Prediction of the Burden of Mental Diseases Using a Microsimulation Model

Andreas Bauer[1*], Christoph Urach[2], Felix Breitenecker[1]

[1]Department of Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstraße 8-10, 1040 Vienna, Austria; [*]*andreas.e101.bauer@tuwien.ac.at*

[2]dwh simulation services, Neustiftgasse 57-59, 1070 Vienna, Austria

SNE Simulation Notes Europe SNE 26(1), 2016, 47-52
DOI: 10.11128/sne.26.tn.10328
Received: January 20, 2016; Revised: February 12, 2016;
Accepted: March 5, 2016;

Abstract. This work aims to predict the burden of mental diseases to provide sufficient capacities for treatment. A microsimulation model is built to simulate the course of events of mentally ill patients. Three scenarios of simulations are defined to test the consequences of using differently detailed patient-level data on result quality. Significant differences in the results are encountered. The overall numbers and times of patients events are analyzed as well as the number of events per patient. The differences between the results for the different scenarios and for the various subpopulations regarding patient parameters are pointed out. For example, psychotic patients tend to have more readmissions. Further analyses regarding the connection between ambulant contacts and readmissions to the hospital are performed. Also, regional differences of Lower Austria compared to the entire Austrian population are analyzed. Finally, an intervention strategy with compulsory ambulant contacts is examined.

Introduction

The prediction of the burden of mental diseases is important to provide sufficient capacities in the hospitals. The overall number of readmissions to hospital is estimated as well as the numbers of events for subpopulations defined by certain patient characteristics to determine the required capacity and its change over time. Also, the influence of outpatient contacts to a psychiatrist on number and times of readmissions is examined.

The consideration of regional aspects is important for the accuracy of the simulation results. So, the situation of patients with mental diseases for Lower Austria is investigated in detail and compared with the situation of entire Austria.

The availability of patient data is often a problem. In these cases privacy protection only allows usage of k-anonymized data. So, it is not certain to get significant results with the given data. Differently detailed patient-level data based on the same set is used to analyze the effect on the quality of the outcome.

1 Survival Analysis

Methods from the field of survival analysis are used to build the statistical model behind the simulation model.

Survival analysis deals with the analysis of data of the time until the occurrence of a particular event. This kind of data is frequently encountered in medical research and referred as survival data.

Survival analysis mainly deals with the estimation of the survival function and the hazard function. The survival function $S(t)$ gives the probability that the events has not occurred until time t and the hazard function $\lambda(t)$ gives the instantaneous rate of occurrence of the event.

The cumulative hazard function Λ is defined as

$$\Lambda(t) = \int_0^t \lambda(x)dx. \tag{1}$$

The Nelson-Aalen estimate is an estimate for the cumulative hazard function Λ [1]. Let d_t and n_t denote the number of people that experience the event at time t respectively are at risk at time t. Let t_i denote the event times. Then Λ can be estimated by

$$\hat{\Lambda}(t) = \sum_{i:t_i \leq t} \frac{d_{t_i}}{n_{t_i}}. \tag{2}$$

1.1 Cox model

The Cox model is a model for the hazard function [2]. It assumes that the ratio of the hazards of different exposure groups remains constant over time. The hazard at time t for individual i with covariate vector X_i is assumed to be

$$\lambda_i(t) = \lambda_0(t)\exp(X_i\beta) \tag{3}$$

where λ_0 is an unspecified nonnegative function called baseline hazard function and β is a vector of regression coefficients.

Stratified Cox model. The stratified Cox model is an extension of the Cox model and allows for multiple strata [3]. The strata divide the subjects into disjoint groups and each subject is member of exactly one stratum. Each of which has a distinct baseline hazard function but common values for the coefficient vector β. The hazard for individual i belonging to stratum k is

$$\lambda_k(t)e^{X_i\beta}. \tag{4}$$

2 Model

2.1 Data

Two datasets are used in this work. Dataset *dataaut* consists of data of patients from Austria and dataset *datanoe* consists of data of patient from Lower Austria. Patient parameters are age, sex, length of stay in the psychiatric department and the diagnosis made during the initial stay at the hospital. Also times of readmissions, ambulant visits to a psychiatrist and deaths are included. It depends on the chosen scenario which events are actually considered in the model. The data samples are used for the parametrization and the sampling of the population of the simulation model.

2.2 Model description

The chosen model type is a microsimulation model. That means that it follows the bottom-up approach and every single individual is modeled. This approach is chosen because not only the cross-sectional analysis is important but also the longitudinal pathways of single individuals. Furthermore, this approach is suitable for the analysis of different policies and scenarios. Another reason is that the characteristics of the individuals are manageable with a bottom-up approach.

The events of a patient are expressed by state changes. The possible ways through the states are described by a transition matrix which can be interpreted as a directed acyclic graph. Every individual starts in state R (released after the first admission to hospital). If the most recent event of the patient was the ith readmission, the patient is in state A_i and if the most recent event was the ith ambulant psychiatrist visit, the patient is in state P_i. The dead patients are in state D. In order to calculate the times of the events respectively the probabilities for the events to occur the hazard and survival functions have to be modeled. The Cox model and the Nelson-Aalen estimate are applied to determine the according statistical models. The hazard functions for the transitions are estimated with a stratified Cox model. The strata represent the transitions [4].

The overall simulation time is fixed. The simulation starts for every patient with the day of the release from the first stay in a psychiatric department of a hospital. The simulation is executed in discrete time steps of one day.

The starting population is sampled from real data described in Section 2.1. It is modeled as a closed cohort, so there is no change in the size of the population except for deaths.

3 Simulations

3.1 Definition of scenarios

The given datasets of full records of patients are used to examine the consequences of using differently detailed patient-level data on result quality. Data at hand are often incomplete or contain only information about specific events due to data protection issues, loss of data and many other reasons. Three scenarios with different number and order of the readmissions that are used for the Cox model are defined. So, the scenarios only differ in terms of the parametrization. Each scenario is executed with and without ambulant contacts to psychiatrists.

In scenario 1, only the first readmission of each patient is considered and all the other readmissions that are available in the data are dismissed. In the simulation, the transition rates from any readmission state A_i to states A_{i+1}, P_{i+1} and D are assumed to be equal for all i. There are two versions of this scenario considering the visits to the psychiatrist, one with ambulant contacts ($1a$) and one without ($1b$).

In scenario 2, the first z readmissions of each patient are considered. All readmissions are considered independently from each other, even if they belong to the same subject. So, there is no order of the readmissions and every readmission is regarded as first readmission. Like in scenario 1, the transition rates from any readmission state A_i to states A_{i+1}, P_{i+1} and D are assumed to be equal for all i.

Between two consecutive admissions up to one contact to a psychiatrist is considered. Again, there are two versions of this scenario, one with contacts to the psychiatrist ($2a$) and one without any contacts ($2b$).

In scenario 3, the first z readmissions of each patient are considered with the same number z as in scenario 2 but in contrast to scenario 2 the readmissions are ordered. That means that for the first z readmissions the transition rates from a readmission state are independent. From the $(z+1)$th readmission on, the rates are assumed to be equal to the transition rates starting from state A_z. In scenario $3a$, at most one contact to a psychiatrist between two consecutive admissions is possible. Therefore, also the psychiatrist contacts are ordered. In scenario $3b$, no psychiatrist contacts are considered.

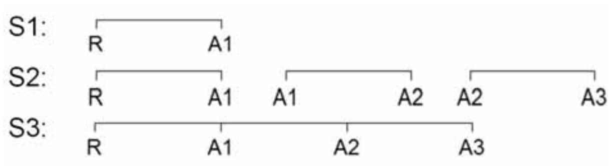

Figure 1: Schematic representation of the information needed in the three scenarios.

In Figure 1, the utilization of data in the three scenarios for a time line with three readmissions is presented.

3.2 Results

Simulations for single scenarios and comparisons of the matching scenarios that only differ in the inclusion of psychiatrist contacts and comparisons of all scenarios with and without contacts to a psychiatrist for populations from Austria and Lower Austria are carried out.

The simulation time is 2 years, because the majority of the readmissions, especially of the first readmissions, which are the most crucial events, happen within this period. A population of 18638 individuals is sampled from datasets *dataaut* and *datanoe*.

An exemplary result for scenario $1a$ is shown in Figure 2. The evolution of the distribution of patients over the states is shown. On the x-axis time is plotted, on the y-axis the percentage of each state is plotted on top of each other. The area under each curve is filled with a distinctive color. The states are coded with colors. Dark green represents state R, the readmission states are assigned to lighter shades of green, the psychiatrist states have shades of red and dark red represents the state death D.

The share of the patients in state R decreases almost exponentially. After two years about 50 percent are remaining in state R. The percentage of deaths increases almost linearly. At the end of the simulation around 3.4% of the population is dead.

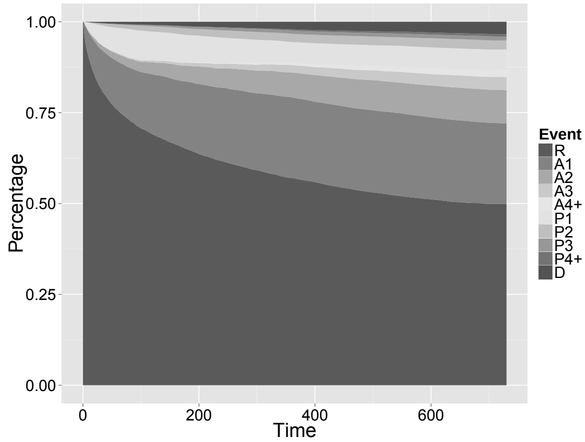

Figure 2: Evolution of patient over the states for scenario $1a$ for a population from Austria.

3.3 Comparison of scenarios

Patients with outpatient contacts to a psychiatrist (OPC) are compared with those without outpatient contacts (non-OPC). In Figure 3, the percentage of patients with readmissions is shown for both groups and all scenarios. Patients with ambulant contacts have a much higher percentage of readmissions, in scenarios $2a$ and $3a$ even twice as much as patients without ambulant contacts.

In Table 1, the percentages of patients with readmissions are compared for all scenarios. It can be seen that scenarios $2a$ and $2b$ have a higher percentage of readmissions. The reason is that the transition probability from state R to state A_1 is higher in scenario 2, because in scenario 1 only the first readmissions from the data are used to fit the rate from state R to A_1 while in scenario 2 all readmission times are treated as first readmission times. Since the times for the later read-

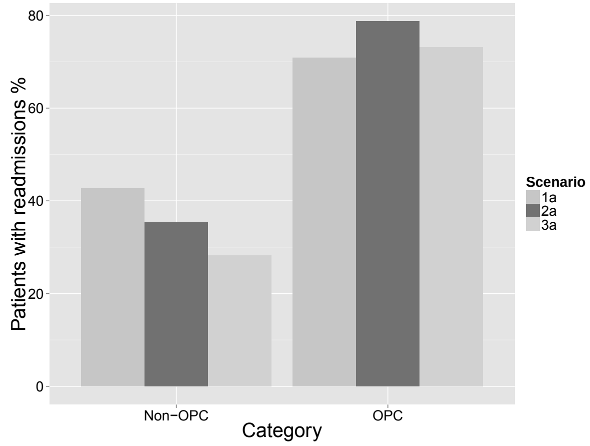

Figure 3: Comparison of the percentages of patients with readmissions between patients with and without outpatient contacts.

missions are shorter in average, the median of the first readmission times drops from 75 days for scenarios 1 and 3 to 63 days for scenario 2. This leads to a higher probability for entering state A_1 and having a readmission.

Scenario	$1a$	$1b$	$2a$	$2b$	$3a$	$3b$
Readmissions (%)	42	43	51	51	42	42

Table 1: Percentage of patients with readmissions.

In order to analyze the results in greater depth typical pathways of patients during the simulation are defined. In addition to the number of readmissions of a patient the times of these are taken into account to classify the pathways.

Nine typical, distinctive pathways are chosen to split the population in roughly equally sized classes. Only the class of patients with no readmission is much bigger than the others.

In Table 2, an overview of the classification for patients without ambulant contacts is given.

In Figure 4, the sizes of the classes for scenarios $1b$, $2b$ and $3b$ are presented. Class 1 is not shown in the plot, because the number of patients without readmissions has already been analyzed and the focus is on patients with readmissions.

In scenarios $1b$ and $2b$ are more than twice as many patients in class 2 than in scenarios $3b$. That means more individuals have exactly one readmission shortly after the release. This can be explained by the fact that

Class	Readmissions	Month of first readmission
1	0	–
2	1	1
3	1	2-6
4	1	7-12
5	1	13-24
6	2-4	1
7	2-4	2-6
8	2-4	7-24
9	> 4	any

Table 2: Classification of patient pathways.

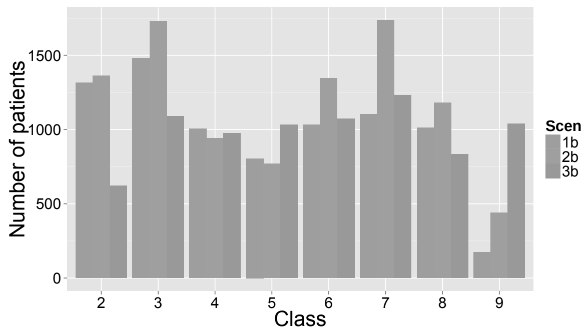

Figure 4: Size of classes for the scenarios without ambulant contacts.

in scenario $3b$ patients in average have more readmissions and in scenario $2b$ more patients have readmissions in general. The number of individuals with more than four readmissions is much higher in scenario $3b$.

3.4 Intervention

An intervention strategy is examined to possibly reduce the number of readmissions. According to this strategy a compulsory visit to an ambulant psychiatrist 30 days after every admission to hospital is implemented. The question is, if this strategy can reduce the number of readmissions to hospital.

Type of Event	No intervention	Intervention
Readmissions	42.2	67.8
OPC	27.7	99.7
Deaths	3.9	3.6

Table 3: Comparison of percentages of the occurrence of events for scenario $3a$ with and without intervention.

In Table 3, the percentages of patients with readmissions, ambulant psychiatrist contacts (OPC) and deaths are compared for scenario $3a$. The percentage of patients with readmissions is much higher with the intervention strategy. This leads to the conclusion that an ambulant contact increases the probability for a readmission. The comparison of OPC and non-OPC patients in Figure 3 already hypothesizes this result. In the intervention scenario almost every patient visits a psychiatrist during the simulation. So, this strategy does not succeed in reducing the number of readmissions.

3.5 Comparison of simulations for Austria and Lower Austria

Results of the simulations for populations from Austria and Lower Austria are compared in terms of number and times of events. The evolutions of the patients distribution over the states for the two populations in scenario $3a$ are presented in Figures 5 and 6. The share of the patients in state R has a similar evolution for both simulations and decreases almost exponentially. About 50 percent of the patients from Austria are remaining in state R at the end of the simulation, about 47 percent of the other population. So, patients from Lower Austria have more readmissions, since the numbers for states P_i and D are very similar for both populations.

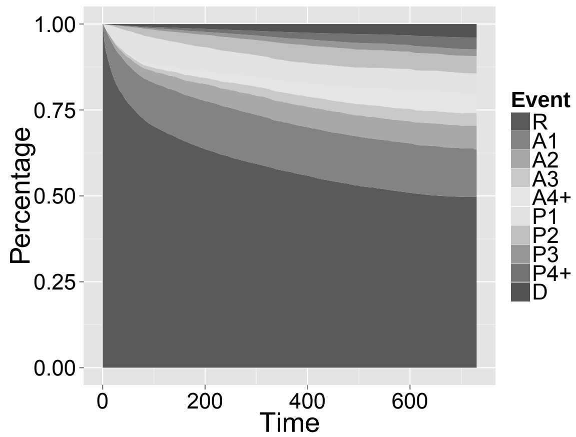

Figure 5: Evolution of the patients distribution over the states in scenario $3a$ for Austria.

The proportions of the two populations with readmissions, ambulant psychiatrist contacts (OPC) and deaths are displayed in Table 4. The population from Lower Austria has more events of every type. This can be linked to a higher percentage of psychotic patients in that population.

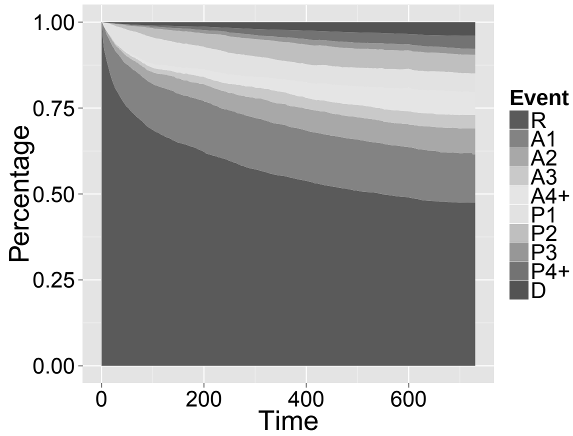

Figure 6: Evolution of the patients distribution over the states in scenario $3a$ for Lower Austria.

Type of Event	Austria	Lower Austria
Readmissions	42.2	44.8
OPC	27.7	28.6
Deaths	3.9	4.0

Table 4: Comparison of the percentages of the populations with readmissions, ambulant contacts and deaths.

3.6 Sensitivity analysis

The influence of the composition of the population on the number and times of readmissions is examined in the course of a sensitivity analysis.

Firstly, a base case with a random subpopulation sampled from dataset *dataaut* is considered. Starting from that population other populations are generated by changing only one parameter of all patients at a time while leaving the other parameters unchanged. 12 populations are generated: all male/female, all five years younger/older, length of stay 50% shorter/longer and all with each of the six diagnosis groups.

The number of patients with readmissions and the deviation of the number from the base case is calculated. In Figure 7, a tornado plot for the number of patients with readmissions for all populations is presented. This diagram is a bar chart with bars listed vertically and ordered by length. The vertical line at 0 marks the base case with no deviation. The bar for each parameter reaches from the deviation of the highest value to the deviation of the lowest value of the populations where that particular parameter is changed.

The populations with single diagnosis groups have the highest deviations ranging from has about 23% less

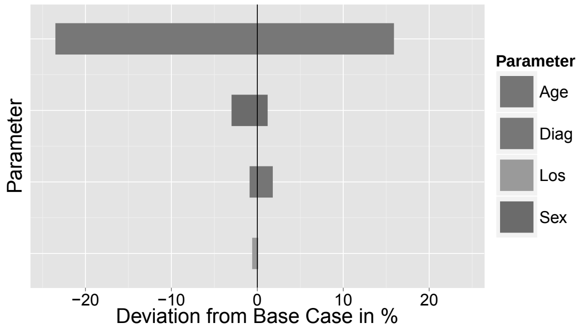

Figure 7: Tornado plot for number of patients with readmissions.

patients with readmissions than the base case to about 15% more. The male population has about 3% less readmissions than the base case, the female has 1.2% more. Also, the older population and the population with longer stays in hospital have a slightly higher number of readmissions compared to the younger patients and the population with shorter stays. All of these deviations are below 2%.

4 Conclusions

The microsimulation model is an appropriate tool to model the course of events of patients. Both the longitudinal analyses of the single patients as well as the cross-sectional analyses can be carried out with little effort.

In general, the results show an exponential decrease over time of the number of patients with no event. Nevertheless, about half of the patients have no readmissions during the simulation. The number of patients with one ambulant contact and no readmission has its peak after half a year and declines afterwards. So, many of these patients have a readmission soon after the visit to the psychiatrist. The percentage of patients with a particular number of readmissions is indirectly proportional to the number of readmissions.

The comparison of the results of different scenarios shows that in scenarios 2*a* and 2*b* the patients have more readmission than in the other scenarios. This is due to an overestimation of the number of readmissions because the order of the readmissions is not considered in these scenarios. So, the results of scenarios with a lower level of data detail show significantly varying results from scenario 3*a* which uses the most detailed data. However, scenario 3*a* requires data of entire patient histories which is rarely available due to data protection issues.

For a more detailed analysis, the population is split into classes defined by times and number of readmissions of patients. In comparison to the other scenarios, the readmissions of patients with only one readmission are later and the average number of readmissions per patients is higher in scenarios 3*a* and 3*b*.

The sensitivity analysis shows that the diagnosis of the population has a dramatic influence on the number of readmissions.

The proportion of patients with readmissions is much higher for patients with previous ambulant psychiatrist visit. Thus, ambulant contacts increase the probability for readmissions and are in most cases an indicator for a worsening of the condition of the patient. This also leads to the fail of the reduction of readmissions by the intervention strategy of compulsory visits to a psychiatrist after a certain time after the last admission.

The comparison of the populations of whole Austria and Lower Austria shows that more patients of the latter have readmissions and also ambulant contacts. This can be the result of the differing composition of the populations regarding the parameter distributions.

References

[1] Kirkwood BR, Sterne JAC. *Essential Medical Statistics*. Essentials. Wiley. 2003.

[2] Cox DR. Regression models and life-tables. *Journal of the Royal Statistical Society Series B (Methodological)*. 1972;34(2):187–220.

[3] Therneau TM, Grambsch PM. *Modeling survival data: extending the Cox model*. Springer Science & Business Media. 2000.

[4] Beyersmann J, Putter H. A note on computing average state occupation times. *Demographic Research*. 2014; 30(62):1681–1696.