Noname manuscript No.
(will be inserted by the editor)

Evaluating Readmission Rates and Discharge Planning by Analyzing the Length-of-Stay of Patients

Wanlu Gu · **Neng Fan** · **Haitao Liao**

Accepted

Abstract The length-of-stay (LOS) is an important quality metric in health care, and the use of phase-type (PH) distribution provides a flexible method for modeling LOS. In this paper, we model the patient flow information collected in a hospital for patients of distinct diseases, including headache, liveborn infant, alcohol abuse, acute upper respiratory infection, and secondary cataract. Based on the results obtained from fitting Coxian PH distributions to the LOS data, the patients can be divided into different groups. By analyzing each group to find out their common characteristics, the corresponding readmission rate and other useful information can be evaluated. Furthermore, a comparison of patterns for each disease is analyzed. We conclude that it is important to offering better service and avoiding waste of sources, by the analysis of the relations between groups and readmission. In addition, comparing the patterns within distinct diseases, a better decision for assigning resources and improving the insurance policy can be made.

Keywords Phase-type distribution · healthcare quality · length-of-stay · Markov chains · readmission rate

1 Introduction

The length-of-stay (LOS) is a term to describe the period of a patient staying in hospital from admission to discharge. It has been employed as a proxy for measuring the utilization of hospitalization. Lots of work on modeling and evaluating LOS has been conducted. In El-Darzi et al. (1998), the flow of patients through geriatric hospitals was described in terms of short-stay, medium-stay, and long-stay states, and the effect of LOS blocks is assessed to assist in understanding bed occupancy. In Xie et al.

W. Gu · N. Fan (✉)
Department of Systems and Industrial Engineering, University of Arizona, Tucson, AZ 85721
E-mail: nfan@email.arizona.edu

H. Liao
Department of Industrial Engineering, University of Arkansas, Fayetteville, AR 72701

(2005), a Markov model was developed for the LOS of elderly people moving within and between residential home and nursing home.

In this paper, we will utilize phase-type (PH) distributions to analyze the collected patient flow information in Banner University Medical Center Tucson - Main Campus and South Campus. A PH distribution is a probability distribution constructed by a convolution or mixture of exponential distributions. As a special version of PH distributions, the "Erlang method of stages" was proposed by Erlang (1917) to model the time for the analysis of telephone networks. The general theories of PH distributions were initially established by Neuts (1975), and since then the PH distributions have been widely used in many applications, such as telecommunications, finance, queueing theory, insurance risk, survival analysis, reliability theory, drug kinetics (see Fackrell, 2003), as well as diverse areas of stochastic modeling, e.g., Neuts (1981) and Asmussen (1992) for queueing theory, Asmussen and Rolski (1992), Asmussen et al. (1996), Bladt (2005) and Vatamidou et al. (2014) for risk theory, and Ruiz et al. (2008), Cumani (1982), Liao et al. (2017) in reliability. Neuts (1981) later constructed a series of theories for PH distributions and paved the foundation for further research. Kao (1988) extended the work of Neuts by presenting a procedure for computing the renewal and related functions of PH renewal processes. Based on the prior work, Asmussen (2008) showed that PH distributions can approximate any distribution arbitrarily close for nonnegative random variable. For a thorough understanding of PH distributions, readers are referred to Neuts (1981), O'Cinneide (1990), Fackrell (2003) and Bodrog et al. (2008).

The application of PH distributions in the healthcare field has been increasing over time to interpret healthcare systems and improve the efficiency of healthcare delivery. The survey by Fomundam and Herrmann (2007) showed that the healthcare processes can be viewed as queueing systems, in which patients arrive, wait for service, obtain service, and then depart. Fomundam and Herrmann (2007) summarized a range of subareas in healthcare as waiting time and utilization analysis, system design, and appointment systems.

Although the applications of PH distributions vary widely in scope and scale, the main goal of the research includes resources utilization (staff, facility, bed), patients (waiting time, service time, departure destination) and cost. McClean and Millard used a two-term mixed exponential distribution to fit data on LOS in departments of geriatric medicine (McClean and Millard, 1993). Faddy and McClean (1999) interpreted the phases in terms of increased severity of illnesses being treated, and also analyzed two covariates as the age of patient at admission and the year of admission against the phases. In Marshall and McClean (2003), the conditional PH distribution was used for modeling the LOS of a group of elderly patients. Marshall and McClean (2004) considered the use of Coxian PH distributions for modeling patient LOS for elderly patients and investigated the classification issue based on the resulting distribution. They classified the patients into different groups according to their LOS and identified common characteristics among these groups. The identified characteristics offer better insights into the overall management and bed allocation. Tang et al. (2012) demonstrated the application of Coxian PH stochastic regression models to hospital LOS, and proposed a Reversible Jump Markov chain Monte Carlo algorithm to dynamically select the number of phases. This method allows estimation of mean LOS, median and other percentiles. Kim et al. (2012) and Al Hanbali et al. (2012) both considerd queueing system with the batch Markovian arrival process and phase-type service time, and corresponding algorithms were developed. In Turgeman et al. (2015), the Coxian PH distribution was used to fit the LOS data of patients with congestive heart failure. They also analyzed the connections among patient social, clinical, and historical characteristics within each group and estimated the associated readmission risk.

In the literature, the PH distributions are mainly applied to analyze the LOS of patients with elder diseases (see Xie et al., 2005) who usually need long-term care. In this paper, we apply the PH distributions in a different way by studying five different types of diseases at the same time to identify the LOS groups within each disease and then present the changes of patterns from disease to disease. In particular, the LOS information was collected in Banner University Medical Center Tucson - Main Campus and South Campus for patients with diseases of headache, liveborn infant, alcohol abuse, acute upper respiratory infection, and secondary cataract.

Scher et al. (1998) presented the first US-based study describing the prevalence and characteristics of frequent headache in the general population. The overall prevalence of frequent headache was 4.1%. Headaches are commonly seen in neurology practices and headache subspecialty centers. Since there are various causes of this disease, it is hard to describe the headaches of all types or to tell clearly the symptoms of each individual patient. The International Classification of Headache Disorders (IHS, 2013) is a widely recognized classification when the diagnosis is uncertain. In practice, the situation of headaches is complex, and thus it is necessary and important to study the LOS in a hospital.

Maisels and Kring (Maisels and Kring, 1998) evaluated the effect of postnatal age at the time of discharge on the risk of readmission to hospital with specific reference to readmission for hyperbilirubinemia. Bisquera et al. (2002) and Melnyk et al. (2006) both investigated low birth weight premature infants. The former advocated additional research into preventive measures to reduce the incidence of Necrotizing Enterocolitis (NEC) by showing the impact of NEC on the LOS and hospital charges. The latter measured parent-infant interaction during the Neonatal Intensive Care Unit (NICU) and NICU LOS in order to evaluate the efficacy of an educational-behavioral intervention program that was designed to enhance parent-infant interactions and parent mental health outcomes.

Alcohol abuse is a heterogeneous set of behaviors that include any pattern of ethyl alcohol intake that causes medical and social complications (Cloninger et al., 1981). It influences children from both genetic and environmental aspects. In Saitz et al. (1997), it was observed that having an alcohol-related diagnosis is associated with more use of intensive care, longer inpatient stays, and higher hospital charges.

Acute respiratory infection is familiar to most people. It may interfere with normal breathing, affecting either upper respiratory system (sinuses to vocal chords) and lower respiratory system (vocal chords to lungs). Some research, e.g., Terry et al. (1995) and Nieman et al. (1990), has been conducted on the effects of some treatments on specific patients group. Eccles (2002) discussed some ideas concerning the seasonality of acute upper respiratory tract viral infections and put forward their own hypothesis that it is due to cooling of the nasal airway. In this paper, we will investigate the acute upper respiratory infection LOS data to better understand and improve the treatment system.

Secondary cataract formation is the most common postoperative complication following cataract surgery, and up to 40-50% of all patients need some clinical follow-up or treatment due to this complication (Lundgren et al., 1992). Much effort has been taken to prevent the development of secondary cataract on new drugs and experimental study (e.g., Ismail et al., 1996). Fife and Rappaport (1976) studied the effects of construction noise on LOS for simple cataract surgery, which may influence the incidence of secondary cataract directly. Kuchle et al. (1997) analyzed the impact of pseudoexfoliation (PEX) on secondary cataract following cataract extraction. They concluded that eyes with PEX have a higher frequency of secondary cataract, that is a common-seen disease but is special as it is related to a previous surgery.

In this paper, the patients are divided into different groups based on the results obtained from fitting Coxian PH distributions to LOS data of five diseases separately. Then the analysis of discharge direction and readmission is applied to each disease and among different diseases. It is obvious that for different diseases, the number of groups and the probability distributions are not the same. Some disease, such as "headache", have data centered around one group, while the "alcohol abuse" has two peaks at group 2 (less than 4 hours) and group 5 (less than 11 hours). For the readmission rates, the "liveborn infant" has a large number of not being readmitted for all groups, which is different from other four diseases. In addition, the "secondary cataract" has very similar values of three readmission rates, which may convey that the LOS group makes no difference on readmission in this disease.

The remainder of this paper is organized as follows. In Section 2, we will introduce the properties of Coxian PH distributions after a short review of the exponential distribution and continuous time Markov chain. Then, we will further address a method of fitting PH distributions and how to classify LOS groups. Section 3 contains the fitting results and further analysis of five diseases, and provides the corresponding analysis between LOS group and discharge direction as well as readmission rates. Finally, Section 4 concludes the paper.

2 Phase-type Distributions

A continuous phase-type distribution is the distribution of the time from the initial state until absorption in the absorbing state in a Continuous-Time Markov Chain (CTMC) (Neuts, 1981). In the following, the background of CTMC will be introduced with some basic concepts for PH distributions and models.

2.1 Coxian PH distributions

The exponential distribution $Exp(\lambda)$ plays an essential role in PH distributions. Its Probability Density Function (PDF) is $f(x|\lambda) = \lambda exp(-\lambda x),\ x \geq 0$, where $\lambda > 0$ is a rate parameter and the corresponding random variable exhibits the "memoryless property" (Fackrell, 2009). The CTMCs are a class of stochastic processes $\{X(t)\}_{t\geq 0}$

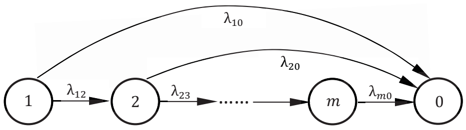

Fig. 1: State Transition Diagram of a CTMC

with a discrete state space $S = \{0, 1, 2, \cdots, m\}$, in which the time between transitions follows an exponential distribution (Neuts, 1981). The continuous-time variable $t \in [0, \infty)$, $m \geq 1$ is a finite number, and the states $1, \cdots, m$ are transient states and state 0 is an absorbing state.

Let the random variable Y denote the time from the initial state until absorption in the absorbing state. Then, Y is said to have a (continuous) PH distribution (Neuts, 1981), and the phase corresponds to state in the CTMC. One special type of PH distributions is the Coxian PH distribution (see Fig. 1), which can overcome difficulty in parameter estimation for the general form of PH distributions by ensuring that the transient states (or phases) of the model are ordered (Fackrell, 2009; Marshall and McClean, 2004). As illustrated in Fig. 1, the stochastic process begins from the first transient state and may either move sequentially or enter the absorbing state 0 directly. The time spent in state i is exponentially distributed with parameter λ_i, which is also interpreted as the average rate moving out of state i. The rate of moving out from state i to state $i+1$ is $\lambda_{i,i+1}$, and the rate to absorbing state 0 is λ_{i0}. According to the balance equations for the Markov chain (Serfozo, 2009), we have $\lambda_i = \lambda_{i,i+1} + \lambda_{i0}$ for $i = 1, \cdots, m-1$ and $\lambda_m = \lambda_{m0}$.

For the Coxian PH distribution, the initial state distribution is $\pi = [\mathbf{1}, \mathbf{0}, \cdots, \mathbf{0}]_{\mathbf{1 \times m}}$, and the transition matrix T consists of transition rates for $\{X(t)\}_{t \geq 0}$,

$$\mathbf{T} = \begin{bmatrix} -\lambda_1 & \lambda_{12} & 0 & \cdots & 0 \\ 0 & -\lambda_2 & \lambda_{23} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -\lambda_{m-1} & \lambda_{m-1,m} \\ 0 & 0 & \cdots & 0 & -\lambda_m \end{bmatrix}_{m \times m}, \quad \mathbf{q} = \begin{bmatrix} \lambda_{10} \\ \lambda_{20} \\ \vdots \\ \lambda_{m-1,0} \\ \lambda_{m,0} \end{bmatrix}_{m \times 1}$$

where $\mathbf{q}$ consists of the rates to the absorbing state. The PDF, cumulative density function (CDF), mean, and variance of absorbing time Y can be expressed as

$$\begin{aligned} f(y) &= -\mathbf{T}\pi e^{\mathbf{T}y}\mathbf{e} = \mathbf{q}\pi exp(\mathbf{T}y), \\ F(y) &= P(Y \leq y) = 1 - \pi exp(\mathbf{T}y)\mathbf{e}, \\ E[Y] &= \int_0^\infty y f(y) dy = \int_0^\infty y \mathbf{q}\pi exp(\mathbf{T}y) dy = -\pi \mathbf{T}^{-1}\mathbf{e}, \\ Var[Y] &= E[Y^2] - E[Y]^2 = 2\pi\mathbf{T}^{-2}\mathbf{e} - (\pi\mathbf{T}^{-1}\mathbf{e})^2, \end{aligned}$$

respectively, where $\mathbf{e} = (1, 1, \cdots, 1)^T_{m \times 1}$.

The m states may then be used to describe stages of a process which terminates at some point (Marshall and McClean, 2004). They also pointed out that the Coxian PH distributions have been successfully applied to modeling patient LOS in a hospital. In practice, the LOS data can be divided into different groups based on their various values. Early in Sorensen (1996), the meaning of groups of LOS was identified as patient groups requiring similar care levels and similar resource. In analyzing of hospital stay, the $E[Y]$ is actually the average time spent in hospital from admission to discharge for a certain type of patients. The $Var[Y]$ measures how far this set of LOS data are spread out from $E[Y]$.

2.2 Fitting method

In Section 2.1, it is already shown that the Coxian PH distributions depend on the parameters in π and $\mathbf{T}$. Then $\theta = (\pi, \mathbf{T})$ is called the representation of a PH distribution, which needs to be estimated in statistical fitting. Maximum Likelihood Estimation (MLE) is the most popular method used to fit data and approximate distributions with PH distributions (Fackrell, 2009). Fackrell also presented a brief overview of some distribution approximation algorithms and their developments and extensions. Esparza et al. (2011) studied the detailed information about parameter estimation for PH distributions using MLE in different cases.

Let $Y = [Y_1, Y_2, \cdots, Y_n]$ be an independent and identically distributed sample from a population with probability density function $f(y; \theta_1, \cdots, \theta_k)$. The likelihood function is defined by

$$\mathbf{L}(\theta|\mathbf{y}) = L(\theta_1, \cdots, \theta_k | y_1, \cdots, y_n) = \prod_{i=1}^{n} f(y_i|\theta_1, \cdots, \theta_k) = \prod_{i=1}^{n} \pi exp(\mathbf{T} y_i)\mathbf{q}.$$

For each sample point $\mathbf{y}$, we use $\hat{\theta}(\mathbf{y})$ to denote the MLE of the parameter θ, at which $\mathbf{L}(\theta|\mathbf{y})$ attains its maximum. Finding the global maximum of $\mathbf{L}(\theta|\mathbf{y})$ is equivalent to finding the maximum of the log-likelihood function $\ln \mathbf{L}(\theta|\mathbf{y}) = \sum_{i=1}^{n} \ln(f(y_i))$, where $f(y_i) = \pi exp(\mathbf{T} y_i)\mathbf{q}$. If the log-likelihood function is differentiable (in θ_i), we can find the possible candidates by solving

$$\frac{\partial}{\partial \theta_i} \ln \mathbf{L}(\theta|\mathbf{x}) = 0, \; i = 1, \cdots, k. \tag{1}$$

The Expectation-Maximization (EM) algorithm has been applied to fit the Coxian PH distributions. Asmussen et al. (1996) developed an extended EM algorithm to minimize the information divergence (maximize the relative entropy), and it is the main method implemented in program EMpht[1]. There are usually two steps in each iteration of the EM algorithm. Before that, the program first starts with initial values $\theta^0 = (\pi^{\mathbf{0}}, \mathbf{T}^{\mathbf{0}})$ which are usually randomly chosen. The two steps are: 1) E-step: calculating the function $h : \theta \to E_{\theta^k}(\ln \mathbf{L}(\theta|\mathbf{y})|\mathbf{y})$; 2) M-step: The new estimators are expressed as $\theta^{k+1} = \arg\max_{\theta} h(\theta)$. By iteratively applying these two steps, the likelihood function is monotonically increasing. The program stops when there is no significant improvement in likelihood, and the resulting estimated parameters will be used for further analysis.

[1] http://home.math.au.dk/asmus/pspapers.html

2.3 LOS groups

As shown in Sections 2.1 and 2.2, the transition rates $\lambda_{i,j}$ and $\lambda_{i0}, i, j = 1, \cdots, m$ can be obtained from the estimated $\mathbf{T}$. In this paper, Akaike Information Criterion (AIC) was used to decide the most appropriate number of states (Equation (2)) by taking into consideration the number of parameters k.

$$AIC = 2k - 2\max_{\theta} \ln \mathbf{L}(\theta|\mathbf{y}). \tag{2}$$

Let $P_i,\ i = 1, \cdots, m$ denote the probability of transferring from state i to absorbing state 0. Then P_i can be calculated by integrating the corresponding density functions. For example, P_1 is the probability from state 1 to state 0. According to the balance equation $\lambda_1 = \lambda_{12} + \lambda_{10}$, the time spent in state 1 follows $exp(\lambda_{12} + \lambda_{10})$. But the time contributes to two directions, one transferring to state 2 with rate λ_{12} and another one for being absorbed to state 0 with rate λ_{10}. Then, the proportion of being absorbed to state 0 from state 1 directly is $\frac{\lambda_{10}}{\lambda_{10}+\lambda_{12}}$. Similarly, all P_i can be obtained by the formula in Equations (3).

$$\begin{aligned} P_1 &= \int_0^\infty \lambda_{10} \exp^{-(\lambda_{12}+\lambda_{10})t}\, dt = \frac{\lambda_{10}}{\lambda_{10}+\lambda_{12}}, \\ P_2 &= \int_0^\infty \lambda_{20} \exp^{-(\lambda_{23}+\lambda_{20})t}\, dt \int_0^\infty \lambda_{12} \exp^{-(\lambda_{12}+\lambda_{10})t}\, dt \\ &= \frac{\lambda_{12}}{\lambda_{10}+\lambda_{12}} \times \frac{\lambda_{20}}{\lambda_{20}+\lambda_{23}}, \\ P_m &= \frac{\lambda_{12}}{\lambda_{10}+\lambda_{12}} \times \frac{\lambda_{23}}{\lambda_{20}+\lambda_{23}} \times \cdots \times \frac{\lambda_{m-1,m}}{\lambda_{m-1,0}+\lambda_{m-1,m}}. \end{aligned} \tag{3}$$

In data processing (Section 3.1), the LOS data is sorted in an ascending order to determine the LOS groups. Fig. 2 shows the framework of the corresponding analysis between LOS group and the discharge directions as well as readmission rate groups. The first group has the shortest LOS, and the mth one has the longest LOS. In Section 3.2, analytic results for the five diseases will be presented separately.

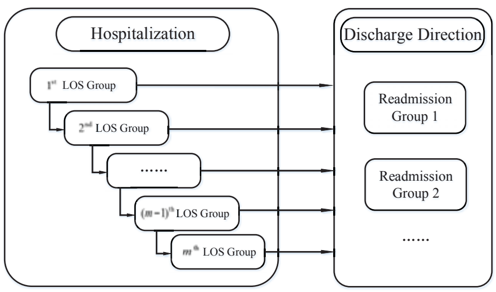

Fig. 2: State Transition Diagram of a LOS

3 Fitting PH distributions to LOS Data

3.1 Data processing

The methods and models in Section 2 are applied to the patient flow data collected from 2013 through 2016 in Banner University Medical Center Tucson - Main Campus and South Campus. The ICD-10 codes[2] for the five diseases and their corresponding approximate conversions from "2018 ICD-10-CM CMS General Equivalence Mappings" are shown in Table 1. The size of records and the number of patients for each disease are also listed. These diseases are common ailments with distinct characteristics and treatment patterns.

Table 1: Disease Types

ICD-10	Approximate conversion	# of records	# of patients
R51	Headache	12,866	5,170
Z38.00	Single liveborn infant, delivered vaginally	1,546	1,443
F10.10	Alcohol abuse	1,906	522
J06.9	Acute upper respiratory infection	4,639	2,174
H26.499	Other secondary cataract, unspecified eye	359	185

The transition records include the information about patience ID, record ID, diagnosis, admission time, bed ID, department, discharge time, and discharge destination. In one transition record of a patient, the time interval between the admission time and the discharge time is treated as one LOS data point. In order to unify the unit of time interval and ensure the sensitivity of fitting results, the time unit is set as minute.

The records unrelated to the disease studied in each case are filtered out during data processing. For example, for the disease "liveborn infant", patients are readmitted with common pediatric diseases. For other diseases, the situation is similar. The readmitted records in the case of "headache" all contain diagnoses similar or related to "headache", such as "dizziness and giddiness" or "fever". The filtering follows the 2018 ICD-10-CM Coding Rules[3]. By these rules, the readmitted records are guaranteed to be with the same type of diseases. For the patients who have readmission records, the difference between the previous discharge time and the next admission time is the readmission interval. The readmission rate is simply the proportion of readmission records. We use six months of Time to Next Admission (TTNA) as a threshold to identify Long Time Readmission Rate (LTRR) and Short Time Readmission Rate (STRR). Similarly, the No Readmission Rate (NRR) is the proportion of not being readmitted. Considering the three readmission rates, their changing trends will directly show the inherent difference in LOS groups and diseases. Since one patient can be readmitted again and thus has several different transition records, the LOS groups actually are groups of records. For example, a patient can be in group 1 for his first record and be in group 2 for another one.

[2] http://www.icd10data.com/ICD10CM/Codes

[3] http://www.icd10data.com/

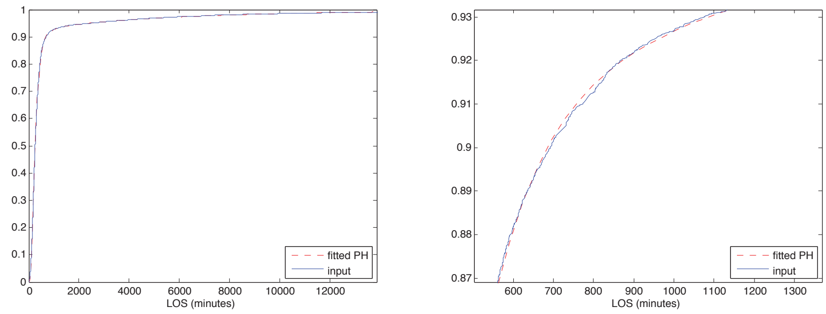

Fig. 3: (a) Fitted PH distribution vs. Input Distribution of R51 (b) Partial Enlarged Detail of R51

For the discharge destinations, which usually contain more than 20 places, we classify them into six main groups, including home, hospital, medical center, health facility, other unspecific facilities, and nursing home.

We fitted a Coxian PH distribution to the LOS data using the EMpht program with maximum of 1000 iterations. The resulting maximum log-likelihood values, minimum AIC for the five diseases, and the number of LOS groups are presented in Table 2.

Table 2: Log-likelihood and Groups

Disease	Maximum $\ln \mathbf{L}(\theta\|\mathbf{y})$	# of Groups	Minimum AIC
R51	-30178.05476	7	60382.10952
Z38.00	-13721.43258	3	27452.86516
F10.10	-14277.45685	7	28580.91370
J06.9	-30293.34526	7	60612.69052
H26.499	-2074.26930	4	4162.53859

3.2 Results

3.2.1 Headache

Fig. 3 illustrates the fitting of the PH distribution to the input LOS data of "headache". The solid line and dashed line are the CDF of input LOS data and the fitted PH distribution, respectively. The fitted distribution follows closely to the input one even in the partially enlarged plot (Fig. 3(b)), indicating that the fitting is adequate.

Table 3 presents the estimated parameters, the probability, the sample size, and the maximum LOS of each state. The first column is the rates from transient states to absorbing state. The next column is the proportion of each state. We want to emphasize that the transient states are not same with the LOS groups, since the probability

Table 3: Results of PH Distributions Fitting of R51

Transition rate		Probability		Group	Maximum LOS	# of records
λ_{10}	0.000176	P_1	0.015084	G_1	39	195
λ_{20}	0.000178	P_2	0.013597	G_2	55	172
λ_{30}	0.005626	P_3	0.382946	G_3	220	4,917
λ_{40}	0.016944	P_4	0.440500	G_4	518	5,673
λ_{50}	0.001652	P_5	0.079340	G_5	1,129	1,025
λ_{60}	0.000182	P_6	0.062366	G_6	17,085	802
λ_{70}	0.000045	P_7	0.006168	G_7	92,450	81

from one transient state to the absorbing state can be 0, which means that the corresponding LOS group does not exist. Group G_i corresponds to the ith LOS group. Finally, the maximum LOS and the size of data points for each state are listed. The patients with headache may be regarded as belonging to seven LOS groups: 0-39 minutes, 40-55 minutes, 56-220 minutes, 221-518 minutes, 519-1129 minutes (less than 1 day), 1130-17085 minutes (1-12 days) and 17085-92450 minutes (12-65 days). Additionally, the expected value of the fitted distribution is 779.82 minutes (average LOS is approximately 13 hours), which is the same as the sample mean of the input LOS data. But for the standard deviation, the fitted one has 3275.6 that is large than the input one (3227.11).

The transition through groups cannot reveal the severity of patients, but does influence the discharge directions and readmission rate. In Table 4, the amount and proportion of each discharge direction are shown with respect to the corresponding group number. In general, most of the patients go back home after treatment. But for those long LOS patients (groups 6 and 7), there are a large number of patients being transferred to other unnamed facilities. Therefore, for patients with headache staying in hospital for more than 17085 minutes (approximately 12 days), it seems difficult for them to recover at home. For the first two groups, with LOS less than 1 hour, none of the patients discharge to other hospitals, health facility or nursing home. These groups of patients may be in good health and their headache may result from getting a cold or other indispositions.

Table 4: Discharge destination of R51

Group	Home		Hospital		Medical center		Health facility		Other		Nursing home	
G_1	187	95.90%	0	0.00%	1	0.51%	0	0.00%	7	3.59%	0	0.00%
G_2	168	97.67%	0	0.00%	1	0.58%	0	0.00%	3	1.74%	0	0.00%
G_3	4,737	96.36%	9	0.18%	50	1.02%	60	1.22%	49	1.00%	11	0.22%
G_4	5,388	94.99%	25	0.44%	85	1.50%	105	1.85%	59	1.04%	10	0.18%
G_5	864	84.38%	11	1.07%	51	4.98%	53	5.18%	35	3.42%	10	0.98%
G_6	604	75.50%	18	2.25%	35	4.38%	36	4.50%	98	12.25%	9	1.13%
G_7	44	55.00%	2	2.50%	2	2.50%	0	0.00%	30	37.50%	2	2.50%

In Table 5, the LTRR, STRR and NRR are listed for all 7 groups. Other information, such as the number of TTNA being longer than 6 months or not is also included. Fig. 4 illustrates the variation trend of three rates with groups. The STRR is the high-

Table 5: Readmission Rate of R51

Group	TTNA >180	TTNA>0	LTRR	STRR	NRR
G_1	38	135	19.49%	49.74%	30.77%
G_2	28	99	16.28%	41.28%	42.44%
G_3	708	2,698	14.40%	40.47%	45.13%
G_4	947	3,350	16.69%	42.36%	40.95%
G_5	140	678	13.66%	52.49%	33.85%
G_6	149	516	18.58%	45.76%	35.66%
G_7	12	48	14.81%	44.44%	40.74%

est for most groups and occupies about half of all rates, while LTRR is relatively stable and always has low values. For the NRR, it is close to STRR in group 2, 3 and 4. For group 3, the NRR is the highest rate and has its own largest value.

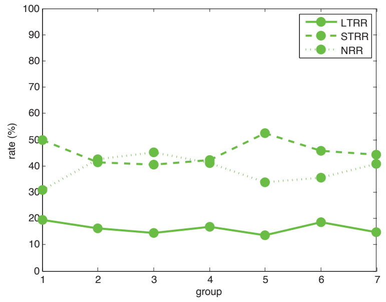

Fig. 4: Readmission Rates vs. Group of R51

3.2.2 Single liveborn infant, delivered vaginally

The fitted PH distribution with 3 states and the distribution of the input data are presented in Fig. 5. Since their general trends change in the same way, the fitting is also adequate.

It is concluded in Maisels and Kring (1998) that newborn infants discharged at any time < 72 hours have significantly high risk for readmission to hospital. Both the newborn infants delivered vaginally and delivered by cesarean were considered. In this paper, we focus on the ones delivered vaginally and give more detailed LOS grouping with reference to readmission risk. In Table 6, the fitted rates $\lambda_{20}, \lambda_{30}, \lambda_{40}, \lambda_{50}$ are almost 0. This means that the probabilities of departure from states 2,3,4, and 5 to the absorbing state (discharge) are all 0. In other words, patients who are in state 2 always stay in the hospital until state 6 and finally to absorbing state. Thus, the three groups of LOS are 0-425 minutes (less than 1 day), 425-6384 minutes (less than 5 days) and 6385-106933 minutes (5-75 days). The LOS mean is about 5371.76 minutes (average LOS is 4 days) for both the fitted and observed distributions, and the

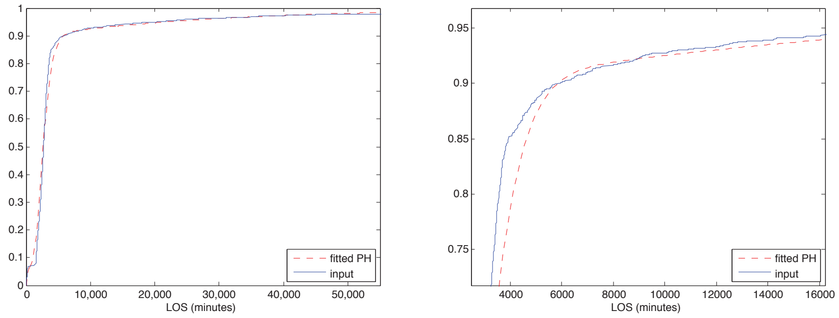

Fig. 5: (a) Fitted PH distribution vs. Input Distribution of Z38.00 (b) Partial Enlarged Detail of Z38.00

sample standard deviation is 12205.69, that is smaller than the standard deviation of the fitted distribution (12440.01).

Table 6: Results of PH Distributions Fitting of Z38.00

Transition rate		Probability		Group	Maximum LOS	# of records
λ_{10}	0.019206	P_1	0.069759	G_1	425	107
λ_{20}	0.000000	P_2	0.000000			
λ_{30}	0.000000	P_3	0.000000			
λ_{40}	0.000000	P_4	0.000000			
λ_{50}	0.000000	P_5	0.000000			
λ_{60}	0.104980	P_6	0.834163	G_2	6,384	1,290
λ_{70}	0.002081	P_7	0.096078	G_3	106,933	149

Table 7 presents the distribution of all four destinations. The proportions of going back home are all higher than 80% and there are very few other discharge directions. In addition, those who stay in hospital from 1-5 days seem to have the best recovery since there is an almost 100% of going home rate. This “disease” is different from others since no bacteria or virus causes damages physically or psychologically. Liveborn infants have a strong ability to grow up healthily as long as they are well fed and taken care of.

Table 7: Discharge Destination of Z38.00

Group	Home		Hospital		Medical center		Other	
G_1	99	92.52%	0	0.00%	6	5.61%	2	1.87%
G_2	1,272	98.60%	0	0.00%	0	0.00%	18	1.40%
G_3	131	87.92%	1	0.67%	0	0.00%	17	11.41%

Table 8 shows that the NRR values for all groups are extremely high compared with other diseases, indicating that most of liveborn infants and their mothers recover

Table 8: Readmission Rate of Z38.00

Group	TTNA >180	TTNA>0	LTRR	STRR	NRR
G_1	9	21	8.41%	11.21%	80.37%
G_2	40	79	3.10%	3.02%	93.88%
G_3	7	15	4.70%	5.37%	89.93%

very well after discharging from hospital. Fig. 6 shows the changing readmission rates as the group varies. The NRR is the highest and has more than 80% for the three groups, while the LTRR and STRR are relatively low and similar. Besides, the three rates are quite stable, showing that the LOS does not influence much on the structure of readmission rates.

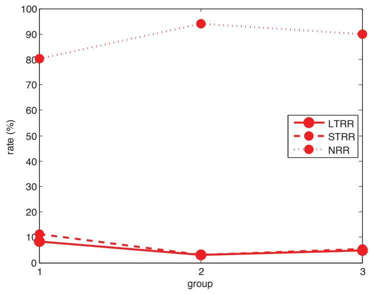

Fig. 6: Readmission Rates vs. Group of Z38.00

3.2.3 Alcohol abuse

Fig. 7 shows the closeness of the fitted distribution to the input data. The points from the fitted distribution fall approximately along with the observed distribution line.

The LOS data can be classified into 7 groups where group 5 with 369-652 minutes (7-11 hours) of staying takes the largest proportion, group 6 comes the second as 653-1418 minutes (about half of one day) (see Table 9). But there is no obvious concentration in one specific group, which means that the LOS data of "alcohol abuse" has a low kurtosis and most patients have LOS within one day. Moreover, patients with LOS longer than 1 day are all in group 7, with LOS from 1-44 days. The mean of LOS is 1116 minutes (average LOS is about 19 hours), the sample standard deviation 3462 is larger than the fitted standard deviation 3084.

In Table 10, the rate of discharging to home is relatively high for first 6 groups but much smaller for group 7. For the longest LOS (the 7th) group, the ratios of going to other hospital, medical center, nursing home or other unspecific facilities are higher compared with those in other groups. Only groups 5 and 7 have a few patients transferred to other hospitals. Interestingly, only group 1 does not have patients going to other hospital, medical center, health facility and nursing home.

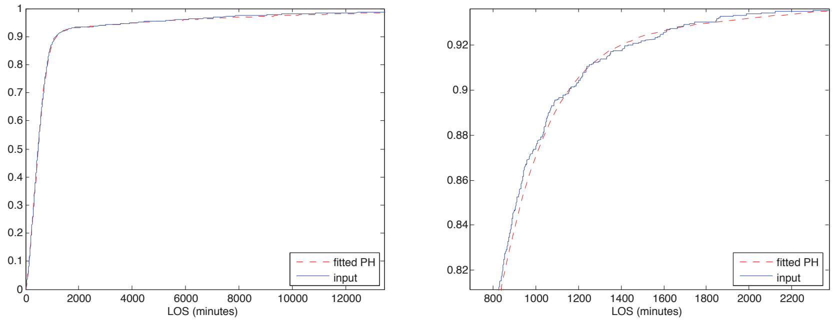

Fig. 7: (a) Fitted PH distribution vs. Input Distribution of F10.10 (b) Partial Enlarged Detail of F10.10

Table 9: Results of PH Distributions Fitting of F10.10

Transition rate		Probability		Group	Maximum LOS	# of records
λ_{10}	0.000145	P_1	0.017339	G_1	41	33
λ_{20}	0.001821	P_2	0.191542	G_2	230	363
λ_{30}	0.001691	P_3	0.127270	G_3	334	243
λ_{40}	0.000570	P_4	0.040171	G_4	368	75
λ_{50}	0.003126	P_5	0.301121	G_5	652	576
λ_{60}	0.008565	P_6	0.242093	G_6	1,418	462
λ_{70}	0.000130	P_7	0.080464	G_7	62,985	154

Table 10: Discharge Destination of F10.10

Group	Home& home visit		Hospital		Medical center		Health facility		Other		Nursing home	
G_1	32	96.97%	0	0.00%	0	0.00%	0	0.00%	1	3.03%	0	0.00%
G_2	320	90.91%	0	0.00%	4	1.14%	23	6.53%	1	0.28%	4	1.14%
G_3	212	87.24%	0	0.00%	6	2.47%	17	7.00%	1	0.41%	7	2.88%
G_4	64	85.33%	0	0.00%	1	1.33%	6	8.00%	4	5.33%	0	0.00%
G_5	517	89.76%	3	0.52%	3	0.52%	34	5.90%	10	1.74%	9	1.56%
G_6	406	87.88%	0	0.00%	7	1.52%	33	7.14%	8	1.73%	8	1.73%
G_7	95	61.69%	2	1.30%	10	6.49%	12	7.79%	30	19.48%	5	3.25%

Table 11: Readmission Rate of F10.10

Group	TTNA >180	TTNA>0	LTRR	STRR	NRR
G_1	3	29	9.09%	78.79%	12.12%
G_2	36	237	9.92%	55.37%	34.71%
G_3	26	159	10.70%	54.73%	34.57%
G_4	3	51	4.00%	64.00%	32.00%
G_5	69	394	11.98%	56.42%	31.60%
G_6	62	341	13.42%	60.39%	26.19%
G_7	18	104	11.69%	55.84%	32.47%

The trend is similar to the one of R51, the “ headache”. The STRR is also the highest one and even larger than that in Section 3.2.1. The LTRR is very low and stable for all groups. We can suppose that for those alcoholics, it is common for them to come to hospital again and again since they are addicted to alcohol. Besides, for patients in group 1, who stay for less than one hour, the STRR is larger and NRR is much smaller than those of other groups. It seems that a quick treatment and hasty leaving will not help the patients get rid of the abuse totally (Table 11, Fig. 8).

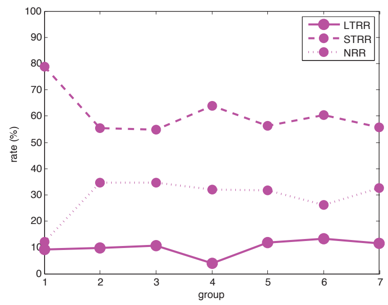

Fig. 8: Readmission Rate vs. Group of F10.10

3.2.4 Acute upper respiratory infection, unspecified

Fig. 9 shows both the fitted distribution and the observed distribution. Two function lines overlap, showing the adequacy of the fitting.

In Table 12, the group 3, as the largest group, accounts for 56.98% of the patients. It has the LOS from 41-217 minutes, approximately 1-4 hours. The second largest groups are groups 4 and 5, which have a range of 4-12 hours. Further more, the sample mean and the expected value of the fitted distribution are both 618.87 minutes (average LOS is about 11 hours). The standard deviations of the two distributions are 3289.72 and 3113.30, respectively.

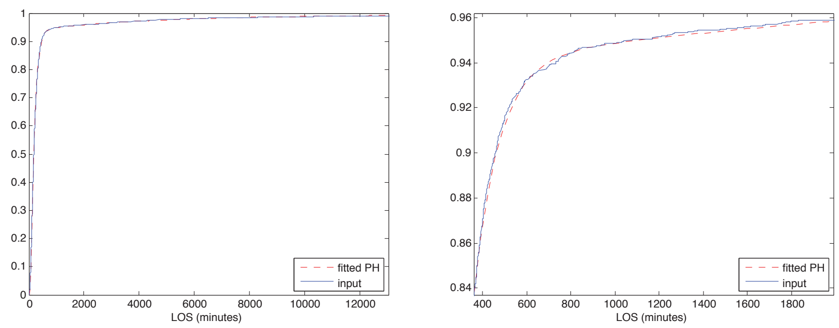

Fig. 9: (a) Fitted PH distribution vs. Input Distribution of J06.9 (b) Partial Enlarged Detail of J06.9

Most patients discharge from hospital within 12 hours. For those who stay in hospital longer than 12 hours, groups 6 and 7 in Table 13, there is a lower proportion of going back home. Especially for group 7, only 68.63% can return home while 23.53% go to other facilities and 5.88% to nursing home. The longest LOS can be related with people having bad physical condition since they do not have a strong recovery ability (needing more treatments) and cannot take care of themselves.

Table 12: Results of PH Distributions Fitting of J06.9

Transition rate		Probability		Group	Maximum LOS	# of records
λ_{10}	0.000098	P_1	0.008724	G_1	23	39
λ_{20}	0.000515	P_2	0.027552	G_2	41	125
λ_{30}	0.020370	P_3	0.569752	G_3	217	2,641
λ_{40}	0.013621	P_4	0.184537	G_4	319	860
λ_{50}	0.006278	P_5	0.150469	G_5	744	699
λ_{60}	0.000218	P_6	0.048088	G_6	11,075	224
λ_{70}	0.000057	P_7	0.010877	G_7	105,311	51

Additionally, in Table 14 and Fig. 10, the NRR for shorter LOS groups, group 1-4, is larger than the other two rates. The STRR increases with group number becoming larger. This may due to the severity of disease, the longer the LOS is, the severer the infection the patient had. Another phenomenon worth to notice is that group 6 has the lowest NRR, which means that almost all patients in this group need to be readmitted to hospital. But there are very few of patients discharge to nursing home. Our guess is that patients in group 6 have a very bad infection but their physical conditions are still good. However, the group 7 contains some elder patients not only got caught with severe infection but also had a bad health condition, and there is a high probability that some of them may pass away only by this common disease. That is why the NRR is even higher than that of group 5 and group 6.

Table 13: Discharge Destination of J06.9

Group	Home& home visit		Hospital		Medical center		Health facility		Other		Nursing home	
G_1	38	97.44%	0	0.00%	0	0.00%	1	2.56%	0	0.00%	0	0.00%
G_2	121	96.80%	0	0.00%	0	0.00%	1	0.80%	3	2.40%	0	0.00%
G_3	2,601	98.49%	3	0.11%	16	0.61%	15	0.57%	5	0.19%	1	0.04%
G_4	827	96.16%	2	0.23%	17	1.98%	10	1.16%	2	0.23%	2	0.23%
G_5	651	93.13%	6	0.86%	22	3.15%	12	1.72%	5	0.72%	3	0.43%
G_6	193	86.16%	1	0.45%	4	1.79%	7	3.13%	17	7.59%	2	0.89%
G_7	35	68.63%	0	0.00%	1	1.96%	0	0.00%	12	23.53%	3	5.88%

3.2.5 Other secondary cataract, unspecified eye

Figure 11 illustrates the fitted distribution and the observed distribution overlap, indicating the adequacy of the fitting.

Table 14: Readmission Rate of J06.9

Group	TTNA >180	TTNA>0	LTRR	STRR	NRR
G_1	10	19	25.64%	23.08%	51.28%
G_2	15	55	12.00%	32.00%	56.00%
G_3	367	1,201	13.90%	31.58%	54.52%
G_4	148	488	17.21%	39.53%	43.26%
G_5	113	478	16.17%	52.22%	31.62%
G_6	48	161	21.43%	50.45%	28.13%
G_7	8	32	15.69%	47.06%	37.25%

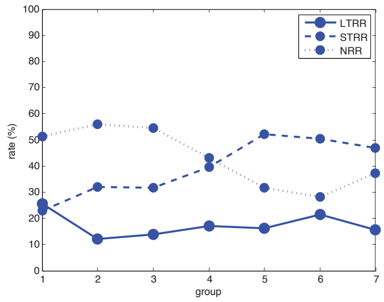

Fig. 10: Readmission Rate vs. Group of J06.9

Similar to “liveborn infant” in Section 3.2.2, the rates of being absorbed from first three states, $\lambda_{10}, \lambda_{20}, \lambda_{30}$, are almost 0 (see Table 15). And in state 4, there is only one record of admission with a stay of 15 minutes. It is reasonable to treat it as an outlier. Therefore, after admission, patients always stay in the hospital longer until state 5 and then choose to stay or discharge. Hence, the final model consists of groups 2, 3 and 4. Among these three groups, group 2 has the LOS from 0-257 minutes (within 5 hours), group 3 has the LOS 258-625 minutes (within 11 hours) and group 4 has the LOS 626-7460 minutes (no more than 6 days). Additionally, group 2 takes the largest proportion as 90.84% among the three groups. The mean of the LOS for observed data and fitted distribution are 203.94 minutes (average LOS is around 4 hours), while the sample standard deviation is 559.34 and the fitted standard deviation is 563.80.

In Table 16, the returning home rates for the three groups are very close to 100%. The variation of discharging destination is also very limited. It may due to that the “secondary cataract” itself is not a severe disease and does not need any complex treatment.

For the readmission rates in Table 17 and Fig. 12, the LTRR for three groups are similar. The STRR increases and the NRR decreases with the group number becoming larger. For short LOS, STRR occupies more than half. But for long LOS, NRR becomes the largest rate.

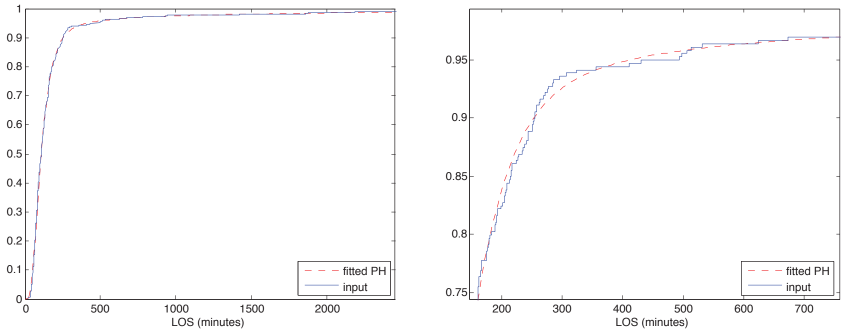

Fig. 11: (a) Fitted PH distribution vs. Input Distribution of H26.499 (b) Partial Enlarged Detail of H26.499

Table 15: Results of PH Distributions Fitting of H26.499

Transition rate		Probability		Group	Maximum LOS	# of records
λ_{10}	0.000000	P_1	0.000000			
λ_{20}	0.000000	P_2	0.000000			
λ_{30}	0.000000	P_3	0.000000			
λ_{40}	0.000002	P_4	0.000028	G_1	15	1
λ_{50}	0.063863	P_5	0.908426	G_2	257	325
λ_{60}	0.002725	P_6	0.059956	G_3	625	22
λ_{70}	0.000476	P_7	0.031590	G_4	7,460	11

Table 16: Discharge Destination of H26.499

Group	Home		Health facility		Other		Medical center	
G_2	323	99.38%	0	0.00%	2	0.62%	0	0.00%
G_3	20	90.91%	1	4.55%	0	0.00%	1	4.55%
G_4	9	81.82%	1	9.09%	0	0.00%	1	9.09%

Table 17: Readmission Rate of H26.499

Group	TTNA >180	TTNA>0	LTRR	STRR	NRR
G_2	45	152	13.85%	32.92%	53.23%
G_3	4	12	18.18%	36.36%	45.45%
G_4	2	7	18.18%	45.45%	36.36%

4 Conclusions

This paper investigates the patient flow data collected in Banner University Medical Center Tucson - Main Campus and South Campus. Based on the numerical experiments, the PH distributions can approximate very well as a first step in the analysis of the LOS. Besides clustering of patients by their LOS information, this paper also presents a direct view of the diversity of patterns from disease to disease. Additionally, by fitting a Coxian PH distribution to patients' LOS data, we can identify the LOS groups with discharge directions and readmission rates.

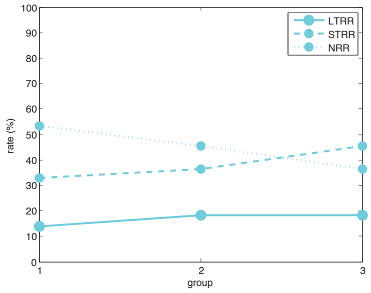

Fig. 12: Readmission Rate vs. Group of H26.499

From the results, the number of LOS groups and the data distribution within groups vary from disease to disease. Some of these diseases have data centered around one group, e.g., "headache" has most patients in groups 3 and 4, while some have data distributed in another way, e.g., the "alcohol abuse" has two peak at group 2 and group 5 separately. For the readmission rates, in general, the LTRR is stable for the five diseases. But for STRR and NRR, the observations are different. For example, the "liveborn infant" is quite different from others with a large number of NRR for all groups. For the disease "secondary cataract", all three readmission rates of three groups have similar values, and this may convey that the LOS group makes no difference on readmission for this disease.

Each LOS group is a collection of patients sharing same characteristics. The analysis of the relations between group and readmission will offer better service to a new patient and avoid sources wasting by reviewing patients' previous medical records. In addition, comparing the readmission structure among distinct diseases and describing the inherent differences in patterns can help make better decision in assigning resources as well as improving the insurance policy.

Acknowledgement. We would appreciate the University of Arizona Center for Biomedical Informatics & Biostatistics Department of Biomedical Informatics Services for providing the data. This material is based upon work supported by National Science Foundation Grants CMMI #1634282 and #1635379.

References

Al Hanbali, A., de Haan, R., Boucherie, R. J., van Ommeren J. C. W. (2012) Time-limited polling systems with batch arrivals and phase-type service times. *Annals of Operations Research*, 198(1), 57-82.

Asmussen, S. (1992) Phase-type representations in random walk and queueing problems. *The Annals of Probability*, 20(2), 772-789.

Asmussen, S. (2008) *Applied Probability and Queues*. Springer Science and Business Media.

Asmussen, S., Rolski, T. (1992) Computational methods in risk theory: a matrix-algorithmic approach. *Insurance: Mathematics and Economics*, 10(4), 259-274.

Asmussen, S., Nerman, O., Olsson, M. (1996) Fitting phase-type distributions via the EM algorithm. *Scandinavian Journal of Statistics*, 23(4), 419-441.

Bisquera, J. A., Cooper, T. R., Berseth, C. L. (2002) Impact of necrotizing enterocolitis on length of stay and hospital charges in very low birth weight infants. *Pediatrics*, 109(3), 423-8.

Bladt, M. (2005) A review on phase-type distributions and their use in risk theory. *ASTIN Bulletin: The Journal of the International Actuarial Association*, 35(1), 145-161.

Bodrog, L., Horvath, A., Telek, M. (2008) Moment characterization of matrix exponential and Markovian arrival processes. *Annals of Operations Research*, 160(1), 51-68.

Cloninger, C. R., Bohman, M., Sigvardsson, S. (1981) Inheritance of alcohol abuse. Cross-fostering analysis of adopted men. *Archives of General Psychiatry*, 38(8), 861-868.

Cumani, A. (1982) On the canonical representation of homogeneous Markov processes modelling failure-time distributions. *Microelectronics Reliability*, 22(3), 583-602.

Eccles, R. (2002) An explanation for the seasonality of acute upper respiratory tract viral infections. *Acta Otolaryngol*, 122(2), 183-191.

El-Darzi, E., Vasilakis, C., Chaussalet, T., Millard, P. H. (1998) A simulation modelling approach to evaluating length of stay, occupancy, emptiness and bed blocking in a hospital geriatric department. *Health Care Management Science*, 1(2), 143-149.

Erlang, A. (1917) Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges. *Elektrotkeknikeren*, 13.

Esparza, L. J. R., Nielsen, B. F., Bladt, M. (2011) Maximum likelihood estimation of phase-type distributions. Kgs. Lyngby, Denmark: Technical University of Denmark (IMM-PHD-2010-245).

Fackrell, M. W. (2003) Characterization of matrix-exponential distributions. PhD Thesis, School of Applied Mathematics, University of Adelaide, South Australia.

Fackrell, M. (2009) Modelling healthcare systems with phase-type distributions. *Health care Management Science*, 12, 11-26.

Faddy, M. J., McClean, S. I. (1999) Analysing data on lengths of stay of hospital patients using phase-type distributions. *Applied Stochastic Models in Business and Industry*, 15(4), 311-317.

Fife, D., Rappaport, E. (1976) Noise and hospital stay. *American Journal of Public Health*, 66(7), 680-681.

Fomundam, S., Herrmann, J. W. (2007) A Survey of Queuing Theory Applications in Healthcare. Technical Report TR 2007-24. Institute for Systems Research, University of Maryland.

Headache Classification Committee of the International Headache Society (IHS) (2013) The International Classification of Headache Disorders, 3rd edition (beta version). *Cephalalgia*, 33(9), 629-808.

Ismail, M. M., Alio, J. L., Jose, M. (1996) Prevention of secondary cataract by antimitotic drugs: experimental study. *Ophthalmic Research*, 28(1), 64-69.

Kao, E. P. (1988) Computing the Phase-type renewal and related functions. *Technometrics*, 30(1), 87-93.

Kim, C. S., Mushko, V. V., Dudin, A. N. (2012) Computation of the steady state distribution for multi-server retrial queues with phase type service process. *Annals of Operations Research*, 201(1), 307-323.

Kuchle, M., Amberg, A., Martus, P., Nguyen, N. X., Naumann, G. O. (1997) Pseudoexfoliation syndrome and secondary cataract. *British Journal of Ophthalmology*, 81(10), 862-866.

Liao, H., Zhang, Y., Guo, H. (2017) An Erlang-Coxian-based method for modeling accelerated life testing data. *In Advances in Through-life Engineering Services*, 165-186. Springer International Publishing.

Lundgren, B., Jonsson, E., Rolfsen, W. (1992) Secondary cataract. *Acta Ophthalmologica*, 70(S205), 25-28.

Maisels, M. J., Kring, E. (1998) Length of stay, jaundice, and hospital readmission. *Pediatrics*, 101(6), 995-998.

Marshall, A. H., McClean, S. I. (2003) Conditional phase-type distributions for modelling patient length of stay in hospital. *International Transactions in Operational Research*, 10(6), 565-576.

Marshall, A. H., McClean, S. I. (2004) Using Coxian phase-type distributions to identify patient characteristics for duration of stay in hospital. *Health Care Management Science*, 7(4), 285-289.

McClean, S., Millard, P. (1993) Patterns of length of stay after admission in geriatric medicine: an event history approach. *The Statistician*, 42(3), 263-274.

Melnyk, B. M., Feinstein, N. F., Alpert-Gillis, L., Fairbanks, E., Crean, H. F., Sinkin, R. A., Gross, S. J. (2006) Reducing premature infants' length of stay and improving parents' mental health outcomes with the Creating Opportunities for Parent Empowerment (COPE) neonatal intensive care unit program: a randomized, controlled trial. *Pediatrics*, 118(5), e1414-1427.

Neuts, M. F. (1975) Probability distributions of phase-type. Department of Mathematics, University of Louvain, Louvain, 173-206.

Neuts, M. F. (1981) *Matrix-Geometric Solutions in Stochastic Models*. The Johns Hopkins University Press, Baltimore.

Nieman, D. C., Nehlsen-Cannarella, S. L., Markoff, P. A., Balk-Lamberton, A. J., Yang, H., Chritton, D. B. W., Lee J. W., Arabatzis, K. (1990) The effects of moderate exercise training on natural killer cells and acute upper respiratory tract infections. *International Journal of Sports Medicine*, 11(06), 467-473.

O'Cinneide, C. A. (1990) Characterization of phase-type distributions. *Stochastic Models*, 6(1), 1-57

Ruiz-Castro, J. E., Perez-Ocon, R., Fernandez-Villodre, G. (2008) Modelling a reliability system governed by discrete phase-type distributions. *Reliability Engineering and System Safety*, 93(11), 1650-1657.

Saitz, R., Ghali, W. A., Moskowitz, M. A. (1997) The impact of alcohol-related diagnoses on pneumonia outcomes. *Archives of Internal Medicine*, 157(13), 1446-1452.

Scher, A. I., Stewart, W. F., Liberman, J., Lipton, R. B. (1998) Prevalence of frequent headache in a population sample. *Headache: The Journal of Head and Face Pain*, 38(7), 497-506.

Serfozo, R. (2009) *Basics of applied stochastic processes*. Springer Science and Business Media.

Sorensen, J. (1996) Multi-phased bed modelling. *Health Services Management Research*, 9(1), 61-67.

Tang, X., Luo, Z., Gardiner, J. C. (2012) Modeling hospital length of stay by Coxian phase-type regression with heterogeneity. *Statistics in Medicine*, 31(14), 1502-1516.

Noah, T. L., Henderson, F. W., Wortman, I. A., Devlin, R. B., Handy, J., Koren, H. S., Becker, S. (1995) Nasal cytokine production in viral acute upper respiratory infection of childhood. *The Journal of Infectious Diseases*, 171(3), 584-592.

Turgeman, L., May, J., Ketterer, A., Sciulli, R., Vargas, D. (2015) Identification of readmission risk factors by analyzing the hospital-related state transitions of Congestive Heart Failure (CHF) patients. *IIE Transactions on Healthcare Systems Engineering*, 5(4), 255-267.

Vatamidou, E., Adan, I. J., Vlasiou, M., Zwart, B. (2014) On the accuracy of phase-type approximations of heavy-tailed risk models. *Scandinavian Actuarial Journal*, 2014(6), 510-534.

Xie, H., Chaussalet, T. J., 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.