Computers & Industrial Engineering 112 (2017) 675–689

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Diagnostic and modeling of elderly flow in a French healthcare institution

Fatima E. Hamdani [a,b,c], Malek Masmoudi [a,b,c,*], Ahmad Al Hanbali [d], Fatima Bouyahia [e], Abdellah Ait Ouahman [e]

[a] Université de Lyon, F-42023 Saint Etienne, France
[b] Université de Saint Etienne, Jean Monnet, F-42000 Saint-Etienne, France
[c] LASPI, F-42334 IUT de Roanne, France
[d] Faculty of Behavioural, Management and Social Sciences, Dep. Industrial Engineering and Business Information Systems, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
[e] Ecole Nationale des Sciences Appliquées – Marrakech, Laboratoire du Génie Electrique et Conception des Systèmes (LGECOS), Marrakech, Morocco

ARTICLE INFO

Article history:
Available online 7 May 2017

Keywords:
Markov model
Coxian
Statistical techniques
Likelihood estimation
Elderly flow
Length of stay

ABSTRACT

One of the highest priorities in the French health care system is to deal with the continuous growth of the percentage population older than 65 years, expected to reach 31% in 2030. This development poses enormous challenges to the operations of the health care system, especially, related to elder patients. The elderly flow in the hospital services is typically uncertain and subject to variations on the length of stay in each stage and on the path or sequence of stages followed by the patient. For that reason, we propose to model the patient flow in a hospital as a continuous-time Markov chain with an absorbing state representing the elderly discharge from the hospital. Three Markov chains are provided with different levels of details and computation complexity. The first model called aggregated provides a prediction of the length of stay per service, the second model called Coxian provides a reliable prediction of the total length of stay, and the third model called detailed provides a prediction of the length of stay per class of elderly. A classification of elderly based on multiple correspondence technique is considered before the application of the third model. Our models are fitted with the data collected from Roanne Hospital, a typical French health care structure.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

In France, it is expected that the rate of ageing, the percentage of the population older than 65 years, will increase from 21% in 2011 to 31% in 2030 and the number of dependent elderly in 2040 will exceed 1.2 million (European Commission, 2015). Facing with this continuous growth of the elderly population and their life expectancy increase, the efficiency improvement of the health care system operations becomes a social and an economic necessity for the French health institutions.

The empirical part of this paper considers the city Roanne located in the Rhône-Alpes region in La Loire department in France. The rate of elderly older than 75 years in Raonne is around 12% of the city population. This is higher than the departmental rate of La Loire, around 11%, the regional rate of Rhone Alpes, around 8%, and the national rate in France, around 9%. For this reason, the French health authority has decided to conduct an exploration project specifically designated for the Hospital Center of Roanne. In this paper, we contribute to the research question on the statistical analysis of elderly flow inside the Hospital Center based on realistic accurate models. This study is as a preliminary step to provide several decision support tools on the capacity and operational planning of the required resources in the hospital.

The modeling of the elderly pathway in health care institutions has been so far investigated in several papers in the literature. In Irvine, McClean, and Millard (1994), a two-state Markov model is proposed to capture the movement of patients through geriatric hospitals in which the first state is devoted to acute/rehabilitative patients and the second is for long-stay patients. A Coxian distribution incorporating a Bayesian belief network to integrate the patient data is developed in Marshall and McClean (2003, 2004). In Faddy and McClean (2005) the Coxian distribution is fitted to the data, using the maximum likelihood method, on the time spent by a geriatric patient in the hospital. This is done with a Coxian-type model with multiple absorbing states representing discharge to home, discharge to private nursing home, and the patient death.

* Corresponding author at: Université de Saint Etienne, Jean Monnet, F-42000 Saint-Etienne, France.
E-mail address: mas_malek@yahoo.fr (M. Masmoudi).

http://dx.doi.org/10.1016/j.cie.2017.05.009
0360-8352/© 2017 Elsevier Ltd. All rights reserved.

This model is applied to a real data set of stroke patients from the Belfast City Hospital (McClean, Garg, Barton, & Fullerton, 2010). In Spector, Mutter, Owens, and Limcango (2012), a model is proposed with two categories of care; home care and institutional care. The patients admitted to long-term care are more likely to be rejected or die, especially for people aged 75 years and older. A continuous-time Markov chain to model the length of stay of elderly people moving within and between residential home and nursing home is proposed in Xie, Chaussalet, and Millard (2005). In Xie, Chaussalet, and Millard (2006), a model-based approach is proposed to extract from a routinely health care dataset the high-level length-of-stay patterns of residents in long-term care. The modeling of the patient flow in the health care systems using a closed queuing network is given (Chaussalet, Xie, & Millard, 2006), and a semi-open queuing network in Xie, Chaussalet, and Rees (2007). The application of these models to a geriatric department in the UK shows their usefulness in helping the managers in gaining a better understanding of the patients flow within the department. To the best of our knowledge, no equivalent study has been so far considered within the French context.

In this paper, we provide three models to understand the dynamics of elderly pathways admitted in a French healthcare organization, namely the Roanne Hospital. These models are provided with different levels of details and computation complexity. The first model called Coxian provides a reliable prediction of the total length of stay, the second model called aggregated provides a prediction of the length of stay per service, and the third model called detailed provides a prediction of the length of stay per class of elderly. We collect data from the Roanne Hospital to fit these models. We use the data on the length of stays to calculate the stay length distribution per service and the patients' pathway to calculate the transition rates. The other qualitative collected data such as patients' frailty, pathologies, and lifestyles are used to first classify and then calculate the length of stay per type of elderly (Avila-Funes et al., 2008; D'Avignon & Mareschal, 1989; Loones, David-Alberola, & Jauneau, 2008).

This paper is organized as follows. Section 2 contains an overview of the French healthcare system and a description of the dataset collected from the Roanne Hospital center. Section 3 contains three models for the elderly pathway in hospital, namely; the aggregated model, the Coxian type model, and the detailed model. Note, for the latter model a classification technique of patients type is needed. In Section 4, we apply the models to Roanne Hospital center and discuss the results. In Section 5, we give possible applications of our models. We conclude that paper in Section 6 and give a future research direction.

2. French healthcare system for elderly people

The French healthcare system of elderly people consists of several institutions. On one hand, we have the extra-hospital institutions like the residential home, the nursing home, and the home support services with home and nursing service. On the other hand, we have the intra-hospital institutions with short, medium, and long stay. These (intra) institutions are organized as follows:

- The short stay supports the patients coming most of the time from emergency services. It comprises three services: SAU[1] for hosting and emergencies, UHCD[2] for short term hospitalization unit, and MCO[3] for medical obstetric surgery. The SAU and UHCD are the two services of the Emergency department (ED).
- The medium stay, referred to as SSR,[4] provides rehabilitation and reintegration.
- The long stay USLD[5] provides accommodation and care for elderly requiring constant medical supervision. This service can be outside hospital, which is the case for the Roanne Hospital Center.

In this study, we focus particularly on short and medium stay institutions within the Roanne Hospital Center. At the short stay, the elder patient is usually admitted at the SAU for consultation by an emergency doctor then, according to the elderly case, the patient is transferred to the short-term hospitalization unit UHCD to be under supervision and to ensure the recovery. If the elder patient requires additional care, she/he is transferred to medical surgery and obstetrics MCO unit. In the medium stay SSR unit, the patient is rehabilitated and reintegrated together with similar patients. The long stay offers housing and care for elderly who are independent in their daily life but requires a medical supervision. Typically, the majority of patients passes through SAU and receives a short-term and/or medium term hospitalization. In our case study, we focus on this stream of elderly as the project interest lies in the challenging pathway of patients admitted at the SAU.

For the data acquisition and in consultation with a team of gerontologists and clinical research doctors a form is jointly created. This form is used by various medical actors that follow the elderly from their admission in the SAU until their discharge from the hospital. The collected data are digitized and analyzed to get realistic figures representing the current situation in Roanne Hospital center. Note, based on the collected data we know which services have been visited by the patient and how much time is spent in each of them. The form is consistent and includes information on the elderly characteristics such as age, sex, area of origin, type of residence, and medical record home aids, reported comorbidities, admission requirements, hospitalization motives, and fragility. We also know the patient coming from the nursing home on the elder faller and elder with cognitive troubles. We were able to collect data of 241 elderly over 74 years following the intra-hospitalization process in the periods between February 23 and March 1, and March 30 and 5 April 2015. This category of patients represents 70% of all admissions in Roanne Hospital center arriving to the SAU in these periods. Among these patients, 45.2% are over 84 years old, 53.5% are females. These patients come from different geographic departments: 80.9% from La Loire, 10.8% from the Rhône, 7.8% from Rhône et Loire. Around 79% come from their private home, 15.5% from nursing home, and 5.8% from residential home. A percentage of 39% are living alone at home and the family environment is present in 85.9% of the cases. Several stochastic models are provided in the next section. The collected data is used to fit these models.

3. Modeling the pathway of elderly persons in hospital with phase type distribution

The time spent by an elderly in a service department in the hospital is typically uncertain and subject to variations. In fact, inside each service the patient can go through different phases (care's stages) depending on his medical case. The patient movement (transition) from a phase to another is random and varies from a patient to another. As a result, the process that keeps track of the location of an elder patient at a certain time is a stochastic process. Due to the limited number of phases in each service in a hospital the state space of this process is finite. In addition, after certain

[1] SAU: abbreviation of the French expression "Service d'Accueil des Urgences".
[2] UHCD: abbreviation of the French expression "Unité d'Hospitalisation de Courte Durée".
[3] MCO: abbreviation of the French expression "Médecine Chirurgie Obstétrique".
[4] SSR: abbreviation of the French expression "Soins de Suite et de Réadaptation".
[5] USLD: abbreviation of the French expression "Unité de Soins Longue Durées".

time the patient is discharged from the hospital. There are several reasons for patient discharge. In the case of Roanne Hospital Center, the reasons are patient death, need of medical supervision in USLD, release to the nursing home, residential home, home with support, or his home without support. The event of a patient discharge can be modelled as the transition to an absorbing (discharge) state. Therefore, the location of a patient in a hospital is an absorbing stochastic process.

In the literature, one class of the widely used stochastic processes is the Markov process. The key property of a Markov process is the so-called Markov property. According to this property, the probability of transition to a specific state depends on the current state of the process and, otherwise, is independent of the transitions history. This motivates us to assume that the location of a patient in a hospital is an absorbing Markov process that is continuous in time. In the following, we consider and analyze three absorbing Markov chains to represent patient length of stay: (1) Coxian model with no differentiation between the type of patients in which only the total length of stay in the hospital is modelled, (2) aggregated absorbing Markov model with no differentiation between the type of patients in the different services with each modelled as a set of phases, (3) a detailed absorbing Markov chain with patient differentiation depending on its type.

3.1. Coxian-type Markov chain model

A Coxian-type Markov chain is a special case of phase-type Markov chain. Namely, one-step transitions are only possible from state i to i + 1 or to absorbing (discharge) phase. It is well used in literature and provides a good prediction of the length of stay (LOS) (Marshall & McClean, 2004).

Let $\{X(t) : t \geqslant 0\}$ denote the continuous-time Markov chain model, with a finite state space {phase 1, phase 2,..., phase n, phase m}, phase m represents the absorbing state of discharge. Note, in this model we do not link the phases to the services. Therefore, it is only possible to give the estimate of the total length of stay in the hospital. In Fig. 1, we give the transition diagram of the Coxian-type Markov chain. Let β_{ij} denote the transition rate between phase i and phase j, $(i \neq j)$.

An elderly spends an exponentially distributed length of time in phase i with parameters, μ_i, $i = 1, 2, \ldots, n$, and once this time elapses she moves to the next phase or she leaves the system. The transition rate from phase i to phase i + 1 is $\beta_{ii+1}, i = 1, \ldots, n-1$, and the discharge rate is $\beta_{im}, i = 1, \ldots, n$. Let T the random variable of the time to reach the discharge state m. The distribution of T is a Coxian distribution, denoted as PH (π, D), whereD is the sub-matrix of transition rates restricted to the transient phases matrix of dimension n, and d the column vector of transition rates from the transient phases to the discharge state. The number of parameters in a Coxian distribution is equal to $2n - 1$. The corresponding density functions, $t > 0$, $f(t) = \pi \exp(Dt)d$, where π is a vector of probabilities defining in which phase the chain starts. In this model, we consider that all the patients start in phase 1, i.e., $\pi = (1, 0, \ldots, 0)$. The statistical approach that we use to estimate the parameters of the Coxian model is based on the maximum-likelihood estimator and the Expectation-Maximization Algorithm (EMA). For more details we refer to Section 3.2.

3.2. Aggregated Markov chain model

In this section, we consider an aggregated absorbing Markov chain model for the movement of elderly patient within and between the hospital's services. Generally, in France the hospitals centers are composed of the ED (SAU, UHCD), MCO, and SSR. The elderly patient starts at the ED, she spends a period of time after that she can move to another service for more intensive care for example to either MCO, SSR, or discharge from hospital. Being in MCO and SSR there is also a chance that the patient is discharged from the hospital. Note, the length of stay per service is modelled as a set of phases. Let X(t) represents the location of a patient at time t, i.e., in which phase the patient is currently present. Eventually, every patient will be discharged after a certain time. Therefore, the process $\{X(t) : t \geqslant 0\}$ is a continuous-time absorbing Markov process. The transient state space is finite and is represented as follows $\{phase1, phase2, \ldots, phase\ v\}$. Let α_{ij} denote the instantaneous transition rate between phase i and j, such that $\sum_j \alpha_{ij} = 0$. In our model (Fig. 2), we group the phases into clusters representing the hospital services. Therefore, the length of stay distribution per service consists of multiple phases. For ease of notation, we use E, M, and S, and D, to refer to ED, MCO, SSR, and Discharge, respectively. Given the structure of the AMC, the Generator matrix Q of the model can be given as

$$Q = \begin{bmatrix} Q_{EE} & Q_{EM} & 0 & Q_{ED} \\ 0 & Q_{MM} & Q_{MS} & Q_{MD} \\ 0 & 0 & Q_{SS} & Q_{SD} \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where Q_{ab} is the submatrix of the transition rates between the phases of clusters a and $b \in \{E, M, S, D\}$. Note Q_{aa} represents the transitions between the phases of cluster a. The length of stays in a is denoted by t_a.

The probability densities of the length of stay in cluster a (service a) and in cluster a before jumping to cluster b are given by

$$f_a(t) = -\pi_a \exp(Q_{aa} t_a) Q_{aa} 1,$$
$$f_{ab}(t) = \pi_a \exp(Q_{aa} t_a) Q_{ab} 1,$$

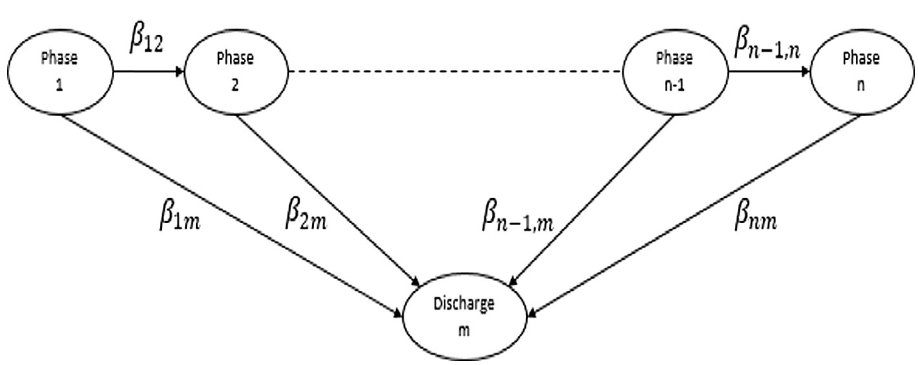


Fig. 1. Coxian Markov chain model for the intra-hospital path way of elderly patients.

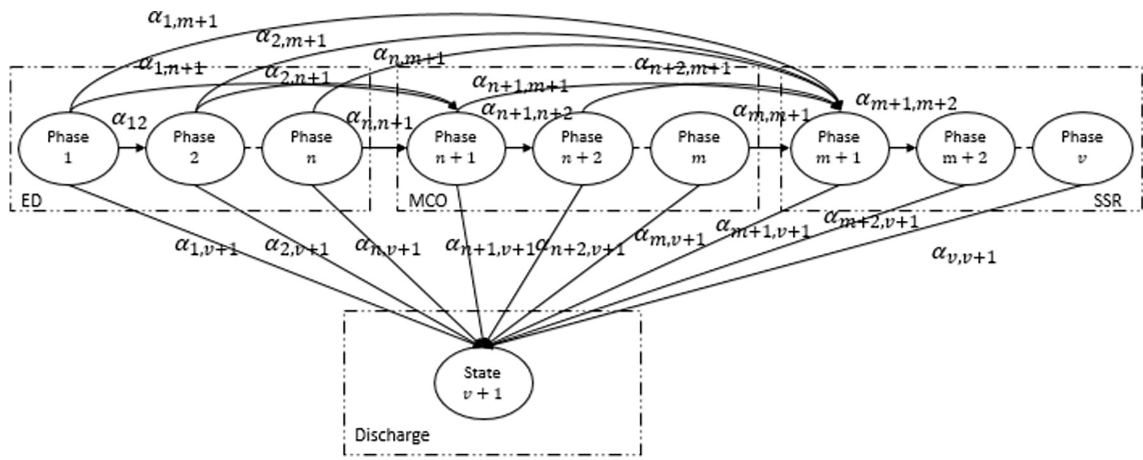


Fig. 2. Aggregated Markov model for the intra-hospital pathway of elderly patients.

where $\exp(Q_{aa}t_a)$ is the exponential of the matrix Q_{aa} multiplied by the time t_a, 1 is column vector of one's of appropriate length, π_a isas a row vector of the probabilities for the process to start in a phases in cluster a.The survival probability of the length of stay in cluster a (service a) is given by

$$\overline{F}_a(t) = \pi_a \exp(Q_{aa}t_a)1.$$

We use statistical approach based on the maximum-likelihood estimator and the Expectation-Maximization Algorithm (EMA), to estimate the parameters α_{ij}. This is because since no known closed-form solution this numerical search approach is typically used to solve the system of equations derived by setting to zero the partial derivatives of the likelihood function. Solving these equations with the Newton-Raphson method turns out to be very complicated, see (McLachlan & Krishnan, 2008).

The input parameters of the EMA are the sequence of services visited by the elderly patient $i, i = 1, \ldots, n$, and the duration in the services before leaving the hospital or data collection period stops whatever occurs first, t_l^i, $l \in \{E, M, S\}$. Let θ denote the vector of parameters α_{ij} to be estimated. The log-likelihood function of n samples is given as

$$\mathcal{F}(\theta) = \sum_{i=1}^{n} F_i(\theta),$$

where $F_i(\theta)$ is the log-likelihood function of the ith sample. Note, during the data collection period there are some patients who joined and are discharged from the hospital and others who joined and are still in the hospital at the end of the collection period. The former patients can follow one of the following sequences: E-D, M-D, S-D, E-M-D, E-M-S-D. However, the latter patients who are still in the hospital can follow one the followings: E-E, M-M, S-S, E-M, M-S, E-M-S. For example, M-M means a patient joined the ED before the start of the collection period and at the end of collection she was still in MCO. Therefore, the total time spent in MCO is greater than the observed time. This gives that the likelihood function of this patient is expressed as a function of the survival probability. The M-D sequence means a patient joined the ED before the start of the collection period and before the end of collection she was discharged. In this case, the likelihood function of this patient is expressed as a function to the density probability. In total there are eleven possible sequences which gives that $F_i(\theta)$ can be written as

$$F_i(\theta) = \sum_{k=1}^{11} \gamma_{i,k} \log(f_{i,k}(\theta)), \quad i = 1, \ldots, n$$

where

$$\gamma_k = \begin{cases} 1 & \textit{if patient i passes through the sequence k}\} \\ 0 & \textit{otherwise} \end{cases}$$

and, $f_{i,k}(\theta)$, $k = 1, \ldots, 11$, is the likelihood function of the different sequences. The sequences are ordered as follows: EE, ED (Emergency department-discharge), MM, MD, SS, SD, EM, EMD, EMSD, MS, EMS. The likelihood functions then read

$$\begin{aligned}
f_{i,1}(\theta) &= f_{EE}(\theta) = \pi_E \exp(Q_{EE}t_E^i)1 \\
f_{i,2}(\theta) &= f_{ED}(\theta) = \pi_E \exp(Q_{EE}t_E^i)Q_{ED}1 \\
f_{i,3}(\theta) &= f_{MM}(\theta) = \pi_M \exp(Q_{MM}t_M^i)1 \\
f_{i,4}(\theta) &= f_{MD}(\theta) = \pi_M \exp(Q_{MM}t_M^i)Q_{MD}1 \\
f_{i,5}(\theta) &= f_{SS}(\theta) = \pi_S \exp(Q_{SS}t_S^i)1 \\
f_{i,6}(\theta) &= f_{SD}(\theta) = \pi_S \exp(Q_{SS}t_S^i)Q_{SD}1 \\
f_{i,7}(\theta) &= f_{EM}(\theta) = \pi_E \exp(Q_{EE}t_E^i)Q_{EM} \exp(Q_{MM}t_M^i)1 \\
f_{i,8}(\theta) &= f_{EMD}(\theta) = \pi_E \exp(Q_{EE}t_E^i)Q_{EM} \exp(Q_{MM}t_M^i)Q_{MD}1 \\
f_{i,9}(\theta) &= f_{EMSD}(\theta) = \pi_E \exp(Q_{EE}t_E^i)Q_{EM} \exp(Q_{MM}t_M^i)Q_{MS} \exp(Q_{SS}t_S^i)Q_{SD}1 \\
f_{i,10}(\theta) &= f_{MS}(\theta) = \pi_E \exp(Q_{MM}t_M^i)Q_{MS} \exp(Q_{SS}t_S^i)1 \\
f_{i,11}(\theta) &= f_{EMS}(\theta) = \pi_E \exp(Q_{EE}t_E^i)Q_{EM} \exp(Q_{MM}t_M^i)Q_{MS} \exp(Q_{SS}t_S^i)1
\end{aligned}$$

The objective is to maximize the log-likelihood function of n individual using the EMA which is based on an iterative procedure (Asmussen, Nerman, & Olsson, 1996). The first iteration of the algorithm starts with an initial guess of the parameters vector θ^0. Based on θ^0, we try to estimate a new value of the parameters θ^1. This is done in a two steps: Expectation step (E-Step) and Maximization step (M-Step). In the E-step, given θ^0 we compute the expectation of $\mathcal{F}(\theta)$. In the M-Step, we find θ^1, a new estimate of θ, by maximizing the expectation of $\mathcal{F}(\theta)$. These steps are repeated until the algorithm converges. For more details we refer the reader to Asmussen et al. (1996) and Olsson (1998).

3.3. Detailed absorbing Markov chain model

In this section, we consider a detailed absorbing Markov chain to model the LOS of patients taking into account their characteristics. The Markov models described in the previous two sections assume that the transition rates are the same for all elderly patients. In reality, these rates are dependent on the degree of illness and fragility of a patient. Therefore, the classification of elderly patients according to their features gives better insight in modeling the patient flow in the hospital. Classification can be performed using the expertise of geriatric doctors and the data analytics method.

In Fig. 3, we show the transition diagram of the detailed Markov chain model for a single class. For this detailed Markov chain, we consider that the LOS in a service depends on the patient class $f, f = 1, \ldots, \mathrm{F}$, and the type of care pathway $p, p = 1, \ldots, \mathrm{P}$, she/he follows. For P = 3, we can differentiate between short pathway stay consisting of ED, mid-stay of ED and MCO, and long stay of ED and MCO and SSR. Note, we can have multiple phases per pathway. Therefore, in this analysis the number of services is kept equal to three: ED, MCO, and SSR. However, we impose that the stay in a service is dependent on the patient class and his pathway. A path p, $p = 1, \ldots, \mathrm{P}$, is followed with a probability p_{fp} and has s_p phases. We noted the transition rate between phases i and j, in class f, by α_{ij}^{fp} .Then, the transition rate to be discharge from phase i, by α_{id}^{fp}.

Given the structure of the Markov chain model in Fig. 3, the corresponding infinitesimal generator matrix between the transient states Q_f is as follows:

$$Q_f = \begin{bmatrix} G_{f1} & 0 & \ldots & 0 & 0 \\ 0 & G_{f2} & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & G_{fP} \end{bmatrix},$$

where G_{fp} for $\mathrm{p} = \{1, \ldots, \mathrm{P}\}$ are the generator matrices of the transient path p for a class f. Let the columns vectors h_{fp} denote the transition rates from the transient phases of path p to the discharge.

We denote ψ the class of distribution formed by a mixture of a set Coxian distribution. Ψ is a linear convex combination of the density functions $g_{f1}(t), g_{f2}(t), \ldots, g_{fP}(t)$ for a class f, which represent, respectively, the density functions of the first, second, and so on, until the P^{th} path for each class f. The density function of is ψ given by the expression:

$$\psi_f(t) = \sum_{p=1}^{P} p_{fp} g_{fp}(t), \text{ with } g_{fp}(t) = \pi_{fp} \exp(G_{fp} t) h_{fp};$$

$$\sum_{p=1}^{P} p_{fp} = 1;\; p_{fp} > 0,\; t > 0,$$

where π_{fp}, $\mathrm{p} = \{1, \ldots, \mathrm{P}\}$, denotes the initial probabilityvector of the path p in the class f. The density function of an arbitrary type of patients $\varphi(t)$ is alinear convex combination of the densities $\psi_f(t)$ given as

$$\varphi(t) = \sum_{f=1}^{F} \sum_{p=1}^{P} p_{fp} g_{fp}(t); \quad \sum_{p=1}^{P} p_{fp} = 1; p_{fp} > 0. t > 0$$

To estimate the parameters of the detailed aggregated Markov model, we apply the EMA to each class. However, before applying the detailed absorbing Markov model, we need to classify patients according to their characteristics; patient's demography, social circumstances, fragility; and comorbidities into a set of different classes.

4. Applications and results

Three Markov chain models are provided in this paper to capture the flow of elderly within and between hospital services. In this section, we apply these three modeling approaches to the Roanne Hospital Center case. Real data of 241 elderly from Roanne Hospital Center are considered to parameterize these three models and show their accuracy.

4.1. Application of the Coxian model

The length of stay of elderly patients in the hospital is fitted using a Coxian distribution. The procedure adopted to find the best fit is sequential in nature consisting of increasing the number of phases until there is very little improvement in the fit quality. The best compromise between model complexity and goodness of fit is obtained with four phases. Table 3 shows the results: the number of phases by applying Akaike information criterion (AIC) and Bayesian information criterion (BIC). According to this table, we find that for four phases AIC and BIC criteria reach their smallest value Therefore, the best fit is obtained with four phases. Note, in this case 'phase five' will represent the discharge state. In Fig. 4, we show the transition rate diagram of the Coxian model with four phases (note that Phase five is the absorbing state which represents the discharge state).

The first phase of the Coxian model is considered the short acute stay with an average LOS of $1/(\beta_{12} + \beta_{15})$ = 0.38659 days

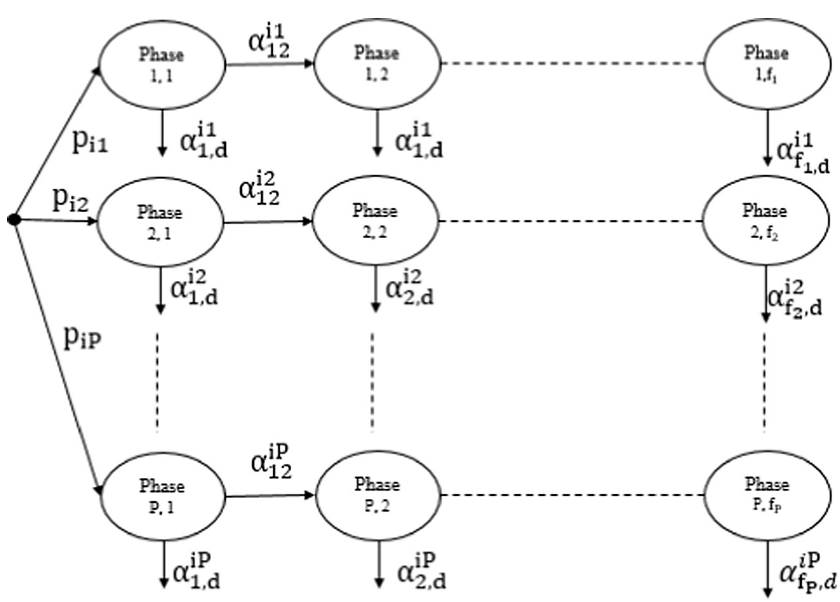


Fig. 3. Markov Chain model for class f of elderly patients.

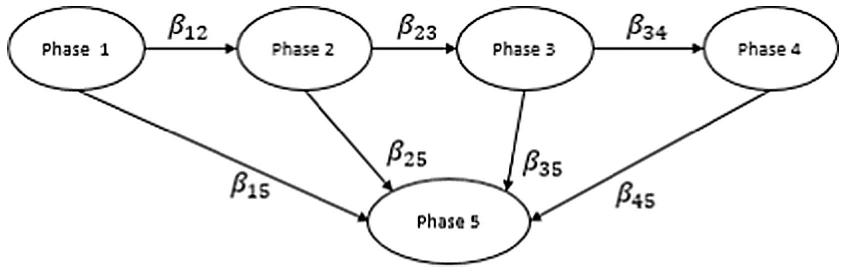


Fig. 4. Transition diagram of the Coxian model fitted to the data of Roanne hospital (phase 5 represents the discharge state).

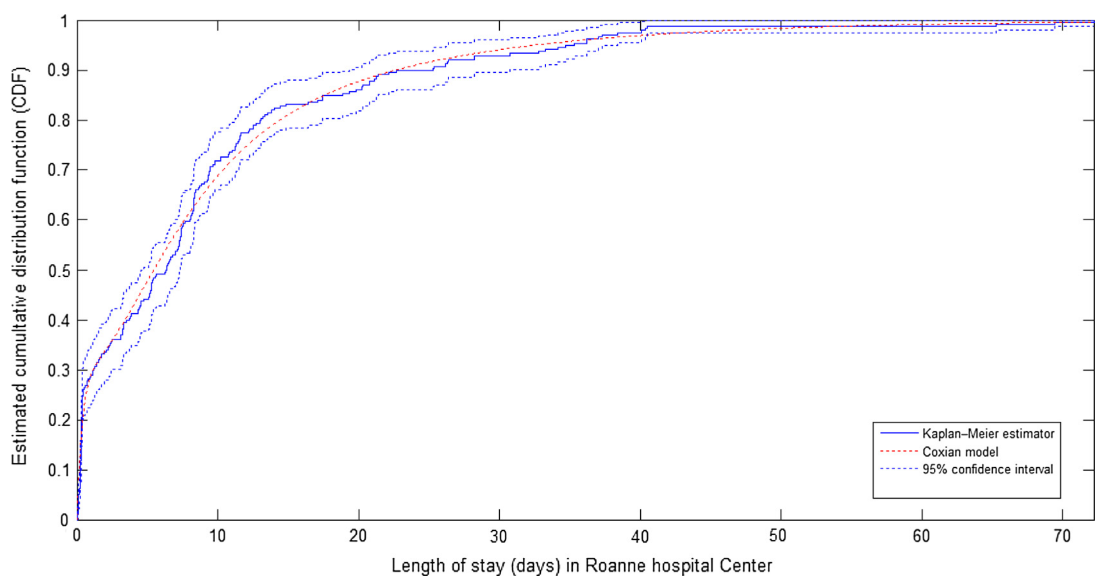


Fig. 5. Kaplan–Meier estimator (complete blue line, 95% confidence blue dotted lines) and Coxian model (red dotted line) cumulative distribution: length of stay of elderly in Roanne hospital. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(9.27 h), about $\beta_{12}/(\beta_{12}+\beta_{15}) = 69\%$ of elderly-patients move to the second phase to pass an average of 3.75 days $(1/(\beta_{23}+\beta_{25})$. The third phase is considered as the second phase in the medium stay, elderly pass on an average an LOS of 3.75 days, 32% $(=\beta_{34}/(\beta_{34}+\beta_{35}))$ of elderly move to the last phase and the rest leave the hospital. The last phase is considered as the long stay, the elderly-patient pass an average 15.67 days $(=1/\beta_{45})$, before being discharged from the hospital. The average total time in hospital is 9.09 days.

In Fig. 5, we show the cumulative distribution of the LOS of the fitted Coxian model to the data of Roanne hospital compared to the non-parametric Kaplan-Meier estimator (Kaplan & Meier, 1958). The root mean square error between the model and the estimator is equal to 0.0252. We conclude the Coxian model is accurate and gives interesting results at the global level on the length stay in the hospital. However, the model does not provide much detail on the length of stay per service. In the following section, we will analyze the aggregated model with more detailed view of the LOS of a patient per service.

Table 1
The number of phases of the Coxian model and its transition rates for the elderly in Roanne Hospital Center.

Number of Phases	AIC	BIC	Transition rate	Value
P = 1	3425.77525	3427.05337	β_{12}	1.796553
P = 2	3268.505	3271.06123	β_{15}	0.790173
P = 3	3254.70562	3258.53997	β_{23}	0.266546
P = 4	**3239.97882**	**3242.53505**	β_{25}	0
P = 5	3247.97934	3253.0918	β_{34}	0.085511
P = 6	3247.41537	3255.08406	β_{35}	0.181035
P = 7	3256.21569	3265.1625	β_{45}	0.063798

Bold values represent the smallest AIC and BIC criteria.

4.2. Application of the aggregated absorbing Markov model

The Aggregated Markov chain model was fitted to the Roanne hospital data in two septs. In the first step, we determine the num-

Table 2
The number of phases per service department of the aggregated Markov model applied to Roanne Hospital Center.

Service	Num. of phases	AIC	BIC
ED	p = 1	479.403135	480.626204
	p = 2	367.268965	369.640009
	p = 3	342.335675	345.777495
	p = 4	**271.811699**	**277.154653**
	p = 5	279.747059	284.180273
MCO	p = 1	2308.07143	2309.20248
	p = 2	2178.58116	2180.71948
	p = 3	2158.90451	2161.92068
	p = 4	**2156.57172**	**2160.33034**
	p = 5	2166.15612	2170.51544
	p = 6	2161.31458	2166.12612
SSR	p = 1	302.928334	304.206449
	p = 2	288.847745	291.403975
	p = 3	**286.503041**	**290.337385**
	p = 4	287.915359	293.027818
	p = 5	289.961542	296.352115

Bold values represent the smallest AIC and BIC criteria.

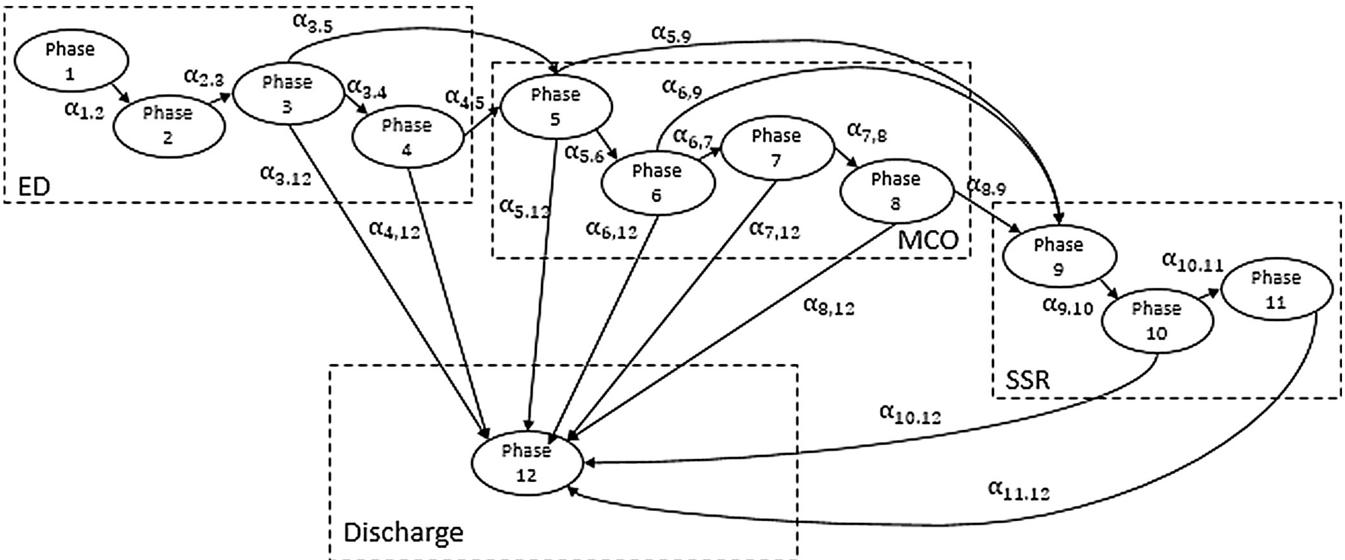


Fig. 6. Transition diagram of the aggregated Markov model fitted to the data of Roanne hospital center (Phase 12 represents the discharge state).

ber of phases in each service. The chosen number of phases is the one with the smallest AIC and BIC criteria (see Table 1 for the results). We find that we need four phases for the LOS in the ED service, four phases for the LOS in the MCO service, and three phases for the LOS in the SSR service. In the second step, we proceed to fit the overall length of stay (LOS) with the mixture exponential distribution to determinate the transition rate of the phases (see Table 2 for the results). In Fig. 6, we show the transition rate diagram of the aggregated Markov model (note that phase 12 represents the discharge state).

The best compromise between the AIC and BIC for the ED service was obtained with 4 phases. The average LOS is about 0.10 days respectively, in the first phase $(= 1/\alpha_{1,2})$, second $(1/\alpha_{2,3})$ and the third $(= 1/(\alpha_{3,4} + \alpha_{3,5} + \alpha_{3,12}))$ phases of ED. In the third phase, 33.51% $(\alpha_{3,4}/(\alpha_{3,4} + \alpha_{3,5} + \alpha_{3,12}))$ and 66.38% $(= \alpha_{3,5}/(\alpha_{3,4} + \alpha_{3,5} + \alpha_{3,12}))$ are the probability to move to phase four in ED and to the first phase in MCO (Phase 5), respectively. The rest, 0.11%, leave the hospital. In the last phase of ED, 0.34% $(= \alpha_{4,12}/(\alpha_{4,12} + \alpha_{4,5}))$ leave the hospital and the rest pass to MCO, i.e., to phase 5. The average LOS in ED is 0.56 days $(= 1/\alpha_{1,2} + 1/\alpha_{2,3} + 1/(\alpha_{3,4} + \alpha_{3,5} + \alpha_{3,12}) + \alpha_{3,4}/[(\alpha_{3,4} + \alpha_{3,5} + \alpha_{3,12}) * (\alpha_{4,12} + \alpha_{4,5})])$.

For MCO we have 4 phases. In the first phase of MCO, 65% $(= \alpha_{5,6}/(\alpha_{5,6} + \alpha_{5,9} + \alpha_{5,12}))$, 1% $(= \alpha_{5,9}/(\alpha_{5,6} + \alpha_{5,9} + \alpha_{5,12}))$, and

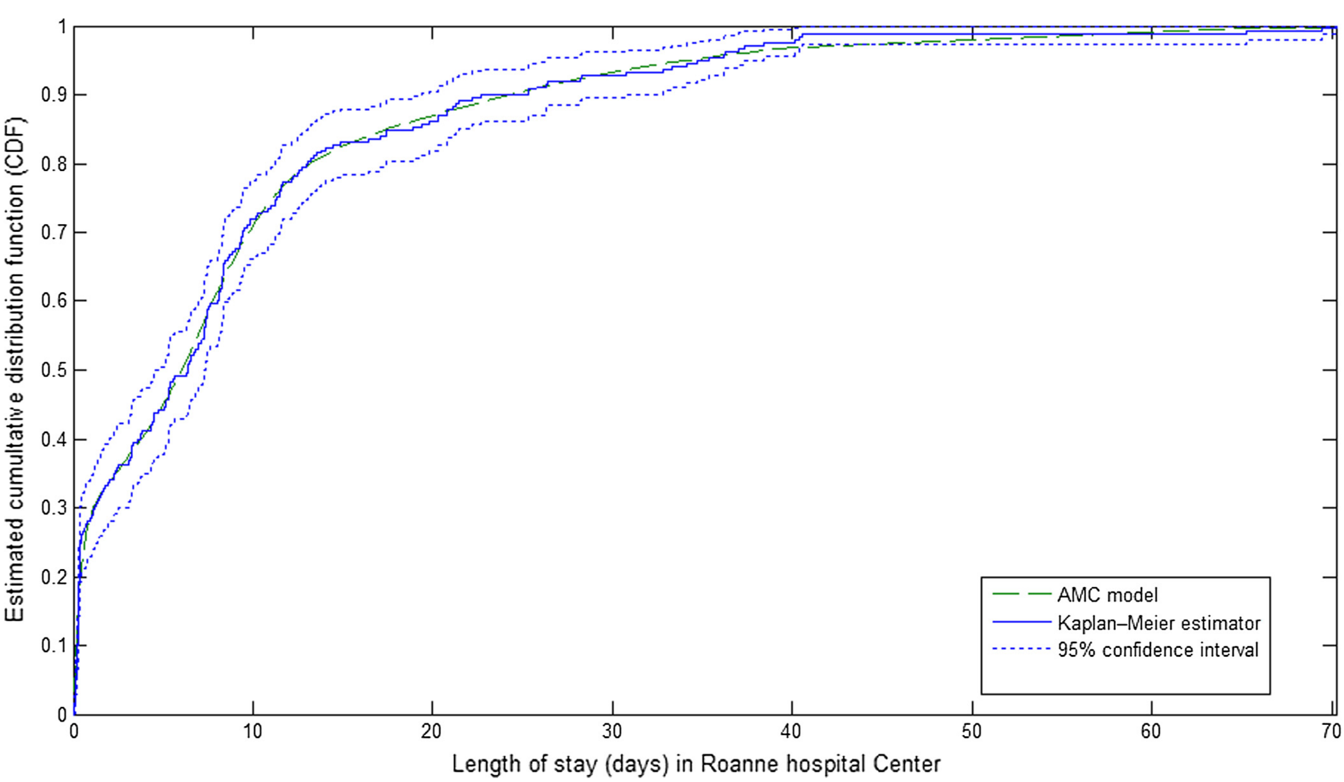


Fig. 7. Kaplan–Meier estimator and aggregated Markov model cumulative distribution: length of stay of elderly in Roanne hospital.

Table 3
Estimated parameters of the aggregated Markov model in Roanne Hospital Center.

Service	Average Los (days)	Transition rate	Value
ED	0.10004385	$\alpha_{1,2}$	9.995617
	0.10004385	$\alpha_{2,3}$	9.995617
	0.10004385	$\alpha_{3,4}$	3.349766
		$\alpha_{3,5}$	4.386262
		$\alpha_{3,12}$	2.259589
	0.78475750	$\alpha_{4,5}$	0.841024
		$\alpha_{4,12}$	0.433255
MCO	6.36719621	$\alpha_{5,6}$	0.137342
		$\alpha_{5,9}$	0.002891
		$\alpha_{5,12}$	0.072285
	1.26019498	$\alpha_{6,7}$	0.616332
		$\alpha_{6,9}$	0.023112
		$\alpha_{6,12}$	0.154083
	1.26017433	$\alpha_{7,8}$	0.168627
		$\alpha_{7,12}$	0.624914
	1.25945695	$\alpha_{8,9}$	0.126031
		$\alpha_{8,12}$	0.667962
SSR	9.42693653	$\alpha_{9,10}$	0.106079
	10.5260942	$\alpha_{10,11}$	0.031667
		$\alpha_{10,12}$	0.063335
	8.10696928	$\alpha_{11,12}$	0.123351

34% ($= \alpha_{5,12}/(\alpha_{5,6} + \alpha_{5,9} + \alpha_{5,12})$),) are, respectively, the probability to move to the second phase in MCO, SSR and discharge from hospital. In the second phase of MCO, Phase 6, with an average LOS of 1.26 days ($= 1/(\alpha_{6,7} + \alpha_{6,9} + \alpha_{6,12})$) and probability 80%, 3%, and 17%, respectively, to move to the third phase in MCO (Phase 7), pass to the first phase in SSR (Phase 9), and to discharge from hospital. Then, in the third phase in MCO 77% move to the last phase (Phase 8), and 23% are discharged from the hospital. In the last phase of MCO 16% move to the first in SSR (Phase 9) and 84% leave the hospital. The average LOS in MCO is about 10 days and 28 days in SSR.

The aggregated continuous Markov chain model gives interesting results and provides a good tool to model the LOS, but it does not differentiate between the types of patients. To examine the model fit quality, in Fig. 7 we plot the LOS cumulative distribution of the AMC model fitted to the data of Roanne Hospital center using the estimated parameters in Table 3 and the non-parametric Kaplan-Meier estimator with 95% confident interval. The close agreement between the AMC and the non-parametric Kaplan and Meier estimator shows the model accuracy and its performance to capture the behavior of the LOS in the hospital. The root mean square error between the AMC model and the estimator is equal to 0.0191. We conclude that AMC model is more accurate than the Coxian model.

4.3. Application of the detailed absorbing Markov model

In our case study a part of the collected data are qualitative. We use the Multiple Correspondence Analysis (MCA) (Greenacre & Blasius, 2006) to convert the qualitative data to quantitative values. This operation must be done before the classification. The Hierarchical Agglomerative Clustering (HAC) (Greenacre & Blasius, 2006; Guojun, Chaoqun, & Jianhong, 2007), a non-monitored method, is then applied to classify the elderly patient representing a similar behavior, especially, based on their fragility criteria. The main idea of this method is to calculate a table of Euclidean distance between individuals (patients) in order to define the existing similarities among the different criteria of the elder patient fragility. The algorithm starts with the attribution of an elderly patient to a class. Then, for each iteration, it starts to constitute number of classes by grouping the nearest individuals from a partition using the minimum jump (WARD criteria). This

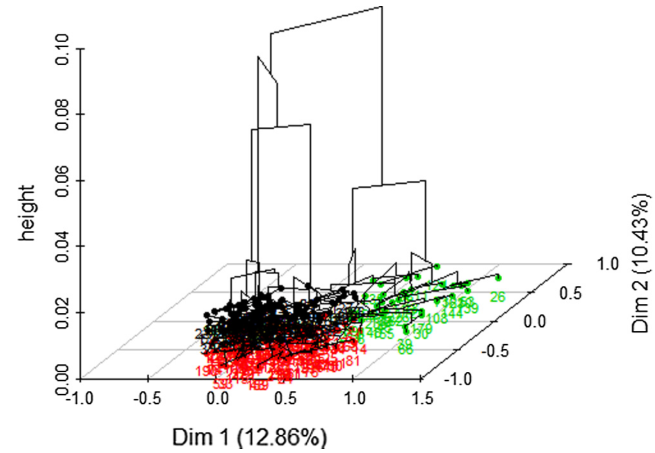


Fig. 8. Hierarchical clustering of elderly patients into three classes using MCA and HAC in the statistical R tool.

criterion allows the detection of similarities among groups of individuals. The algorithm stops after getting one class.

From MCA and HAC we get a hierarchical tree that represents the different partitions. Fig. 8 shows the results; it is divided on three classes depending on the inertia gain. The resulted eigenvalue gave us the rate of inertia associated to each dimension; the first Dim1 explains 12.86% of global inertia, and the second 10.43%. The first class represents 48% from the global data, the second 39%. Finally, the third one represents 13%. Multivariate logistic regression is exploited to explain the classification results and understanding the patient's features effect on this classification.

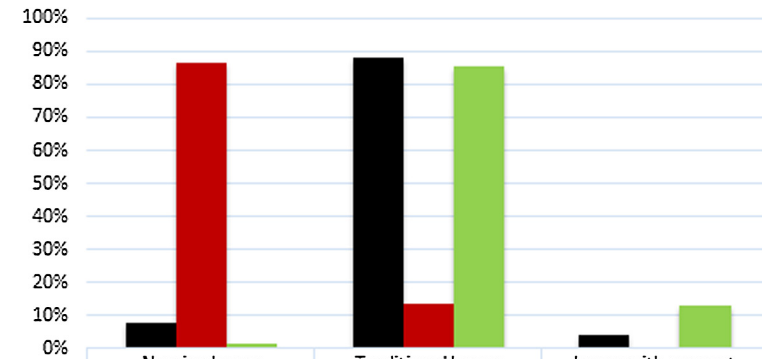
We apply the procedure to fit all the variables integrated in the classification as explicative variables and the clustering results (class 1, class 2, or class 3) as dependent variable. The p-value of the likelihood ratio test for an elderly, is considered to be non-significant if $p > 0.20$. Following the Table 4, we can summarize according to Khi2 tests probability, the influence of variables in the clustering decision. The variables age influence in the clustering decision for the class 1 and the class 2. The patients belonging to class 3 tend to be elder than those in class 1: the category D (age $\geq$ 90) with 11%, C (85 $\leq$ age < 90) with 25%, B (80 $\leq$ age < 85) with 30%, and A (75 $\leq$ age < 80) with 34%. For class 3, the category D presents > 33% and A around 19%.

The residence place is the modality the more important in the clustering. The totality of persons belonging to class 3 lives in traditional home. However, no one is in the nursing home. Class 1 has the majority from home with support, but has also some persons from the nursing home (see Fig. 7). We note that the totality of persons belonging to class 3 is living in a traditional home. The majority of persons in Class 1 are from home with support, but some are from the nursing home (see Fig. 9). Persons in Class 2 are the elderly with behavioral problems, cognitive difficulties, and neurology problems. The majority in class 2 lives alone without family (92.30%). The persons belonging to class 3, have neurology problem, Polypharmacy, idling problems, acute renal failure, or are readmitted in hospital.

We now apply the detailed absorbing Markov model to the case study. We consider 3-paths per class of elderly. The first one corresponds to the short stay and represents the ED service. The second path corresponds to the middle term stay, composed of a short stay in ED, and then followed by a middle stay in the MCO service. Finally, the last path, long stay, consists of the short stay, the middle stay, and the SSR service. The length of stay of elderly patients in the hospital is modelled using a mixture of a set of Coxian models. The procedure adopted to find the best fit is sequential in nat-

Table 4
Cross tabulation of classification and regression results.

	Variables	Modality	Effectifs	%	Class 1			Class 2			Class 3		
					Pr > Wald	Pr > Khi^2	%	Pr > Wald	Pr > Khi^2	%	Pr > Wald	Pr > Khi^2	%
Applicant's demography	Age	$75 \leq a < 80$	64	26.556	0.09510313		33.588	0.911		7.692	0.115		19.048
		$80 \leq b < 85$	68	28.216		0.701	29.771		0.904	30.769		0.726	23.810
		$85 \leq c < 90$	62	25.726		0.374	25.191		0.553	34.615		0.484	23.810
		$90 \leq d$	47	19.502		0.061	11.450		0.578	26.923		0.059	33.333
	Gendre	Female	129	53.527	0.251		47.328	0.270		73.077	0.697		60.317
		Male	112	46.473		0.251	52.672		0.270	26.923		0.697	39.683
Social circumstances	Lieu Résidence	Home with support	14	5.809	0.456		3.817	0.000		15.385	<0.0001		12.698
		Traditional home	190	78.838		0.354	89.313		0.672	84.615		0.258	85.714
		Ehpad	37	15.353		0.211	6.870		0.005	7.692		<0.0001	1.587
	Living mode	With children	10	4.149	0.137		6.107	0.348		92.308	0.212		3.175
		Not alone	91	37.759		0.154	48.092		0.151	3.846		0.256	25.397
		Alone	140	58.091		0.615	45.802		0.188	96.154		0.795	71.429
	Social environment	Fragile	33	13.693	0.282		9.924	0.462		96.154	0.176		26.984
		Present	207	85.892		0.167	90.076		0.236	3.846		0.081	73.016
		Nonexistent	1	0.415		0.272	87.786		0.977	30.769		0.685	90.476
Fragility	Cancer	No	215	89.212	0.931		12.214	0.985		69.231	0.957		9.524
		Yes	26	10.788		0.931	97.710		0.985	88.462		0.957	80.952
	Neurology-problem	No	206	85.477	0.001		2.290	0.013		11.538	0.001		19.048
		Yes	35	14.523		0.001	98.473		0.013	53.846		0.165	77.778
	Mal Rhum.	No	219	90.871	0.004		1.527	0.244		46.154	0.003		22.222
		Yes	22	9.129		0.004	68.702		0.244	69.231		0.003	41.270
	Acute renal failure	No	144	59.751	0.007		31.298	0.775		30.769	0.006		58.730
		Yes	97	40.249		0.007	86.260		0.775	26.923		0.006	61.905
Comorbidity	Falls	No	190	78.838	0.010		13.740	0.937		73.077	0.009		38.095
		Yes	51	21.162		0.010	72.519		0.937	26.923		0.009	25.397
	Marches (steps)problems	No	130	53.942	0.004		27.481	0.238		73.077	0.007		74.603
		Yes	111	46.058		0.004	94.656		0.238	76.923		0.007	77.778
	Cognitifs-troubles	No	199	82.573	0.000		5.344	0.004		23.077	0.098		22.222
		Yes	42	17.427		0.000	88.550		0.004	61.538		0.098	82.540
	Humor-troubles	No	206	85.477	0.929		11.450	0.332		38.462	0.583		17.460
		Yes	35	14.523		0.929	99.237		0.332	69.231		0.583	96.825
	Behavioral problems	No	225	93.361	0.109		0.763	0.024		30.769	0.791		3.175
		Yes	16	6.639		0.109	88.550		0.024	92.308		0.791	61.905
	Polypharmacy	No	189	78.423	0.003		11.450	0.257		7.692	0.021		38.095
		Yes	52	21.577		0.003	89.313		0.257	7.692		0.021	85.714
	Readmission	No	215	89.212	0.098		10.687	0.144		30.769	0.175		14.286
		Yes	26	10.788		0.098	33.588		0.144	34.615		0.175	19.048



	Nursing home	Traditional home	home with support
Class 1	8%	88%	4%
Class 2	86%	14%	0%
Class 3	2%	85%	13%

Fig. 9. The residence place modality for each class (class 1 in black, class 2 in red, and class 3 in green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ure consisting of increasing the number of phases until there is very little improvement in the fit quality using the AIC and BIC. The best compromise between model complexity and goodness of fit are indicated in Table 5. In the class 1, we find the three pathways, with four phases, three phases and three phases, respectively. However, the patients belonging to the second class have only the possibility to pass by two paths; the short stay and the middle stay. Class 3 has three pathways with three phases for first two and two phases for the last.

4.3.1. Class 1 results

Following the maximum likelihood and the EMA in Table 6 we show the transition rates and the probability of class 1 patients to follow path p (=1, 2, 3). In Fig. 10, we show the transition rate diagram of class 1 of the detailed Markov chain. Note, the vertical transitions pointing to the bottom represent the transitions to the discharge state.

Based on the detailed Markov model we find for class 1 that 57% of all admissions pass through the first path corresponding to the short stay, then 33.3% $(= \alpha^{11}_{15}/\lambda^{11}_1)$ will be discharged after passing an average LOS of 0.44 days $(= 1/\lambda^{11}_1)$, and the rest move to the second phase, 'Phase 1,2', See Fig. 10. Note, $\lambda^{11}_1 = \alpha^{11}_{12} + \alpha^{11}_{15}$. In 'Phase 1,2', the patients spend about 1.23 h (=$1/\lambda^{11}_2$), 0.125% $(= \alpha^{11}_{25}/\lambda^{11}_2)$ leave hospital, 0.875% $(= \alpha^{11}_{23}/\lambda^{11}_2)$ pass the third one, 'Phase 1,3', to spend the same average. Then, 7.14% $(= \alpha^{11}_{35}/\lambda^{11}_3)$ discharge and the rest 92.86%, move to the last one, 'Phase 1,4', to spend an average of 0.31 days $(= 1/\lambda^{11}_4)$ before being discharged. In the second path, the probability to follow it is 37% and a patient spends 6.34 days (=$1/\lambda^{12}_1$) on average in the first phase, 23% discharge, 77% pass to the second phase, 'Phase 2,2', and she spends an average of 2.21 days $(= \lambda^{12}_2)$. Then, 63.15% move to last one phase, 'Phase 2,3', to spend an average LOS of 2.25 days. For the third path, the long stay, 6% move through it and pass an average of 39 days $(= 1/\lambda^{13}_1 + 1/\lambda^{13}_2 + 1/\lambda^{13}_3)$ before leaving the hospital.

In Fig. 11, we plot the LOS cumulative distribution function of the LOS of class 1 patients of the detailed model fitted to the data of Roanne hospital center in comparison with the non-parametric Kaplan and Meier estimator, the Coxian Model, and the AMC model. According to Fig. 11, all the three models fit well the LOS of class 1 patients. The root mean square error of the cumulative function of class 1 detailed model is 0.0256, AMC model is 0.0318, and of Coxian model is 0.0351. In this case, we conclude that the class 1 detailed model is the most accurate, then the AMC model, and the Coxian model is the least accurate.

4.3.2. Class 2 results

Following the maximum likelihood and the EMA in Table 7 we show the transition rates and the probability that a class 2 patient follows path $p(= 1, 2)$. Note, the probability that an elderly belongs to class 2 is 39%. In Fig. 12, we show the transition rate diagram of class 2 patients of the detailed Markov chain. Note, the vertical transitions pointing to the bottom represent the transitions to the discharge state.

The probability to follow path 1 is 48% with an average LOS of 0.30 days (=$1/\lambda^{21}_1$) and the rest, 52%, pass through the medium stay, path 2. A patient spends an average of 4 days (=$1/\lambda^{22}_1$) in the first phase, 'Phase 2,1'. Then, the patient moves to the second phase, 'Phase 2,2', to spend the same average, but in this phase there is 36% $(= \alpha^{22}_{24}/(\alpha^{22}_{24} + \alpha^{21}_{24}))$ chance to leave the system. In the last phase, 'Phase 2,3', the average LOS is also 0.30 days.

In Fig. 13, we plot the LOS cumulative distribution function of class 2 patients of the detailed model fitted to the data of Roanne hospital center in comparison with the non-parametric Kaplan and Meier estimator, the Coxian Model, and the AMC model. According to Fig. 13, Coxian and AMC models do not fit well the LOS of class 2 patients. The root mean square error of the cumulative function of class 2 detailed model is 0.0559, Coxian model is 0.1439, and of AMC model is 0.1455. In this case, we conclude that the class 1 detailed model is the most accurate, then the Coxian model, and the AMC model is the least accurate. Note, the difference between the Coxian and AMC models is small and they both overestimate LOS class 2 patients.

4.3.3. Class 3 results

Following the maximum likelihood and the EMA in Table 8 we show the transition rates and the probability that a class 3 patient follows path $p(= 1, 2, 3)$. Note, the probability that an elderly belongs to class 3 is 13%. In Fig. 12, we show the transition rate diagram of class 2 patients of the detailed Markov chain. Note, the vertical transitions pointing to the bottom represent the transitions to the discharge state.

The probability to follow path 1 is 21.31% and the patients pass an average 0.14 days (=1/ λ^{31}_1) in the first phase, 'Phase 1,1', after that they move to the second phase, 'Phase 1,2', for passing the same average. The discharge rate from the second phase is

Table 5
The number of phases for elderly pathway per class of detailed Markov model applied to Roanne Hospital.

Class	Path		AIC	BIC
Class1	Path 1	p = 1	121.476065	122.514179
		p = 2	92.9699078	94.6928037
		p = 3	93.8147478	95.8230048
		p = 4	**70.8464782**	**70.7151661**
		p = 5	82.4149995	83.5674775
		p = 6	92.1075815	90.0017526
	Path 2	p = 1	1143.42407	1144.53317
		p = 2	1084.01193	1085.98845
		p = 3	**1076.74946**	**1079.33008**
		p = 4	1078.84098	1081.73805
		p = 5	1085.47491	1088.37341
		p = 6	1092.98284	1093.65153
	Path 3	p = 1	176.677445	177.494021
		p = 2	168.667674	169.405722
		p = 3	**168.408834**	**167.576511**
		p = 4	173.339908	168.166652
		p = 5	185.656067	170.04664
		p = 6	217.099283	172.767971
Class2	Path 1	**p = 1**	**14.5874952**	**15.6751337**
		p = 2	32.8072028	34.7076944
		p = 3	41.6465419	44.0571571
		p = 4	48.4733875	51.0595304
		p = 5	54.6339467	57.0245200
	Path 2	p = 1	203.128477	204.180177
		p = 2	191.268861	193.040777
		p = 3	**189.952399**	**192.072457**
		p = 4	192.835114	194.883743
		p = 5	197.210484	198.712169
		p = 6	202.571921	202.984795
Class3	Path 1	p = 1	24.3508503	25.6289650
		p = 2	23.5487308	26.1049601
		p = 3	**22.464494**	**26.298838**
		p = 4	29.4599269	34.5723856
		p = 5	33.4627591	39.8533324
		p = 6	41.5231331	49.1918210
	Path 2	p = 1	658.674314	659.952429
		p = 2	624.385967	626.942196
		p = 3	**622.030099**	**625.864443**
		p = 4	624.603183	629.715642
		p = 5	629.453044	635.843617
		p = 6	631.890631	639.559319
	Path3	p = 1	135.694577	136.972691
		p = 2	**131.158493**	**133.714723**
		p = 3	131.691459	135.525803
		p = 4	126.094515	128.650744
		p = 5	133.437173	138.549631
		p = 6	137.304272	143.694845

Bold values represent the smallest AIC and BIC criteria.

74.86% (=$\alpha_{24}^{31}/(\alpha_{24}^{31}+\alpha_{23}^{31})$). The rest pass to the last phase to pass an average of 1 day. The probability to follow path 2, the medium stay, is 72.13% and on average a patient spends 11.57 days. Finally, 8.19% patients spend an average LOS of 43.74 days in path 3, the long stay, before being discharged (see Fig. 14)

In Fig. 15, we plot the LOS cumulative distribution function of class 3 patients of the detailed model fitted to the data of Roanne hospital center in comparison with the non-parametric Kaplan and Meier estimator, the Coxian Model, and the AMC model. According to Fig. 15, Coxian and AMC models do not fit well the LOS of class 3 patients. The root mean square error of the cumulative function of class 3 detailed model is 0.0230, AMC model is 0.0.0972, and of Coxian model is 0.1019. In this case, we conclude that the class 1 detailed model is the most accurate, then the AMC model, and the Coxian model is the least accurate. Note, the difference between the Coxian and AMC models is small and they both underestimate the LOS of class 3 patients.

According to the expert opinion, doctors from Roanne Hospital center, the obtained results of the detailed Markov model are representative of reality.

5. Models utilization

In this paper, we focused our study on the elderly pathway diagnostic and modeling. We collected data that concern elderly patients (241 elderly over 74 years) following the intra-hospitalization process in the periods between February 23 and March 1, and March 30 and 5 April 2015. This category of patients represents 70% of all admissions in Roanne Hospital Center arriving to the SAU in these periods. We proposed three models with different level of details (1) per elderly in the hospital (Coxian model), (2) per elderly per service (Aggregated model), and (3) per elderly class per service (Detailed model). These three models are complementary and provide accurate estimation of the LOS distribution per phase, per service and per class of elderly. The first model, Coxian model, is inspired from the literature and we apply it to our case study to get a preliminary idea of the number of phases in the care pathway and the global LOS of an elderly in the hospital. Based on the AIC and BIC criteria we find that four main phases are needed to model the global LOS of an elderly. The Coxian model accuracy is measured by comparison of the model's LOS cumulative function with the non-parametric Kaplan-Meier estimator. The Second model, aggregated Markov model, allows us to get the LOS per service for all elderly in the considered population.

Table 6
Estimated parameters of elderly pathway of class 1 of the detailed Markov model applied to Roanne Hospital.

	Path	Probability	Rate parameter	Value	Transition rate	Value
Class1	Path 1	0.57	$\lambda_1^{11}=\alpha_{12}^{11}+\alpha_{15}^{11}$	2.2666502	α_{12}^{11}	1.51110013
					α_{15}^{11}	0.75555007
			$\lambda_2^{11}=\alpha_{23}^{11}+\alpha_{25}^{11}$	19.427977	α_{23}^{11}	2.42849713
					α_{25}^{11}	16.9994799
			λ_3^{11}	19.427977	α_{34}^{11}	1.38771264
					α_{35}^{11}	18.0402644
			λ_4^{11}	3.1942797	α_{45}^{11}	3.19427977
	Path 2	0.37	λ_1^{12}	0.157632	α_{12}^{12}	0.03621276
					α_{14}^{12}	0.12141924
			λ_2^{12}	0.451675	α_{23}^{12}	0.16644065
					α_{24}^{12}	0.28526842
			λ_3^{12}	0.443712	α_{34}^{12}	0.443712
	Path 3	0.06	λ_1^{13}	0.076084	α_{12}^{13}	0.076084
					α_{14}^{13}	0
			λ_2^{13}	0.075081	α_{23}^{13}	0.075081
					α_{24}^{13}	0
			λ_3^{13}	0.076084	α_{34}^{13}	0.076084

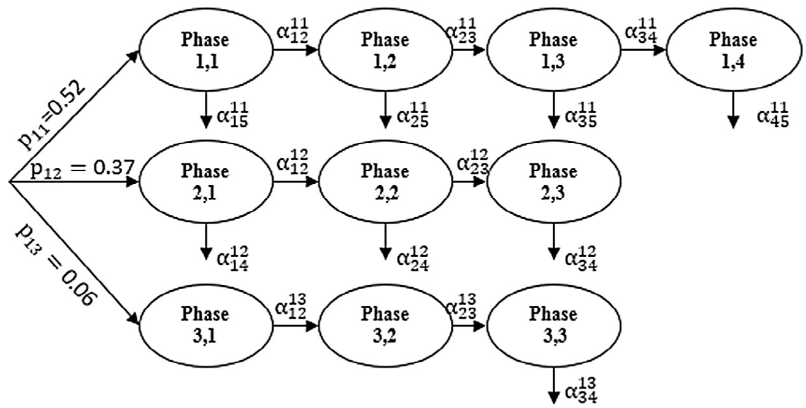


Fig. 10. Transition diagram of class 1 patients of the detailed Markov model fitted to the data of Roanne hospital center.

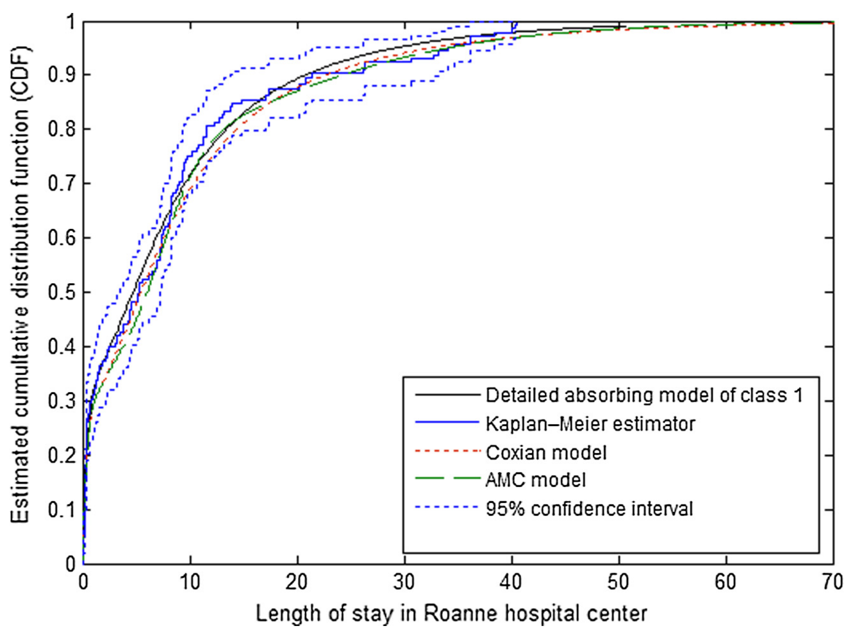


Fig. 11. Kaplan–Meier estimator and detailed absorbing Markov cumulative distribution of class 1 patients: length of stay of elderly in Roanne hospital center.

Table 7
Estimated parameters of elderly pathway of class 2 of the detailed Markov model applied to Roanne Hospital center.

	Path	Probability	Rate parameter	Value	Transition rate	Value
Class 2	Path 1	0.48	λ_1^{21}	3.233839	α_{12}^{21}	3.233839
	Path 2	0.52	λ_1^{22}	0.294118	α_{12}^{22}	0.294118
					α_{14}^{22}	0
			λ_2^{22}	0.294118	α_{23}^{22}	0.09049785
					α_{24}^{22}	0.20362015
			λ_3^{22}	0.294118	α_{34}^{22}	0.294118

The obtained LOS cumulative function is compared to the Kaplan-Meier estimator and we find that the model is more accurate than the Coxian model even though a larger number of parameters is required (11 parameters compared to 4 parameters). The third model, detailed absorbing Markov model, allows differentiating between the elderly classes in the hospital services. Based on their several characteristics, the elderly are classified into three classes. The detailed Markov model allows estimating the LOS per service for each class of elderly in the considered population. The obtained LOS cumulative function is compared to the Kaplan-Meier estimator and we find that the model is accurate even though a larger number of parameters is required (36 parameters in total).

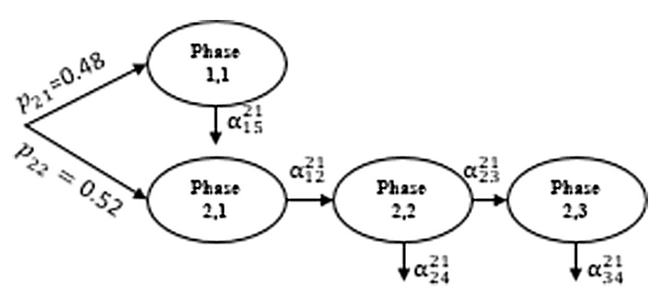


Fig. 12. Transition diagram of class 2 patients of the detailed Markov model fitted to the data of Roanne hospital center.

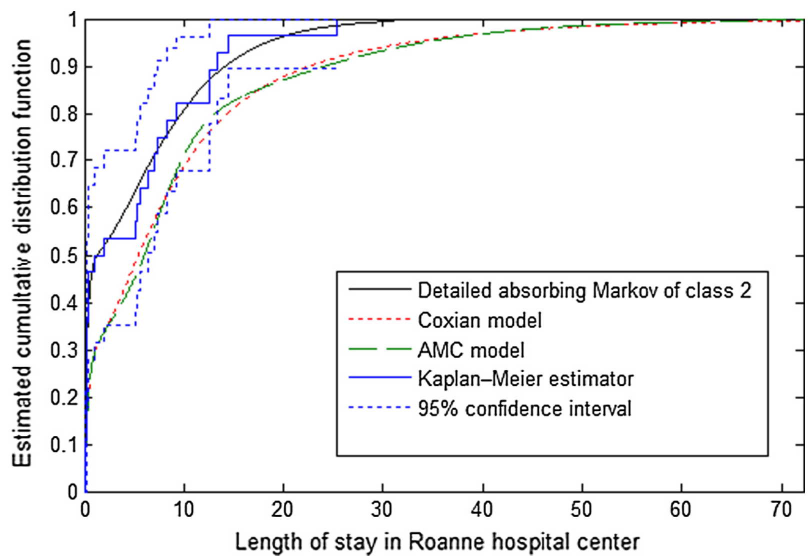


Fig. 13. Kaplan–Meier estimator and detailed absorbing Markov cumulative distribution of class 2 patients: length of stay of elderly in Roanne hospital center.

Table 8
Estimated parameters of elderly pathway of class 3 patient of the detailed Markov model applied to Roanne Hospital.

Class	Path	Probability	Rate parameter	Value	Transition rate	Value
Class3	Path 1	0.2131	λ_1^{31}	7.069870	α_{12}^{31}	7.069870
					α_{14}^{31}	0
			λ_2^{31}	7.069870	α_{23}^{31}	1.777184
					α_{24}^{31}	5.2926
			λ_3^{31}	1.010442	α_{34}^{31}	1.010442
	Path 2	0.7213	λ_1^{32}	0.201101	α_{12}^{32}	0.201101
					α_{14}^{32}	0
			λ_2^{32}	0.199813	α_{23}^{32}	0.199813
					α_{24}^{32}	0
			λ_3^{32}	0.631571	α_{34}^{32}	0.631571
	Path3	0.0819	λ_1^{33}	0.045722	α_{12}^{33}	0.045722
					α_{15}^{33}	0
			λ_2^{33}	0.045722	α_{25}^{33}	0.045722

Proof of concept on other possible utilization of our models in practice: Our Markov chain based approach is generic and can be applied to the whole categories of patients. This can allow us to study the performance of the hospital services: to determine the resources capacity in the hospital, for example the number of beds per service, to reduce the waiting time and by consequence the LOS of patients. This needs the data of all patients including non-elderly patients, which are not the focus of our study. The collection of data of non-elderly patients could be part of a future study. Whenever, this is done a simulation model using the results of the aggregated Markov model can be used to determine the best beds allocation in the MCO and SSR services. In practice, several scenarios in the simulation could be considered by changing the assignment of available beds to different services. In our case of study the number of beds in the ED, MCO and SSR are equal to 12, 25 and 76, respectively. A simple way to generate these scenarios can be done by increasing and decreasing the number of beds in MCO and SSR by one unit while considering the same number of beds in total. Moreover, nowadays, the hospitals are facing an increase in the demand and are looking for optimizing their activities cost. To reduce the LOS in MCO, and cover the maximum of the demand, the discharge of elderly to nursing home before SSR is one of the possible decisions that could be taken based on the elderly characteristics. Therefore, a simulation model based on the results of the detailed Markov chain could be developed to quantify the impact of discharge probability by class of elderly on the LOS. In practice, several scenarios could be considered by changing the discharge rate before the stay in SSR (after the stay in the MCO) for the three classes of elderly. In our case of study

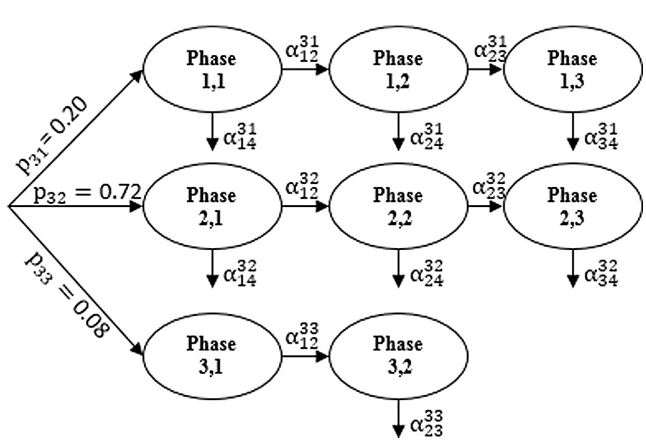


Fig. 14. Transition diagram of class 3 patients of the detailed Markov model fitted to the data of Roanne hospital center.

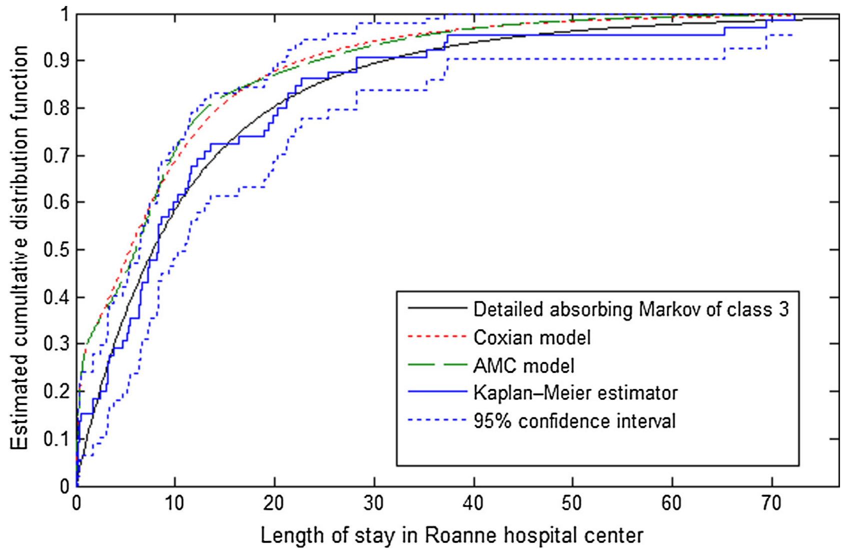


Fig. 15. Kaplan–Meier estimator and detailed absorbing Markov cumulative distribution of class 3 patients: length of stay of elderly in Roanne hospital center.

the percentage of discharge is 40% and 60% for the first class, and third class, respectively. The simulation scenarios can be generated by increasing and decreasing the percentage of discharge by 1%. The simulation could be performed with several replications for each scenario to find the best scenario yielding the lowest global LOS on average. In addition, this simulation model could be used to make a cost study. For example, in the French refund system the MCO cost depends on the activity and the SSR receives annual endowment (Or & Renaud, 2009). The funding calculation of the hospital is based on the Diagnosis Related Groups (DRG) with a basic refund amount, and lower and upper bounds with the related adjustment according to the LOS. The aim of this sort of funding is to give incentive to the hospitals to reduce the patient LOS in the service without sacrificing the care quality. For each patient, based on his/her corresponding DRG, we can calculate the stay refund and compare it to the money spent by each service during the patient stay. The two changes of the discharge rate after MCO and the capacities of MCO and SSR could be repeated here to generate different scenarios similarly to the scenarios considered before. We keep the best scenario with the capacity of MCO and SSR and the discharge rate for the first and third classes of elderly, providing the highest extra revenue (refunding revenues minus total cost).

6. Conclusion and perspective

In this paper, we are interested in the elderly pathway diagnostic and modeling. We propose three ways to model the elderly pathway. The first model, called aggregated model, is a continuous-time finite size Markov chain. This model allows us to estimate the length of stay of elderly in the hospital services and their pathways. In the second model, the length of stay in the hospital is fitted to a Coxian distribution. The third model, called detailed model, is a continuous-time Markov chain with a mixed of Coxian distribution. In this model we differentiate between the types of patients. Therefore, as a preliminary step we classify the elderly based on frailty, lifestyle, and pathology. By knowing the patient characteristics we first identify her class of fragility and the probability to follow a specific pathway. Here, we differentiate between short, middle, and long stay within an elderly class. We apply our models to a French healthcare system. Comparing these models the aggregate model is richer in information such length of stays per service and more accurate in predicting the length of stay in the hospital than the simple Coxian model. The detailed model is the richest in information since we differentiate between patient types and length of stays.

The provided models are of interest to managers of French Healthcare structures to master the elderly pathways and evaluate the performance of intra-hospital services. Moreover, the local authorities in France face real challenges when it comes to the annual funding prediction for healthcare structures. In fact, the length of stay and the healthcare cost of existing patients in addition to the unknown number of future admissions make the budgeting of local authority complex. Our study is a preliminary step toward the development of decision support system that helps in making more accurate cost prediction, to decide the patient admission strategy, and to meet decisions concerning resource allocation. Future work will focus on these aspects.

References

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

Avila-Funes, J. A., Helmer, C., Amieva, H., Barberger-Gateau, P., Le Goff, M., Ritchie, K., ... Dartigues, J. F. (2008). Frailty among community-dwelling elderly people in France: The three-city. *Journal of Gerontology: Medical Sciences, 63*(10), 1089–1096.

Chaussalet, T., Xie, H., & Millard, P. (2006). A closed queuing network approach to the analysis of patient flow in health care systems. *Methods of Information in Medicine, 45*(5), 492–497.

D'Avignon, G., & Mareschal, B. (1989). Specialization of hospital services in Quebec: An application of the promethee and GAIA methods. *Mathematical and Computer Modelling, 12*(10), 1393–1400.

European Commission. (2015). The 2015 Ageing Report Economic and budgetary projections for the 28 EU Member States (2013–2060) (pp. 1725–3217). <http://ec.europa.eu/economy_finance/publications/european_economy/2015/pdf/ee3_en.pdf>.

Faddy, M., & McClean, S. (2005). Markov chain modelling for geriatric patient care. *Methods of Information in Medicine, 44*(3), 369–373.

Greenacre, M., & Blasius, J. (2006). *Multiple correspondence analysis and related methods*. Chapman & Hall/CRC Statistics in the Social and Behavioral Sciences. ISBN 9781584886280.

Guojun, G., Chaoqun, M., & Jianhong, W. (2007). *Data clustering: Theory, algorithms and applications. ASA-SIAM series on statistics and applied mathematics*. 978-0-89871-623-8.

Irvine, V., McClean, V., & Millard, P. (1994). Stochastic models for geriatric in-patient behavior. *IMA Journal of Mathematics Applied in Medicine and Biology, 3*(3), 207–216.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association, 53*(282), 457–481.

Loones, A., David-Alberola, E., & Jauneau, P. (2008). *La fragilité des personnes âgées: Perceptions et mesures*. Centre de Recherche pour l'Étude et l'Observation des Conditions de Vie. cahier de recherche.

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

Marshall, A., & McClean, S. (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., Garg, L., Barton, M., & Fullerton, K. (2010). Using mixed phase-type distributions to model patient pathways. In *IEEE 23rd international symposium on computer-based medical systems (CBMS), Perth* (pp. 172–177).

McLachlan, G. J., & Krishnan, T. (2008). Wiley series in probability and statistics. In *The EM algorithm and extensions*, pp. 359). Hoboken, NJ, USA: John Wiley & Sons Inc.. http://dx.doi.org/10.1002/9780470191613.scard.

Olsson, M. (1998). *The EMpht-programme, technical report*. Department of mathematics, Chalmers University of Technology <http://www.math.lth.se/matstat/staff/asmus/pspapers.htm>.

Or, Z., & Renaud, T. (2009). *Principes et enjeux de la tarification A l'activité A l'hôpital (T2A). Enseignements de la théorie économique et des expériences étrangères*. Institut de recherche et documentation en économie de la santé <http://www.irdes.fr/EspaceRecherche/DocumentsDeTravail/DT23PrincipEnjeuxTarificActiviteHopital.pdf>.

Spector, W., Mutter, R., Owens, P., & Limcango, R. (2012). *Transitions between nursing homes and hospitals in the elderly population 2009 HCUP statistical brief #141*. Rockville: Agency for Healthcare Research and Quality <http://www.hcup-us.ahrq.gov/reports/statbriefs/sb141.pdf>.

Xie, H., Chaussalet, T., & Millard, P. (2005). A continuous time Markov model for the length of stay of elderly people in institutionallong-term care. *Journal of the Royal Statistical Society: Series A. Statistics in Society, 168*(1), 51–61.

Xie, H., Chaussalet, T., & Millard, P. (2006). A model-based approach to the analysis of patterns of length of stay in institution along-term care. *IEEE Transactions on Information Technology in Biomedicine, 10*(3), 512–518.

Xie, H., Chaussalet, T., & Rees, M. (2007). Semi-open queueing network approach to the analysis of patient flow in healthcare systems. In *Twentieth IEEE international symposium on computer-based medical systems, Maribor* (pp. 719–724).