Demand Forecasts for Chronic Cardiovascular Diseases Medication Based on Markov Chains

Pronósticos de Demanda de Medicamentos para Enfermedades Cardiovasculares Crónicas Basadas en Cadenas de Markov

Armando Hermosilla[a,*], Richard Carmagnola[a], Carlos Sauer[b], Eduardo Redondo[a,c] and Luis Centurion[a]

[a] *Industrial Engineering Department, FIUNA – Universidad Nacional de Asuncion, Paraguay*
[b] *Applied Mathematics Department, FIUNA – Universidad Nacional de Asuncion, Paraguay*
[c] *Laboratory of Mechanics and Energy, FIUNA – Universidad Nacional de Asuncion, Paraguay.*
[*] *Corresponding author: ahermosilla@fiuna.edu.py*

Abstract—In this work, we propose and evaluate models to predict the demand for cardiovascular drugs using Markov chains. The models use transactional data of patient medication delivery to identify consumption levels. These levels are considered as Markov chain states. Four model configurations were evaluated, differing on the arrival/departure nature of patients to the system and the inclusion of an idle state. The models were trained with 12 months of real data and tested with a four-month horizon. Also, the models were sequentially applied to an 18-month Losartan consumption data, simulating thus the chained implementation in a real scenario. The MAPE of two months ahead forecasts ranged from 3.92% to 5.55% in three of the four evaluated models. As well, our results showed that variations of the consumption level could be modeled using Markov chains, and in low inventory levels situations, these tools are usable to prioritize patients with higher levels of consumption.

Keywords—Markov Chains, Inventory Management, Demand Planning, Stochastic Processes, Healthcare Supply Chain.

Resumen— En este trabajo proponemos y evaluamos modelos para predecir la demanda de fármacos cardiovasculares utilizando cadenas de Markov. Los modelos utilizan datos transaccionales de entrega de medicamentos a pacientes para identificar niveles de consumo. Estos niveles se consideran estados de la cadena de Markov. Se evaluaron cuatro configuraciones del modelo, que diferían en la naturaleza de llegada/salida de los pacientes al sistema y la inclusión de un estado inactivo. Los modelos fueron entrenados con 12 meses de datos reales y probados con un horizonte de cuatro meses. Además, los modelos se aplicaron secuencialmente a datos de consumo de Losartán de 18 meses, simulando así la implementación encadenada en un escenario real. Las previsiones del MAPE a dos meses vista oscilaron entre 3,92% y 5,55% en tres de los cuatro modelos evaluados. Además, nuestros resultados mostraron que las variaciones del nivel de consumo podrían modelarse utilizando cadenas de Markov y, en situaciones de bajos niveles de inventario, estas herramientas se pueden utilizar para priorizar a los pacientes con mayores niveles de consumo.

Palabras Claves— Cadenas de Markov, Administración de Inventarios, Planeación de Demanda, Procesos Estocásticos, Cadena de Suministro en sector Salud.

I. INTRODUCTION

IN Paraguay, the purchase of healthcare products and medical supplies corresponded to 20% of public procurement spending from 2014 to 2018 [1]. Yet, the inventory of medicines in public health facilities has been shown to reach levels of scarcity in several occasions [2,3]. Therefore, developing more precise inventory managing models might be conceived as an urgent and strategic need to be tackled.

As chronic diseases allow to be expressed in terms of health conditions or states, the concept of homogenous Markov chains is commonly used to analyze the clinical evolution of patients during their treatments [4]. In fact, in the literature, we might find that this type of chains has been widely used to forecast the prevalence of specific illness for long-time horizons as in the case of [5], who forecasted for the prevalence of coronary heart diseases for a ten-year period and concluded that Markov chains could be openly used to model the progression of non-communicable diseases. Also, the results of this research

were used to: first, prepare an adequate prevention work. And second, to plan the need of specialized physicians for this type of disease.

The application of the Markov chain theory in healthcare resource management has shown to be diverse, going from forecasting healthcare expenditures using intrinsic patients' attributes such as diagnostic, gender, and age [6], to controlling the intermittent demand for certain products [7]. The popularity of this approach, besides lying in the convenient "memory-less" characteristic, might be due to the fact that Markov chains are flexible enough to manage different types of criteria, as several authors have pointed out. Nevertheless, few works have addressed the forecasting of healthcare resources for specific chronic diseases during short periods of time. Thus, based on the applicability of this approach to healthcare resource management and demand forecasting, this research aims to fulfill the mentioned gap by modeling the levels of consumption of a given drug as Markov chain states, looking forward to calculating the number of patients in each level of consumption and using it to forecast cardiovascular medication demand. To this end, a police hospital was selected to be the case study. Potential and involved patients only include the national police force and their first-degree relatives (parents, offspring, and spouse). This means that the reachable population remains considerably stable through time, and certain assumptions regarding the entrance of new patients could be used for the analysis of medication consumption.

Based on data from the National Direction of the Public Procurement of Paraguay, and data from the analyzed hospital, cardiovascular medication has the highest average cost per subtype of product. Also, in the last four years, more than 20% of the patients benefitted by medication were patients who were benefitted with cardiovascular medication. Among all types of cardiovascular medication, four drugs benefit the highest number of patients (2,178 to 4,735 patients). This group includes patients been treated with Atorvastatin, Losartan potassium, Enalapril, and Amlodipine Besylate. Therefore, these drugs were chosen to apply the forecasting models.

The proposed models use the Jenks optimization algorithm to determine the levels of consumption to be used as Markov Chain states [8]. Twelve months of demand history are used to obtain the probabilistic transition matrix. To cover variants regarding the arrival/departure nature of patients, four different model configurations are proposed.

All in all, the proposed model describes the transition between consumption levels of patients in the studied hospital to forecast the aggregated demand for each medicine. Decision makers can use this forecast for procurement planning and prioritize patients in higher consumption levels during low-level inventory periods. Also, the quantitative results suggest that the simplest of the studied models tend to project more precise demand forecasts.
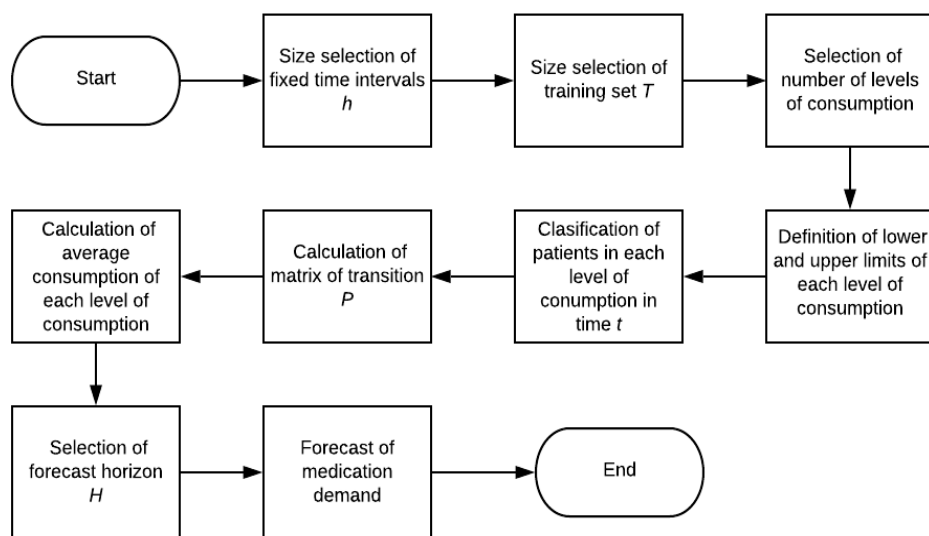

Fig. 1. Proposed Forecasting Model Process

II. METHODS AND PROCEDURES

The diagram in Figure 1 describes the steps composing the followed framework to construct the proposed forecast models.

Markov chains are one type of stochastic model that describes a stochastic process. Their main characteristic

is that their values for a defined time t are conditioned by the value of the process in the time $t-1$. In this research, the states of the Markov chain represent the different levels of consumption on which one patient could be. To analyze how a given patient moves through these levels over time, the transition matrix P is used. Every component P_{ij} represents the probability of transitioning from the level of consumption i to the level

of consumption j in a fixed time interval h.

To define the lower and upper limits of each level of consumption, Jenks optimization algorithm was used. The algorithm is fed by the quantities delivered to patients in the historical data T. The method classifies data into groups by minimizing the absolute deviation of data within a group from the average of the same group.

We define the state vector $x_t = (m_{1,t}, m_{2,t}, \dots, m_{n,t})$ where $m_{n,t}$ corresponds to the number of patients in state n at time t. The average consumption of a state n over the training period T is given by:

$$AvgCon_n = \frac{\sum_T Con_{n,t}}{\sum_T m_{n,t}}$$

where $Con_{n,t}$ represents the consumption by patients that were in state n at time t.

The probability P_{ij} of transitioning from state i to state j within the time interval h, is obtained as:

$$P_{ij} = \frac{Tr_{i,j,h}}{\sum_{j=1}^{n} Tr_{i,j,h}}$$

where $Tr_{i,j,h}$ is the number of patients that transitioned

from state i to state j within the time interval h.

In this work, four configurations of models were studied. To analyze them, it is necessary to define the following vectors:

$$x_t = (m_{1,t}, m_{2,t}, \dots, m_{n,t})$$
$$x'_{t+1} = (m'_{1,t+1}, m'_{2,t+1}, \dots, m'_{n,t+1})$$
$$E_t = (e_{1,t}, e_{2,t}, \dots, e_{n,t})$$
$$O_t = (o_{1,t}, o_{2,t}, \dots, o_{n,t})$$

The vectors E_t and O_t represent the entrance and exit of patients. Here, $e_{n,t}$ represents the number of new patients that enter to the delivery system in the state n at time t and $o_{n,t}$ represents the number of patients that do not receive any medication at time t and were at state n

at time $t-1$. While, in general, we consider three states with the obtained range of consumption levels, to represent the temporary exit of patients, a fourth state is considered, for patients who occasionally do not receive any medication. Table 1 summarizes the main features and particularities of each model configuration.

TABLE 1. Summary of evaluated configurations

Base model	Model 1	Model 2	Model 3
Matrix of transition P stays constant during the time horizon of the training and testing sets.			
Three levels of consumption exist. Each level has an average consumption per person that stays constant during the time horizon of the training and testing sets.			
The variation of the population in the analyzed system stays null.	The net number of incoming patients in each state is projected. This number is the difference between the new patients and the patients who stop receiving medication.	The state of 0 consumption projects the number of patients that stop receiving medication in the next time step.	
		The number of new incoming patients is projected multiplicatively.	The number of new incoming patients is projected additively.
If one patient stops receiving medication for two months, he/she is treated as a new patient the next time he/she receives medication.		If one patient stops receiving medication for 12 months, he/she is treated as a new patient the next time he/she receives medication.	
It does not use an expansion factor	The expansion factor is assumed to stay constant over time. It is also assumed that this factor could be calculated as a two-term moving average.		
Base model	**Model 1**	**Model 2**	**Model 3**
$n = 3$ states	$n = 3$ states	$n = 4$ states	$n = 4$ states
$x'_{t+1} = x_t P$	$x'_{t+1} = x_t P + F'_{t+1}$	$x'_{t+1} = x_t P \, F'_{t+1}$	$x'_{t+1} = x_t P + F'_{t+1}$
-	$F'_{t+1} = \dfrac{F_t + F_{t-1}}{2}$		
-	$F_t = E_t - O_t$	$F_t = \dfrac{\sum_{n=1}^{n=4} m_{n,t}}{\sum_{n=1}^{n=4} m_{n,t+1}}$	$F_t = E_t = (e_{1,t}, e_{2,t}, \dots, e_{n,t})$

While both the base model and the model 1 represent the system with three states of consumption, the model 1 accounts for a net income of patients to the system. Again, both models 2 and 3 include a state for temporary exits of patients, and they differ in the nature of the system expansion: additive or multiplicative.

Finally, to calculate the expected demand:

$$D'_{t+1} = x'_{t+1} AvgCon$$

where:

$$AvgCon = (AvgCon_1, \ldots, AvgCon_n)$$

Diverse techniques to examine and evaluate forecasting error measurements might be found in the literature, including the traditional ones: mean absolute deviation (MAD), mean square error (MSE) [9], mean absolute scaled error (MASE) [10], mean average percentual error (MAPE) [10,11], and some new suggested error and bias measurements: periods in stock (PIS) and number of shortages (NOS) [12]. As aggregate consumption data were used and therefore small or close to zero denominators were not included, we decided to adopt the mean absolute percentual error (MAPE) to measure forecast quality. The advantage of using this method is that it allows us to have a perspective of the errors as absolute values, thus avoiding that the positive and negative variations cancel each other out. Furthermore, to the best of our knowledge, this is one of the most widely used methods for measuring errors due to its scale-independency and interpretability, fact that makes it an appropriate tool for environments in which multidisciplinary actors interact.

To calculate the inventory levels:

$$Inv'_{t+1} = Inv_t + Q_{t+1} - D'_{t+1}$$

where Inv'_{t+1} represents the projection of the inventory level for $t + 1$, Inv_t represents the real inventory levels, Q_{t+1} the quantity to procure and D'_{t+1} the demand forecast for the period $t + 1$.

Also, the models were sequentially tested with Losartan consumption data in a period of 20 months to simulate a real implementation of the models. The process followed is ($s = s_0$, and $e = e_0$ respectively represent the starting and ending month of the training set):

1. Train the model with the interval between s and e.

2. Forecast for a time horizon H.

3. Evaluate results

4. $s = s + 1, e = e + 1$

5. Repeat from step 1 until all months of data available is covered

III. EXPERIMENTAL SETTING

The data provided by the Central Police Hospital of Paraguay consists of the Losartan Potassium consumption from January 2017 to July 2018, and the Atorvastatin, Amlodipine, and Enalapril consumptions from January 2018 to April 2019.

Because of limited data availability, the training period T was chosen to be 12 months. For more flexibility, fixed time steps h were chosen to be equivalent to periods of 2 months. The forecast horizon H is chosen to be $2h = 4$ months. For one iteration of the proposed forecasting model, the Jenks optimization algorithm defined the following states of consumption. Tables 2 and 3 show respectively the distribution of consumptions and average consumption levels per drug, measured in number of delivered pills.

TABLE 2. Levels of two-month consumption measured in number of pills

Losartan Training: Jan 2017-Dec 2017			**Amlodipine** Training: Jan 2018-Dec 2018		
States	**Lower level**	**Upper level**	**States**	**Lower level**	**Upper level**
1	1	43	1	1	41
2	44	92	2	42	80
3	93	210	3	81	180
Enalapril Training: Jan 2018-Dec 2018			**Atorvastatin** Training: Jan 2018-Dec 2018		
States	**Lower level**	**Upper level**	**States**	**Lower level**	**Upper level**
1	1	42	1	1	44
2	43	80	2	45	75
3	81	180	3	76	150

TABLE 3. Average consumption per person on each level of consumption measured in number of pills

	State 0	State 1	State 2	State 3
Losartan	0.0	25.9	61.9	122.9
Enalapril	0.0	25.9	60.0	112.9
Amlodipine	0.0	28.0	60.0	103.4
Atorvastatin	0.0	28.3	59.9	97.7

IV. RESULTS

All models were evaluated with real drug historical data. For Losartan and Amlodipine, model 2 offered the best results in their testing sets. For Enalapril and Atorvastatin, both the base model and model 3 showed better results in the testing sets. Figure 2 summarizes the results of one iteration of training and testing.

Given the availability of 18 months data for losartan, all models were applied successively forecasting and updating monthly to simulate a real scenario. For the first iteration, the starting month was January 2017 and the ending month December 2017 and a forecasting horizon H of 2 months was used. The procedure was applied 6 times, advancing the training period one month at a time. When applying this successive forecasting, Figure 3 shows that the base model and model 3 have, on average, the lowest absolute percentual errors. This is obtained when measuring the precision of one-step forecasts (a period of two months in advance). All models, except model 2, were essentially unbiased when forecasting with the given losartan data.
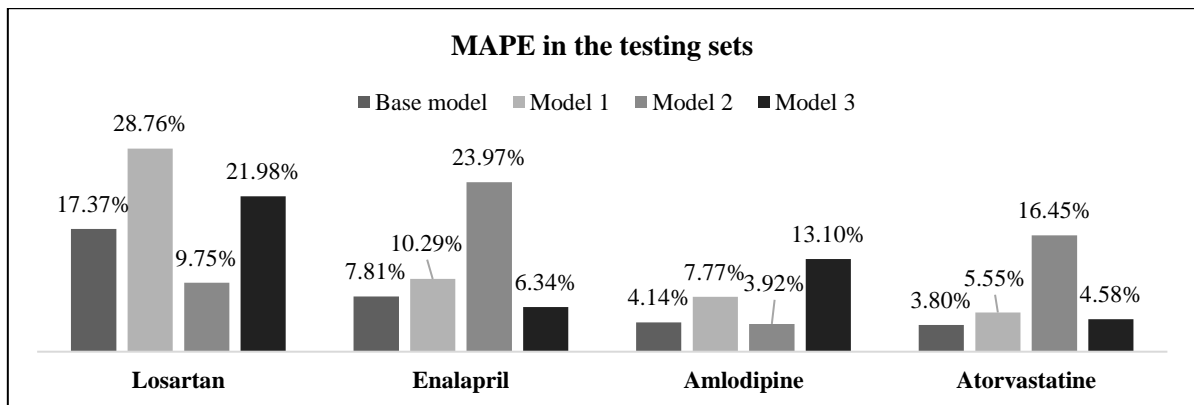

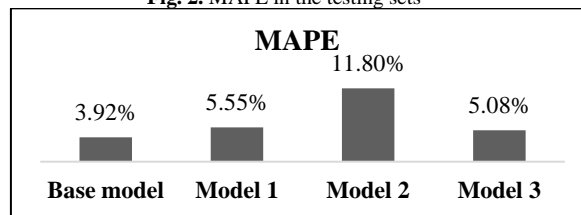**Fig. 2.** MAPE in the testing sets

Fig. 3. One-step forecasts MAPE from sequentially testing with Losartan data

Also, model 3 capability of forecasting inventory levels was tested. An inventory shortage was projected six months in advance. A right level of adherence between the forecasted and real inventory levels was observable. The Figure 4 help us to visualize the resulting forecast for Losartan Potassium for 2018.
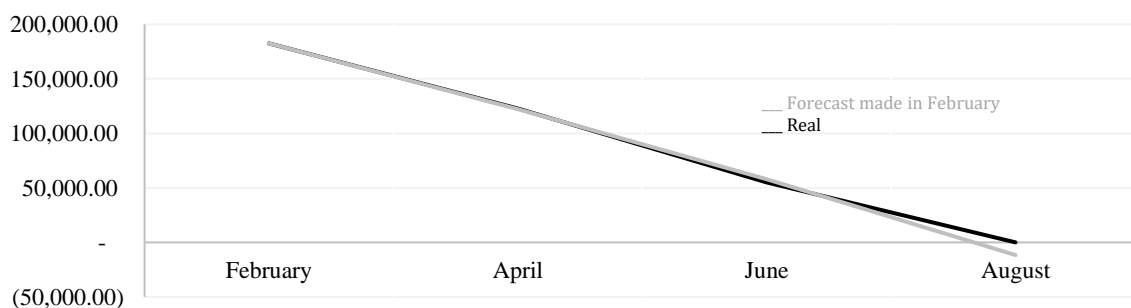

Fig. 4. Inventory level of Losartan Potassium in the year 2018 measured in number of pills

V. DISCUSSION

Although the literature on Markov chain models applied to prescribe drugs in healthcare has not been significantly addressed, several works on inventory control of products with similar irregular and intermittent demand behavior were found [e.g. 10, 13, 14]. Also, different Markov chain models combined with curve fitting techniques were used for predictions in various contexts [9, 10, 11, 15].

In this work, different assumptions were made and tested using four models. On the one hand, a simpler model in which Markov chain states model patients' transitions through the different consumption states. On the other hand, three other models based on different assumptions related to how the population increases or decreases over time.

In general, the simplest model performed the best when forecasting different drugs. This fact was surprising because the referred model considered a constant level of the patient's population, which might not always be true. After all, patients can enter and exit the studied healthcare system.

The inventory levels might have been forecasted using curve fitting techniques or other time series related approaches [9, 15]. Nevertheless, the method proposed in this work goes deeper in analyzing patients' transition through the different consumption states, orienteering the model towards a patient-oriented management model. Modeling the system like this aims to improve the managerial actions' definition based on each patient's state of consumption.

All in all, our method gave predictions with good out of sample accuracy and more information to understand how patients generate the demand for cardiovascular products. The latter aspect is necessary for a comprehensive understanding of how patients and the healthcare supply chain interact, allowing healthcare providers to make decisions towards improving the studied healthcare service.

VI. CONCLUSIONS

Medication delivery to patients is a dynamic system that changes over time since not only patient consumption varies over the years, but also new patients will eventually arrive at or leave the system. For this reason, the use of Markov chains with various expansion factors of the population in the system was proposed and used.

Further research can focus on the application of the models to a broader set of historical data. Because of data availability, this was not possible in the analyzed healthcare center, and we plan to test the application of these models in other hospitals. As well, optimization parameters such as training and testing horizons, intervals' sizes, and the number of states can be further explored. The data used for this research was transactional at the hospital pharmacy, free for patients.

However, real demand data (including unsatisfied), which was not available, would be more appropriate to test the proposed models. Moreover, recalling the fact that the population level evolves, it is recommended to explore how different assumptions on this expansion factor impact the accuracy and interpretability of models and predictions.

REFERENCES

[1]. Dirección Nacional de Contrataciones Públicas, 2019. *Buscador De Contratos*. [online] Contrataciones.gov.py. Available at: <https://www.contrataciones.gov.py/buscador/contratos.html?nro_nombre_licitacion=&categorias_licitacion%5B%5D=33&fecha_desde=&fecha_hasta=&tipo_fecha=&proveedor%5Bproveedor%5D=&estado_contrato=&codigo_contratacion=&catalogo%5Bcodigos_catalogo_n4%5D=&provee> [Accessed 27 August 2020].

[2]. ADN. (2018). Reconocen crisis por falta de medicamentos - ADN Paraguayo. Retrieved 27 August 2020, from http://www.adndigital.com.py/reconocen-crisis-falta-medicamentos/

[3]. RDN. (2020). Médicos anuncian huelga por desabastecimiento de medicamentos e insumos. Retrieved 27 August 2020, from https://www.rdn.com.py/2019/06/01/medicos-anuncian-huelga-por-desabastecimiento-de-medicamentos-e-insumos/.

[4]. Wan, L., Lou, W., Abner, E., & Kryscio, R. J. (2016). A comparison of time-homogeneous Markov chain and Markov process multi-state models. *Communications in Statistics: Case Studies, Data Analysis and Applications*, *2*(3-4), 92-100.

[5]. Ivanchuk, P., & Ivanchuk, M. (2017). One Example of Using Markov Chain Monte Carlo Method for Predicting in Medicine. *Cardiovascular Research*, *1*(4), 113-116.

[6]. McGhee, C. R., Glasser, J. H., Chan, W., Pomeroy, N., & Chan, W. (2003). Forecasting Health Care Expenditures and Utilization Based on a Markov Process and a Deterministic Cost Function in Managed Care Settings. *Lecture Notes-Monograph Series*, 229-238.

[7]. Kocer, U. U. (2013). Forecasting intermittent demand by Markov chain model. *International Journal of Innovative Computing, Information and Control*, *9*(8), 3307-3318.

[8]. Jenks, G. F., & Caspall, F. C. (1971). Error on choroplethic maps: definition, measurement, reduction. *Annals of the Association of American Geographers*, *61*(2), 217-244.

[9]. Papana A., Folinas D. and A. Fotiadis. (2012). Forecasting the consumption and the purchase of a drug, 2nd International Conference on Supply Chain, October, Greece.

[10]. Kocer, U. U. (2013). Forecasting intermittent demand by Markov chain model. *International Journal of Innovative Computing, Information and Control*, *9*(8), 3307-3318.

[11]. Hassan, M. R., & Nath, B. (2005, September). Stock market forecasting using hidden Markov model: a new approach. In *5th International Conference on Intelligent Systems Design and Applications (ISDA'05)* (pp. 192-196). IEEE.

[12]. Wallström, P., & Segerstedt, A. (2010). Evaluation of forecasting error measurements and techniques for intermittent demand. *International Journal of Production Economics*, *128*(2), 625-636.

[13]. E. A. Silver, Operations research in inventory management: A review and critique, *Operational Research, vol.29*, pp.628-645, 1981.

[14]. T. R. Willemain, C. N. Smart and H. F. Schwarz, A new approach to forecasting intermittent demand for service parts inventories, *International Journal of Forecasting, vol.20*, pp.375-387, 2004.

[15]. Wettermark, B., Persson, M. E., Wilking, N., Kalin, M., Korkmaz, S., Hjemdahl, P., ... & Regional Drug Expert Consortium in Stockholm County Council. (2010). Forecasting drug utilization and expenditure in a metropolitan health region. *BMC health services research*, *10*(1), 128.