A multistate model to evaluate COPD progression integrating drugs consumption data and hospital databases

Nicola Bartolomeo[1], Paolo Trerotoli[1], Gabriella Serio[1]

ABSTRACT

BACKGROUND: Dealing with the increased costs related to the diagnosis and treatment of chronic-degenerative diseases requires a better knowledge of patients' true care pathways. The study objective was to explore, using multi-state modeling, how analyses of drug prescriptions and data obtained from hospital discharge sheets can be used in combination to build a model of patients' health care pathways in a non experimental setting. The model was applied to Chronic Obstructive Pulmonary Disease (COPD).

METHODS: Based on the GOLD guidelines, access to hospitalization for COPD and prescription pharmaceuticals were awarded to seven transient, theoretically progressive states. The intensity of transitions was estimated with the non-parametric method proposed by Aalen and Johansen for multi-state Markov models, non-homogeneous in time.

RESULTS: The COPD patients included in the study totaled 111190. Patients admitted with a diagnosis of COPD without exacerbation had a growing probability over time of needing prescriptions for inhaled corticosteroids (ICS) or the set combination of long-acting beta-agonists (LABA) and ICS; they also had a rising probability of an exacerbation. The use of ICS alone or in combination with LABA delays hospital admission for acute respiratory failure by about 6 months, as compared to short-acting beta-agonists or anticholinergics.

CONCLUSION: The probabilities of a transition and their distribution in relation to time, sex, age and clinical status can be a helpful tool to guide those operating in the health care sector who are called upon to carry out decisions from the standpoints of both efficacious clinical management and an efficient use of resources.

Key words: COPD. Drug Prescriptions. Health Administrative Databases. Markov model. Multistate models.

(1) Department of Biomedical Science and Human Oncology, Chair of Medical Statistics, University of Bari. Bari. Italy.

CORRESPONDING AUTHOR: *Nicola Bartolomeo - Department of Biomedical Science and Human Oncology, Chair of Medical Statistics, University of Bari - Policlinico, Piazza Giulio Cesare, 11 70124 Bari (Italy)*
e-mail: nicola.bartolomeo@uniba.it - tel / fax(+39) 080 5478479

DOI: ***10.2427/11145***
Accepted on May 14, 2015

RESEARCH FUND: *the research was funded with Scientific Research Fund of the University of Bari Aldo Moro.*

INTRODUCTION

In more economically advanced nations the growing average age of the population, together with alterations in habits and lifestyles, have brought about a change in the health profiles of the populations, featuring a progressive rise in the prevalence of chronic-degenerative type diseases [1].

The total costs associated with the care of patients with chronic diseases, that requires multidisciplinary medical approaches for diagnosis and therapy, are rising. They already account for about 70% of public health expenditure [2] and this percentage will likely increase as the population continues to age, bringing about further rises in pharmaceutical costs. A better knowledge of the diagnostic and therapeutic pathways patients actually do undergo could ensure a more efficient management of health care services and, in particular, a greater efficacy of patients health promotion.

The steps in the diagnostic and therapeutic pathway can be reconstructed by combined analyses of hospital admissions/discharge data and drug prescriptions. The events reported in such databases can be interpreted as the outcomes (multiple events) of a longitudinal statistical model.

To identify the "typical" pathways a longitudinal model needs to be built, taking into account the need to assess the time intervals between each subsequent step in the pathway, that can correspond to different outcomes. The majority of longitudinal studies reported in epidemiological and clinical literature rely on classic "survival" analysis methods such as Cox's proportional hazards model [3].

Survival analysis methods ('time-to-event'), designed to analyze the time to a single endpoint, may not be suitable. In fact, in order to analyze multiple-events data, separate survival analytical models need to be implemented for each event of interest, and subjects are typically censored at the time of occurrence of the other "competing" events [4,5]. However, the use of single endpoint methods does not allow researchers to test for possible differences between the effects of the same prognostic factor on the risks of different events. Alternative approaches such as "competing risks" models [6,7] or time-dependent covariates [8] are not well suited to accounting for the complexity of real-life processes of disease progression, in which both mutually exclusive endpoints and sequences of different events may be observed.

To account for such complexities, the methodological literature suggests that the use of Markov multi-state models, designed specifically for modeling more than one outcome, has inherent theoretical advantages [9,10,11]. Because they are a generalization of classical survival analyses, they are commonly used to describe the history of an individual, who may at any time be in one of a finite number of states [11].

In multi-state models an event can be considered as a transition from one state to another. These states can be either absorbing states from which further transitions

cannot occur (such as death), or transient states which may be followed by

subsequent transitions to other states.

For more than one transient state, where study subjects can pass from one state to another, Kay (1982) [12] extended the model that involves J + 2 'disease' states (the J event endpoints, the right censoring, and the absorbing state) to a more general multi-state stochastic process allowing for several transient disease states.

This model is highly suitable for the study of multistate stages transitions as they may occur in chronic diseases. In our study we have applied this model, that includes several recurrent states and lacks an absorbing state.

Various multi-state models have been reported in the literature. For example, Zhang et al. [13] demonstrated that multistate models are a useful tool for the evaluation of the transition intensity and relative risk of transition for rheumatic disease studies. Price et al. [14], in a recent paper, applied a multistate model to evaluate the risk of pelvic inflammatory disease.

These application deal with pathways that have a limited number of states and a specific evolution over time.

The aim of the current study is to explore, using Markov multi-state modeling, how analyses of drug prescriptions and data obtained from hospital discharge sheets can be used in combination to build a model of patients health care pathways in a non experimental setting. For this purpose, the health care administrative data banks of hospital admissions and of drug prescriptions were linked using suitable Record Linkage

techniques. The chosen disease on which to apply the statistical method was Chronic Obstructive Pulmonary Disease (COPD).

METHODS

Study Population

The study was conducted using the Hospital Discharge Sheets (HDS) obtained from all Apulian hospitals for the years 2006-2010, and the Drug Prescriptions (DP) made in Apulia in the same period. The study population corresponds to those subjects hospitalized with a principal or secondary diagnosis code of COPD (ICD-9-CM: without acute exacerbation 490.20; with acute exacerbation 490.21) or respiratory failure (ICD-9-CM 518.81) in the period between 01/01/2006 and 31/12/2010. For this reason, all patients who had been hospitalized at least once with a diagnosis of COPD in the year 2005 were excluded, as were subjects who appeared in the DP database, in 2005, with at least one prescription for one of the Anatomical Therapeutic Chemical classifications (ATC) shown in Table 1.

All subjects aged less than 40 years were also excluded, as well as asthmatic subjects, i.e. all patients aged less than 55 years who had been hospitalized in the period considered (2006-2010) exclusively with a discharge diagnosis of asthma (code ICD9-CM 493.--).

TABLE 1

ATC GROUPS BY "PRESCRIPTION" STATES

STATE	CLASS	ATC	ATC NAME
1 - PRESCRIPTION SABA SA+AS SAAN	Short-acting beta agonists (SABA)	R03AC02	Salbutamol
		R03AC03	Terbutalin
		R03AC04	Fenoterol
	Short-acting set combinations + other anti-asthma drugs (SA+AS)	R03AK03	Fenoterol and other drugs for obstructive airways syndromes
		R03AK04	Salbutamol and other drugs for obstructive airways syndromes
	Short-acting Anticholinergics (SAAN)	R03BB01	Ipratropium bromide
		R03BB02	Oxitropium bromide
2 - PRESCRIPTION LABA XANT LAAC	Long-acting beta agonists(LABA)	R03AC12	Salmeterol
		R03AC13	Formoterol
		R03AC16	Procaterol
	Xanthinic Derivatives (XANT)	R03DA04	Theophylline
	Long-acting Anticholinergics (LAAC)	R03BB04	Tiotropium bromide
3 - PRESCRIPTION ICS ICS+LABA LEUCO	Inhaled Corticosteroids (ICS)	R03BA01	Beclomethasone
		R03BA02	Budesonide
		R03BA03	Flunisolide
		R03BA04	Betamethasone
		R03BA05	Fluticasone
	Set combination of long-acting beta agonists + ICS (ICS+LABA)	R03AK06	Salmeterol and other drugs for obstructive airways syndromes
		R03AK07	Formoterol and other drugs for obstructive airways syndromes
	Leucotrienics (LEUCO)	R03DC01	zafirlukast
		R03DC03	montelukast
4 - PRESCRIPTION SC	Systemic Corticosteroids (SC)	H02A---	

From the DP database, all subjects who had been prescribed only "Systemic corticosteroids" (ATC H02---) in the period considered were excluded.

Staging of COPD

Identified cases of COPD were retrieved in the DP database, on the basis of a record of prescription of at least one drug recommended for the treatment of COPD according to the GOLD guidelines (Global initiative for chronic Obstructive Lung Disease) [15]. By grouping drugs according to the recommendations for pharmacological treatment of subjects affected by COPD, and depending on their duration of action and administration method, four progressive transitory "prescription" states were defined (table 1).

The progression of COPD can be followed through the patients' hospitalizations as from the disease onset up to the more acute phases. On the basis of the Hospital Discharge Sheets (HDS), three phases can be discerned:

- COPD without acute exacerbation (diagnostic code ICD9-CM 490.20);
- COPD with acute exacerbation (diagnostic code ICD9-CM 490.21);
- acute respiratory failure (diagnostic code ICD9-CM 518.81).

The three states indicated above are all possible transitory states in a multi-state model suggesting a progressive diagnostic pathway.

Variables studied

The demographic variables included in the study were gender and age, subdivided into two classes: under 65 years; over 65 years of age. The sequence of hospitalizations and drug prescriptions of a patient affected by COPD can also depend on the severity of the overall clinical picture of the subject, measured by Charlson's Comorbidity Index [16]. Taking into account that the index is a sum of the weight related to the death risk, it has been classified in three groups: without risk (CCI=0), CCI within the interval 1-3, CCI>3 [17]. The "time" variable, serving to determine the probability of transition, corresponds in the model applied to the weeks between the transition from one state to enter the next state.

Statistical Analyses

To assess the Markov multi-state model with non homogeneous time, seven states were identified (figure 1), three related to possible hospitalizations and four corresponding to the ATC groups prescription:

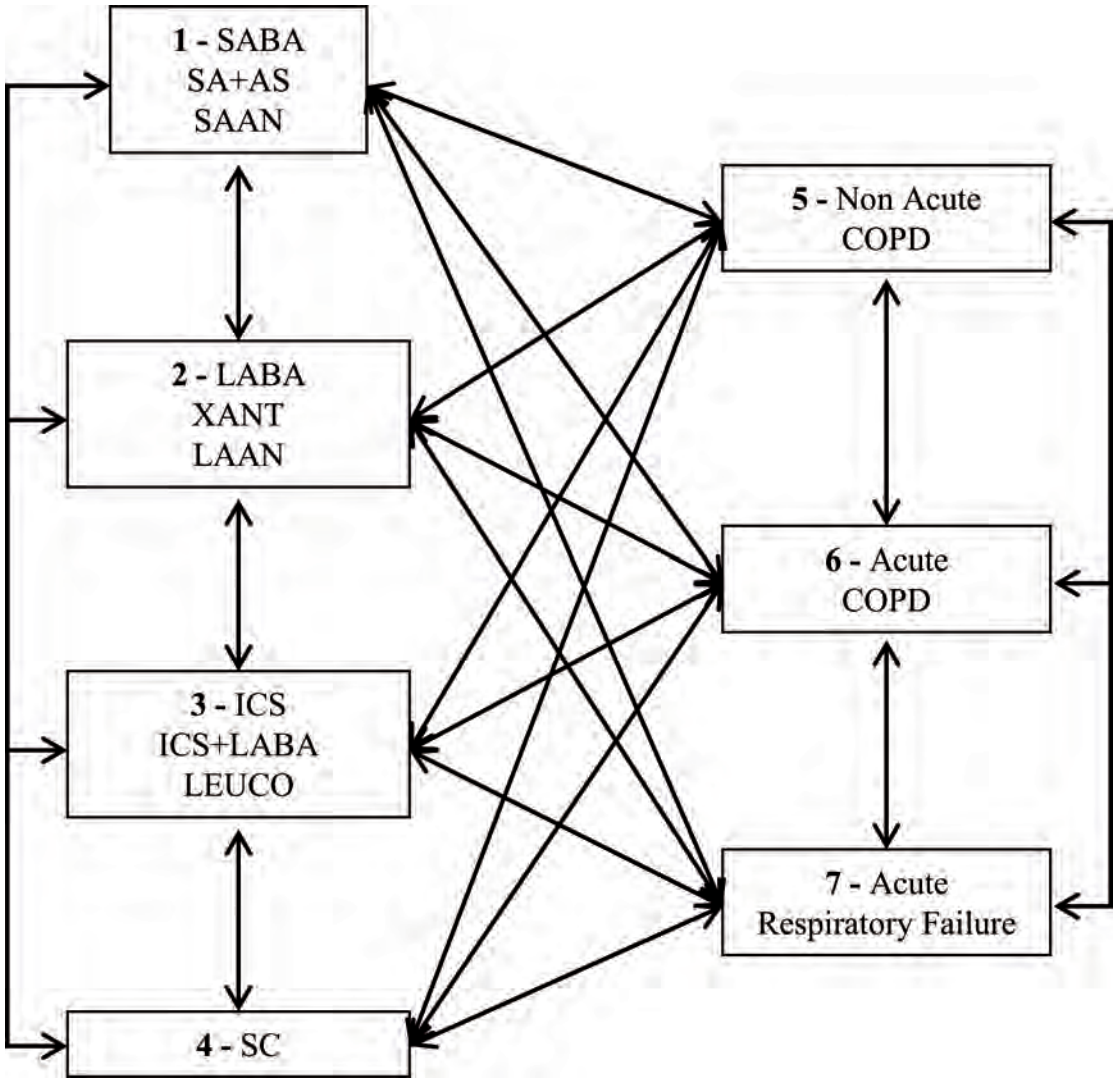
- State 1 defined by drug prescription of Short-acting beta agonists (SABA), or Short-acting set combinations plus other anti-asthma drugs (SA+AS), or Short-acting Anticholinergics (SAAN);
- State 2 defined by drug prescription of Long-acting beta agonists (LABA), or Xanthinic Derivatives (XANT), or Long-acting Anticholinergics (LAAC);
- State 3 defined by drug prescription of Inhaled Corticosteroids (ICS), or Set combination of long-acting beta agonists plus ICS (ICS+LABA), or Leucotrienics (LEUCO);
- State 4 defined by drug prescription of Systemic Corticosteroids (SC);
- State 5 defined by hospitalization for COPD w/o exacerbation
- State 6 defined by hospitalization for COPD with acute exacerbation
- State 7 defined by hospitalization for Acute Respiratory Failure.

Since the main aim was to assess a model to evaluate repeated hospitalizations for COPD and the role of drug prescriptions, if any, in the care pathway, a model was applied to estimate the probabilities for any type of transitions between states, without an absorbing state. Death, the strong endpoint to consider as an absorbing state, was not included in the model because it occurs after a long disease course in COPD patients and the Death Register for the time period subsequent to 2010 was not available. Subjects were considered censored at the end of the observation period (31/12/2010) or at the time of the last available information.

Given the high number of states, the intensity of transition was estimated with the non parametric method proposed by Aalen and Johansen [18] for generic multi-state models with a finite number of states. This model is particularly well suited to large quantities of data on subjects whose observations are censored to the right. The censoring process was assumed to be independent; this assumption allows the censoring process to depend only on the state currently occupied, in the sense that an

FIGURE 1

MULTI-STATE MODEL APPLIED TO COPD



additional knowledge of censorings before any time t does not alter the risk of failure at t [19,20,21,22].

In COPD, as in all chronic diseases, the assumption of homogeneity over time is implausible because there is expected to be a notable variability of the risks of some or all of the transitions during the follow-up. In such cases, to take into account this variability over time of the intensity of transitions, the time axis was subdivided into consecutive intervals, assuming a constant intensity within each interval, but varying between them [23].

The states of the estimate model, shown in Figure 1, are all transitory because they can be followed by another state; the transition hazard from state i ∈ S to state j ∈ S, i ≠ j is defined as

$$\alpha_{ij}(t)dt = P(X_{t+DT} = j \mid Xt = i) \quad (1)$$

The cumulative transition hazards are then defined as

$$A_{ij}(t) = \int_0^t \alpha_{ij}(u)du,\ A_{ii}(t) = -\sum_{j \neq i} A_{ij}(t) \quad (2)$$

Let

$$TPij(s,t) = P(X_t = j \mid X_s = i),\ \text{per } i,j \in S,\ s \leq t \quad (3)$$

be the probability that an individual who is

in state *i* at time s will have moved to state *j* at later time *t*, and we write TP(*s,t*) for the (K+1) x (K+1) matrix of those transition probabilities. TP(*s,t*) can be recovered from the transition hazards through product integration [18,24]:

$$\mathrm{TP}(s,t) = \prod_{(s,t]}(\mathrm{I} + d\mathrm{A}(u)) \qquad (4)$$

where A(*t*) is the matrix of cumulative hazards.

Let $N_{ij}(t)$ be the number of observed direct transitions from state *i* to state *j* up to time *t* and $Y_i(t)$ be the number of individuals under observation in state i just before time *t*.

The non-diagonal entries of the matrix of cumulative transition hazards A(*t*) may be estimated by the Nelson-Aalen estimator [25,26]

$$\hat{A}_{ij}(t) = \int_0^t \frac{dN_{ij}(u)}{Y_i(u)},\ i \neq j,\ e\ A_{ii}(t) = -\sum_{j \neq i}\hat{A}_{ij}(t) \qquad (5)$$

So the estimator of the matrix of cumulative risks of transition becomes

$$\widehat{\mathrm{TP}}(s,t) = \prod_{(s,t]}(\mathrm{I} + d\widehat{\mathrm{A}}(u)) \qquad (6)$$

As $\widehat{\mathrm{A}}(t)$ is a matrix of step-functions with a finite number of jumps on (*s,t*), the product integral can be written as a finite matrix product

$$\widehat{\mathrm{TP}}(s,t) = \prod_{(s<t_k\leq t]}(\mathrm{I} + \Delta\widehat{\mathrm{A}}(t_k)) \qquad (7)$$

where the product is taken over carried over all observed transition times in (*s,t*].

The non-diagonal entries (*i,j*), $i \neq j$, of $\mathrm{I}+\Delta\widehat{\mathrm{A}}(t_k)$ are the number of observed direct $i \rightarrow j$ transitions, divided by the number of individuals under observation in state *i* just prior to time t_k. The diagonal entries of $\mathrm{I}+\Delta\widehat{\mathrm{A}}(t_k)$ are such that each row equals 1.

State changes and corresponding transition probabilities respect to time are shown in graphs; pointwise confidence intervals for estimators of transition probabilities follow the method described in Andersen et al. [25]. The authors selected state changes that suggest a possible contribution of drug prescriptions in the description of disease progression.

To assess the covariate effects on transition probabilities, survival curves were estimated for each group, and comparison was done with log rank test.

Analysis were performed with R [27] using ETM package [28].

RESULTS

From 2006 to 2010, 111,190 subjects satisfied the inclusion criteria; mean (sd) follow-up was approximately 153.9 weeks (sd 73.2). Mean age at the start of follow-up was 72.6 years (sd 11.6). There were 64,924 males, equal to 58.4% of the total.

In total, there were 111,803 hospital admissions with at least one discharge diagnosis of COPD without acute exacerbation, COPD with acute exacerbation, or Acute Respiratory Failure, showing a peak in 2009 (23,912 hospitalizations). Male subjects accounted for 60% of hospital admissions (67,057). COPD without acute exacerbation was most common, and showed a peak in 2008 (11,695 admissions); whereas the peaks for both COPD with acute exacerbation presentation and Acute Respiratory Failure were in 2009 (8,821 and 5,802, respectively).

There were 302,516 ATC referable to COPD prescribed in Apulia between 2006 and 2010 for the study subjects in the follow-up, showing a peak in 2010 (79,578).

In general, the drugs most often prescribed in the entire period were long-acting Anticholinergics (class ATC R03BB04), accounting for 46,705 prescriptions, and 15.44% of the total, followed by the set combination of Salmeterol and inhaled Corticosteroids (ATC R03AK06), accounting for 42,391 prescriptions equal to 14.01% of the total.

Of the entire study population, 39.1%, i.e. 43,457 subjects, started follow-up in state 5 (COPD without acute exacerbation), 20.7% in state 6 (COPD with acute exacerbation) and 16.3% in state 7 (Acute Respiratory Failure). The percentage of subjects who began the pathway with a drug prescription is lower: most had been firstly prescribed inhaled corticosteroids (ICS) or the set combination of inhaled corticosteroids and long-acting beta-agonists (ICS+LABA) (14,165 subjects, 12.7% of the total). Only 2,340 subjects (2.1%) began with a prescription of LABA, xanthinic derivatives (XANT) or long-acting anticholinergics (LAAC).

Table 2 shows the subjects characteristics related to the state they occupied at the time of entering the study.

According to the model shown in Figure 1, there were 307,519 transitions of state, including transitions to the "censored" state. Most subjects had started in state 3, i.e. a

TABLE 2

SUBJECTS CHARACTERISTICS RELATED TO THE STATE THEY OCCUPIED AT THE TIME OF ENTERING THE STUDY

STATE	N	AGE, MEAN (SD)	AGE >65, N (%)	MALE, N (%)	CI > 3, N (%)
1- SABA SA+AS SAAN	3424	71.5 (11.5)	2469 (72.11%)	1929 (56.34%)	-
2 - LABA XANT LAAC	2340	73.9 (10.8)	1860 (79.49%)	1473 (62.95%)	-
3 - ICS ICS+LABA LEUCO	14165	71.1 (11.4)	10028 (70.79%)	8170 (57.68%)	-
4 - SC	6648	70.7 (11.7)	4616 (69.43%)	3709 (55.79%)	-
5 - COPD WITHOUT EXACERBATION	43457	72.5 (11.2)	32695 (75.24%)	27344 (62.92%)	879 (2.02%)
6 - COPD WITH EXACERBATION	22987	74.2 (11.5)	18162 (79.01%)	13078 (56.89%)	496 (2.16%)
7 - ACUTE RESPIRATORY FAILURE	18169	73.0 (12.7)	13512 (74.37%)	9221 (50.75%)	740 (4.07%)

SABA (Short-acting beta agonists); SA+AS (Short-acting set combinations + other anti-asthma drugs); SAAN (Short-acting Anticholinergics); LABA (Long-acting beta agonist); XANT (Xanthinic Derivatives); LAAC (Long-acting Anticholinergics); ICS (Inhaled Corticosteroids); ICS+LABA (Set combination of long-acting beta agonists + ICS); LEUCO (Leucotrienics); SC (Systemic Corticosteroids).

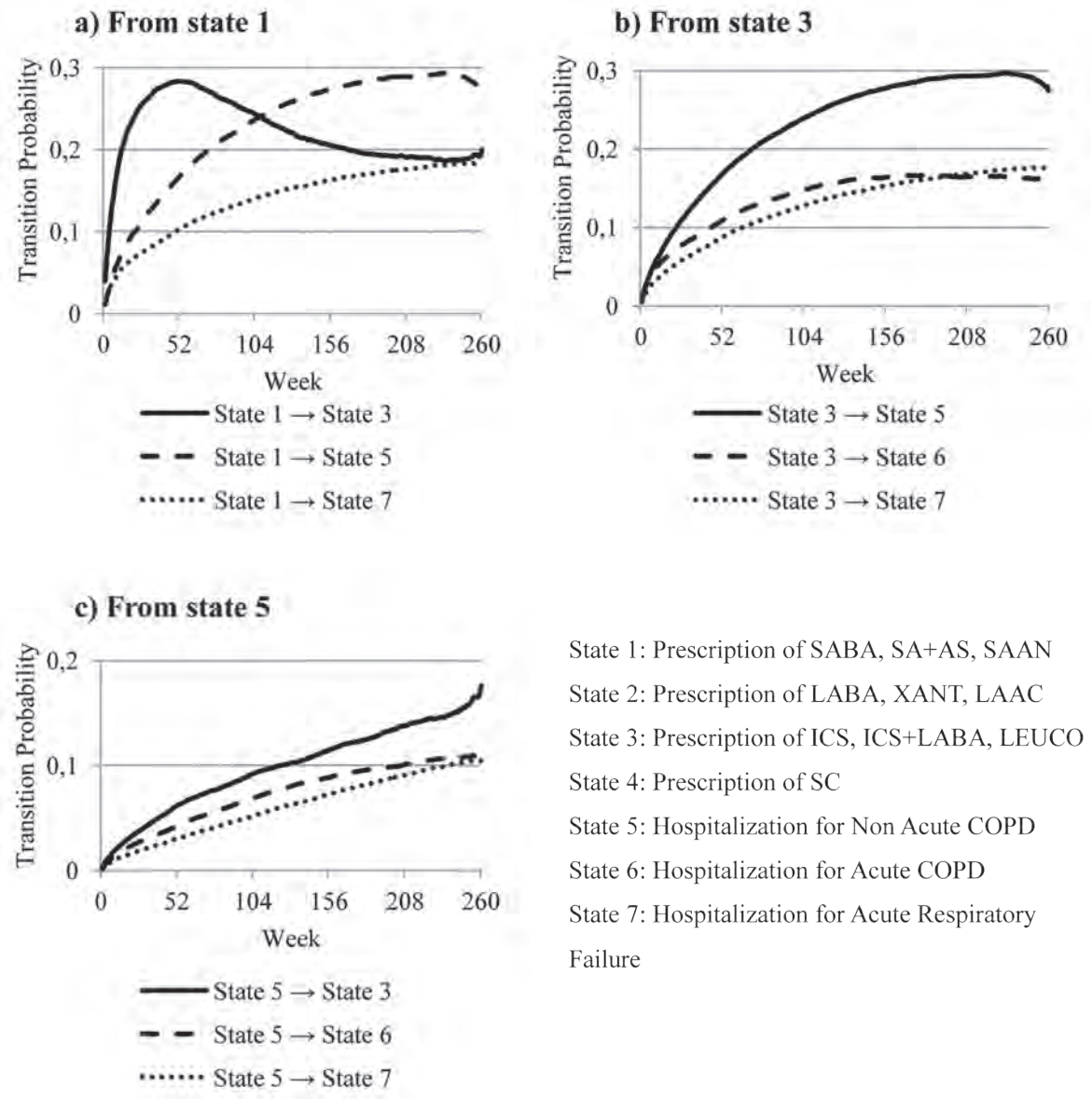
prescription for inhaled corticosteroids alone or in set combination with LABA.

Through the transition probability matrix, estimated using the Aalen-Johansen estimator, between the beginning time (0) and end time (260 weeks) we have found that the subjects admitted and discharged with a diagnosis of COPD without acute exacerbation had a transition probability TP=0.45 (95%CI, 0.43-0.47) of remaining in the same state; about 17.6% (95%CI, 16.2%-19.0%) of them were later prescribed an ICS or set combination of ICS+LABA; 10.8% (95%CI, 10.4%-11.3%) suffered a documented acute exacerbation and hospitalization and 10.4% (95%CI, 10.1%-10.8%) were admitted with Acute Respiratory Failure (ARF). Patients with an acute exacerbation, as compared to those without, had a lower probability of remaining in the same state (29.2%; 95%CI, 28.3%-30.2%) and higher probability of progressing to ARF (13.6%; 95%CI, 13.1%-14.1%). A fairly high percentage of subjects (22.3%; 95%CI, 21.3%-23.2%) returned to the COPD without acute exacerbation state: this transition may be a sign of possible state misclassification errors caused by coding errors arising in the health care administrative databases. There were smaller differences between the probabilities of transition of the subjects from the four prescription states to a disease state: for example, COPD without acute exacerbation was preceded by state 2 in 27.7% (95%CI, 26.6%-28.8%) of subjects and by state 4 in 27.0% (95%CI, 26.0%-28.1%). COPD with acute exacerbation and ARF were prevalently preceded by state 1.

Figure 2a shows the transition probabilities from state 1 (prescription of SABA, SA+AS, SAAN) to states 3, 5 and 7: in the first year there was a sharp rise in the probability of passing from state 1 to the use of inhaled corticosteroids, alone or in set combination with LABA (TP=0.283 at 52 weeks, 95%CI, 0.277-0.290). Subsequently this probability declined, and was exceeded, by about the 108th week, by the probability of being hospitalized with a diagnosis of COPD without acute exacerbation (state 5). The probability of a transition to ARF (state 7) is growing over time. The probability of a transition to a new acute exacerbation (state 6) remained slightly higher than that of

FIGURE 2

STATE CHANGE PLOTS FOR MULTI-STATE MODEL APPLIED TO COPD



State 1: Prescription of SABA, SA+AS, SAAN

State 2: Prescription of LABA, XANT, LAAC

State 3: Prescription of ICS, ICS+LABA, LEUCO

State 4: Prescription of SC

State 5: Hospitalization for Non Acute COPD

State 6: Hospitalization for Acute COPD

State 7: Hospitalization for Acute Respiratory Failure

a transition to hospitalization for ARF (state 7) until about 190 weeks whereas the probability of a transition to the use of LABA/XANT/LAAC (state 2) or OS (state 4) was constantly below 10% (data not shown).

There were no substantial differences in the probabilities of transition from state 2 to the other states (data not shown).

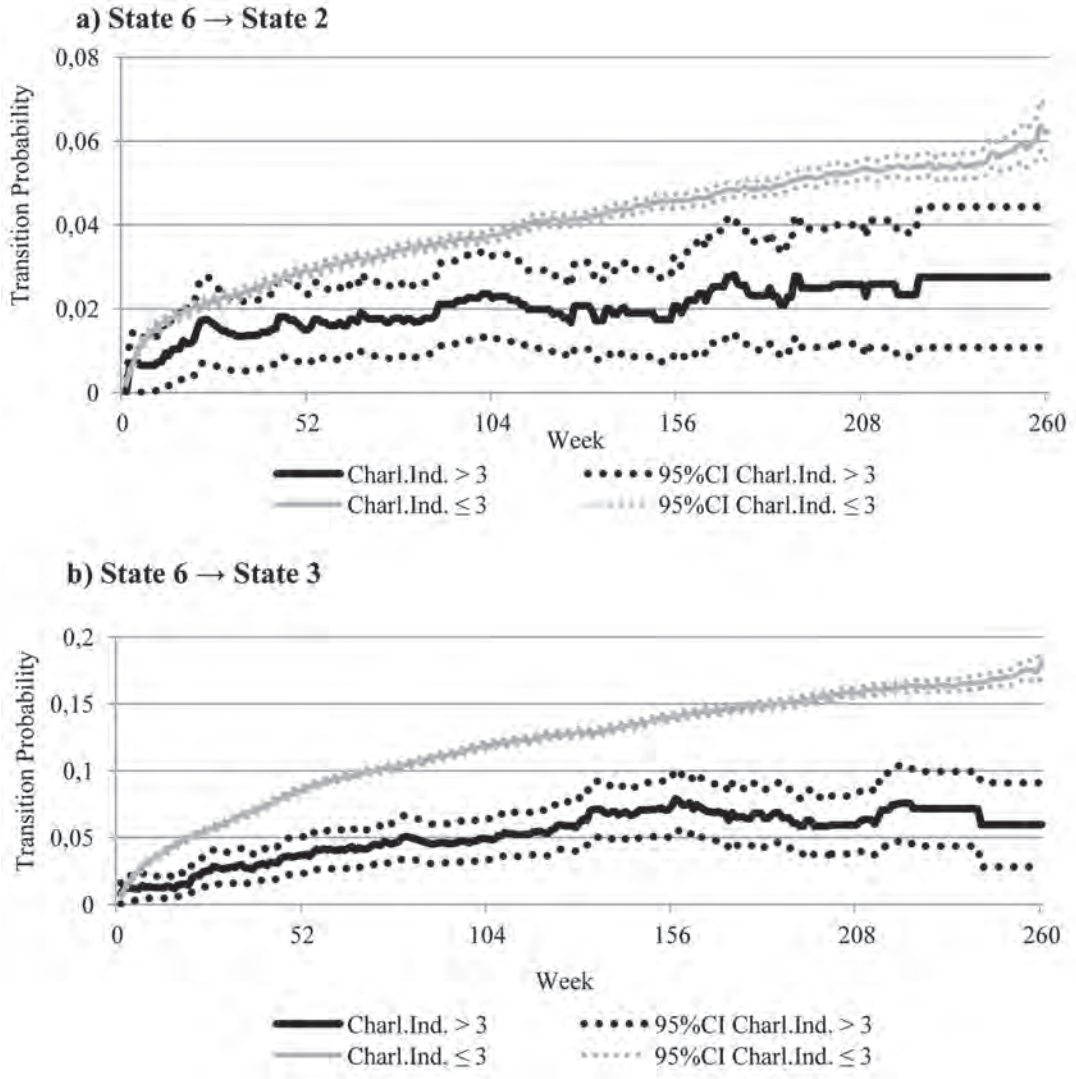
After about 2 years, about 25% (TP=0.24; 95%CI, 0.236-0.0.244) of subjects prescribed ICS or the set combination ICS+LABA (state 3) were hospitalized with a diagnosis of COPD without acute exacerbation (Figure 2b).

Patients admitted and discharged with a diagnosis of COPD without acute exacerbation (state 5) had a growing probability over time of needing prescriptions for ICS or ICS+LABA (state 3); they also had a rising probability of an acute exacerbation requiring hospitalization (state 6) (Figure 2c).

Comparing the state change 1→7 with the state change 3→7 it can be seen that the use of inhaled corticosteroids alone or in combination with LABA (state 3), delays a hospital admission

FIGURE 3

PROBABILITY OF TRANSITION STRATIFIED BY CHARLSON'S INDEX.
3a) STATE 6 -> STATE 2; 3b) STATE 6 -> STATE 3



for ARF (state 7) by about 6 months, as compared to short-acting beta-agonists (SABA) or anticholinergics (SAAN) (state 1).

Subjects with acute exacerbation of COPD and a Charlson Index >3 had a lower probability of a transition from state 6 to state 2 (prescription of LABA, XANT, LAAC) (p<.0001) and from state 6 to state 3 (prescription of ICS, ICS+LABA, LEUCO) (p<.0001) as compared to those with a Charlson Index of 3 or less (Figure 3). In general, subjects with CCI>3 had a significantly lower probability of transition toward states that suggest a lower severity of COPD.

Figure 4 shows two transitions to state 3 for which stratification by age class was the main determinant establishing the time trend. In the first six months, the probability of passing from the use of LABA/XANT/LAAC (state 2) to that of ICS or ICS+LABA (state 3) grows at the same rate, whereas in the next six months the probability remains the same for subjects over 65 years of age but grows further for subjects under 65. In the remaining follow-up period the difference between the two age classes remains constant at about 5% (figure 4a); the difference between the two survival curves was statistically significant (p=0.017) . However, in the transition from hospitalization for ARF (state 7) and a prescription for ICS or ICS+LABA (state 3) the rising curve shows

different slopes (p<.0001). At the end of the period the TP for subjects under the age of 65 is about twice (18% vs. 9%) that of subjects aged over 65 (figure 4b).

Comparison by gender highlights a significantly greater TP for males in almost all state transitions, except for the passage from state 2 to state 7 (p<.0001), from state 3 to state 4 (p=0.003) and from state 3 to state 7 (p<.0001) that shows a significantly greater TP for females. The difference of TP between genders is not significant for transitions involving state 1 and state 4.

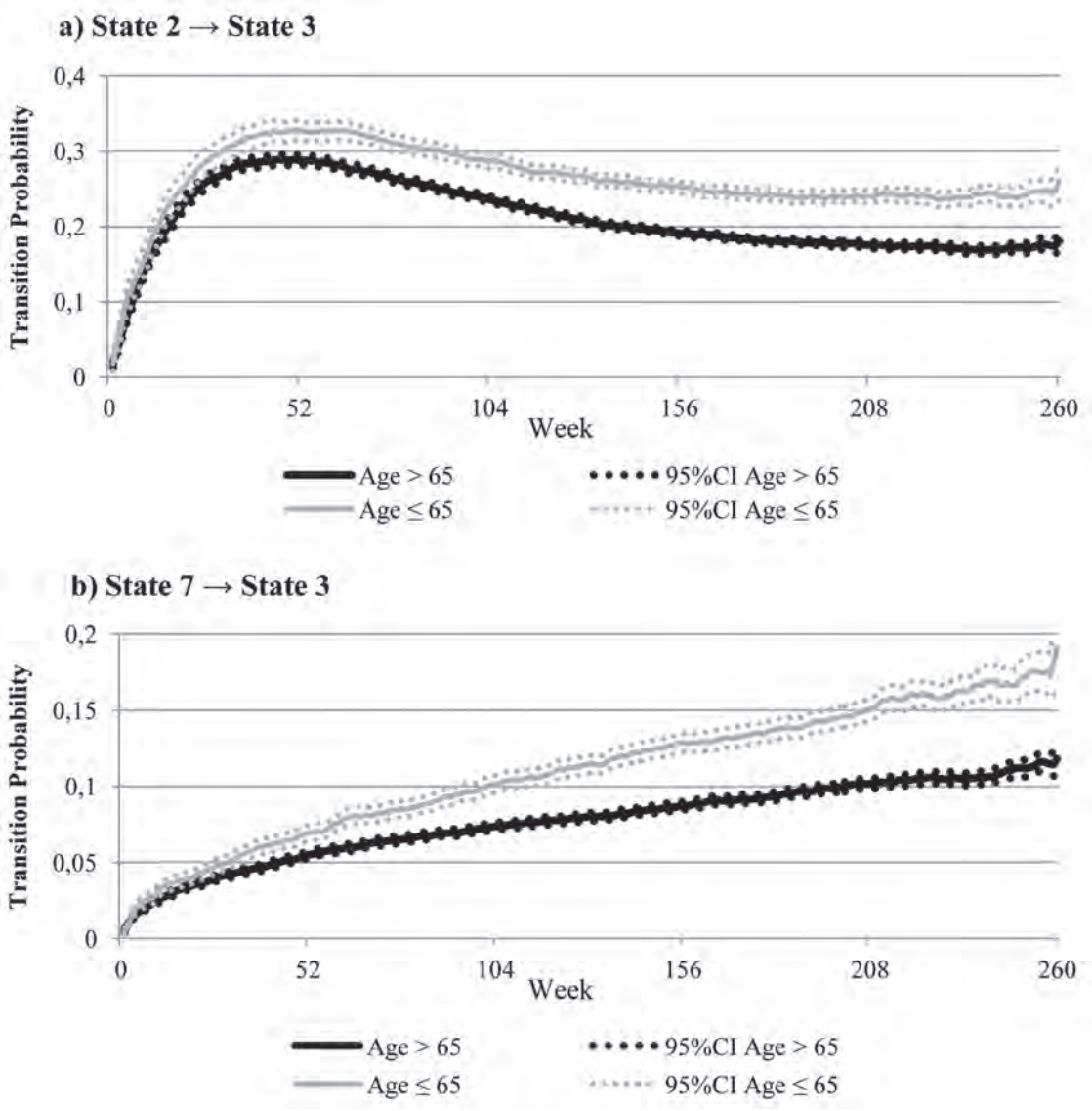
DISCUSSION

The analysis presented herein has the aim of verifying whether the application of stochastic multi-state techniques to data obtained as information flows from hospital discharge sheets and drug prescription records can describe the health care pathways of patients with a chronic-degenerative disease. The transitions of state according to hospitalizations were completed with what we define as "prescription" states.

The achievement of this aim depended above all on the choice of the multi-state model

FIGURE 4

PROBABILITY OF TRANSITION STRATIFIED BY AGE CLASS.
4a) STATE 2 -> STATE 3; 4b) STATE 7 -> STATE 3



to apply, that in turn depends on the type of data available. In fact, since no buffer state was available all the observations were censored to the right, whereas because the target was the capture of all incident cases of COPD in order to follow the disease process from the start, there was no censoring to the left. The presence of such data features needs to be treated with some caution, especially when it is necessary to build a likelihood function for the model [20]. Klein and Moeschberger [29] and Commenges [30] dealt in detail with how to treat incomplete observations. As in many cases reported in the literature, in the model applied in this work, too, the observations censoring mechanism was assumed to be independent of the process under study. This assumption implies that the time to censoring and survival times are independent. In other words, censoring is independent of an unusually high or low risk for occurrence of an event, therefore survival time for censored and uncensored individuals is the same and the removal of censored individuals from the analysis would yield an unbiased estimate of time to event.

Another essential aspect when choosing a multi-state model is the assumption of the invariability over time of the intensity of transition that regulates the stochastic process. Because it was necessary to choose a Markov process that is not homogeneous over time, we decided to perform non parametric modeling of the stochastic process. This was done to take into account the right-censored observations, despite the assumption of the independence of the observations censoring mechanism from the process under study, that should be considered with caution when building the likelihood function. In addition the complexity of the hypothesized multi-state model underlined the frequent convergence of problems in parametric estimation methods, that are also due to the computational difficulties of MLE. The Estimator described by Nelson-Aalen, used to estimate the Empirical Probabilities of Transition, is a consistent estimator that allowed us to analyze a complex model with seven transitory states. The effect of the covariates on the probabilities of transition was verified by stratified analysis.

The decision to implement so complex a multi-state model stems from the need to adapt administrative data to the study of the typical evolution of a chronic disease. In the chosen disease, specifically COPD, the GOLD guidelines [15] for staging and treatment were adopted.

The stage of individual subjects affected by COPD was identified on the basis of hospital discharge records and drug prescriptions that are only indirectly correlated to the stage and hence the degree of severity of the disease, as defined by the GOLD guidelines (stages I to IV) based on clinical parameters.

Therefore, the model we propose to describe the progression of COPD can only partially capture the true evolution of the disease. Another cause of poor correspondence is the high percentage of hospitalizations for non acute COPD, observed by studying the administrative databases but not justifiable as a stage of disease evolution and so rather to be considered to reflect inappropriate patient management.

Administrative databases have been used in recent studies to assess both the efficacy of treatment and the cost-benefit ratio. The primary end-points were largely new hospitalizations and total expenses for the services provider [31,32,33]. In our study, too, the impact of further hospitalizations was evident, but the model we employ is very different from the traditional logistic model or the multivariate models frequently used by other authors, because our intent was to identify the treatment pathways and possible variations.

The model in this paper does not take into account the misclassification of patients due to incomplete information in the two administrative databases used. A possible solution could be the application of a hidden Markov model as in a previous experience [citaz] with liver cirrhosis and liver cancer, but there were few states and transitions, an absorbing state was included, and the authors used only one source of data, with an easier model to fit.

In 2012, Lodewijckx and other authors [34] conducted a study with a similar intent, but in that case they designed a cRCT (cluster Randomized Clinical Trial) and the definition of the pathway was more precise, because the data source and study design were based on prospective collection of clinical data.

In the multi-state model implemented by Sun and other authors in 2011 [35], the Markov process was applied to compare the effectiveness of a therapeutic association. However, also in this case the data were not retrieved from an administrative database, but rather from works in literature, and the transitions of state were

conditioned by what could be deduced from these works. Limitations included the fact that different settings contributed to generate the model estimates, as well as the fact that the different studies from which they were extrapolated often referred to brief observation periods.

It seems clear that structuring a multi-state model on administrative data can make it possible to provide health management staff with useful data for cost-benefit assessments or simply to monitor the appropriateness of treatment. This avoids the need for complex prospective or experimental studies and, provided the database has been correctly administered, allows them to refer to their own organizational settings and evaluate quite long periods of time.

CONCLUSIONS

The present multi-state analysis applied to longitudinal data obtained from health care administrative databases allowed us to evaluate the diagnostic-therapeutic pathway faced by subjects affected by a chronic degenerative disease.

The results of this application highlighted the valuable contribution of data on drugs consumption. It could be seen, for instance, that in the case of COPD some subjects whose disease became manifest at hospitalization had already been taking inhaled corticosteroids (ICS) and above all the set combination of ICS+long-acting beta agonists, and that in these subjects ICS replace short-acting beta agonists, anticholinergics, long acting beta agonists and theophylline already after a few weeks.

It was found that the administration of inhaled corticosteroids delays a hospital admission for acute respiratory failure by about 6 months, as compared to short-acting beta agonists (alone or in set combination with antiasthmatics) or anticholinergics. The results of this study show that the probabilities of a transition and their distribution in relation to time, sex, age and clinical status can be a helpful tool for those operating in the health care sector, who are called upon to carry out therapeutic decision-making from the standpoints of both efficacious clinical management and an efficient use of resources.

AUTHORS' CONTRIBUTIONS: *NB conceived the study, conducted the analysis, wrote the manuscript. PT conceived the study, collaborated in the analysis, drafted the manuscript. GS supervised the analysis and reviewed the manuscript. All authors read and approved the final manuscript.*

COMPETING INTERESTS: *The authors declare that they have no competing interests.*

ACKNOWLEDGEMENTS: *We thank Sabrina Triggiani to help in data management.*

We thank Mary Victoria Pragnell, B.A. for language help.

References

[1] Ilmarinen J, Costa G. L'invecchiamento dei lavoratori nell'Unione Europea. Med Lav 2000; 91(4):279-295.

[2] Progetti e campagne salute. Available from: http://www.cittadinanzattiva.it/comunicati/salute/altro.html [Accessed 26 September, 2014]

[3] Altman DG, De Stavola BL, Love SB, Stepniewska KA. Review of survival analyses published in cancer journals. Br J Cancer 1995;72:511-8. [http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2033978/pdf/brjcancer00042-0259.pdf]

[4] Pepe MS, Mori M. Kaplan-Meier, marginal or conditional probability curves in summarizing competing risks failure time data? Stat Med 1993;12:737-51.

[5] Lin DY. Non-parametric inference for cumulative incidence functions in competing risks studies. Stat Med 1997;16:901-10. [http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-0258(19970430)16:8%3C901::AID-SIM543%3E3.0.CO;2-M/pdf]

[6] Lunn M, McNeil D. Applying Cox regression to competing risks. Biometrics 1995;51:524-32.

[7] Tai BC, Machin D, White I, Gebski V. Competing risks

analysis of patients with osteosarcoma: a comparison of four different approaches. Stat Med 2001;20:661-84.

[8] Dancourt V, Quantin C, Abrahamowicz M, Binquet C, Alioum A, Faivre J. Modeling recurrence in colorectal cancer. J Clin Epidemiol 2004;57:243-51. [http://www.sciencedirect.com/science/article/pii/S0895435603003470]

[9] Andersen PK, Keiding N. Multi-state models for event history analysis. Stat Methods Med Res 2002;11:91-115.

[10] Commenges D. Multi-state models in epidemiology. Lifetime Data Anal 1999;5:315-27.

[11] Hougaard P. Multi-state models: a review. Lifetime Data Anal 1999;5:239-64.

[12] Kay R. The analysis of transition times in a multistate stochastic process using proportional hazard regression model. Commun. Statist. Theor. Meth. 1982;11:1743–56.

[13] Zhang B, Lin H, Hunter DJ, et al., A Multistate Transition Model for Osteoarthritis Pain Change. Commun Stat Theory Methods. 2009;38(18):3297-306. [http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3039455/]

[14] Price MJ, Ades AE, De Angelis D, et al., . Risk of Pelvic Inflammatory Disease Following Chlamydia trachomatis Infection: Analysis of Prospective Studies With a Multistate Model. Am J Epidemiol. 2013;178(3):484-92 [http://aje.oxfordjournals.org/content/178/3/484.long]

[15] Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management and prevention of COPD (2010). Available from: http://www.goldcopd.org/uploads/users/files/GOLDReport_April112011.pdf [Accessed 26 September, 2014].

[16] Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chron Dis 1987;40:373-383.

[17] Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9- CM administrative databases. J Clin Epidemiol 1992;45:613-619.

[18] Aalen O, Johansen S. An empirical transition matrix for non-homogeneous Markov chains based on censored observations. Scand J Statist 1978;5:141-150.

[19] Prinja S, Gupta N, Verma R. Censoring in Clinical Trials: Review of Survival Analysis Techniques. Indian J Community Med. 2010 Apr;35(2): 217–221. [http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2940174/]

[20] Meira-Machado L, de Uña-Alvarez J, Cadarso-Suárez C, Andersen PK. Multi-state models for the analysis of time-to-event data. Stat Methods Med Res. 2009 Apr;18(2):195-222. [http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2692556/]

[21] Sun Y, Lee J. Testing independent censoring for longitudinal data. Stat Sin. 2011;21(3):1315–1339. [http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3749084/]

[22] Ling Lan. Nonparametric Inference for Multistate Model. ProQuest, 2008.

[23] Alioum A, Commenges D. MKVPCI: A computer program for Markov models with piecewise constant intensities and covariates. Comput Methods Programs Biomed 2001;64:109-19.

[24] Gill R, Johansen S. A Survey of Product-Integration with a View toward Application in Survival Analysis. Ann Statist 1990;18:1501-55.

[25] Andersen PK, Borgan O¬, Gill RD, Keiding N. Statistical Models Based on Counting Processes. Springer-Verlag, New-York; 1995.

[26] Allignol A, Beyersmann J, Schumacher M. mvna: An R Package for the Nelson-Aalen Estimator in Multistate Models. R News 2008;8:48-50.

[27] R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [http://www.R-project.org/]

[28] Allignol A, Schumacher M, Beyersmann J. Empirical Transition Matrix of Multi-State Models: The etm Package. Journal of Statistical Software 2011;38(4):1-15. [http://www.jstatsoft.org/v38/i04/]

[29] Klein JP, Moeschberger ML. Survival Analysis: Techniques for Censored and Truncated Data. (2nd ed) Springer-Verlag, NewYork; 2003.

[30] Commenges D. Inference for multi-state models from interval-censored data. Stat Methods Med Res 2002;11:167-182. [http://www.sciencedirect.com/science/article/pii/S0169260700000948]

[31] Menzin J, Boulanger L, Marton J, et al., The economic burden of chronic obstructive pulmonary disease (COPD) in a U.S. Medicare population. Respir Med 2008;102:1248-1256. [http://www.sciencedirect.com/science/article/pii/S0954611108001431]

[32] Dalal AA, Shah M, D'Souza AO, Crater GD. Rehospitalization risks and outcomes in COPD patients receiving maintenance pharmacotherapy. Respir Med 2012;106:829-837. [http://www.sciencedirect.com/science/article/pii/S0954611111003994]

[33] Wang Y, Stavem K, Humerfelt S, Dahl FA, Haugen T: Readmissions for COPD. propensity case-matched comparison between pulmonary and non-pulmonary departments. Clin Respir J 2013;7:375-381.

[34] Lodewijckx C, Decramer M, Sermeus W, Panella M, Deneckere S, Vanhaecht K. Eight-step method to build the clinical content of an evidence-based care pathway: the case for COPD exacerbation. Trials 2012;13:229. [http://www.trialsjournal.com/content/13/1/229]

[35] Sun SX, Marynchenko M, Banerjee R, et al.,Cost-effectiveness analysis of roflumilast/tiotropium therapy versus tiotropium monotherapy for treating severe-to-very severe COPD. J Med Econ 2011;14:805-15.