identified between physical activity duration and this transition intensity, while individuals with a “normal weight” BMI classification exhibit reduced transition intensities relative to other BMI categories.

**Table 6.** Summary of the final GLM for the transition intensities  $\mu_u^{12}$ .

Variable	Estimate	Std. Error	t-Value	Pr(>  t )
(Intercept)	−6.87333	0.08792	−78.181	$<2 \times 10^{-16}$
Age [45,50)	1.32306	0.13766	9.611	$<2 \times 10^{-16}$
Age [50,55)	2.02961	0.11867	17.104	$<2 \times 10^{-16}$
Age [55,60)	2.69289	0.10464	25.735	$<2 \times 10^{-16}$
Age [60,65)	3.41294	0.09685	35.240	$<2 \times 10^{-16}$
Age [65,70)	4.13175	0.09738	42.427	$<2 \times 10^{-16}$
Age [70,75)	4.87628	0.09334	52.243	$<2 \times 10^{-16}$
Age [75,80)	5.55835	0.09352	59.432	$<2 \times 10^{-16}$
Age [80,85)	6.22471	0.09291	66.999	$<2 \times 10^{-16}$
Age [85,90)	6.93904	0.09352	74.201	$<2 \times 10^{-16}$
Age [90,95)	7.58535	0.09616	78.886	$<2 \times 10^{-16}$
Age [95,101)	8.07791	0.13251	60.959	$<2 \times 10^{-16}$
Sex Female	−0.54602	0.03026	−18.041	$<2 \times 10^{-16}$
BMI Obese	0.45861	0.03906	11.742	$<2 \times 10^{-16}$
BMI Overweight II	0.25769	0.03602	7.155	$1.01 \times 10^{-12}$
Smoking	0.59303	0.07884	7.522	$6.81 \times 10^{-14}$
Exercise 3 to less than 5 h	−0.14942	0.03980	−3.755	$1.76 \times 10^{-4}$
Exercise 5 or more hours	−0.45964	0.03991	−11.518	$<2 \times 10^{-16}$



**Figure 3.** Effects of risk factors on the response variable in the GLM for transition intensities  $\mu_u^{12}$ .

3.2.2. Further Transition Intensities

An analogous procedure is applied to the further transition intensities, using the corresponding definitions of the response variables  $i_u$  and exposures  $e_u$ . As in  $\mu_x^{12}$ , each transition intensity is modelled using an overdispersed Poisson regression with a log-link function, incorporating  $\log(e_u)$  as an offset term.

The parameter estimates of the fitted GLMs for all transition intensities  $\mu_u^{ij}$  included in matrix (1) are presented in Table 7.

**Table 7.** Estimates of the final GLMs for the transition intensities in matrix (1).

Variable	$\mu_u^{12}$	$\mu_u^{13}$	$\mu_u^{14}$	$\mu_u^{21}$	$\mu_u^{23}$	$\mu_u^{24}$	$\mu_u^{32}$	$\mu_u^{34}$
(Intercept)	-6.87333	-7.08794	-6.24542	-7.73021	-7.24478	-6.50972	-7.2253	-5.25677
Age [45, 50)	1.32306	-7.08794	1.55028	1.13871	1.80614	1.31045	-7.2253	2.26929
Age [50, 55)	2.02961	1.27903	1.55028	2.21250	1.80614	2.09526	-7.2253	2.26929
Age [55, 60)	2.69289	1.86906	2.18784	2.21250	1.80614	2.09526	-7.2253	3.20371
Age [60, 65)	3.41294	2.32632	2.56409	3.23600	2.81912	2.71701	2.4750	3.82183
Age [65, 70)	4.13175	3.08132	2.98246	3.85452	3.46311	3.02606	2.4750	4.51543
Age [70, 75)	4.87628	3.60239	3.49946	4.49781	4.15368	3.44995	3.5056	5.17514
Age [75, 80)	5.55835	4.11340	4.05923	5.20484	4.81944	3.94461	3.5056	5.75366
Age [80, 85)	6.22471	4.74307	4.52411	5.86117	5.37452	4.35849	4.6662	6.49562
Age [85, 90)	6.93904	5.37162	4.52411	6.49617	6.08471	4.84102	4.6662	7.09436
Age [90, 95)	7.58535	6.16580	5.31089	7.20018	6.78148	5.29609	4.6662	7.74841
Age [95, 101)	8.07791	6.16580	5.31089	7.66899	6.78148	5.29609	4.6662	8.25315
Sex Female	-0.54602	-0.47488	-0.32332	0.58868	-0.50526	-0.33586	0.4038	-0.50080
Region 2000–4999	—	—	-6.24542	—	—	0.31495	—	0.13896
Region 5000–9999	—	—	-6.24542	—	—	—	—	0.13896
Region 10,000–19,999	—	—	0.25148	—	—	0.18779	—	0.31372
Region 20,000–99,999	—	—	0.25148	—	—	0.18779	—	0.31372
Region >100,000	—	—	0.25148	—	—	0.31495	—	0.50977
BMI Underweight	-6.87333	-7.08794	-6.24542	-7.73021	-7.24478	-6.50972	—	0.12981
BMI Overweight I	-6.87333	-7.08794	-6.24542	-7.73021	-7.24478	-6.50972	—	0.12981
BMI Overweight II	0.25769	0.30254	0.21937	-0.27563	0.21041	0.15863	—	0.33459
BMI Obese	0.45861	0.30254	0.21937	-0.50894	0.44667	0.31333	—	0.49146
Smoking	0.59303	0.48355	0.26558	-0.53043	0.55937	0.31360	—	0.50395
Exercise 2–3 h	-6.87333	-7.08794	-0.12527	-7.73021	—	—	—	-5.25677
Exercise 3–5 h	-0.14942	-7.08794	-0.12527	0.14312	—	—	—	-0.14996
Exercise >5 h	-0.45964	-0.24204	-0.12527	0.45877	—	—	—	-0.38900

Legend: —, not significant.

In terms of risk factors, we observe that the variables age and sex are statistically significant in all models, along with the variables BMI and smoking, except for the transition intensity between the Severely Disabled and Mildly Disabled states. Additionally, the variable region is found to be significant solely for transitions to the Dead state, while the variable exercise does not exhibit significance for transitions from the Mildly Disabled state, except for those directed to the Autonomous state.

Subsequently, we generate predicted values for the response variables based on the fitted GLMs for all 100,000 individuals in the population. From Equation (9), the predicted number of transitions is divided by the total exposure to risk, resulting in risk-adjusted estimated transition intensities for each individual as predicted by the fitted models.

In conclusion, by leveraging the simulated trajectory data and applying the GLM framework, the proposed approach enables the estimation of transition intensities at the individual level, incorporating demographic characteristics. The predicted values obtained from the GLMs allow for the derivation of a functional representation of transition intensities across different risk profiles within the population.

### 3.2.3. Function for Transition Intensities Depending on Risk Factors

In this section, functional forms for the transition intensities, as functions of individuals' age, with the parameters reflecting the risk profile of a given group are derived. To this end, we consider Gompertz–Makeham-type transition intensities (10), and we fit the corresponding non-linear least squares (NLS) regression model to the predicted values obtained from the GLMs.

We start again by analysing the transition intensity,  $\mu_x^{12}$ , from the Autonomous to the Mildly Disabled state, aiming now to establish its functional form with respect to age for the overall population.

By fitting a NLS regression model to the predicted values of the transition intensities obtained from the GLM, using the functional form of Equation (10), and initializing with the parameter values provided in Table 2, the parameters  $\alpha_{ij}$ ,  $\beta_{ij}$ , and  $\gamma_{ij}$  are estimated through the minimisation of the residual sum of squares.

To improve the regression model’s fit at advanced ages, the NLS fitting procedure was restricted to individuals aged 50 and above. This approach aligns with the methodology of Esquivel et al. (2021), whose calibrated parameters—illustrated in Table 2—were derived under similar constraints.

As a result, the following functional form was obtained:

$$\mu_{x, \text{Population}}^{12} = -0.005982 + 10^{0.056215x - 4.952844}, \quad x \geq 50,$$

resulting from the model presented in Table 8, where the last two columns include a 95% confidence interval for the estimated parameters.

Table 8. NLS model for the function for the transition intensities of the whole population ( $x \geq 50$ ).

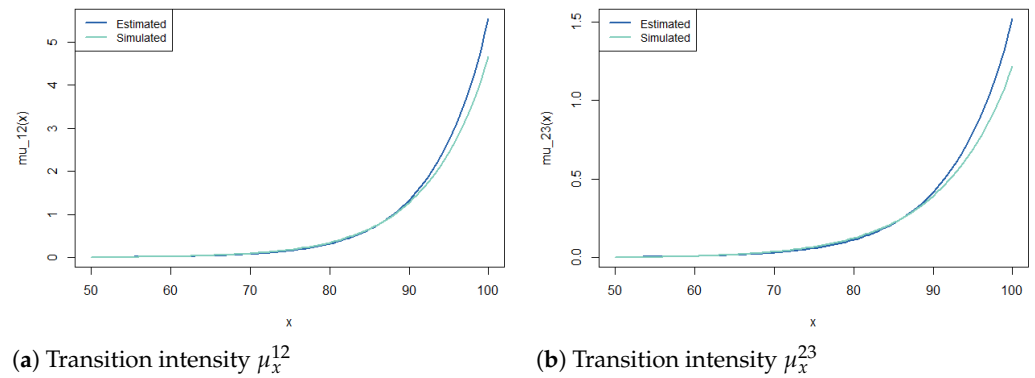
Parameter	Estimate	Std. Error	t-Value	Pr(>  t )	95% Confidence Interval	
					Lower Bound	Upper Bound
$\alpha_{12}$	0.056215	0.000212	265.214	$<2 \times 10^{-16}$	0.055799	0.056630
$\beta_{12}$	−4.952844	0.019500	−253.993	$<2 \times 10^{-16}$	−4.991064	−4.914623
$\gamma_{12}$	−0.005982	0.001130	−5.292	$1.22 \times 10^{-7}$	−0.008197	−0.003766
$\alpha_{13}$	0.050064	0.000177	282.483	$<2 \times 10^{-16}$	0.049716	0.050411
$\beta_{13}$	−5.135247	0.016269	−315.636	$<2 \times 10^{-16}$	−5.167135	−5.103358
$\gamma_{13}$	−0.001484	0.000202	−7.352	$1.98 \times 10^{-13}$	−0.001880	−0.001088
$\alpha_{14}$	0.033596	0.000130	257.81	$<2 \times 10^{-16}$	0.033341	0.033852
$\beta_{14}$	−3.564274	0.011993	−297.19	$<2 \times 10^{-16}$	−3.587781	−3.540767
$\gamma_{14}$	−0.005209	0.000334	−15.59	$<2 \times 10^{-16}$	−0.005864	−0.004554
$\alpha_{21}$	0.054953	0.000201	273.986	$<2 \times 10^{-16}$	0.054560	0.055346
$\beta_{21}$	−5.139355	0.018443	−278.663	$<2 \times 10^{-16}$	−5.175503	−5.103206
$\gamma_{21}$	−0.003686	0.000551	−6.685	$2.33 \times 10^{-11}$	−0.004767	−0.002605
$\alpha_{23}$	0.048841	0.000164	297.74	$<2 \times 10^{-16}$	0.048519	0.049162
$\beta_{23}$	−4.797967	0.015008	−319.70	$<2 \times 10^{-16}$	−4.827382	−4.768552
$\gamma_{23}$	−0.004394	0.000301	−14.61	$<2 \times 10^{-16}$	−0.004984	−0.003805
$\alpha_{24}$	0.035594	0.000122	291.750	$<2 \times 10^{-16}$	0.035355	0.035833
$\beta_{24}$	−3.816267	0.011209	−340.460	$<2 \times 10^{-16}$	−3.838237	−3.794297
$\gamma_{24}$	−0.001572	0.000241	−6.521	$7.07 \times 10^{-11}$	−0.002045	−0.001100
$\alpha_{32}$	0.024318	0.000108	226.01	$<2 \times 10^{-16}$	0.024107	0.024529
$\beta_{32}$	−3.120329	0.009920	−314.54	$<2 \times 10^{-16}$	−3.139773	−3.100885
$\gamma_{32}$	−0.011296	0.000165	−68.53	$<2 \times 10^{-16}$	−0.011619	−0.010973
$\alpha_{34}$	0.055437	0.000217	255.152	$<2 \times 10^{-16}$	0.055011	0.055863
$\beta_{34}$	−3.982169	0.019982	−199.283	$<2 \times 10^{-16}$	−4.021335	−3.943003
$\gamma_{34}$	−0.003552	0.009380	−0.379	$7.05 \times 10^{-1}$	−0.021937	0.014832

For a general model applicable across all ages and all LTC financing models, different NLS regression models of the Gompertz–Makeham type could be fitted for individuals below and above 50 years of age.

To assess the quality of the simulation results, we compare the estimated parameters with those of the transition intensities originally used to simulate the trajectories in Section 3.1. As a reminder, the initial transition intensities were generated by adjusting the baseline parameters from Table 2, using the adjustment factors provided in Table 4,

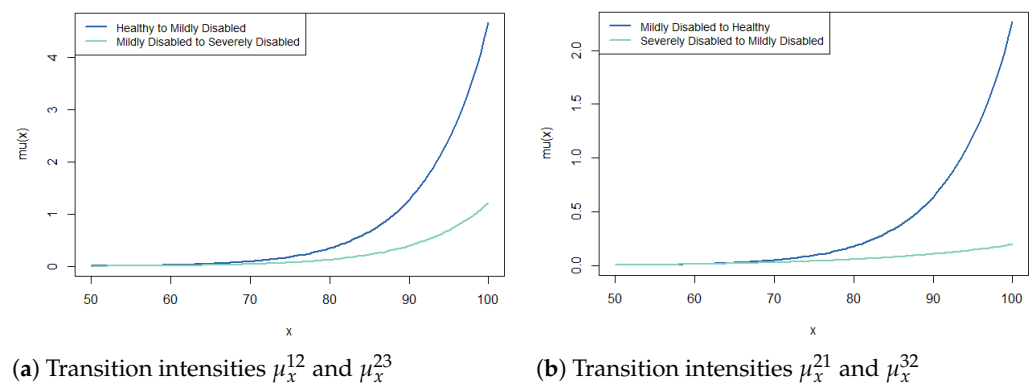
according to the characteristics of each individual. For the purpose of comparison, we compute the mean estimate of each parameter across the entire population.

For illustrative purposes, Figure 4 presents the transition intensities obtained both through the estimation procedures and from the simulation, aggregated over the entire population/portfolio. We focus specifically on the transitions from the Autonomous state to the Mildly Disabled state and from the Mildly Disabled state to the Severely Disabled state, which reflect consistency.



**Figure 4.** Functions for the transition intensities  $\mu_x^{12}$  and  $\mu_x^{23}$  for the whole population.

In Figure 5, we compare the transition intensities corresponding to disability deterioration—i.e., transitions from the Autonomous state to the Mildly Disabled state,  $\mu_x^{12}$ , and from the Mildly Disabled state to the Severely Disabled state,  $\mu_x^{23}$ —, as well as the corresponding recovery intensities from these states. It is observed that, for individuals aged over 75, the value of  $\mu_x^{12}$  increases at a higher rate than that of  $\mu_x^{23}$ , which is consistent with the findings reported in Esquível et al. (2021).



**Figure 5.** Functions for the transition intensities regarding disability deterioration and recovery for the whole population.

In particular, the proposed methodology allow us to obtain a functional form for the transition intensities as a function of age, stratified by risk factors and levels.

As an example, we present the functions for the transition intensity  $\mu_x^{12}$  specific to male and female individuals. We begin by stratifying the population by sex and then fitting group-specific NLS regression models to the predicted transition intensity values (obtained from the GLM). These models follow the functional form presented in (10), with the starting parameters summarised in Table 2. The resulting functions are as follows:

$$\mu_{x,\text{Male}}^{12} = -0.008172 + 10^{0.056355x - 4.838605}, \quad x \geq 50, \tag{11}$$

$$\mu_{x,\text{Female}}^{12} = -0.003818 + 10^{0.056489x - 5.087744}, \quad x \geq 50. \tag{12}$$

Figure 6 presents the functions (11) and (12), alongside the raw estimates of the transition intensities for each age. As a reminder, the crude estimates were obtained, in this case, by dividing the number of transitions between the Autonomous state and the Mildly Disabled state by the waiting time in the Autonomous state, as outlined in (4). It is observed that the obtained functions provide a good fit to the raw estimates for each year of age.

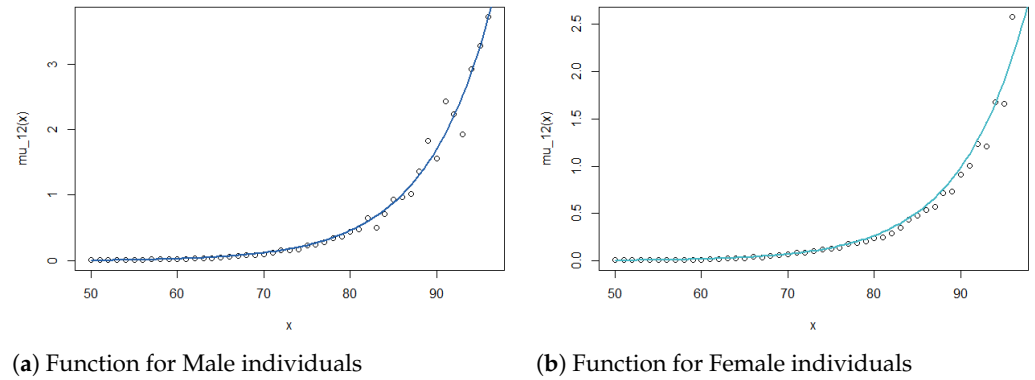


Figure 6. Functions for the transition intensities  $\mu_x^{12}$  and corresponding raw estimates by Sex.

By applying the same procedure to each of the significant risk factor included in the final models, transition intensity functions specific for each risk factor have been obtained. Table 9 presents the parameters to the stratification for each significant risk factor and the corresponding transition intensity functions are illustrated in Figure 7.

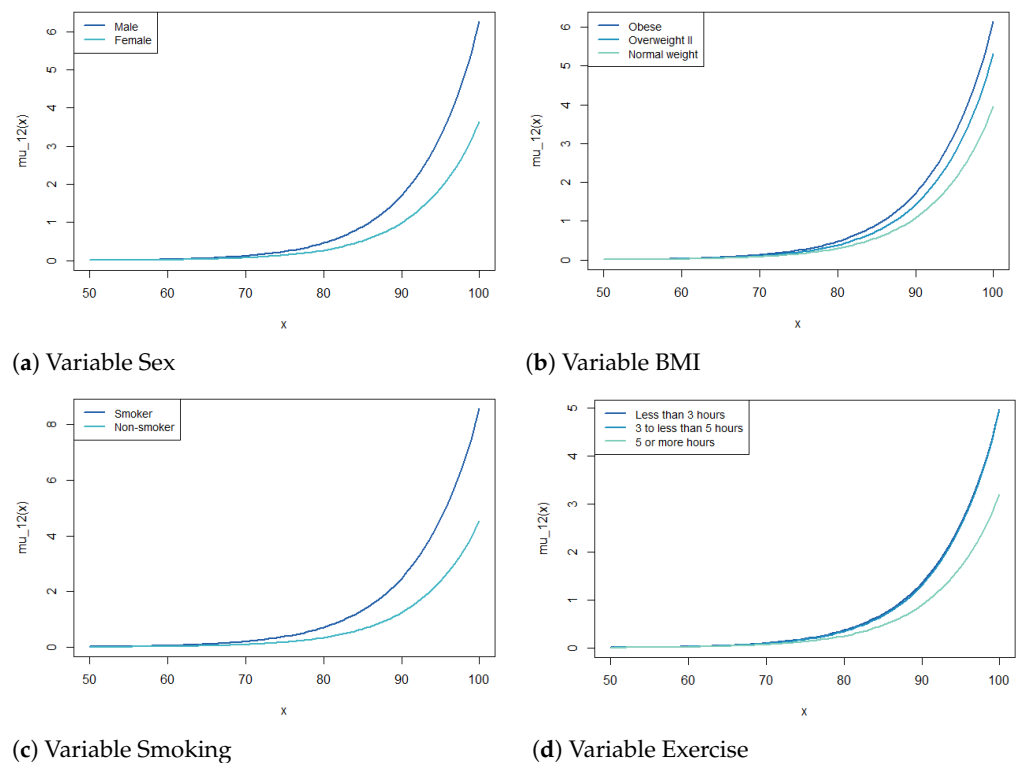


Figure 7. Functions for the transition intensities  $\mu_x^{12}$  differentiated by each risk factor.

**Table 9.** Summary of the NLS models for the function for the transition intensities  $\mu_x^{12}$  by risk factor.

Risk Factor	Level	Parameter	Estimate	Std. Error	t-Value	Pr(> t )
Sex	Male	$\alpha_{12}$	0.056355	0.000215	262.404	$<2 \times 10^{-16}$
		$\beta_{12}$	-4.838605	0.019698	-245.639	$<2 \times 10^{-16}$
		$\gamma_{12}$	-0.008172	0.001339	-6.104	$1.05 \times 10^{-9}$
	Female	$\alpha_{12}$	0.056489	0.000191	295.044	$<2 \times 10^{-16}$
		$\beta_{12}$	-5.087744	0.017594	-289.176	$<2 \times 10^{-16}$
		$\gamma_{12}$	-0.003818	0.000761	-5.017	$7.59 \times 10^{-7}$
BMI	Normal	$\alpha_{12}$	0.056213	0.000198	284.469	$<2 \times 10^{-16}$
		$\beta_{12}$	-5.023406	0.018084	-277.778	$<2 \times 10^{-16}$
		$\gamma_{12}$	-0.003538	0.000748	-4.727	$2.29 \times 10^{-6}$
	Overweight II	$\alpha_{12}$	0.057104	0.000319	179.036	$<2 \times 10^{-16}$
		$\beta_{12}$	-4.984908	0.029242	-170.473	$<2 \times 10^{-16}$
		$\gamma_{12}$	-0.003417	0.001541	-2.217	$2.66 \times 10^{-2}$
	Obese	$\alpha_{12}$	0.055158	0.000315	175.280	$<2 \times 10^{-16}$
		$\beta_{12}$	-4.727131	0.028842	-163.897	$<2 \times 10^{-16}$
		$\gamma_{12}$	-0.009723	0.001943	-5.003	$5.71 \times 10^{-7}$
Smoking	Non-Smoker	$\alpha_{12}$	0.056042	0.000201	279.176	$<2 \times 10^{-16}$
		$\beta_{12}$	-4.947743	0.018471	-267.87	$<2 \times 10^{-16}$
		$\gamma_{12}$	-0.006503	0.001062	-6.124	$9.2 \times 10^{-10}$
	Smoker	$\alpha_{12}$	0.053891	0.000926	58.197	$<2 \times 10^{-16}$
		$\beta_{12}$	-4.455835	0.084891	-52.489	$<2 \times 10^{-16}$
		$\gamma_{12}$	-0.009858	0.007114	-1.386	$1.66 \times 10^{-1}$
Exercise	2-3 h	$\alpha_{12}$	0.056392	0.000168	334.730	$<2 \times 10^{-16}$
		$\beta_{12}$	-4.943397	0.015399	-321.018	$<2 \times 10^{-16}$
		$\gamma_{12}$	-0.002389	0.000700	-3.414	$6.42 \times 10^{-4}$
	3-5 h	$\alpha_{12}$	0.057796	0.000283	204.53	$<2 \times 10^{-16}$
		$\beta_{12}$	-5.085398	0.025737	-197.595	$<2 \times 10^{-16}$
		$\gamma_{12}$	-0.001479	0.000920	-1.607	$1.08 \times 10^{-1}$
	>5 h	$\alpha_{12}$	0.055266	0.000255	216.956	$<2 \times 10^{-16}$
		$\beta_{12}$	-5.021899	0.023156	-216.873	$<2 \times 10^{-16}$
		$\gamma_{12}$	-0.001668	0.000585	-2.853	$4.34 \times 10^{-3}$

From Figure 7, we can observe that the transition intensity  $\mu_x^{12}$  is higher for Males than for Females, particularly at older ages. Regarding the BMI, individuals classified as “Obese” exhibit the highest intensities, followed by “Overweight II”, while those in the “Normal weight” group (which also includes “Underweight” and “Overweight I” individuals) display lower values. As expected, Smoker individuals present higher transition intensities than Non-Smokers, with differences increasing with age. For physical activity, the “Less than 3 h” group (which results from merging the “Less than 2 h” and “2 to less than 3 h” categories) shows higher transition intensities, while the intensity decreases with increasing activity levels, reaching the lowest values for individuals exercising “5 or more hours” per week.

Finally, by applying the same method to groups of individuals sharing a specified set of characteristics across multiple risk factors, it is possible to derive transition intensity functions as a function of age specific to each risk profile.

As an illustration, for transitions from the Autonomous state to the Mildly Disabled state, let us consider the two risk profiles defined in Table 10 as follows: a “Low risk” profile, corresponding to the combination of levels associated with lower transition intensities, and a “High risk” profile, corresponding to the combination of characteristics leading to higher intensities, obtained from the GLM coefficients in Table 7.

**Table 10.** Example risk profiles corresponding to low and high risk.

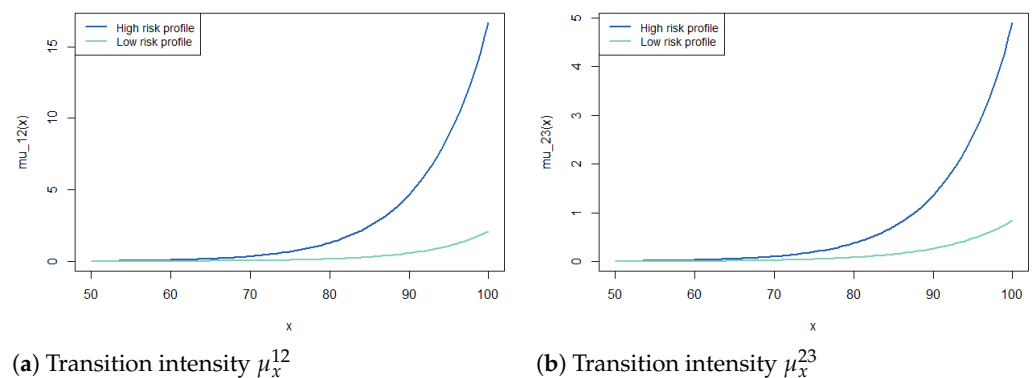
Risk Profile	Sex	BMI	Smoking	Exercise
Low risk	Female	Normal weight	Non-smoker	5 or more hours
High risk	Male	Obese	Smoker	Less than 3 h

As before, a NLS regression model was fitted to the predicted transition intensities (obtained from the GLM) for each group, where each group consists of individuals matching the characteristics of the respective risk profile. Thus, the following functions for the transition intensity  $\mu_x^{12}$  corresponding to the “Low risk” and “High risk” profiles were obtained:

$$\mu_{x, \text{Low risk}}^{12} = -0.002752 + 10^{0.056357x - 5.285772}, \quad x \geq 50, \tag{13}$$

$$\mu_{x, \text{High risk}}^{12} = -0.023962 + 10^{0.055257x - 4.303352}, \quad x \geq 50. \tag{14}$$

Figure 8 illustrates functions (13) and (14), as well as the corresponding functions for high- and low-risk profiles for the transition between Mildly Disabled and Severely Disabled. In both cases, transition intensities are consistently higher for individuals with a high-risk profile, as expected.



**Figure 8.** Functions for the transition intensities  $\mu_x^{12}$  and  $\mu_x^{23}$  by risk profile.

In summary, by incorporating multiple risk factors and stratifying individuals by risk profile, from the proposed methodology, we derive transition intensity functions  $\mu_x^{ij}$  that explicitly depend on age  $x$  and individuals’ characteristics. This framework enables the differentiation of transition intensities, and thus risk-adjusted transition probabilities, supporting the accurate calculation of premiums and reserves, within the context of disability and long-term care insurance, for general multi-state models. For LTC insurers, accurate risk premiums transform risk management from a reactive cost centre into a strategic driver of profitability and sustainability. By embedding advanced analytics into pricing, insurers can achieve more stable portfolios, improved customer segmentation, and long-term market growth, while fulfilling their core mission of providing sustainable LTC coverage.

### 4. Conclusions

Long-Term Care insurance has gained interest due to population ageing and the increasing prevalence of dependency among the elderly in developed countries. In Portugal, despite facing similar socio-economic challenges, the absence of private LTC insurance products is mainly due to the complexity of risk modelling and data scarcity enabling accurate pricing and reserving.

In this paper, we propose a practical application of estimation and risk-adjusted graduation methods for disability and LTC multi-state models, supported by a simulated dataset

based on Portuguese demographic characteristics. A GLM framework was employed to graduate transition intensities while incorporating individual risk profiles, enhancing the model's practical applicability for insurance purposes.

We first introduced the disability/LTC model designed for this study, featuring two degrees of disability/lack of autonomy, and the corresponding transition matrix, and we applied the transition intensity approach for parameter estimation. Under the assumption of constant intensities between consecutive ages, we derived the raw maximum likelihood estimates for the transition intensities. To incorporate the impact of risk factors, we implemented a risk-adjusted graduation technique based on GLMs, using overdispersed Poisson models with Gompertz–Makeham-type predictors.

Extending the approach in [Rensham and Haberman \(1995\)](#), in addition to age, our model incorporates individual risk factors such as sex, region, BMI, smoking habits, and exercise habits, to enhance heterogeneity modelling. In the absence of real portfolio data, we simulated a cohort of 100,000 individuals using Portuguese national statistics as the demographic foundation. Transition intensities were derived from the calibrated parameters in [Esquível et al. \(2021\)](#), with multiplicative adjustment factors defined to incorporate individual risk characteristics. This approach enabled the simulation of complete trajectories across all individuals of the population.

GLMs were fitted for each transition intensity, adjusting the covariate structures to ensure statistical significance. Variables such as age, sex, BMI, and smoking habits consistently proved significant, while region and exercise showed more limited relevance. GLM predictions enabled stratification by risk profiles, and NLS regressions were applied to derive smooth functions of transition intensities as a function of age.

Finally, by combining multiple risk factors, we derived age-dependent transition functions specific to chosen “Low risk” and “High risk” profiles, demonstrating a clear differentiation in expected transitions based on individual characteristics. Similar results may be obtained for any risk profile within the population/portfolio.

Key findings include the following: (1) higher  $\mu_x^{12}$  intensities for males, smokers, obese individuals, and those with lower physical activity; (2) physical activity correlated inversely with transition intensities; (3) region was not a significant factor for most of the transitions; and (4) higher risk profiles consistently exhibited higher transition intensities, particularly at advanced ages.

The methodology proposed in this paper provides a foundation for the development of disability and long-term care insurance products, enabling the construction of robust and risk-sensitive pricing structures tailored to the Portuguese market. While developing this generalizable approach, which allows for the construction of robust and risk-sensitive pricing structures, we specifically addressed Portugal's challenges in long-term care risk modelling (particularly data scarcity) by basing our simulations and model results on studies based on real-world Portuguese data, see [Esquível et al. \(2021\)](#), and official publicly available Portuguese population statistics, see [INE \(2020a, 2021a, 2021b, 2021c, 2024\)](#). This dual focus ensures both methodological innovation and contextually relevant outcomes for insurance product development.

#### *4.1. Policy Implications*

The accurate estimation of transition intensities is fundamental for ensuring actuarial fairness, financial stability, and regulatory compliance in insurance. For insurers, reliable multi-state models and risk-adjusted transition intensities allow for precise risk differentiation, thus preventing adverse selection, optimising premium pricing, and ensuring solvency under regulatory frameworks like Solvency II. In countries like Portugal, where ageing populations and evolving morbidity patterns—such as the increasing prevalence of

chronic diseases—shape longevity, disability, and long-term care risks, robust transition data are critical to design sustainable health-related insurance products.

At the macroeconomic level, these insights aid policymakers in anticipating healthcare demands and addressing inequalities in coverage, for instance, in cases where regional disparities in health service availability are a reality. For private insurers, leveraging real Portuguese data mitigates model risk and fosters innovation (e.g., dynamic pricing for long-term care). Conversely, underestimating transition intensities could compromise portfolios' risk measures, while overestimation might exclude vulnerable groups from affordable coverage. Thus, investing in granular and population-specific modelling is not just a technical imperative but a societal one, bridging actuarial science and public policy in an era of demographic transition.

#### 4.2. Limitations and Future Work

This research offers valuable insights for insurance regulators and social policymakers, particularly in addressing the growing challenges of disability and long-term care needs in ageing populations like Portugal's, while the proposed framework demonstrates strong theoretical foundations and practical potential, we acknowledge several limitations that point to important directions for future work.

The model's adaptability to different institutional contexts, although conceptually robust, requires empirical validation to realise its full potential. A critical next step involves testing the framework with real insurance portfolio data, which would serve two essential purposes as follows: first, to validate the simulated transition patterns and risk factor parameterisation against observed experience; second, to identify potential additional covariates that could enhance the model's predictive accuracy. Such empirical validation would naturally lead to refinements in both the GLM specifications and transition intensity functions, ensuring better alignment with Portugal's unique demographic and risk profiles.

Future research should also explore extensions of this work, including possible integrations with public health strategies and emerging care models, which could foster valuable synergies between actuarial science and policy development. From an implementation perspective, careful attention must be given to validating assumptions, testing calibration choices, and developing guidelines for adapting the framework to diverse subpopulations.

These research directions would significantly strengthen the model's reliability and practical relevance for the Portuguese market, while maintaining the methodological rigour established in the current study. By addressing these challenges, the framework can evolve into a robust tool for insurers and policymakers alike, supporting data-driven decisions in product design, premium setting, and resource allocation for Portugal's evolving social protection needs.

**Author Contributions:** G.R.G. participated in the Conceptualisation, Methodology, Investigation, and Writing—review and editing—and was responsible for Validation. B.A.C. was also responsible for Investigation, as well as for the Software Implementation, Formal Analysis, Data Simulation, and Writing—original draft preparation. M.L.E. participated in the Conceptualisation, Methodology, Investigation, and Writing—reviewing. G.R.G. was responsible for Funding Acquisition. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is funded by national funds through the FCT—Fundação para a Ciência e a Tecnologia, I.P., under the scope of the projects UIDB/00297/2020 (<https://doi.org/10.54499/UIDB/00297/2020>) and UIDP/00297/2020 (<https://doi.org/10.54499/UIDP/00297/2020>)—Center for Mathematics and Applications.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data supporting the conclusions of this article will be made available by the authors upon request. These data were derived from the following resources available in the public domain: for Health statistics, see [INE \(2020a, 2024\)](#) and for Census data see [INE \(2021a, 2021b, 2021c\)](#).

**Acknowledgments:** The authors are grateful to the anonymous referees for their valuable comments, questions, and suggestions on an earlier version of this paper.

**Conflicts of Interest:** The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analysis, or interpretation of the data; in the writing of the manuscript; or in the decision to publish the results.

## References

- Arrighi, Yves, Thomas Rapp, and Nicolas Sirven. 2017. The impact of economic conditions on the disablement process: A Markov transition approach using SHARE data. *Health Policy* 121: 778–85. [\[CrossRef\]](#)
- Baione, Fabio, and Susanna Levantesi. 2014. A health insurance pricing model based on prevalence rates: Application to critical illness insurance. *Insurance: Mathematics and Economics* 58: 174–84. [\[CrossRef\]](#)
- Barreira, Luís Filipe, Abel Paiva, Beatriz Araújo, and Maria Joana Campos. 2023. Challenges to Systems of Long-Term Care: Mapping of the Central Concepts from an Umbrella Review. *International Journal of Environmental Research and Public Health* 20: 1698. [\[CrossRef\]](#) [\[PubMed\]](#)
- Casasnovas, Guillem Lopez, and Catia Nicodemo. 2016. Transition and duration in disability: New evidence from administrative data. *Disability and Health Journal* 9: 26–36. [\[CrossRef\]](#)
- Cui, Xiaodong, Xuexiang Duan, Ching-Ter Chang, and Shuhua Jiang. 2022. Health Transition Probability and Long-Term Care Cost Estimation. *Mathematical Problems in Engineering* 2022: 7980111. [\[CrossRef\]](#)
- Czado, Claudia, and Florian Rudolph. 2002. Application of survival analysis methods to long-term care insurance. *Insurance: Mathematics and Economics* 31: 395–413. [\[CrossRef\]](#)
- Dickson, David C. M., Mary R. Hardy, and Howard Richard Waters. 2019. *Actuarial Mathematics for Life Contingent Risks*. Cambridge: Cambridge University Press.
- Esquível, Manuel Leote, Gracinda Rita Guerreiro, Matilde Castro de Oliveira, and Pedro Corte Real. 2021. Calibration of transition intensities for a multistate model: Application to long-term care. *Risks* 9: 37. [\[CrossRef\]](#)
- Esquível, Manuel Leote, Nadezhda Pavlovna Krasii, and Gracinda Rita Guerreiro. 2024. Estimation–calibration of continuous-time non-homogeneous Markov chains with finite state space. *Mathematics* 12: 668. [\[CrossRef\]](#)
- European Commission: Directorate-General for Employment, Social Affairs and Inclusion. 2021. Long-Term Care Report: Trends, Challenges and Opportunities in an Ageing Society. Volume II, Country Profiles. *Publications Office of the European Union*. Available online: <https://data.europa.eu/doi/10.2767/183997> (accessed on 22 July 2024).
- Fong, Joelle H., Adam Wenqiang Shao, and Michael Sherris. 2015. Multistate actuarial models of functional disability. *North American Actuarial Journal* 19: 41–59. [\[CrossRef\]](#)
- Forfar, David O., John J. McCutcheon, and A. David Wilkie. 1988. On graduation by mathematical formula. *Journal of the Institute of Actuaries* 115: 1–149. [\[CrossRef\]](#)
- Fuino, Michel, and Joël Wagner. 2018. Long-term care models and dependence probability tables by acuity level: New empirical evidence from Switzerland. *Insurance: Mathematics and Economics* 81: 51–70. [\[CrossRef\]](#)
- Gompertz, Benjamin. 1825. 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. *Philosophical transactions of the Royal Society of London* 115: 513–83.
- Guo, Tian, Yunwei Zhang, Gang Xu, Wenxian Liu, Hansheng Ding, and Shaofan Chen. 2024. Activities of Daily Living Disability Transition Patterns in Older Adults with Chronic Diseases: A Four-Year Cohort Study in China. *Healthcare* 12: 2088. [\[CrossRef\]](#) [\[PubMed\]](#)
- Haberman, Steven, and Ermanno Pitacco. 2018. *Actuarial Models for Disability Insurance*. London: Routledge.
- Instituto Nacional de Estatística. 2020a. *Inquérito Nacional de Saúde: 2019*. Lisboa: INE. Available online: [https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine\\_destaques&DESTAQUESdest\\_boui=414434213&DESTAQUESmodo=2](https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_boui=414434213&DESTAQUESmodo=2) (accessed on 22 July 2024).
- Instituto Nacional de Estatística. 2020b. *Projeções de População Residente em Portugal*. Lisboa: INE. Available online: [https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine\\_destaques&DESTAQUESdest\\_boui=406534255&DESTAQUESmodo=2](https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_boui=406534255&DESTAQUESmodo=2) (accessed on 29 August 2024).
- Instituto Nacional de Estatística. 2021a. *População Residente do Sexo Feminino Segundo os Censos: Total e por Grupo Etário*. Lisboa: PORDATA. Available online: <https://www.pordata.pt/pt/estatisticas/populacao/populacao-residente/populacao-residente-por-sexo-e-grupo-etario> (accessed on 22 July 2024).

- Instituto Nacional de Estatística. 2021b. *População Residente do Sexo Masculino Segundo os Censos: Total e por Grupo Etário*. Lisboa: PORDATA. Available online: <https://www.pordata.pt/pt/estatisticas/populacao/populacao-residente/populacao-residente-por-sexo-e-grupo-etario> (accessed on 22 July 2024).
- Instituto Nacional de Estatística. 2021c. *População Residente Segundo os Censos: Total e por Dimensão dos Lugares*. Lisboa: INE. Available online: [https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine\\_qa\\_populacao](https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_qa_populacao) (accessed on 22 July 2024).
- Instituto Nacional de Estatística. 2023. *Estimativas de População Residente: 2023*. Lisboa: INE. Available online: [https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine\\_destaques&DESTAQUESdest\\_boui=645507713&DESTAQUESmodo=2&xlang=pt](https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_boui=645507713&DESTAQUESmodo=2&xlang=pt) (accessed on 29 August 2024).
- Instituto Nacional de Estatística. 2024. *Estatísticas da Saúde: 2022*. Lisboa: INE. Available online: <https://www.ine.pt/xurl/pub/439489924> (accessed on 22 July 2024).
- Kabuche, Doreen, Michael Sherris, Andrés Mauricio Villegas, and Jonathan Ziveyi. 2024. Pooling functional disability and mortality in long-term care insurance and care annuities: A matrix approach for multi-state pools. *Insurance: Mathematics and Economics* 116: 165–88. [CrossRef]
- Lee, Seok-Hwan, Chon Yongho, and Yun-Young Kim. 2023. Comparative Analysis of Long-Term Care in OECD Countries: Focusing on Long-Term Care Financing Type. *Healthcare* 11: 206. [CrossRef]
- Leung, Edward. 2006. A Multiple State Model for Pricing and Reserving Private Long Term Care Insurance Contracts in Australia. *Australian Actuarial Journal* 12: 187–247.
- Lopes, Hugo, Céu Mateus, and Nicoletta Rosati. 2020. Identifying the long-term care beneficiaries: Differences between risk factors of nursing homes and community-based services admissions. *Aging Clinical and Experimental Research* 32: 2099–2110. [CrossRef]
- Lopes, Hugo, Gracinda Rita Guerreiro, Manuel Leote Esquível, and Céu Mateus. 2021. Identifying the Main Predictors of Length of Care in Social Care in Portugal. *Portuguese Journal of Public Health* 39: 21–35. [CrossRef]
- Makeham, William Matthew. 1860. On the law of mortality and the construction of annuity tables. *Journal of the Institute of Actuaries* 8: 301–10. [CrossRef]
- Oliveira, Matilde Castro de. 2017. *Calibração e Simulação de um Modelo de Cadeias de Markov para um seguro Long-Term Care*. Master's dissertation, NOVA FCT, Universidade Nova de Lisboa, Lisbon, Portugal.
- Oliveira, Matilde Castro de, Manuel Leote Esquível, Susana Nascimento, Hugo Ramalheira Lopes, and Gracinda Rita Guerreiro. 2017. Estimation of Markov transition probabilities via clustering—An application to Long Term Care. Paper presented at Symposium on Big Data in Finance, Retail and Commerce, Lisbon, Portugal, November 2–3, pp. 85–90.
- Organization for Economic Co-Operation and Development. 2023. *Health at a Glance 2023: OECD Indicators*. Paris: OECD Publishing.
- Pitacco, Ermanno. 2014. Health insurance. In *Basic Actuarial Models*. New York: Springer.
- Pritchard, Delme John. 2013. Modeling disability in long-term care insurance. *North American Actuarial Journal* 10: 48–75. [CrossRef]
- Renshaw, Arthur Edward. 1991. Actuarial graduation practice and generalised linear and non-linear models. *Journal of the Institute of Actuaries* 118: 295–312. [CrossRef]
- Renshaw, Arthur Edward. 1994. A comparison between the mortality of non-smoking and smoking assured lives in the United Kingdom. *Journal of the Institute of Actuaries* 121: 561–71. [CrossRef]
- Renshaw, Arthur Edward, and Steven Haberman. 1995. On the graduations associated with a multiple state model for permanent health insurance. *Insurance: Mathematics and Economics* 17: 1–17. [CrossRef]
- Ross, Sheldon Michael. 1996. *Stochastic Processes*. Hoboken: John Wiley & Sons.
- Waters, Howard Richard. 1984. An approach to the study of multiple state models. *Journal of the Institute of Actuaries* 111: 363–74. [CrossRef]
- Wolthuis, Henk. 2003. *Life Insurance Mathematics (The Markovian Model)*. Amsterdam: Universiteit van Amsterdam.
- World Health Organization. 2009. *Global Health Risks: Mortality and Burden of Disease Attributable to Selected Major Risks*. Geneva: WHO.
- World Health Organization. 2024. *World Health Statistics 2024: Monitoring Health for the SDGs, Sustainable Development Goals*. Geneva: WHO.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.