The Estimation of Adopted Mortality and Morbidity Rates Using Markov Model and the Phase Type Law: the Turkish Case

Fulya Akat[a], A.Sevtap Selcuk-Kestel[a] and Fatih Tank[b,*]

[a]Middle East Technical University, Institute of Applied Mathematics, 06531 Ankara, Turkey; [b]Ankara University, Faculty of Applied Sciences, Department of Insurance and Actuarial Sciences, Ankara - Turkey

ARTICLE HISTORY
Compiled May 9, 2018

ABSTRACT
This paper aims to estimate mortality rate, morbidity-mortality rates of a chronic disease utilizing phase type law in the frame of two and three state Markov processes. The application on commonly used mortality tables in Turkey are adopted to Markov process to estimate the future mortalities with respect to phase type distribution for the purpose of justifying. Using one absorbing state, two and three state Markov Models calculate the time until absorbing of the death and death by phase type distribution for each gender. Consequently, the 3-state probabilities in estimating the mortality-morbidity rates of IHD for Turkish population yield a significant information on the health management and pricing health insurance products.

KEYWORDS
Phase type distributions; mortality; morbidity; two state markov process; three state markov process; IHD

1. Introduction

Improvement of life standards and health conditions have caused a change in life expectancy. Mortality measures do not always provide sufficient information about the population, as the collection of the complete data may take over 100 years. Because of these reasons, mortality and morbidity estimations, especially that of chronic diseases and disabilities has become more important [13]. Addition to this, morbidity and mortality rates have important effect on actuarial valuations. Its estimation and its efficiency are significant in financial decisions. Precise knowledge in morbidity and mortality rates are also important for Social Security Institutions (SSI) and insurance companies. The estimation of morbidity and mortality supply guiding principle to decision makers for the calculation of premiums and the estimation of reserves [18]. As the estimation of the mortality and morbidity rates require mostly a long term collection of data, many tables in the literature are the results of estimation techniques. The well-known models in estimating the mortality rates are DeMoivre, Gompertz, Makeham and Weibull, which depends on historical realizations of populations [6]. Additionally, recent developments, such as Lee-Carter, the mortality rate using a self-generating set of iterations, can be used when the past information related a population

*Corresponding author email: tank@ankara.edu.tr (F.Tank)

does not fit to the models above.

In the morbidity aspect, incidence and the prevalence are the important indicators where both require individual base observations over long terms. Incidence is a measure of the probability of the occurrence of a given medical condition in a population within a specified period of time. Whereas, prevalence represents the proportion of a population found to have a condition of a specific disease in morbidity studies. Kaplan-Meier is one of the well-known methods in the literature in the estimation of the morbidity rates [13]. A novel approach to estimate mortality and morbidity rates proposed by Lin et al. (2007) and Kay (1986) is Markov model and phase type distribution. This method implements the stochastic structure into the estimation of the rates, as well as, to calculate the adopted probabilities which refer to adaptation of mortality rates to markov process. Additionally it allows introducing new states,such as, causes of the death or illnesses.

As it is well known, as a critical illness, IHD is accounted as the most commonly observed, deathly illness compared to other diseases. Contrary to the mortality, morbidity of a certain disease, such as Ischemic heart disease (IHD) allows more states, such as Healthy-Disease, Disease-Disease, Disease-Death. Therefore, the likelihood of these states are important indicators on the impact of the disease. IHD is selected as the case study due to its remarkable fatality rate, especially, in the western countries [2]. The statistics indicate that 20% of the causes of the total death come from IHD [8].

There exists many studies in literature on morbidity estimation ([22] [11]). Heligman and Pollard propose a mathematical expression for the age pattern of mortality tables by applying to Australian National Mortality data. This model allows mortality comparisons regarding age and sex both among countries and within the same country over time [11]. These studies are also used in this study to develop the methodology on estimation of adopted mortality and morbidity rates. However, to our best knowledge in the recent literature, markov process considering mortality and morbidity of IHD using phase type distribution has not been studied.

Phase type distributions which take into account the duration between states as markov jump process is commonly used in health sciences to embed the influence of the absorbing state of the disease. The advantage in phrasing the transition from one state to another with random duration gains importance especially in observations which do not have long history. The implementation of the phase type distribution to mortality from a certain chronic disease or natural causes is the main motivation in this paper.

Phase type distribution is introduced as the distribution of absorbing times of the Markov process in the earlier studies [4]. Ellis (1978) gives some definitions and properties of two states of Markov process [9]. Wand and Pham (2011) define neutral ageing with degeneration of normal ageing [26]. Winkler introduces a theory of continuous Markov process [27]. Collins and Huzurbazar [7] review multi-state models, focusing on time-homogeneous semi-Markov processes. Hubbard et al. [14] propose models in which the time scale of a non-homogeneous Markov process is transformed to an operational time scale on which the process is homogeneous. Phase type distribution is used in the risk theory by Blandt at which basic definition of the distribution is given and phase type renewal theory is introduced [4]. Faddy uses this distribution to apply to some industrial data for the purpose of determining the failure times [10]. McClean and Millard use this model to determine the relationship between hospital and community cares for the older patients [21]. Xie et al. use continuous time Markov model to analyse transition within and between residential and nursing home cares.

A procedure is introduced to determine the structure of the model and estimating parameters by maximum likelihood [28]. Steinsaltz et al. (2004) illustrate the theory and application of Markov models of aging [23]. Keyfitz and Rogers (1982) show the standard life contingency formulas with matrix analogues [[19]]. Jackson et al. (2003) use multistate Markov model for disease progression with classification error [16]. Hoogenveen et al. (2010) introduce Markov-type multi-state model for deal with effects of changes in risk factors for chronic diseases on morbidity and mortality [12]. Lin and Liu (2007) use Markov process and phase type distribution in the mortality analyses which aims to derive the estimates of model parameters does not account any disease related death [20].

The aim of this study is to implement phase type distribution to estimate adopted mortality rates for existing and currently used mortality tables and three-state morbidity rates for a selected disease. To achieve this aim, an application is performed on Turkish mortality tables, and IHD recordings based on Turkish SSI data base.

This paper illustrates the impact of implementing the methodology proposed in the improvement of mortality and morbidity rates [1]. Phase type distribution which is mostly applied to disease rates, is applied to Turkish population data first time in the literature.

The remainder of the paper is organized as follows: Section 2 introduces the methodology of Markov process and phase type distribution. In Section 3, estimation of mortality rates using two-state markov process is introduced on Turkish Mortality tables. In Section 4, morbidity and mortality estimation of Ischemic heart disease (IHD) are analysed by three-states Markov process, time until exposure to the disease and time until death, that are given with phase type distribution. Finally, the discussion and conclusion are provided in the last section.

2. Preliminaries: Markov Process and Phase Type Distribution

Markov process is a stochastic process that satisfies the Markov property, which is used in many fields for different purposes, such as finance, demography, insurance, industrial sectors and etc. [10, 17, 19]. In literature, there are continuous and discrete, time-homogeneous and non-homogenous, finite state and infinite state versions and combinations of Markov process [13, 24, 27].

Definition 2.1. *Given* $\{X_n : n \geq 0\}$ *with a state space* $S \in \{0, 1, 2, .., k\}$ *is called a discrete-time Markov chain if and only if it has the Markov property as*

$$p_{ij} = \mathbb{P}\{X_{s+t} = j | X_s = i, X_u = x_u, 0 \leq u < s\} = \mathbb{P}\{X_{s+t} = j | X_s = i\} \tag{2.1}$$

The Markov property given above describes the past and future states which are conditionally independent and the time of staying at state i before moving to state j follows an exponential distribution with parameter λ_{ij}, which is denoted as transition rate.

Each transition probability, p_{ij}, represented in matrix form $P(t)$ can be expressed in the form of its transition (hazard) rate, λ_{ij}, matrix $Q(t)$ $(t = 0, 1, 2, ...)$ using Kolmogorov differential equations [17, 20] as follows:

$$P(t) = \exp(Q(t)), t = 0, 1, 2, ... \tag{2.2}$$

Definition 2.2. ***Phase Type Distribution***
Let $\{X(t)\}_{t=0,1,2,...}$ be a Markov jump process on the finite state space $E = \{1, 2, .., k\}$ where the states $1, 2, .., k-1$ are the transient states and state k is the absorbing state. Then $\{X(t)\}_{t=0,1,2,...}$ has an intensity matrix of the form

$$Q = \begin{pmatrix} \mathbf{T} & \mathbf{t} \\ \mathbf{0} & 0 \end{pmatrix} \tag{2.3}$$

$\boldsymbol{T}$ is $(k-1) \times (k-1)$ dimensional matrix, $\boldsymbol{t} = -\boldsymbol{T}e$ where $e' = (1 \quad 1 \dots 1)$, and $\boldsymbol{0}$ is $k-1$ dimensional row vector of zeros. Let $\alpha = (\alpha_1 \quad \alpha_2 \dots \alpha_k)$ denote the initial distribution of $(X(t))_{t\geq 0}$ over the transient states only where $\alpha_i = \mathbb{P}(X_0 = i)$, $\mathbb{P}(X_0 = k) = 0$ [4, 5, 10].

Definition 2.3. *The time until death, τ , defined as*

$$\tau = \inf\{t \geq 0 | X_t = k\} \tag{2.4}$$

is said to have phase type distribution and denoted as

$$\tau \sim PH(\alpha, \boldsymbol{T}) \tag{2.5}$$

The set of parameters $(\alpha, \mathbf{T})$ is said to be a representation of the phase type distribution [20]. The $(\alpha, k-1)$ are the dimension of the phase type distribution. Initial distribution, α, determines the probability mass function on the states [23].

Some special case, such as consideration of mortality require implementing the random factor, ϵ, having impact on the state probability. To cope with this random effect, we modify the phase type distribution with respect to ϵ.

Theorem 2.1. *Let $f(.)$ be the density function, s be the age, α be the initial distribution. If $\boldsymbol{T}$ and $\boldsymbol{t}$ come from the intensity matrix then,*

$$f(s) = \alpha \exp(\boldsymbol{T}s)\boldsymbol{t} \tag{2.6}$$

Theorem 2.2. *Let $S(s)$ be the survival function, then*

$$\begin{aligned} S(s) &= 1 - F(s) \\ &= \alpha \exp(\boldsymbol{T}s)_{ij} e \end{aligned} \tag{2.7}$$

and

$$\lambda(s) = \frac{\alpha \exp(\boldsymbol{T}s)\boldsymbol{t}}{\alpha \exp(\boldsymbol{T}s)e} \tag{2.8}$$

As markov models has the assumption on constant transition rates with respect to time, the age specific transition intensities related to mortality rates are depicted by including random component (ϵ) to the transition rates.

2.1. *Two State Markov Process*

Two state Markov process has one transition with two states of which the last one represents absorbing state. $P(t)$ can be written as [17]

$$P(t) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \exp(-\lambda t) & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

Thus,

$$P(t) = \begin{pmatrix} \exp(-\lambda t) & 1 - \exp(-\lambda t) \\ 0 & 1 \end{pmatrix}$$

Rewriting the transition matrix and the transition rate matrix with phase type formulation given in Equation (2.2), matrix turns out to be

$$P(s) = \exp(Qs) = \begin{pmatrix} \exp(-\mathbf{T}s) & \mathbf{e} - \exp(-\mathbf{T}s)\mathbf{e} \\ 0 & 1 \end{pmatrix}$$

The probability of being state 1 at time t is $p_t = \exp(-\lambda t)$ which is equivalent to the exponential survival function $S(t) = \exp(-\lambda t)$.

2.2. *Three-State Markov Process*

In the three-state Markov process the transition is only in one direction and the last state represents the absorbing state.

By the spectral decomposition, similar to the two-state case, the three state $P(t)$ matrix is expressed as follows:

$$P(t) = \begin{pmatrix} \exp(-(\lambda_{12} + \lambda_{13})t) & \exp(-\lambda_{12}t) & 1 - (1 + \exp(-\lambda_{13}t))\exp(-\lambda_{12}t) \\ 0 & \exp(-\lambda_{23}t) & 1 - \exp(-\lambda_{23}t) \\ 0 & 0 & 1 \end{pmatrix} \tag{2.9}$$

Rewriting the transition matrix and the transition rate matrix with phase type notation, $P(t) = \exp(Q(t)), t = 0, 1, 2, ...$ yields

$$Q = \begin{pmatrix} \mathbf{T} & \mathbf{t} \\ \mathbf{0} & 0 \end{pmatrix} \tag{2.10}$$

$$\mathbf{T} = \begin{pmatrix} -(\lambda_{12} + \lambda_{13}) & \lambda_{12} \\ 0 & -(\lambda_{23}) \end{pmatrix} \tag{2.11}$$

$$\mathbf{t} = \begin{pmatrix} \lambda_{13} \\ \lambda_{23} \end{pmatrix}. \tag{2.12}$$

Here, the probability of being state 1 and 2 at time t can be found as $p_t = \exp(-\lambda \mathbf{t})$.

3. Mortality Estimation: Turkish case

The improvement of medical technology and life standards cause higher expected life time and lower mortality rates leading to increase in population at older ages. Turkey, like other countries, is also exposed to low fertility and longer life time, even though its high rate in younger ages. Mortality rates are the key components to make inferences and predictions on future estimates of some demographic indicators such as population.

The population census official results, which was started in 1927 and then switched to Address Based Population Registration System (ABPRS) in 2008, are summarized in Table 1. The refractions in 1938 and 2007 are significant as the first date refers to the inclusion of Hatay city in the borders of Turkish Republic and the latter corresponds to changed to ABPRS.

Table 1. Turkish Population Data Between 1927 to 2013 [2]

	Year	Population	Male	Female
Census	1927	13,648,270	6,563,879	7,084,391
	1935	16,158,018	7,936,770	8,221,248
	1940	17,820,950	8,898,912	8,922,038
	1945	18,790,174	9,446,580	9,343,594
	1950	20,947,188	10,527,085	10,420,103
	1955	24,064,763	12,233,421	11,831,342
	1960	27,754,820	14,163,888	13,590,932
	1965	31,391,421	15,996,964	15,394,457
	1970	35,605,176	18,006,986	17,598,190
	1975	40,347,719	20,744,730	19,602,989
	1980	44,736,957	22,695,362	22,041,595
	1985	50,664,458	25,671,975	24,992,483
	1990	56,473,035	28,607,047	27,865,988
	2000	67,803,927	34,346,735	33,457,192
ABPRS	2007	70,586,256	35,376,533	35,209,723
	2008	71,517,100	35,901,154	35,615,946
	2009	72,561,312	36,462,470	36,098,842
	2010	73,722,988	37,043,182	36,679,806
	2011	74,724,269	37,532,954	37,191,315
	2012	75,627,384	37,956,168	37,671,216
	2013	76,667,864	38,473,360	38,194,504

CSO 1980 Mortality Table was the most commonly used table till 2010 in insurance sector and social security institutions. Due to regulatory requirements, sector started using custom tailored Turkish Mortality Tables, TRSH2010, TRH2010 and SGK2008. These tables represent insured mortality rates, total population rates and social security beneficiary mortality rates. The tabulation of these tables is out of scope of this paper as it requires dealing with different type of data set and methodology to evaluate these rates [25]. Although the completion of these tables are important progress the effective use of these tables is still an issue in pension and life insurance valuations. The contradicting points are discrepancy in the age range (Table 2). We find it is vital to illustrate the influence of different mortality rates on the framework proposed in

this paper.

Table 2. The Age Range in Turkish Mortality Tables

	CSO 1980	TRSH 2010	TRH 2010	SGK 2008
Age Range	0-99	0-110	0-99	14-110

For these reasons we purpose employing phase type distribution for the mortality estimation from natural and IHD caused death. This approach is regarded as an alternative valuation of Turkish mortality rates.

3.1. *Mortality Estimation with two-state Markov Process*

We estimate the duration time until death and mortality rates using the existing mortality tables (CSO 1980, TRSH 2010, TRH 2010 and SGK 2008) by using phase type distribution. To achieve this, primarily, Turkish population data and selected tables are re-arranged in two states: alive and death where the state 'dead' is taken as the absorbing state. Taking p_x, probability of survival, as reference points, $P(t)$ and $Q(t)$ matrices are built to calculate the survival distribution, $S(x)$, and the λ_x for each age, x.

For any specific age, x, the transition probabilities, p_{12}, indicate the probability of moving from alive to dead state, whereas, the transition rates, λ_{12}, representing the hazard rate of changing the states. As here P matrix is stochastic, $p_{12} = 1 - p_{11}$, and the absorbing state gets $p_{21} = 0$ and $p_{22} = 1$. Because of the Markov property, transition rates become $\lambda_{11} = -\lambda_{12}$ and yielding an absorbing state, $\lambda_{21} = \lambda_{22} = 0$.

Table 3. Mortality and hazard rates with phase type distribution

Probability of Death, p_{12}					
Gender	Age	CSO 1980	TRSH 2010	TRH 2010	SGK 2008
Male	0	0.00370	0.01953	0.01953	-
	25	0.00108	0.00075	0.00092	0.00066
	45	0.00319	0.00278	0.00302	0.00268
	58	0.01102	0.01015	0.01205	0.01077
	65	0.02152	0.01929	0.02407	0.02146
Female	0	0.00245	0.00816	0.00816	-
	25	0.00053	0.00021	0.00028	0.00020
	45	0.00237	0.00080	0.00157	0.00119
	58	0.00635	0.00382	0.00588	0.00421
	65	0.01145	0.00868	0.013220	0.00868
Transition rates, λ_{12}					
Male	0	0.00371	0.01992	0.01992	-
	25	0.00108	0.00075	0.00092	0.00066
	45	0.00320	0.00279	0.00303	0.00268
	58	0.01114	0.01025	0.01220	0.01088
	65	0.02199	0.01967	0.02466	0.02193
Female	0	0.00246	0.00823	0.00823	-
	25	0.00053	0.00021	0.00028	0.00011
	45	0.00238	0.00080	0.00158	0.00119
	58	0.00639	0.00384	0.00591	0.00423
	65	0.01158	0.00875	0.01340	0.00875

The estimated mortality values, given in Table 3, based on phase type distribution are plotted for each table, classified with respect to gender together with the original mortality rates for comparison (Figure 1). It is observed firstly that the phase type distribution agrees with the mortality pattern showing consistency in pattern.

However, it shows deviations for the geriatric ages which is realistic concerning. To determine the efficiency of the approach, mean square error deviations (MSE) are calculated (Table 4) for mortality rate, q_x, and transition rate, λ_x which presentslower variation especially for CSO 1980. TRSH 2010 estimates yield the highest MSE among the others which agrees with mortality rates which are also higher compared to others.

Table 4. Mean Square Errors (MSE) for the adopted Mortality Tables

	q_x		λ_x	
MSE	Male	Female	Male	Female
CSO 1980	0.00066	0.00061	0.00076	0.00061
TRSH 2010	0.00238	0.00238	0.01206	0.01098
TRH 2010	0.00081	0.00090	0.01149	0.01960
SGK2008	0.00082	0.00028	0.01569	0.00215

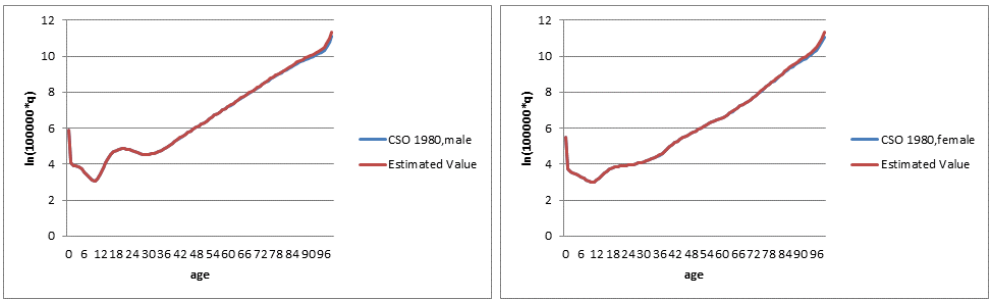
4. Morbidity and Mortality Estimation of IHD

As a chronic type, ischemic heart disease (IHD) causes more deaths and disabilities compared to other diseases in the world. In light of the projection of large increases in IHD throughout the world, IHD is likely to become the most common cause of death worldwide by 2020 [8]. There is no recovery in this disease, and most of the time surgery is a necessary implementation. People even need a medical supervision and treatment annually.

Measuring the individual influence of IHD on state of health and consecutively on economics is important. In case of incomplete historical occurrences, determination of prevalence and incidence of the disease is troublesome. We propose to estimate IHD rates with respect to the states it may require. Two state studies the probabilities and transition rates to change from IHD state to death state, whereas three state case analyzes the same indicators for the states of health, IHD and death.

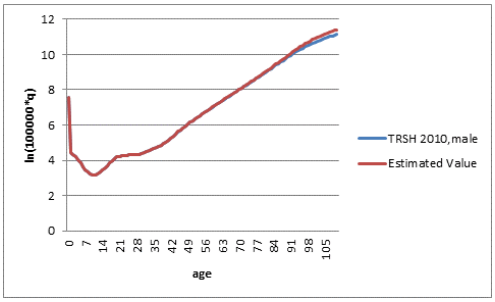
4.1. *Data and Markov-State Framework*

The proposed study is implemented for two- and three-state Markov processes based on the data collected from Turkish SSI for the years 2007-2009. The data contains around 25 million entries representing an insured who is registered to visit an health institute (hospitals, health care centers, clinics) in Turkey. Each entry (insured) contains the age, gender, the date of registration to health institute, the place of birth, location of the health institute, code of the diagnose (ICD 10) based on the uniquely encrypted ID of the insured. It should be noted that, in Turkey, around 90% of the population is covered under social health insurance, as the coverage for one registered worker covers also the belongings and the family of the insured. Among those registrations in the data set around 5 million of them are found to be related to IHD which constitutes the data set employed in the proposed methodology. The comorbidity to other diseases are neglected. The data after 2009, can not be reached due to changes on the data sharing policy of SSI. However, due to the special characteristics of the disease, the time impact is assumed to insignificant on the change of the state of the disease. Because, the disease can not be reversible (chronic), has to be regularly monitored, and disappearance of an IHD patient can either be from leaving the country or death.

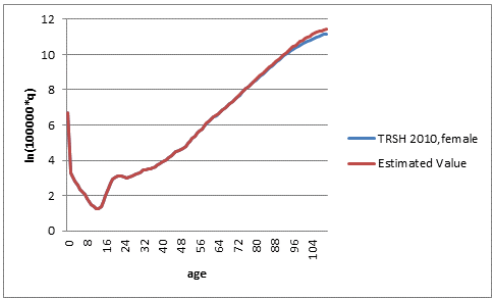


(a) CSO 1980,Male

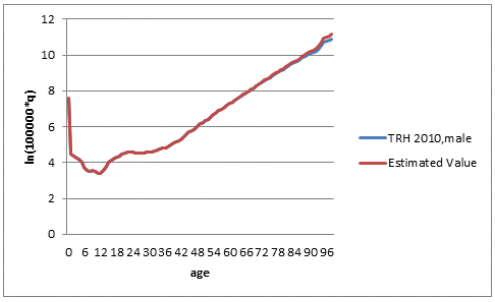
(b) CSO 1980,Female



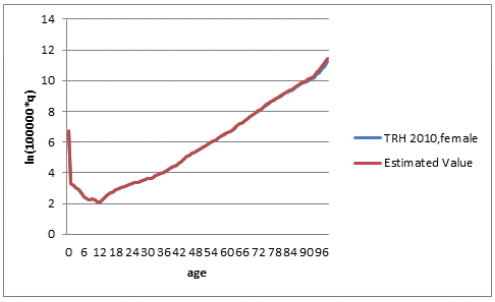
(c) TRSH 2010,Male



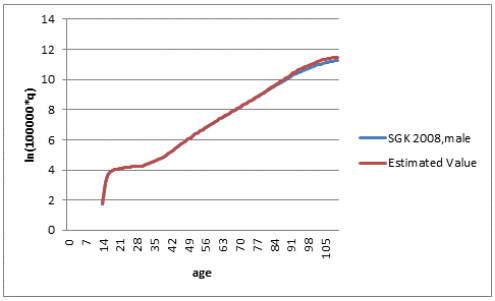
(d) TRSH 2010,Female



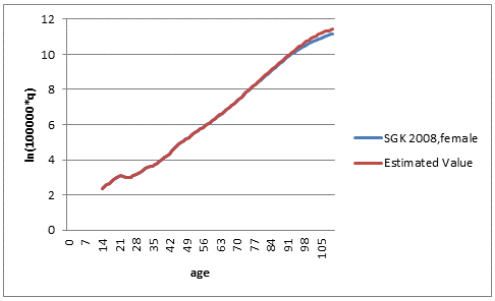
(e) TRH 2010,Male



(f) TRH 2010,Female



(g) SGK 2008,Male



(h) SGK 2008,Female

Figure 1. The comparison of Phase Type Mortality rates (Estimated value) based on selected tables

For this reason, the existing collection of the data is re-structured with respect to

(i) the application to the health institute first time in the year considered
(ii) the frequency of the repeated visits in the same year
(iii) the visit incurred in the consecutive years.

By the help of uniquely encrypted ID numbers, the registration to every health institute visit in the data set is coded with respect to the appearance of the insured in three consecutive years between 2007-2009 to construct a Markov state framework. Among nearly 5 million registrations, of which 25% are female. The age distribution of the IHD patients for each gender is illustrated in Figure 2 for the year 2008. It is noticed that the similar pattern is observed for 2007 and 2009. The highest incidence rate for both genders is between ages 55-65.

Data processing: Each insured is categorized by binary code (1- if IHD diagnosed) if he/she appeared in the registration list for the specific year concerned. The state of disease is determined according to the binary codes in three consecutive years. For instance, 010 code represents that patients appeared at the health institute only in 2008, but not in the years 2007 and 2009. It is known that, a patient suffering from IHD is required to make regular visits to the hospital due to the nature of chronic disease. Therefore, the code 111 refers to the patient appeared in the system for each year taken into account. Additionally, based on the expert opinion received from specialist, 40% of the insured coded as 101 in the data set is assumed either they refuse the treatment at hospitals in 2008 or they might be treated abroad.

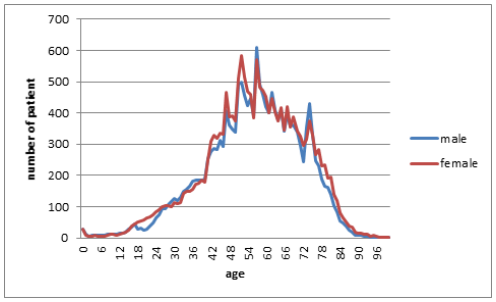
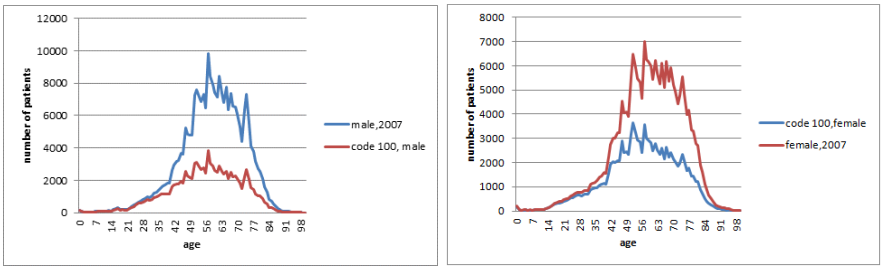


Figure 2. Age distribution of IHD for both Genders in 2008

Then, we count the number of insureds having IHD in 2007 (codes 100, 110, 101 and 111) and in 2008 (codes 010, 110, 011 and 111). Figure 3 shows the number of code 100 fr both genders in year 2007 for illustration.



(a) Male (b) Female

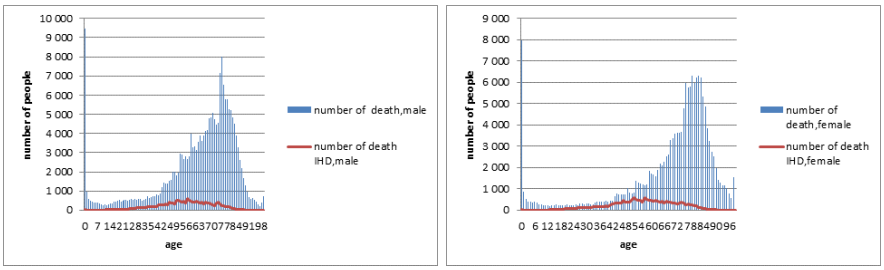
Figure 3. The distribution of the patients with the Code 100 compared to the overall registrations by age

Disappearance in female patients seems to be much higher than male patients and the rate of continuing the treatment is higher in middle ages compared to infants and older ages.

Nevertheless, IHD is known to be in first three lethal diseases (heart diseases, cancer and accident) [8]. Due to its characteristics, expert's opinion and the high percent of population covered by SSI, it is assumed the patients whose registrations disappear in

2009 are regarded as 'died' in the coding.

According to this assumption, the total number of death from IHD in 2009 are given in Figure 4. This figure also illustrates a comparative state of IHD caused-deaths with respect to the number of deaths in 2009 [3].



(a) Male (b) Female

Figure 4. The number of IHD caused deaths with respect to the total number of deaths in 2009.

With respect to proposed framework, we determine two- and three-state Markov process and calculate mortality, morbidity and hazard rates based on the phase type law for each age. For space limitations, the results of proposed framework are presented for the ages, $0, 25, 45, 58$, and 65 to represent infant, young, middle and two alternative retirement ages for Turkey, respectively.

4.2. *Two-State Mortality Estimation for IHD*

The studies in literature state that IHD has %40 of death rate [8]. Two state Markov process is employed at which the first column and row represent having IHD and the second column and row represent death from this illness. Transition matrix, then becomes

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$$

where $p_{ij}, i, j = 1, 2$, defines

p_{11}=$\mathbb{P}$(having IHD in year $t+1$ | having IHD in year t)

p_{12}=$\mathbb{P}$(death in year $t+1$ | having IHD in year t)

p_{21}=$\mathbb{P}$(having IHD in year $t+1$ | death in year t)=0

p_{22}=$\mathbb{P}$(death in year $t+1$ | death in year t)=1

with the condition $\sum p_{ij} = 1$.

Therefore, the probability of being alive with IHD at time t is determined by $p_t = \exp(-\lambda t)$. This approach leads to the probability of death because of IHD given that the patient is already diagnosed, p_{12}, and the transition rates, λ_{12}, under the same conditions to be calculated. Table 5 illustrates these results for the ages selected in gender details. The morbidity rates and transition rates are found to be realistic, high in infant and young ages.

4.3. *Three-State Morbidity and Mortality Estimation for IHD*

This framework require the augmentation of the total population into two parts, a part having IHD and the other not having IHD. Based on the IHD data the partitioning of the total population is done with respect to the mortality tables taken into

Table 5. Two-state Mortality estimation for IHD for some selected ages

	p_{12}		λ_{12}	
Age	Male	Female	Male	Female
0	0.15720	0.15414	0.18651	0.18222
25	0.13829	0.13498	0.16048	0.15604
45	0.08427	0.10350	0.09203	0.11545
58	0.05847	0.07694	0.06210	0.08335
65	0.05315	0.06817	0.05614	0.07316

consideration in the first part of this study. These tables are modified in such a way that, the death probabilities is composed of the rates from non-IHD and IHD groups. we refresh the mortality rates of the total population by separating non IHD people mortality and IHD people mortality rates. Therefore, three state transition matrix

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix}$$

determines the probabilities p_{ij}, $(i, j = 1, 2, 3)$ as follows:
p_{11}=$\mathbb{P}$(non-IHD in year $t+1$ | non-IHD in year t)
p_{12}=$\mathbb{P}$(IHD in year $t+1$ | non-IHD in year t)
p_{13}=$\mathbb{P}$(death in year $t+1$ | non-IHD in year t)
p_{21}=$\mathbb{P}$(non-IHD in year $t+1$ | IHD in year t)
p_{22}=$\mathbb{P}$(IHD in year $t+1$ | IHD in year t)
p_{23}=$\mathbb{P}$(death in year $t+1$ | IHD in year t)
p_{31}=$\mathbb{P}$(non-IHD in year $t+1$ | death in year t)
p_{32}=$\mathbb{P}$(IHD in year $t+1$ | death in year t)
p_{33}=$\mathbb{P}$(death in year $t+1$ | death in year t)
where $\sum p_{ij} = 1$.

Furthermore, as there is no recovery from this disease, $p_{21} = 0$ and the last state is taken as the absorbing state, $p_{31} = 0$, $p_{32} = 0$ and $p_{33} = 1$. Similarly, λ_{12} represents the transition rate from non-IHD state to IHD state. λ_{13} is the transition rate from non-IHD state to death state, λ_{23} is the transition rate from IHD state to death. As there are no recoveries, $\lambda_{21} = 0$ and there are no transitions between death states, λ_{31}, λ_{32} and λ_{33} are equal to 0. In addition, $\lambda_{ii} = -\sum \lambda_{ij}$, $i \neq j$ has to be justified. Thus, $\lambda_{11} = -(\lambda_{12} + \lambda_{13})$ and $\lambda_{22} = -\lambda_{23}$.

In order to define phase type distribution in three-state Markov process, we need to define, α, initial distribution coefficients, according to the proportion of IHD caused deaths with respect to the total deaths for each age and gender in the three consecutive years. Because, the cause of deaths among IHD population may not necessarily be from the illness but other reasons. This study assumes that, IHD patients die only as a consequence of this disease as we do not have any other specific information on the reasons of disappearance. For example, chosing $\alpha = (0.5\ 0.5\ 0)$ implies half of the IHD patients die in the first year whereas the second half disappears the year after.

Implementation of the methodology followed as in the two-state case, the transition probabilities and rates are calculated for each gender and presented in Table 6.

Here, p_{12} and p_{32} values represent the IHD-morbidity rates, IHD-mortality rates for Turkish population, respectively and p_{13} corresponds to the death rate from other reasons. Morbidity rates are much higher in males compared to females, however, the probability of death from IHD are close in both genders. The similar pattern is almost observed in transition rates.

Table 6. Three-state transition probabilities and rates of IHD mortality and morbidity

Gender	Age	p_{12}	p_{13}	p_{23}	λ_{12}	λ_{13}	λ_{23}
Male	0	0.00044	0.02906	0.15414	0.00055	0.01528	0.18651
	25	0.00666	0.00110	0.13496	0.00055	0.00071	0.16048
	45	0.03537	0.00348	0.10350	0.00071	0.00248	0.09203
	58	0.11881	0.01499	0.07694	0.00148	0.01097	0.06210
	65	0.11933	0.02613	0.06817	0.00175	0.01813	0.05614
Female	0	0.00055	0.01505	0.15719	0.00044	0.02993	0.18222
	25	0.00057	0.00072	0.13829	0.00631	0.00104	0.15604
	45	0.00075	0.00255	0.08427	0.03222	0.00234	0.11545
	58	0.00168	0.01136	0.05847	0.08437	0.0039	0.08335
	65	0.00220	0.01898	0.05315	0.04956	0.00279	0.07316

The values refer to the transition rates from non-IHD state to IHD state. To illustrate the influence of the IHD-morbidity transition rate on the population represented by different mortality tables, λ_{ij} $(i = 1, 2 \text{ and } j = 2, 3)$ values are calculated and given in Table 6. Being in exposure to IHD makes the transition to death faster which can be seen in the values of λ_{23}. Infant and young age mortality rates for IHD are too high because of the frailty of this age period.

5. Concluding Comments

This study employs Markov model and phase type distribution which is commonly used approach in health studies but rarely employd in morbidity estimation, to determine the improvement in the mortality and morbidity rates based on an application to Turkish data. This study questions if mortality tables can be improved and if morbidity rate due to a certain disease can be quantified using phase type distribution. Additionally, it extends two-state model to a three-state for the estimation of mortality resulting from the specified disease in total population. The approach proposed is applied to Turkish mortality tables and data collected from SSI on IHD. The results indicate that two-state approach to estimate the mortality has a good agreement on young, middle and selected retirement ages. However, changes in older ages which will give a better capture in longevity risks. Additionally, IHD caused morbidities are estimated by proposed approach for Turkey first time. Extension to three-state in IHD morbidity and mortality enables researchers to predict which portion of the mortality rates correspond to IHD-caused deaths.

References

[1] Akat, F. 2015. The Estimation of Adopted Mortality and Morbidity Rates Using Markov Model and Phase Type Law: Turkish Case, Unpublished M.Sc. Thesis, Middle East Technical University, Institute of Applied Mathematics.

[2] Turkish Statistical Institute, 2011. Statistical indicators 1923-2010, Technical report.

[3] Turkish Statistical Institute, 2014. Turkish demographic structure and its future, 2010-2050, Technical report.

[4] Bladt, M. 2005. A review on phase-type distributions and their use in risk theory, Astin Bulletin, 35(1):145-167.
[5] Bobbio, A. 2003. A. Horváth, M. Scarpa, and M. Telek, Acyclic discrete phase type distributions: Properties and a parameter estimation algorithm, Performance evaluation, 54(1):1-32.
[6] Bowers, N.L., Gerber, H.U., Hickman, J.C., Jones, D.A., and Nesbitt, C.J. 1986. Actuarial mathematics. Vol. 2, Society of Actuaries Itasca, Ill.
[7] Collins, D.H. and Huzurbazar, A.V. 2013. Multi-state stochastic processes: A statistical flowgraph perspective, International Statistical Review, 81(1):78-106.
[8] : Longo, D.L, Fauci, A.S, Kasper, D.L, Hauser, S.L, Jameson, J.L and Loscalzo, J, eds. 2012. Harrison's Principles of Internal Medicine. 18th ed. New York, NY: McGraw-Hill. 2012.
[9] Ellis, M. H. 1978. Distances between two-state Markov processes attainable by Markov joinings. Transactions of the American Mathematical Society, 241:129-153.
[10] Faddy, M. 1995. Phase-type distributions for failure times, Mathematical and computer modelling, 22(10):63-70.
[11] Heligman, L. and Pollard, J.H. 1980. The age pattern of mortality, Journal of the Institute of Actuaries, 107(01):49-80.
[12] Hoogenveen, R.T., van Baal, P.H. and Boshuizen, H.C. 2010. Chronic disease projections in heterogeneous ageing populations: approximating multi-state models of joint distributions by modelling marginal distributions, Mathematical medicine and biology, 27(1):1-19.
[13] Hosgood, G.L. 2002. Markov models to estimate and describe survival time and experience in cohorts with high euthanasia frequency, Ph.D. thesis, Purdue University.
[14] Hubbard, R.A., Inoue, L. and Fann, J. 2008. Modeling nonhomogeneous markov processes via time transformation, Biometrics, 64(3):843-850.
[15] Construction of Turkish Morbidity Tables:Phase I,Report No:2010-1, 2010.
[16] Jackson, C.H. , Sharples, L.D., Thompson, S.G., Duffy, S.W. and Couto, E. 2003. Multi-state markov models for disease progression with classification error, Journal of the Royal Statistical Society: Series D (The Statistician), 52(2):193-209.
[17] Jones, B. 1994. Actuarial calculations using a markov model, Transactions of the Society of Actuaries, 46:69-77.
[18] Kay, R. 1986. A markov model for analysing cancer markers and disease states in survival studies, Biometrics, 42(4): 855-865.
[19] Keyfitz N. and Rogers A. 1982. Simplified multiple contingency calculations, Journal of Risk and Insurance, 49(1):59-72.
[20] Lin, X.S. and Liu, X. 2007. Markov aging process and phase-type law of mortality, North American Actuarial Journal, 11(4):92-109.
[21] McClean, S. and Millard, P. 2007. Where to treat the older patient? can markov models help us better understand the relationship between hospital and community care?, Journal of the Operational Research Society, 58(2):255–261.
[22] Newell, C. 1990. Methods and models in demography, Guilford Press.
[23] Steinsaltz, D. and Evans, S.N. 2004. Markov mortality models: implications of quasistationarity and varying initial distributions, Theoretical Population Biology, 65(4):319–337.
[24] Taylor, G., McClean, S. and Millard, P. 1997. Continuous-time markov models for geriatric patient behaviour, Applied stochastic models and data analysis, 13(3-4):315–323.
[25] Turkish Mortality Tables. 2011. Hacettepe University: 2010 detailed tables. Accessed April 8, 2018. http://www.aktuerya.hacettepe.edu.tr/TurkiyeHayatTablolari.php.
[26] Wang, B. and Pham, T.D. 2011. Estimating neuronal ageing with hidden markov models, in 2011 International Symposiun on Computational Models For Life Sciences (CMLS-11), Vol. 1371:110–117, AIP Publishing.
[27] Winkler, W. 1978. Continuous parameter markov processes, The Annals of Probability, 6(3):459–468.
[28] Xie, H., Chaussalet, T.J. and Millard, P.H. 2005. A continuous time markov model for the length of stay of elderly people in institutional long-term care, Journal of the Royal

Statistical Society: Series A (Statistics in Society), 168(1):51–61.