# Case studies of clinical decision-making through prescriptive models based on machine learning

William Hoyos[a,b], Jose Aguilar[b,c,d], Mayra Raciny[a], Mauricio Toro[b]

[a]*Grupo de Investigaciones Microbiológicas y Biomédicas de Córdoba, Universidad de Córdoba, Montería, Colombia*
[b]*Grupo de Investigación en I+D+i en TIC, Universidad EAFIT, Medellín, Colombia*
[c]*Centro de Estudios en Microelectrónica y Sistemas Distribuidos, Universidad de Los Andes, Merida, Venezuela*
[d]*IMDEA Networks Institute, Madrid, Spain*

---

**Abstract**

*Background:* The development of computational methodologies to support clinical decision-making is of vital importance to reduce morbidity and mortality rates. Specifically, prescriptive analytic is a promising area to support decision-making in the monitoring, treatment and prevention of diseases. These aspects remain a challenge for medical professionals and health authorities.

*Materials and Methods:* In this study, we propose a methodology for the development of prescriptive models to support decision-making in clinical settings. The prescriptive model requires a predictive model to build the prescriptions. The predictive model is developed using fuzzy cognitive maps and the particle swarm optimization algorithm, while the prescriptive model is developed with an extension of fuzzy cognitive maps that combines them with genetic algorithms. We evaluated the proposed approach in three case studies related to monitoring (warfarin dose estimation), treatment (severe dengue) and prevention (geohelminthiasis) of diseases.

*Results:* The performance of the developed prescriptive models demonstrated the ability to estimate warfarin doses in coagulated patients, prescribe treatment for severe dengue and generate actions aimed at the prevention of geohelminthiasis. Additionally, the predictive models can predict coagulation indices, severe dengue mortality and soil-transmitted helminth infections.

*Conclusions:* The developed models performed well to prescribe actions aimed to monitor, treat and prevent diseases. This type of strategy allows supporting decision-making in clinical settings. However, validations in health institutions are required for their implementation.

*Keywords:* Prescriptive model, Clinical decision-making, Predictive model, Artificial intelligence

## 1. Introduction

Prescriptive analytic is an area of data analytic that is concerned with generating actions that lead to desired outcomes in modeled systems [1]. In healthcare, prescriptive modeling has established itself as a promising area for the improvement of healthcare systems. With the development and implementation of prescriptive modeling, it is expected to achieve greater speed and accuracy in the monitoring, treatment and prevention of disease, as well as an improvement in the quality of health care.

In this work, we are interested in developing methodologies to generate prescriptive models to support decision-making focused on treatment, follow-up and prevention of diseases. The development of methodologies for clinical decision-making has generated much interest in recent years. Machine learning (ML), computational intelligence and clinical decision analysis have been widely used for this purpose. However, there are some limitations or disadvantages associated with the use of such approaches. The complexity of the models for medical professionals to understand is a disadvantage, because they consider ML and computational intelligence models as a "black box"[2]. With respect to clinical decision models, specifically, decision trees do not take into account recurrent events and require individuals with similar characteristics. Markov models have been developed to overcome the problems presented by decision trees. However, Markov models ignore the interaction between individuals and consider few health states. Another important problem is the computational complexity; probability evaluations in Markov decision processes can increase with the complexity of the problem or system to be modeled [3, 4]. Finally, another limitation is that clinical decision analysis requires more data than other stochastic modeling techniques due to variations in transition probabilities at each decision stage [5]. Based on these problems, it is necessary to develop methodologies that generate prescriptive models that are explainable to medical professionals, that are computationally efficient regardless of the complexity of the problem, and that have a minimally acceptable performance with small datasets.

In this study, we propose an approach to generate prescriptive models to support decision-making in clinical settings. Our approach is capable of generating prescriptive models that sug-

gest prescribing actions for treatment, follow-up and prevention of diseases. The combination of fuzzy cognitive maps (FCMs) –explainable method– and genetic algorithms (GAs) allowed the development of a methodology for the generation of prescriptive models. The ease of construction and interpretation of FCMs brings an added value different from the models reported in the literature. Our approach starts with FCM creation and subsequent characterization of the FCM using the nature of the concepts. Each concept is discriminated in two layers: *system* and *action*. In the first case, they are all those variables measurable in patients such as demographic variables, signs, symptoms and laboratory tests. While the action variables are all those related to actions aimed at the treatment, follow-up and prevention of diseases. The second stage of our approach consists of the initial instantiation of the system, where the medical user sets the desired state for the system variables. Finally, an optimization algorithm (GAs) is used to find the optimal action values that through the FCM inference system leads to the desired state of the variables related to the system. The proposed approach is tested on three case studies with specific datasets that were collected in previous research. The constructed models are used to make specific prescriptions for each patient according to their sociodemographic, clinical, genetic and laboratory characteristics. The results obtained in this research demonstrate the ability of the prescriptive models designed to generate prescriptions with high accuracy and low error.

The remainder of this paper is organized as follows: Section 2 shows a literature review about the last trends in prescriptive modeling in medical settings. Section 3 describes the methodology used to generate the prescriptive models. The next section presents three case studies with the datasets for each case study, and the configuration of experiments. Section 5 shows the results based on case studies. Section 6 discusses the results and shows a comparison with previous works. Finally, Section 7 concludes the paper.

## 2. Related work

Prescriptive analytic is responsible for the generation of prescriptive models that support decision-making [1]. In this context, the prescription is a set of actions that the decision-maker executes to achieve a given outcome [6]. Prescriptive models can be categorized into three main areas: i) prescriptive modeling using ML, ii) prescriptive modeling using computational intelligence, and

iii) prescriptive modeling using clinical decision analytics. Below, we show some studies related to each of these categories.

### *2.1. Prescriptive modeling using ML*

Prescriptive analytic has attracted much interest due to its potential application in medical environments. The use of ML has been widely extended for the development of prescriptive models to support decision-making in clinical or medical settings [7–10]. For example, Bertsimas et al [7] proposed and implemented two ML methods (prescriptive optimal tree and prescriptive support vector machines) to generate prescriptive models that generate recommendations to reduce the risk of readmission after surgery. The authors used red blood cell transfusion as an actionable feature. The models developed by Bertsitmas et al. have the ability to reduce the risk of readmission by 12% and the results are interpretable because the models allow the identification of variables that influence the prescription made. Harikumar et al. [8] developed a prescriptive analytic solution that uses ML approaches to recommend actions in diabetes, heart attack, and stroke. The goal was to find the smallest change within the actionable characteristics to achieve the change from an undesirable to a desirable class. The capability of the developed models was tested on Center for Disease Control and Prevention (CDC) datasets using logistic regression, k-nearest-neighbor (KNN) and random forest (RF). The most favorable results were for KNN on the stroke dataset (88% accuracy), and for the other datasets the results are very similar. Hosseini et al [9] proposed an algorithm to optimize decision variables with respect to a variable of interest. The developed algorithm used Bayesian networks to reduce diabetes mortality rates, by prescribing the optimal combination of drugs for disease control. The algorithm was tested on a dataset of patients with diabetes and had the particularity of generating interpretable prescriptive models because the variables influencing the prescription could be identified. The models generated by Hosseini et al, obtained an accuracy of 88.75% and an area under the curve of 71.15%.

### *2.2. Prescriptive modeling using computational intelligence*

Computational intelligence is a subarea of artificial intelligence where fuzzy logic, artificial neural networks and evolutionary algorithms are combined. Such approaches have been used for

the development of prescriptive models in clinical settings [11, 12]. For example, Hoyos et al. [11] implemented an autonomous cycle of data analysis tasks where they combined artificial neural networks and GAs to optimize decision-making in the clinical management of dengue. Dengue is a disease that has no cure and its treatment is based on alleviating symptoms and avoiding complications. The models created had the ability to classify dengue and follow the recommendations given by the WHO for the treatment of each type of dengue. Chalmers et al. [12] proposed a prescription approach to optimize the treatment of adolescent idiopathic scoliosis. The goal was to identify optimal orthotic corrections that would reduce disease progression using fuzzy logic. The developed model had the ability to recommend actions that adjust the orthosis and reduce disease progression by 26%.

### *2.3. Prescriptive modeling using clinical decision analysis*

Clinical decision analysis is a quantitative approach widely used to optimize decision-making in healthcare settings [13]. This approach has been extensively implemented to establish or determine the optimal expected utility of treatments or interventions as healthcare strategies to reduce costs, morbidity, or mortality rates [14, 15]. The main techniques within decision analysis comprise decision trees, Markov decision processes and partially observable Markov decision processes.

Clinical decision trees allow the optimization of strategies aimed at screening and treatment of diseases. This approach has been used to quantify the utility of treatments or strategies based on transition probabilities. For example, Kurisu et al. [14] developed a clinical decision analysis with decision trees to quantify the utility of various antipsychotic treatment options (risperidone, haloperidol, olanzapine, amisulpride, ziprasidone and quetiapine) in patients with delirium. Sensitivity analysis showed that quetiapine is the best antipsychotic treatment option for patients with delirium. Keikes et al [15] implemented decision trees to convert colorectal cancer diagnosis and treatment recommendation guidelines into a computational tool for clinical decision support. The decision trees developed and implemented generated recommendations for the diagnosis, follow-up and treatment of colorectal cancer with a concordance of 81% when compared to recommendations suggested by an interdisciplinary team of experts.

Markov chains are a stochastic approach that allows sequential processes to be modeled [16]. Due to the complexity present in clinical decision-making, Markov models are a useful tool to compare the effectiveness and utility of available treatment combinations, optimize screening policies, and prevent disease-related complications [17–21]. For example, Habu [17] conducted a clinical decision analysis using Markov modeling to evaluate the efficacy of two treatment strategies (proton pump inhibitor vs. potassium-competitive acid blocker) for gastroesophageal reflux. The results of the analysis yielded a superiority of the competitive acid blocker with respect to cost-effectiveness and the number of days required to treat the disease. These findings were confirmed by the sensitivity analysis implemented in the study. Similarly, Shen et al [18] compared the efficacy of various combinations of interventions for stroke patients in the convalescent stage. The main strategies used for modeling were rehabilitation therapy, use of traditional Chinese medicine, and acupuncture treatment. The Markov decision model had the ability to recommend the best possible combination of treatments for stroke patients in different stages of recovery. Eghbali-Zarch et al [19] modeled the drug treatment of type 2 diabetes to determine the optimal treatment policy to decrease adverse medication reactions that increase the economic burden of the disease and decrease quality-adjusted life years. The Markov model could recommend treatment options that involve a minimum amount of medication with acceptable expected quality of life.

Dumlu et al [20] proposed a partially observable Markov decision model to establish the optimal screening policy in the preclinical stages of Alzheimer's disease. The model aims to maximize the quality-adjusted life years and recommends the time when the patient should be screened. The results of the cost-effectiveness analysis show that implementing the optimal policies recommended by the model reduced costs. Prayogo et al [21] formulated models based on partially observable Markov sequential processes for the evaluation of screening policies for early diagnosis of lung cancer. Early detection of this type of disease through screening is crucial to decrease mortality rates. The research results demonstrated the ability of the proposed model to recommend an optimal screening policy that guarantees higher quality-adjusted life years.

## 3. Methodology

In this section, we present the methodology to generate prescriptive models. First, we briefly explain the approach used to generate the prescriptive model, which includes the construction of a predictive model. Then, we present three case studies with their datasets and their preprocessing prior to model creation. Fig. 1 shows a schematic representation of the general methodology to achieve the objective of this study. According to the methodology, the first step is data preparation and analysis (cleaning, normalization and balancing). Next, a classical FCM is built to predict using particle swarm optimization (PSO), which is then used by our prescriptive-FCM to assess the actions it could prescribe, in such a way as to find the most appropriate ones.

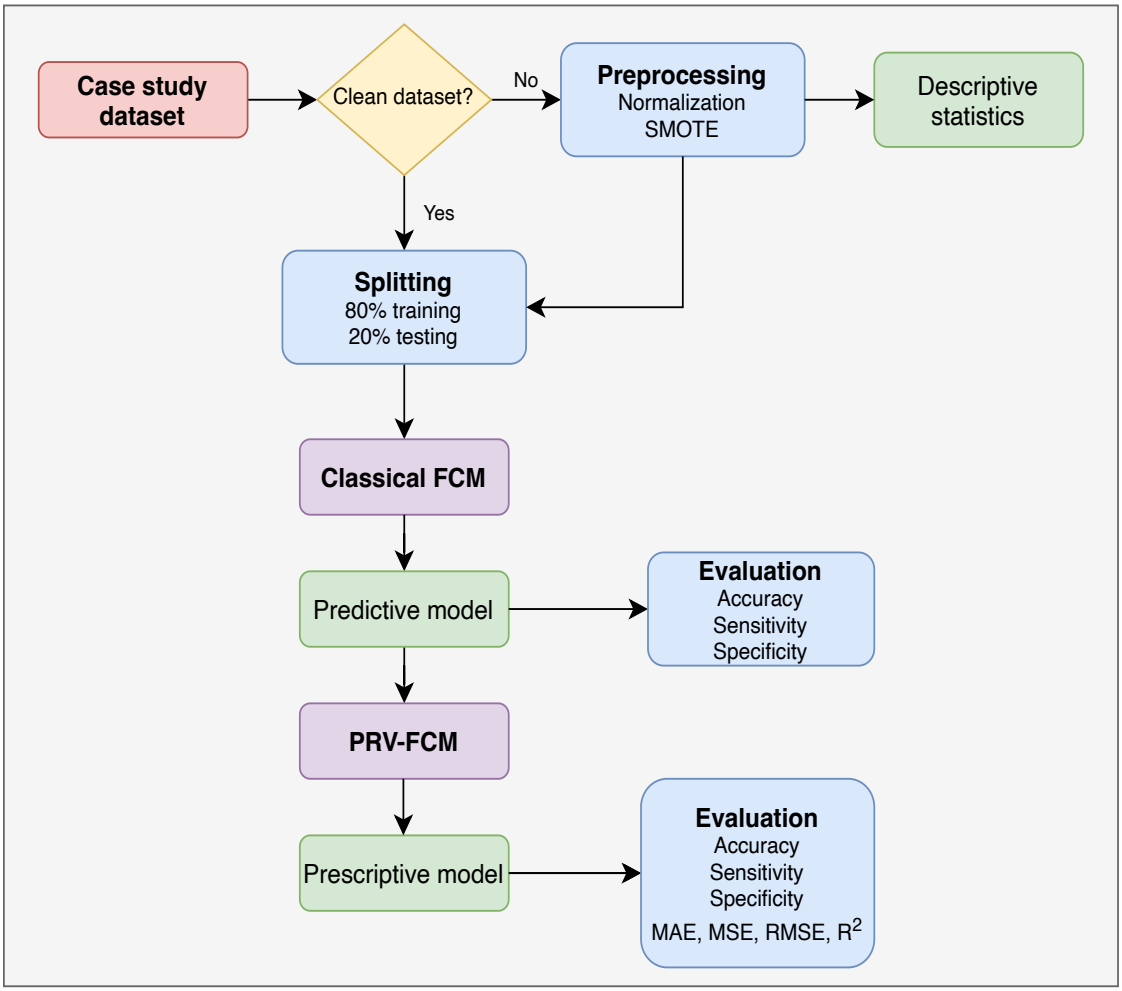


**Fig. 1.** General methodology used in this study.

### 3.1. Descriptive analysis

The descriptive analysis consists of examining data to interpret past behavior and learn about data distribution, such that we can describe things like, for example, that the classes of a label are unbalanced, and if there are variables with a lot of noise. In this case, we use descriptive statistics to extract information from the datasets in each case study. We used measures of central tendency and dispersion to understand the behavior of quantitative data. For qualitative data, we used frequency distribution.

### 3.2. Generation of the predictive models

The predictive models were generated using a data-driven PSO-FCM approach. The predictive model is used by the prescriptive-FCM to propose several sets of actions (each one is a different prescription), and requires a model/function that determines the quality of the proposed prescriptions. The predictive model was used for these tasks.

#### 3.2.1. Data-driven PSO-FCM

FCM is a technique of computational intelligence that allows modeling systems using concepts and relationships. The concepts correspond to the variables of the system to be modeled and the relationships correspond to the influence that exists between them [22–26]. FCMs are composed of a 5-element tuple ($\Psi$) where $n$ is the number of concepts or variables to be modeled, $v$ is an initial or activation vector, $W$ is the weight matrix, and $f()$ is an activation function to keep the concept values in a desired range $r$. Eq. 1 shows the main elements of an FCM. The most commonly used activation functions for FCMs are shown in Table 1.

$$\Psi = \langle n, v, W, f() \rangle \tag{1}$$

FCMs can be built by experts using their knowledge and experience. They can also be built with algorithms that extract the relationships from historical data. The relationships are stored in square matrices to be used in the inference process. Eq. 2 shows an example of an extracted matrix and Fig. 2 shows the FCM constructed with this matrix. In this study, FCMs were constructed using the PSO algorithm due to its superior performance when extracting relationships from the

**Table 1**

Most commonly used activation functions in FCMs.

| Activation function | Equation | Range |
|---|---|---|
| Sigmoid | $f(x) = \frac{1}{1+e^{-\lambda \times x}}$ | $f(x) \in [0, 1]$ |
| Hyperbolic tangent | $f(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$ | $f(x) \in [-1, 1]$ |

data [27–29]. In addition, the lack of experts in each domain limited the creation of FCMs using expert knowledge and experience.

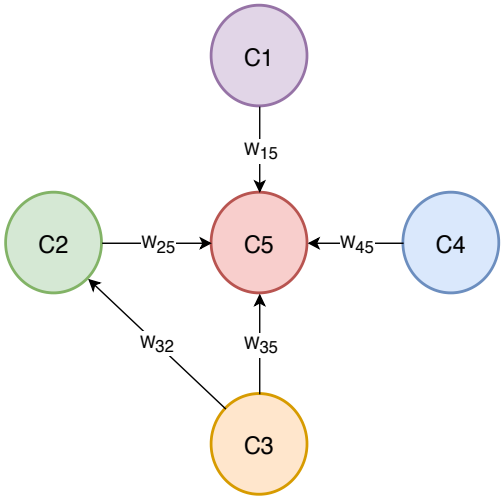


**Fig. 2.** Example of FCM with five concepts and five relationships.

$$\mathbf{W} = \begin{array}{c} \\ \boldsymbol{C_1} \\ \boldsymbol{C_2} \\ \boldsymbol{C_3} \\ \boldsymbol{C_4} \\ \boldsymbol{C_5} \end{array} \begin{array}{c} \begin{array}{ccccc} \boldsymbol{C_1} & \boldsymbol{C_2} & \boldsymbol{C_3} & \boldsymbol{C_4} & \boldsymbol{C_5} \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 & w_{15} \\ 0 & 0 & 0 & 0 & w_{25} \\ 0 & w_{32} & 0 & 0 & w_{35} \\ 0 & 0 & 0 & 0 & w_{45} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \tag{2}$$

PSO is an optimization technique that simulates the behavior of particles in nature [30]. This technique can be used for the construction of FCMs and optimization of their weight matrices

**Table 2**

Inference functions used for inference in FCMs.

| Inference function | Equation | Main characteristics |
|---|---|---|
| *Kosko* [22] | $v_j(t+1) = f\left(\sum_{i=1,i\neg j}^{n} W_{ij}v_i(t)\right)$ | The FCM has no memory capacity because it does not take into account the previous iteration ($v_j(t)$) during inference. |
| *Modified Kosko* [31] | $v_j(t+1) = f\left(\sum_{i=1,i\neg j}^{n} v_j(t) + W_{ij}v_i(t)\right)$ | The FCM has memory capacity because it takes into account the previous iteration ($v_j(t)$) during inference. |
| *Rescaled* [32] | $v_j(t+1) = f\left(\sum_{i=1,i\neg j}^{n} (2\times v_j(t)-1) + W_{ij}(2\times v_i(t)-1)\right)$ | It disables null initial values ($v_j = 0$) that are activated when passed by the activation function. |

(PSO-FCM) [29]. In this way, an optimized FCM is obtained that can be used to predict a response variable. In this case, each FCM is a particle $i$ and the weight matrix ($W_i$) is its position. The algorithm first updates the particle velocity and then its position. Eq. 3 and Eq. 4 show the optimization process with PSO.

$$v_i(t+1) = v_i(t) + r_1 \cdot (W_i^{best} - W_i(t)) + r_2 \cdot (W_i^{gbest} - W_i(t)) \tag{3}$$

$$W_i(t+1) = W_i(t) + v_i(t) \tag{4}$$

where $v_i$ is the particle velocity, $r_1$ and $r_2$ are random values with uniform distribution; $W_i^{best}$ is the best position obtained by a specific particle, while $W_i^{gbest}$ is the best position obtained by any particle in the swarm.

After the construction of the FCM and the optimization of its weight matrix, the FCM was ready to make predictions using inference rules or functions. To date, several inference functions have been reported in the literature, which are used depending on the problem to be solved. Table 2 shows the most commonly used inference functions reported in the literature.

### *3.3. Generation of the prescriptive models*

To generate prescriptive models, we developed a methodology, called prescriptive-FCM. This methodology is an extension of FCMs for prescriptive modeling. In the following, we will explain the proposed approach. Prescriptive-FCM is a prescriptive modeling approach that uses FCMs

and GA to generate prescriptions or optimal actions that achieve a desired outcome in the modeled system. Before explaining our approach, we will explain the elements that compose Prescriptive-FCM. FCMs were briefly explained in the previous subsection, and a brief explanation of GAs follows.

### *3.3.1. GA*

A GA is an optimization technique inspired by the general theory of biological evolution. This technique reflects natural selection where the fittest individuals are selected to reproduce and generate new offspring [33]. Fig. 3 shows the methodological framework for a GA. The first steps in the development of GAs are problem definition and fitness functions. GAs start with a random initial population, whose fitness is calculated using functions that depend on the proposed objective (minimization or maximization). Subsequently, this initial population is subjected to selection, crossover and mutation processes. These procedures are carried out to vary the composition of each of the individuals of the initial population. The individuals with the best fitness are selected and the process is repeated until a certain stop condition is reached.

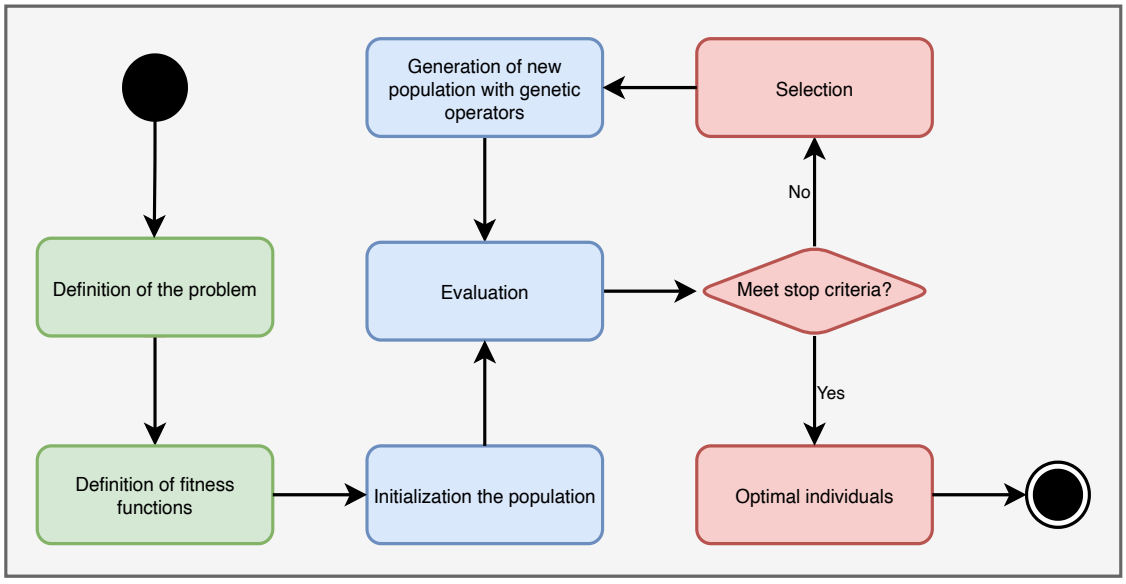


**Fig. 3.** Methodological framework for a GA.

### 3.3.2. *Prescriptive-FCM*

In this study, we propose a methodology called Prescriptive-FCM to generate prescriptive models. Prescriptive-FCM uses three stages for the generation of prescriptive models (see its architecture in Fig. 4). The first stage consists of the characterization of the concepts of the problem to be solved. With these concepts is built the FCM with two layers, according to the nature of the concepts. Thus, these two layers constitute the system concepts and the action concepts. The former is related to the system to be modeled. For example, in a disease, the concepts related to the system could be the symptoms present in the patients. The action concepts, also called prescriptive concepts, are actions that, when executed, modify the system concepts. For example, in a medical problem, an analgesic could be an action concept. Changes in this variable will generate changes in the system variables, in this case, the patient's symptoms. Particularly, the first layer is defined by the previously built predictive model.

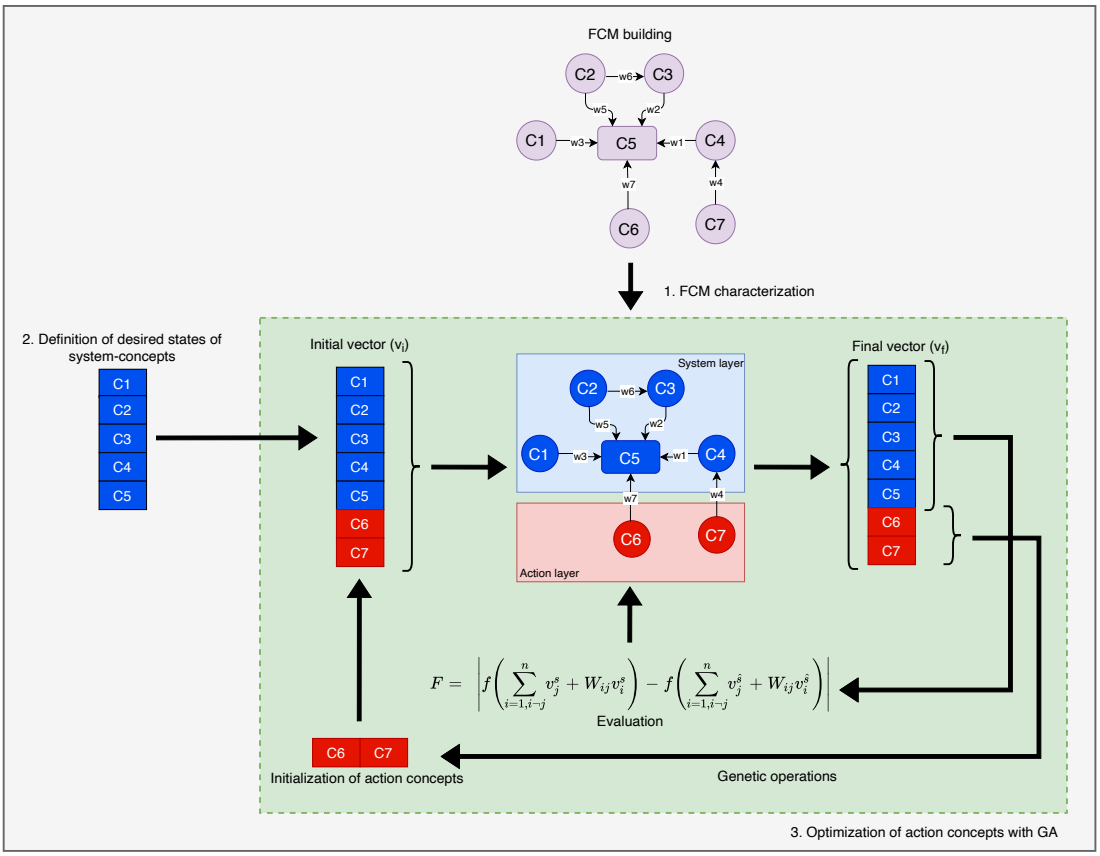


**Fig. 4.** Architecture of Prescriptive-FCM approach.

The second stage of Prescriptive-FCM consists of the definition of the desired state. In this stage, the decision maker defines the desired values of the system concepts. For example, if the physician wants to lower the fever, then she/he will set this concept to a value of 0 because the goal is to minimize the fever as much as possible. The final stage consists of the optimization of the action concepts such that using the inference process of the FCMs leads to the desired system concepts. For this last stage, a GA is used that selects, crosses and mutates the values of the action concepts. The FCM inference process generates a vector corresponding to system concepts and action concepts. The former is used for evaluation with a fitness function (see Eq. 5), while the latter are the prescribed variables. The latter is the ones generated by our proposed methodology.

$$F = \left| f\left( \sum_{i=1, i\neg j}^{n} v_j^s + W_{ij} v_i^s \right) - f\left( \sum_{i=1, i\neg j}^{n} v_j^{\hat{s}} + W_{ij} v_i^{\hat{s}} \right) \right| \tag{5}$$

where $v^s$ is the vector representing the value of the desired concepts, $v^{\hat{s}}$ is the vector representing the values generated by Prescriptive-FCM, $W_{ij}$ is the weight matrix of the characterized FCM. Finally, $f$ is a function that holds the values in the desired range.

## 4. Experiments

### *4.1. Data preparation*

For the validation of our approach, we used three case studies related to the monitoring, treatment and prevention of diseases in public health. These datasets were chosen from public repositories since they contained variables (action concepts) that could be used in medical prescription/recommendation tasks. When analyzing these datasets, we realized that they do not contain information on concomitant diseases or the different stages of disease development. However, we did not find any dataset that could be used in prescription tasks (with action variables), and that, in addition, would incorporate these other variables or all the stages of the development of a disease.

Specifically, they correspond to the estimation of Warfarin dose in anticoagulated patients, treatment of severe dengue (SD) and prevention of soil-transmitted helminth infections. Each case study contained a dataset, which was preprocessed using data cleaning technique. First, rows with missing data were removed to decrease bias. The normalization process of the variables

was performed to scale the variables within the same range and thus improve the speed of model training. In electronic health records, it is very common to find class imbalance in the objective variables. For this reason, we used synthetic minority oversampling technique (SMOTE) to balance the classes before feeding the predictive and prescriptive algorithms. The characteristics of the variables in each of the datasets are described in each case study. For the internal validation process of the models, each dataset was divided into 80% for training and validation and 20% for testing. We used 10-replicate cross-validation to find the optimal hyperparameters of the best model.

### *4.2. Configuration of hyperparameters*

In the development of ML models, it is common to use a combination of hyperparameters, and thus find the optimal values that represent the best model to be used in the test set. We used a 10-fold cross-validation technique to find the best hyperparameters in each model. For this study, we used different hyperparameter values from similar studies reported in the literature depending on the nature of the data in each case study.

For the optimization of FCM matrices with PSO, we use a grid of random values for initial population and iteration steps. For the first case, we use values between 10 and 200, for the second hyperparameter, values between 10 and 800. The inference process of FCMs involves activation functions and their slope, and inference functions. We established a combination of these hyperparameters to find the best model. We implemented the activation functions and inference algorithms described in Table 1 and Table 2, respectively. Finally, the slope of the activation functions was established with a grid of random values between 0.1 and 1000.

The search method used in Prescriptive-FCM was a GA. For this case, we used different combinations of initial population size, crossover and mutation probabilities. The hyperparameter grid for the initial population contained random values between 10 and 400 individuals. For the probabilities, we used a grid of random values between 0 and 1.

### *4.3. Evaluation metrics*

We evaluated the quality of the developed models using several metrics. We use accuracy, sensibility and specificity to measure the quality of classification-type predictive models. We

also use classification metrics to assess the quality of prescriptive models when the prescriptive variables are qualitative in nature. When the prescriptive variables are quantitative in nature, we use mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE) and $R^2$ metrics. The following is a brief description of each of the metrics used to evaluate the performance of the models developed.

- *Accuracy*: percentage of correctly classified examples among the total number of classified examples. Greater accuracy means a greater performance of the model.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{6}$$

  where $TP$ are the true positives, $TN$ are true negatives, $FN$ are false negatives, and $TN$ are true negatives.

- *Sensitivity*: measures the ability of the classifier to predict positive cases to those actually positive.

$$Sensitivity : \frac{TP}{TP + FN} \tag{7}$$

- *Specificity*: measures the ability of the classifier to predict negative cases to those actually negative.

$$Specificity : \frac{TN}{TN + FP} \tag{8}$$

- *MAE*: calculated as an average of absolute differences between the correct prescriptive concepts values and prescriptions.

$$MAE = \frac{1}{m} \sum_{i=1}^{m} \left| v_i^a - \hat{v}_i^a \right| \tag{9}$$

  where $m$ is the number of records in testing set, $v_i^a$ is the actual prescriptive value and $\hat{v}_i^a$ is the prescribed value.

- *MSE*: measures the average square error of our prescriptions. For each point, it calculates the square difference between the prescriptions and the prescriptive concepts, and then, averages those values.

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(v_i^a - \hat{v}_i^a)^2 \tag{10}$$

- *RMSE*: is the squared root of the error described above.

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(v_i^a - \hat{v}_i^a)^2} = \sqrt{MSE} \tag{11}$$

- $R^2$: Coefficient of determination.

$$R^2 = \frac{\sum_{i=1}^{m}(\hat{v}_i^a - \bar{v}_i^a)^2}{\sum_{i=1}^{m}(v_i^a - \bar{v}_i^a)^2} \tag{12}$$

where $\bar{v}_i^a$ is the mean of actual prescriptive values.

### *4.4. Case study 1: warfarin dose estimation*

Warfarin is the most frequently used anticoagulant worldwide to prevent thromboembolism and thrombosis. Establishing the dose of Warfarin is important because a higher dose than necessary may increase the risk of bleeding and a lower dose may decrease protection against thrombotic processes [34]. For coagulation monitoring, physicians use a laboratory test known as the international normalized ratio (INR). The INR value in normal patients is usually 1; however, in patients on anticoagulant therapy, INR levels may be between 2 and 3, a range that generally indicates appropriate anticoagulation for most cases. For patients with values above 3, they present a high risk of bleeding or hemorrhage, while values below 2 represent a risk of thrombosis or thromboembolism [35]. To test our proposed approach, we used a dataset published by *The International Warfarin Pharmacogenetics Consortium (2009)* [36]. Table 3 and Table 4 show the variables used in this dataset. For this case, we used sociodemographic variables such as age and race; anthropometric variables such as height and weight; and the next genetic variables: cytochrome P450, family 2, subfamily C, polypeptide 9 (CYP2C9), and vitamin K epoxide reductase complex, subunit 1 (VKORC1). Additionally, we used INR as a target variable and Warfarin dose as an action

variable. The INR variable was categorized due to the importance of establishing Warfarin doses that maintain INR values between 2 and 3. For this reason, INR was established as *controlled INR* (between 2 and 3) and *altered INR* (lower than 2 or higher than 3). After the data preprocessing described in subsection 4.1, the dataset had 3385 records corresponding to 2085 patients with *controlled INR* and 1800 patients with *altered INR*.

**Table 3**

Descriptive statistics of numerical variables of case study 1.

| Concept | Concept type | Variable name | Median (Interquartile range) |
|---|---|---|---|
| C1 | System | Age (years) | 65.0[55.0-75.0] |
| C2 | System | Height (m) | 1.70[1.61-1.78] |
| C3 | System | Weight (Kg) | 78[65.30-92.30] |
| C8 | Prescriptive | Warfarin | 31.25[22.50-42.0] |

### *4.5. Case study 2: Treatment of SD*

Dengue is a disease caused by a virus and transmitted by the bite of a mosquito of the genus *Aedes spp*. The most severe phase of the disease is known as severe dengue, and represents the main cause of death from dengue [37]. Studies have reported a mortality rate of over 20% when treatment is inadequate or delayed [38]. Currently, dengue has no definitive cure and its treatment is based on the relief of signs and symptoms. In addition, treatment is aimed at considerably reducing the complications that the virus causes during its stay in the patient's body [39]. Establishing the optimal treatment policy for severe dengue is important to avoid complications and reduce mortality rates associated with the disease. To test the proposed methodology, we used a dataset of mortality data from patients with dengue. The data correspond to 398 patients from Córdoba, Colombia. The variables used for the generation of the models are shown in Table 5. In this case, we used 4 variables related to severe dengue such as, extravasation, shock, hemorrhage and organ failure. While 4 treatment related variables were used to find the optimal values to minimize mortality. In this dataset, all variables used had values of 0 for absence and 1 for presence. For the target variable, surviving patients were coded to 0 while deceased patients were coded to 1.

**Table 4**

Descriptive statistics of categorical variables of case study 1.

| Concept | Concept type | Variable name | Category | N | Percentage (%) | CI 95% |
|---|---|---|---|---|---|---|
| C4 | System | Race | | | | |
| | | | White | 1207 | 49.37 | 47.39-51.35 |
| | | | Asian | 424 | 17.34 | 15.84-18.84 |
| | | | Black | 328 | 13.42 | 12.07-14.77 |
| | | | Other | 486 | 19.87 | 18.29-21.45 |
| C5 | System | Amiodarone | | | | |
| | | | No | 2286 | 93.50 | 92.52-94.48 |
| | | | Yes | 159 | 6.50 | 5.52-7.48 |
| C6 | System | Vkorc1 | | | | |
| | | | A/A | 587 | 24.01 | 22.32-25.70 |
| | | | A/G | 937 | 38.32 | 36.39-40.25 |
| | | | G/G | 921 | 37.67 | 35.75-39.59 |
| C7 | System | Cyp2c9 | | | | |
| | | | *1/*1 | 1780 | 72.80 | 71.04-74.56 |
| | | | *1/*2 | 379 | 15.50 | 14.07-16.93 |
| | | | *1/*3 | 215 | 8.79 | 7.67-9.91 |
| | | | Other | 71 | 2.90 | 2.23-3.57 |
| C9 | Target | INR | | | | |
| | | | Controlled INR | 2085 | 53.70 | 52.12-55.26 |
| | | | Altered INR | 1800 | 46.30 | 44.73-47.86 |

After preprocessing of the data, defined in subsection 4.1, there were 210 surviving patients and 188 deceased patients.

### *4.6. Case study 3: Prevention of geohelminthiasis*

Soil-transmitted helminth infection or geohelminthiasis is a disease characterized by the ingestion of embryonated eggs of parasites or by penetration through the skin of their infective larvae present in humid and warm soils [40]. These infections are facilitated by poverty, illiteracy, lack of drinking water and hygienic habits [41]. Prevention of this type of infection is important due to the high morbidity that impacts human health leading to stunting, vitamin deficiencies and poor cognitive function [42]. It is necessary to establish prevention strategies to reduce the morbidity rates associated with these types of infections. Based on these issues, we tested our prescriptive

**Table 5**

Descriptive statistics of variables in case study 2.

| Concept | Concept type | Variable name | Category | N | Percentage (%) | CI 95% |
|---|---|---|---|---|---|---|
| C1 | System | Extravasation | 0 | 277 | 69.60 | 65.08-74.12 |
| | | | 1 | 121 | 30.40 | 25.88-34.92 |
| C2 | System | Shock | 0 | 276 | 69.35 | 64.82-73.88 |
| | | | 1 | 122 | 30.65 | 26.12-35.18 |
| C3 | System | Bleeding | 0 | 161 | 40.45 | 35.63-45.27 |
| | | | 1 | 237 | 59.55 | 54.73-64.37 |
| C4 | System | Organ failure | 0 | 268 | 67.34 | 62.73-71.95 |
| | | | 1 | 130 | 32.66 | 28.05-37.27 |
| C5 | Prescriptive | Transfusion | 0 | 276 | 69.35 | 64.82-73.88 |
| | | | 1 | 122 | 30.65 | 26.12-35.18 |
| C6 | Prescriptive | Cristalloid solutions | 0 | 277 | 69.60 | 65.08-74.12 |
| | | | 1 | 121 | 30.40 | 25.88-34.92 |
| C7 | Prescriptive | Colloid solutions | 0 | 161 | 40.45 | 35.63-45.27 |
| | | | 1 | 237 | 59.55 | 54.73-64.37 |
| C8 | Prescriptive | ICU | 0 | 107 | 26.88 | 22.52-31.24 |
| | | | 1 | 291 | 73.12 | 68.76-77.48 |
| C9 | Target | mortality | Survivor | 210 | 52.76 | 51.19-54.33 |
| | | | Dead | 188 | 47.24 | 45.67-48.81 |

approach to generate a model with optimal recommendations that will lead to disease prevention and thus minimize the occurrence of parasite infections. The dataset used to test the prescriptive approach corresponded to demographic and epidemiological data of 130 school-aged children in a rural area of the department of Córdoba, Colombia. The variables used for model generation are shown in Table 6. Seven variables are classified as variables directly related to the disease, while two variables related to prevention were considered action variables. The target variable indicated the clinical condition of the children with respect to the presence or absence of geohelminths. After preprocessing of the data, the cleaned and sorted dataset contains 64 healthy or uninfected children and 66 infected children.

**Table 6**

Descriptive statistics of variables in case study 3.

| Concept | Concept type | Variable name | Category | N | Percentage (%) | CI 95% |
|---|---|---|---|---|---|---|
| C1 | System | Sex | | | | |
| | | | F | 397 | 52.1 | 48.55-55.65 |
| | | | M | 365 | 47.9 | 44.35-51.45 |
| C2 | System | Weight | | | | |
| | | | <20 | 71 | 9.32 | 7.26-11.38 |
| | | | 20-40 | 552 | 72.44 | 69.27-75.61 |
| | | | 40-60 | 137 | 17.98 | 15.25-20.71 |
| | | | >60 | 2 | 0.26 | -0.1-0.62 |
| C3 | System | Indigenous | | | | |
| | | | No | 576 | 75.59 | 72.54-78.64 |
| | | | Yes | 186 | 24.41 | 21.36-27.46 |
| C4 | System | Source of drinking water | | | | |
| | | | 1 | 22 | 2.89 | 1.7-4.08 |
| | | | 2 | 4 | 0.52 | 0.01-1.03 |
| | | | 4 | 185 | 24.28 | 21.24-27.32 |
| | | | 5 | 514 | 67.45 | 64.12-70.78 |
| | | | 6 | 37 | 4.86 | 3.33-6.39 |
| C5 | System | Floor of the house | | | | |
| | | | 1 | 675 | 88.58 | 86.32-90.84 |
| | | | 2 | 26 | 3.41 | 2.12-4.7 |
| | | | 3 | 60 | 7.87 | 5.96-9.78 |
| | | | 5 | 1 | 0.13 | -0.13-0.39 |
| C6 | System | Disposal of human excreta | | | | |
| | | | 1 | 280 | 36.75 | 33.33-40.17 |
| | | | 2 | 187 | 24.54 | 21.48-27.6 |
| | | | 3 | 295 | 38.71 | 35.25-42.17 |
| C7 | Prescriptive | Child wears closed shoes | | | | |
| | | | 1 | 203 | 26.64 | 23.5-29.78 |
| | | | 2 | 240 | 31.5 | 28.2-34.8 |
| | | | 3 | 319 | 41.86 | 38.36-45.36 |
| C8 | System | Child washes his hands after defecating | | | | |
| | | | 1 | 234 | 30.71 | 27.43-33.99 |
| | | | 2 | 209 | 27.43 | 24.26-30.6 |
| | | | 3 | 319 | 41.86 | 38.36-45.36 |
| C9 | Prescriptive | Child washes his hands before eating | | | | |
| | | | 1 | 317 | 41.6 | 38.1-45.1 |
| | | | 2 | 191 | 25.07 | 21.99-28.15 |
| | | | 3 | 254 | 33.33 | 29.98-36.68 |
| C10 | Target | Geohelminthiasis | | | | |
| | | | Negative | 429 | 56.29 | 54.73-57.85 |
| | | | Positive | 333 | 43.71 | 42.15-45.27 |

## 5. Results

In this section, we show the results of the models generated. Each subsection describes the results of the descriptive statistics, prescriptive model (and its underlying predictive model) for each case study.

### *5.1. Case study 1: warfarin dose estimation*

#### *5.1.1. Descriptive statistics*

Descriptive statistics for this case study are summarized in Table 3 and Table 4. For the statistical description of the data, measures of central tendency such as median with interquartile ranges were used for variables C1, C2 and C3, which had median with interquartile ranges of 65.0[55.0-75.0], 1.70[1.61-1.78] and 78[65.30-92.30], respectively. For categorical variables, the relative frequency with 95% confidence intervals (95% CI) was used. In this study, the majority of individuals were white, with a relative frequency of 49.37% (95% CI = 39-51.35), and 93% (95% CI = 92.52-94.48) of patients reported not taking the antiarrhythmic agent amiodarone. The variables related to the genotypic conditions of the patients, such as C6 - Vkorc1 with category A/G was the most frequent with 38.32% (95% CI = 36.39-40.25) and C7 - Cyp2C9 in category *1/*1 showed higher relative frequency than the other categories 72.8% (95% CI = 71.04-74.56).

#### *5.1.2. Predictive model*

We developed a predictive model using INR as the target variable. This model based on FCM has the ability to predict INR, and is built by adjusting the weights of the FCM using PSO (initial population = 80 individuals, iterations = 120). This FCM is used by Prescriptive-FCM to evaluate the quality of a prescription.

Table 7 shows the performance of the developed predictive models and the optimal hyperparameters of the best model for each case study. Regarding the case study of the warfarin dose estimation, the performance of the model developed with the classical FCM approach obtained values of 0.65, 0.51 and 0.77 for accuracy, sensitivity and specificity, respectively.

**Table 7**

Performance and optimal hyperparameters of the predictive models developed in this work for all case studies.

| Case study | Optimal hyperparameters | Accuracy | Sensitivity | Specificity |
|---|---|---|---|---|
| Warfarin dose | Activation function = sigmoid | 0.65 | 0.51 | 0.77 |
| Treatment of SD | Slope = 10 | 0.74 | 0.79 | 0.68 |
| Prevention of geohelmintiasis | Inference function = Modified Kosko | 0.74 | 0.76 | 0.73 |

### *5.1.3. Prescriptive model*

We developed a prescriptive model that formulated the dose of warfarin for anti-coagulated patients. The GA using Prescriptive-FCM optimized the action concept, which in this case is the warfarin dose. Because warfarin dose was a numerical variable, the performance of the model generated with Prescriptive-FCM was evaluated using MAE, MSE, RMSE, obtaining values of 2.76, 14.8 and 3.8, respectively. We used $R^2$ as a measure of agreement between the actual data and that prescribed by the generated model. Fig. 5 shows a plot with the corresponding $R^2$ value and the significance value of the analysis. For this case study, the $R^2$ value expressed as a percentage was 96%.The optimal hyperparameters for this model were initial population of 50 individuals, crossover and mutation probabilities of 0.1 and 0.3, respectively.

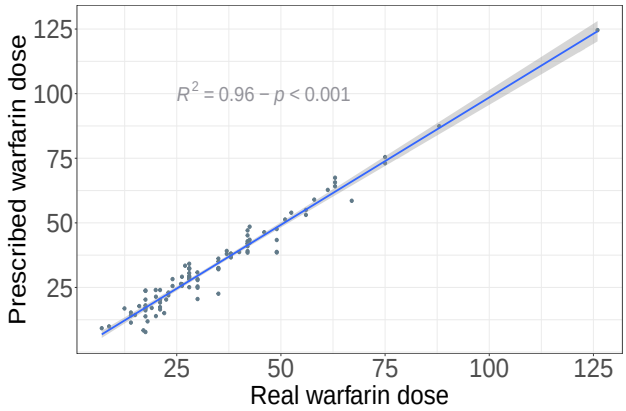


**Fig. 5.** Relationship between the warfarin values prescribed in the dataset and the warfarin values prescribed by our approach.

### 5.2. Case study 2: treatment of SD

#### 5.2.1. Descriptive statistics

Descriptive statistics for this case study are summarized in Table 5. In this dataset, all variables were qualitative. The frequency distribution shows that variable C8 was the most frequent variable in the group of patients who presented SD. The least frequent variables in this category were C1 - extravasation and C6 - use of crystalloid solutions, both with frequencies of 30.40% (95% CI = 25.88-34.9). The opposite case occurred in the group of patients who did not present SD, these two variables C1 and C6 were the most frequent with respect to the others, in both cases the relative frequency was 69.60% (95% CI = 65.08- 74.12). ICU stays within this group only occurred in 26.88% (95% CI = 22.52 - 31.24).

#### 5.2.2. Predictive model

The mortality rate for SD can reach 20% if the clinical management of the disease is not done in an ideal way [38]. For this case study, we developed a model to predict mortality by SD. As in the previous case study, this procedure was performed by adjusting the weights of the FCM constructed by PSO (initial population = 70 individuals, iterations = 140). Table 7 shows the performance of the model developed to predict mortality by SD. The developed model had the ability to predict whether the patient dies or not with an accuracy of 0.74, sensitivity of 0.79 and specificity of 0.68.

#### 5.2.3. Prescriptive model

Prescribing treatment in SD is of vital importance to prevent patient death. We developed a model for prescribing treatment actions aimed at preventing patient death by SD. Four treatment options were used to generate the prescriptive model (see Table 5). Due to the binary nature of these actions, we used accuracy as a metric to evaluate the performance of the developed model. Table 8 shows that the prescriptive model generated with Prescriptive-FCM for the formulation of treatment actions for SD has an accuracy greater than 0.81. The best performance of this model was for the prescription of colloid solutions with an accuracy, sensitivity and specificity of 1. The optimal hyperparameters for this model were initial population of 100 individuals, and crossover and mutation probabilities of 0.5 and 0.5, respectively.

**Table 8**

Performance of the prescriptive model for the treatment of SD.

| Case study | Prescriptive concept | Variable name | Accuracy | Sensitivity | Specificity |
|---|---|---|---|---|---|
| Treatment of SD | C5 | Red blood cells transfusion | 0.81 | 0.64 | 1.00 |
| | C6 | Crystalloid solutions | 0.87 | 0.80 | 0.93 |
| | C7 | Colloid solutions | 1.00 | 1.00 | 1.00 |
| | C8 | Intensive care unit | 0.84 | 0.87 | 0.83 |

### *5.3. Case study 3: Prevention of geohelminthiasis*

#### *5.3.1. Descriptive statistics*

The results of the nine categorical variables that make up this dataset allowed describing it statistically using relative frequencies with 95% CI. 52.1% (95% CI = 48.55-55.65) of the individuals in the dataset were women with weights between 20-40 kg in 72.4% (95% CI = 69.27-75.61) and between 40-60 kg in 17.9% (95% CI = 15.25-20.71). Only 24.4% (95% CI = 21.6-27.46) of the participants reported belonging to an indigenous ethnicity. Variables C4, C5, C6 and C8, all of them from the system and related to epidemiological aspects, showed that the origin of the water for cooking is mainly from wells in 67.4% (95% CI = 64.12-70.78) or from a river or stream in 24.2% (95% CI = 21.24-27.32). Dirt floors predominate in 88.5% (95% CI = 86.32-90.84) of the dwellings of these subjects, and excreta disposal is done in toilets without connection in 38.7% (95% CI = 35.25-42.17) or connected to a septic tank in 36.7% (95% CI = 33.33-40.17) mainly. After defecation few participating subjects washed their hands, 30.7% (95% CI = 27.43-33.99) said they always washed their hands, while 41.8% (95% CI = 38.36-45.36) said they never washed their hands. The two prescriptive variables of the dataset (C7 and C9) showed as results that the use of closed footwear is not a common practice among the study subjects, 41.8% (95% CI = 38.36-45.36) reported never using this type of footwear, likewise, a similar percentage of subjects, 41.6% (95% CI = 38.1-45.1) stated that they washed food before consumption.

#### *5.3.2. Predictive model*

We developed a predictive model with PSO-FCM (initial population = 50 individuals, iterations = 150) to predict the presence of geohelminths infections using demographic and epidemiological variables. The performance of this model can be seen in Table 7. The model predicted the

**Table 9**

Performance of the prescriptive model for the prevention of geohelminthiasis.

| Case study | Prescriptive concept | Variable name | Accuracy | Sensitivity | Specificity |
|---|---|---|---|---|---|
| Prevention of | C12 | Child wears closed shoes | 0.74 | 0.80 | 0.74 |
| geohelminthiasis | C14 | Child washes hands before eating | 0.67 | 0.78 | 0.55 |

parasitosis with an accuracy of 0.74, sensitivity of 0.76 and specificity of 0.73.

### *5.3.3. Prescriptive model*

The prevention of geohelminthiasis is important to avoid the spread of parasites in communities. We developed a model to prescribe two crucial actions in the prevention of geohelminthiasis. The results show the model's ability to prescribe these actions with accuracies between 0.67 and 0.74. The developed model had greater sensitivity than specificity for the two prescriptive variables used (see Table 9). The optimal hyperparameters for this model were initial population of 50 individuals, and crossover and mutation probabilities of 0.5 and 0.5, respectively.

## 6. Discussion

In this study, we developed prescriptive models (and its underlying predictive model) to support decision-making in clinical settings. We used three case studies: the first, related to the estimation of warfarin doses for anticoagulated patients. The second case study related to the treatment of dengue fever to reduce mortality rates. Finally, the third case was focused on the prevention of soil-transmitted parasitic infections.

### *6.1. Warfarin dosing*

The estimation of the warfarin dose is crucial to avoid both bleeding and the presence of clots in patients with coagulation disorders. The developed predictive model used demographic and genetic variables to obtain an acceptable performance (see Table 7). The results are expected due to the lack of clinical and laboratory variables necessary for careful monitoring since there is a wide variation in dose response explained by baseline clinical conditions, lifestyles and food consumption. Including variables such as comorbidities (diabetes and arterial hypertension), would be

useful because these types of diseases have been reported as risk factors for hemorrhagic complications in patients receiving warfarin. Aggregation of these types of variables will possibly allow better prediction of the INR. Another variable to take into account when considering the dose of warfarin is the intake of vitamin K, since it actively participates in the blood coagulation process. To prescribe the appropriate dose of warfarin to maintain a well-controlled INR, it is necessary to consider the measurement of vitamin K in the meals eaten by anticoagulated patients, since any variation in this may change the amount of warfarin to be taken. [43]. Other variables such as lifestyle changes, discontinuation of warfarin, falls or serious injuries, consumption of two or more alcoholic beverages per day, becoming pregnant or breastfeeding may affect the INR [35]. Therefore, it is important to consider some of these changes as variables within the predictive models developed.

Regarding the prescriptive model for estimation of warfarin dose, the results were satisfactory due to very low error values such as MAE below 2.8, exceeding the performance of previous works. Table 10 shows a comparison of the models developed to estimate warfarin dose with the dataset used in the present work. The International warfarin Pharmacogenetics Consortium developed two models using a clinical and pharmacogenetic algorithm, obtaining values of MAE 9.9 and 8.5, respectively [36]. Considering the R2 that measures the degree of agreement between the actual warfarin values in the dataset and the value prescribed in the developed model, our model had a superior performance with values of 0.96. The models developed by this consortium obtained maximum values of 0.43. Another work developed by Chen et al [44], proposed a weighted learning method to estimate warfarin dose on the same dataset used in this study. The results of the model generated with the methodology proposed by Chen obtained an R2 of 0.36. Our model performed better than the models developed and reported in the literature.

### *6.2. SD treatment*

In the second case study, the results demonstrated a good capacity both to predict mortality by SD and prescribe treatment options to prevent the patient's death. The predictive model performed well with accuracy values above 74%. The variables defining SD have functional dependencies with mortality. Several studies have demonstrated the influence of shock, extravasation, bleeding

**Table 10**

Comparison of models developed to estimate warfarin doses.

| Reference | Model | MAE | $R^2$ |
|---|---|---|---|
| [36] | Clinical | 9.9 | 0.26 |
| [36] | Pharmacogenetic | 8.5 | 0.43 |
| [44] | Predictive | - | 0.36 |
| Our work | Prescriptive | 2.7 | 0.96 |

and multiorgan failure on dengue death [45–47]. However, other variables considered as warning signs of dengue may be more influential in the prediction. Among these variables are abdominal pain, hepatomegaly, which consists of an increase in liver size due to fluid accumulation in the abdominal region; small mucosal hemorrhages and edema, which consists of fluid accumulation in the tissues underlying organs.

The prescriptive model for the treatment of SD consisted of prescribing treatment options according to WHO indications. The results showed a good performance of the developed models reaching values between 81% and 100% accuracy. Our model has the capacity to prescribe actions aimed at reducing the dengue mortality rate. The scarcity of works on prescriptive modeling makes it difficult to compare our work with previous studies. To date, there is no prescriptive model for the treatment of SD. An important work to highlight in the palliative treatment of dengue is the one performed by Hoyos et al [11] In this work, a prescriptive model was developed using autonomous cycles of data analysis tasks based in GAs; however, the work was focused on the three types of dengue. In addition, the model developed was validated in specific scenarios and not in a complete dataset.

### *6.3. Geohelminthiasis prevention*

The prevention of soil-transmitted helminth infections is of public health importance. The predictive model generated performed well only using demographic and epidemiological data. However, other epidemiological, clinical and laboratory variables could improve the prediction performance. These variables could be, for example, maternal or caregiver schooling. In the prevention of geohelminthiasis, it is important that those responsible for the care of children have

adequate levels of education since it is possible that people with more schooling are more aware of the importance of adopting healthy practices, such as boiling water or washing hands before handling food; in addition, these people are more capable of transmitting this knowledge to their families. Clinically, geohelminthiases are polymorphic and do not present pathognomonic signs and symptoms, many of them are asymptomatic, so the measurement of clinical variables is related to the presence of a particular parasitic agent; however, among the general symptoms and/or signs are anemia, weight loss and growth retardation. When these symptoms become evident, the parasitic infection is in progress, being useful these clinical variables in the prevention of the course of the intensity of the infection towards severity [41]. In endemic areas for these parasitic infections, the necessary diagnostic tools are often not available and the local epidemiology is unknown, overlooking the performance of laboratory tests that yield diagnostics. Often the results of a blood count, which shows laboratory variables such as hemoglobin and eosinophil count useful in the prediction of geohelminthiasis, are available. These parasites affect nutritional status by various mechanisms by feeding on host tissues, particularly blood, which causes a loss of iron and protein. Likewise, by activating TH2 lymphocytes (T helper type 2), they stimulate the secretion of IgE, producing an increase in the levels of eosinophils in blood, becoming the main cause of eosinophilia in pediatric age [48].

The prescriptive model generated to prescribe geohelminthiasis prevention actions performed acceptably with average accuracy values of 70.5%, perhaps for the reasons mentioned above. Additionally, a small sample size in categorical variables does not allow finding functional dependencies between these variables and the target variable. Despite having used SMOTE to generate new training examples of prescriptive variables, the variability of the data is very low and does not allow finding the necessary patterns to make a prescription with greater accuracy. According to our literature review, to date, no predictive models have been proposed to detect at individual-level geohelminthiasis. Previous work has focused mainly on estimating prevalence over a 5-year period during a disease control program [49]. Another work has been developed to determine the status and distribution of geohelminths in specific regions [50]. In addition, several studies have focused on determining the factors that most influence the disease to develop control strategies [51, 52]. To the best of our knowledge, this is the first work to report a predictive model to detect

geohelminthiasis using only demographic and epidemiological variables.

### *6.4. Comparison with previous approaches*

To situate the proposed methodology within the existing body of research, we conducted a comparative analysis using qualitative criteria to understand the novelty and advantages of our approach over other techniques and frameworks. Table 11 shows the qualitative criteria used for the comparison of approaches proposed in the literature versus our approach. The criteria are:

1. The approach generates and evaluates recommendations to achieve a desired outcome,
2. The approach is simple and easy to understand by medical professionals,
3. The approach was tested in several case studies to demonstrate its generalizability, and
4. The approach uses an explainable or interpretable technique.

The work of Bertsimas et al. [7] only meets criterion 1 because, although it generates recommendations with excellent performance, it is an approach that uses support vector machines, which are complex techniques for medical professionals to understand. In addition, the approach was not evaluated in several case studies to assess its generalizability. The work by Kovalchuk et al. [53] meets criteria 1, 2, and 4 because this approach was based on a three-step process using reference guidelines combined with explainable techniques to improve prediction results and suggested recommendations. The work by Zoubi et al. [54] meets criteria 2 and 4 because, although it is an approach that uses interpretable techniques and is easy to understand by medical professionals, the study does not evaluate the recommendations suggested by the proposed approach and the approach was not evaluated on different datasets to assess its generalizability. The work of Dumlu et al. [20] only meets criterion 1 because it generates recommendations or prescriptions with good performance; however, the mathematical complexity of the Markov decision models is a limitation for interpretability and ease of use by the medical professionals.

This is made possible by FCMs, a highly interpretable technique that simplifies understanding by medical professionals due to its accessible nature in both construction and interpretation process. In addition, our approach demonstrates a solid performance in the generation and evaluation

**Table 11**

Qualitative comparison between previous prescriptive approaches and our proposed approach.

| Criteria | Study | | | | |
|---|---|---|---|---|---|
| | Bertsimas et al [7] | Kovalchuk et al [53] | Zoubi et al [54] | Dumlu et al [20] | Our study |
| 1 | ✓ | ✓ | ✗ | ✓ | ✓ |
| 2 | ✗ | ✓ | ✓ | ✗ | ✓ |
| 3 | ✗ | ✗ | ✗ | ✗ | ✓ |
| 4 | ✗ | ✓ | ✓ | ✗ | ✓ |

of recommendations and prescriptions, excelling in crucial areas such as treatment, monitoring and prevention of diseases of public health relevance. Finally, the validation of our approach encompasses a diverse range of datasets addressing multiple contexts in clinical settings. This evaluation process not only highlights its strong generalizability, but also demonstrates the remarkable achievements in various aspects, including disease treatment, monitoring and prevention.

## 7. Conclusions

### *7.1. General considerations*

In recent years, the development of computer-aided strategies to support decision-making in clinical settings has increased. The objective of this work was to develop prescriptive models to support decision-making in scenarios related to the treatment, follow-up and prevention of diseases of public health interest. We used the Prescriptive-FCM methodology which consists of characterizing a problem into concepts defined as system concepts and action concepts, by using predictive and prescriptive models. The goal was to optimize the action concepts leading to desired outcomes of the system concepts. To train and test the models, we used datasets that included specific variables for each case study, whose data were collected in previous studies. The results demonstrated the ability of the developed models to predict INR values and estimate warfarin dosage in patients on anticoagulation therapy. In addition, we proved the ability to generate models that predict mortality from SD and prescribe treatment actions to avoid fatalities. Finally, we were able to demonstrate that prescriptive models generate actions aimed at the prevention of geohelminth infection.

In summary, our study demonstrated the ability of our Prescriptive-FCM methodology to generate prescriptive models that can be applied to any medical problem, whether for treatment, follow-up or prevention of public health events.

**Table 12**

Concomitant diseases that could influence the follow-up and treatment process of the diseases represented in the case studies [55–57].

| Concomitant disease | SD treatment | Follow-up with warfarin | Prevention of geohelminthiasis |
|---|---|---|---|
| Liver disease | Impaired liver function can compromise the body's ability to handle dengue infection. | Impaired liver function may influence warfarin metabolism and require dose adjustments to avoid bleeding or clotting | Impaired liver function may affect immune response to parasitic infections |
| Chronic kidney disease | Renal insufficiency may hinder the elimination of dengue virus breakdown products and increase the risk of renal complications. | Compromised renal function may influence warfarin excretion and require close monitoring to prevent side effects. | Impaired renal function may influence the excretion of parasites and their eggs in the feces. |
| Autoimmune diseases | Autoimmune disorders can affect the immune response to dengue and complicate the course of the disease. | Immunosuppressive drugs used to treat autoimmune diseases can interact with warfarin and increase the risk of bleeding | Immunosuppressive drugs may influence the immune response to parasitic infections |
| Type 2 diabetes | Changes in blood glucose levels may influence metabolic response to dengue and complicate disease management | Changes in blood glucose levels may affect response to warfarin and require dosage adjustments to avoid complications | Type 2 diabetes can affect susceptibility to parasitic infections and immune response |

### *7.2. Limitations and future work*

This work is not without limitations. Below, we show each of the limitations encountered and future opportunities for research. First, for the construction of the predictive models, we did the characterization of the FCM concepts manually; however, the characterization of these variables could be done automatically, speeding up the creation and training of the models. Second, for the generation of the prescriptive models, we only used one algorithm (GA) that optimized the action concepts for each case study. Other optimization algorithms could improve the quality of the developed models. Third, learning the FCMs (for prediction and prescription) with PSO was performed in a single stage, using system concepts and action concepts together. In this case, using two-stage learning could be more beneficial, because the influence of system variables is different from action variables. Fourth, the present research is a retrospective study where the data were previously collected and the researchers could not choose which variables to add to build the prescriptive models. Data availability is a common limitation when building predictive and

prescriptive models, mainly in the health field due to the sensitivity of the data used. For the validation of our methodology, we used the data available in the datasets with specific variables in each case study. It is important to clarify that the data used belonged to specific populations and that the models developed are not applicable to other populations with different characteristics. If data from other populations are available, our methodology can generate new models that fit the data of interest. Fifth, there are some factors that can positively or negatively affect the efficacy of the treatment and follow-up process at particular stages of disease development. Within this group, we find concomitant diseases, genetic factors, and environmental factors, among others, which unfortunately were not found in all the datasets used for training and testing of the predictive and prescriptive models. Table 12 shows some examples of concomitant diseases that could affect or influence the treatment or follow-up process of the diseases represented in each case study. Sixth, the datasets only had information on two disease states (healthy vs sick or sick vs dead). However, in reality, there are different disease states (see Table 13), which due to the available data we were unable to assess. The addition of important information to the datasets such as the presence of concomitant diseases and the different disease states would allow the development of more robust models that allow a more complete analysis on the process of prevention, treatment and follow-up of diseases of public health interest. Finally, another limitation of our study is the size of the datasets of some case studies for training and testing the models. Currently, the availability of data with prescriptive variables is a major limitation due to the low availability of data related to patient treatment and follow-up in repositories for free use. Collecting more patient records could improve the quality of the models. It has been widely demonstrated that increasing the number of data could improve the quality of predictions and prescriptions using ML.

Despite the limitations present in our study, our proposed methodology is a starting point for the development of models that support decision-making with respect to the prevention, follow-up and treatment of diseases of public health interest. The combination of FCMs with GA is a valuable approach for the development of models to support decision-making in clinical settings. Validation of these models with larger datasets supplemented with important factors, such as concomitant diseases and different disease states, is necessary for their applicability in real clinical settings. In general, our approach is scalable to the incorporation of more variables (such as con-

**Table 13**

Main stages of a disease [58].

| Stage of disease | Description |
|---|---|
| Underlying | This stage refers to genetic predisposition or risk factors that increase the likelihood of developing a disease in the future. In this stage, there are no symptoms or signs of the disease, but underlying factors may be present. |
| Susceptible | At this stage, a person is exposed to causative agents (such as viruses, bacteria or toxins) that could cause disease. Susceptibility may be influenced by genetic, environmental and lifestyle factors. |
| Subclinical | During this stage, the disease is present but no overt clinical symptoms are evident. However, changes in biomedical parameters or medical test results may occur that indicate the presence of the disease. |
| Clinical | At this stage, the characteristic symptoms and signs of the disease become evident. Clinical diagnosis is possible and medical measures can be taken to treat the disease and alleviate the symptoms. |
| Recovery/disability/death | This stage marks the outcome of the disease. There may be complete recovery, long-term disability or death of the individual, depending on the severity of the disease and the effectiveness of treatment. |

comitant diseases), or more disease states (classes).


## Acknowledgements

This study was partially funded by Colombian Administrative Department of Science, Technology and Innovation - COLCIENCIAS (grant number 111572553478) (M. Toro) and Colombian Ministry of Science and Technology *Bicentennial PhD Grant* (W. Hoyos). Also, the authors want to thank the HAMADI 4.0 project, code 22-STIC-06, of the STIC-AmSud regional program, for the support of this work.


## Conflict of interest

The authors declare no conflict of interest.

## CRediT authorship contribution statement

**William Hoyos:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Validation, Visualization & Writing – original draft. **Jose Aguilar:** Conceptualiza-

tion, Formal analysis, Resources, Supervision, Writing – reviewing & editing. **Mayra Raciny:** Conceptualization, Formal analysis, Investigation, Resources, Writing – reviewing & editing. **Mauricio Toro:** Conceptualization, Resources, Supervision, Writing – reviewing & editing.

## References


[1] K. Lepenioti, A. Bousdekis, D. Apostolou, G. Mentzas, Prescriptive analytics: Literature review and research challenges, International Journal of Information Management 50 (2020) 57–70. doi:10.1016/J.IJINFOMGT.2019.04.003.

[2] W. Hogendoorn, F. L. Moll, B. E. Sumpio, M. G. Hunink, Clinical Decision Analysis and Markov Modeling for Surgeons: An Introductory Overview, Annals of Surgery 264 (2016) 268–274. doi:10.1097/SLA.0000000000001569.

[3] D. K. Owens, R. D. Shachter, R. F. Nease, Representation and analysis of medical decision problems with influence diagrams, Medical Decision Making 17 (1997) 241–262. doi:10.1177/0272989X9701700301.

[4] L. N. Steimle, B. T. Denton, Markov decision processes for screening and treatment of chronic diseases, in: International Series in Operations Research and Management Science, volume 248, 2017, pp. 189–222. doi:10.1007/978-3-319-47766-4_6.

[5] A. J. Schaefer, M. D. Bailey, S. M. Shechter, M. S. Roberts, Modeling Medical Treatment Using Markov Decision Processes, in: Operations Research and Health Care, Springer, Boston, MA, 2005, pp. 593–612. doi:10.1007/1-4020-8066-2_23.

[6] H. Harikumar, S. Rana, S. Gupta, T. Nguyen, R. Kaimal, S. Venkatesh, Prescriptive analytics with differential privacy, International Journal of Data Science and Analytics 13 (2022) 123–138. doi:10.1007/s41060-021-00286-w.

[7] D. Bertsimas, M. L. Li, I. C. Paschalidis, T. Wang, Prescriptive analytics for reducing 30-day hospital readmissions after general surgery, PLoS ONE 15 (2020) e0238118. doi:10.1371/journal.pone.0238118.

[8] H. Harikumar, S. Rana, S. Gupta, T. Nguyen, R. Kaimal, S. Venkatesh, Prescriptive analytics through constrained bayesian optimization, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 10937 LNAI, Springer Verlag, 2018, pp. 335–347. doi:10.1007/978-3-319-93034-3_27.

[9] M. M. Hosseini, M. Zargoush, F. Alemi, R. E. Kheirbek, Leveraging machine learning and big data for optimizing medication prescriptions in complex diseases: a case study in diabetes management, Journal of Big Data 7 (2020). doi:10.1186/s40537-020-00302-z.

[10] W. Hoyos, J. Aguilar, M. Toro, Dengue models based on machine learning techniques: A systematic literature review, Artificial Intelligence in Medicine 119 (2021) 102157. URL: https://www.

sciencedirect.com/science/article/pii/S0933365721001500. doi:https://doi.org/10.1016/j.artmed.2021.102157.

[11] W. Hoyos, J. Aguilar, M. Toro, An autonomous cycle of data analysis tasks for the clinical management of dengue, Heliyon 8 (2022) e10846. doi:10.1016/j.heliyon.2022.e10846.

[12] E. Chalmers, D. Hill, V. Zhao, E. Lou, Prescriptive analytics applied to brace treatment for AIS: A pilot demonstration, Scoliosis 10 (2015) 2–5. doi:10.1186/1748-7161-10-S2-S13.

[13] J.-M. Bae, Clinical Decision Analysis using Decision Tree, Epidemiology and Health (2014) e2014025. doi:10.4178/epih/e2014025.

[14] K. Kurisu, K. Yoshiuchi, Comparison of Antipsychotics for the Treatment of Patients With Delirium and QTc Interval Prolongation: A Clinical Decision Analysis, Frontiers in Psychiatry 12 (2021) 1–7. doi:10.3389/fpsyt.2021.609678.

[15] L. Keikes, M. Kos, X. A. Verbeek, T. Van Vegchel, I. D. Nagtegaal, M. J. Lahaye, A. Méndez Romero, S. De Bruijn, H. M. Verheul, H. Rütten, C. J. Punt, P. J. Tanis, M. G. Van Oijen, Conversion of a colorectal cancer guideline into clinical decision trees with assessment of validity, International Journal for Quality in Health Care 33 (2021) 1–8. doi:10.1093/intqhc/mzab051.

[16] N. Bazrafshan, M. M. Lotfi, A finite-horizon Markov decision process model for cancer chemotherapy treatment planning: an application to sequential treatment decision making in clinical trials, Annals of Operations Research 295 (2020) 483–502. doi:10.1007/s10479-020-03706-5.

[17] Y. Habu, Vonoprazan versus Lansoprazole for the Initial Treatment of Reflux Esophagitis: A Cost-effectiveness Analysis in Japan, Internal Medicine 58 (2019) 2427–2433. doi:10.2169/internalmedicine.2535-18.

[18] Y. Shen, M. Hu, Q. Chen, Y. Zhang, J. Liang, T. Lu, Q. Ma, R. Ma, Comparative Effectiveness of Different Combinations of Treatment Interventions in Patients with Stroke at the Convalescence Stage Based on the Markov Decision Process, Evidence-based Complementary and Alternative Medicine 2020 (2020) 1–9. doi:10.1155/2020/8961341.

[19] M. Eghbali-Zarch, R. Tavakkoli-Moghaddam, F. Esfahanian, A. Azaron, M. M. Sepehri, A Markov decision process for modeling adverse drug reactions in medication treatment of type 2 diabetes, Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine 233 (2019) 793–811. doi:10.1177/0954411919853394.

[20] Z. Önen Dumlu, S. Sayin, I. H. Gürvit, Screening for preclinical Alzheimer's disease: Deriving optimal policies using a partially observable Markov model, Health Care Management Science (2022). doi:10.1007/s10729-022-09608-1.

[21] N. Prayogo, D. Pirayesh Neghab, S. Mahbub Rafayal, M. Cevik, Partially Observable Markov Chain Models for Evaluating Lung Cancer Screening Policies, in: Proceedings of the 32nd Annual International Conference on Computer Science and Software Engineering, 2022, pp. 81–90. doi:10.5555/3566055.3566065.

[22] B. Kosko, Fuzzy cognitive maps, International Journal of Man-Machine Studies 24 (1986) 65–75. doi:10.1016/S0020-7373(86)80040-2.

[23] J. Aguilar, A fuzzy cognitive map based on the random neural model, Lecture Notes in Computer Science 2070 (2001) 333—-338. doi:10.1007/3-540-45517-5_37.

[24] J. Aguilar, J. Contreras, The fcm designer tool, in: Glykas (Ed.), Fuzzy Cognitive Maps: Advances in Theory, Methodologies, Tools and Application, Springer, 2010, pp. 71–88.

[25] H. Sanchez, J. Aguilar, O. Teran, J. Gutierrez, Modeling the process of shaping the public opinion through multilevel fuzzy cognitive maps, Applied Soft Computing 85 (2019) 105756. doi:https://doi.org/10.1016/j.asoc.2019.105756.

[26] J. Aguilar, Multilayer cognitive maps in the resolution of problems using the fcm designer tool, Applied Artificial Intelligence 30 (2016) 720–743. doi:10.1080/08839514.2016.1214422.

[27] M. Glykas, Fuzzy Cognitive Maps: Advances in Theory, Methodologies, Tools and Applications, volume 247, 2005. arXiv:arXiv:1011.1669v3.

[28] J. L. Salmeron, S. A. Rahimi, A. M. Navali, A. Sadeghpour, Medical diagnosis of Rheumatoid Arthritis using data driven PSO–FCM with scarce datasets, Neurocomputing 232 (2017) 104–112. doi:10.1016/j.neucom.2016.09.113.

[29] T. G. Altundoğan, M. Karaköse, An approach for online weight update using particle swarm optimization in dynamic fuzzy cognitive maps, in: 2018 3rd International Conference on Computer Science and Engineering (UBMK), 2018, pp. 1–5. doi:10.1109/UBMK.2018.8566363.

[30] J. Kennedy, R. Eberhart, Particle swarm optimization, in: Proceedings of ICNN'95 - International Conference on Neural Networks, volume 4, 1995, pp. 1942–1948 vol.4. doi:10.1109/ICNN.1995.488968.

[31] C. D. Stylios, P. P. Groumpos, Modeling Complex Systems Using Fuzzy Cognitive Maps, IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans. 34 (2004) 155–162. doi:10.1109/TSMCA.2003.818878.

[32] E. I. Papageorgiou, A new methodology for Decisions in Medical Informatics using fuzzy cognitive maps based on fuzzy rule-extraction techniques, Applied Soft Computing Journal 11 (2011) 500–513. doi:10.1016/j.asoc.2009.12.010.

[33] D. E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, volume Addison-We, 1989. doi:10.1007/s10589-009-9261-6.

[34] D. Molina Castaño, M. Campos Quesada, A. Núñez Guerrero, History of anticoagulants and their clinical use in the present, Revista Medica Sinergia 5 (2020) e365. doi:10.31434/rms.v5i2.365.

[35] G. Mora-Pabón, Arritmias y anticoagulación, in: Anticoagulación, primera ed., Sociedad Colombiana de Cardiología y Cirugía Cardiovascular, 2005, pp. 29–35.

[36] T. I. W. P. Consortium, Estimation of the warfarin dose with clinical and pharmacogenetic data, The New

England journal of medicine 360 (2009) 753. doi:10.1056/NEJMOA0809329.

[37] N. T. N. Phuong, D. H. Manh, S. P. Dumre, S. Mizukami, L. N. Weiss, N. Van Thuong, T. T. N. Ha, L. H. Phuc, T. Van An, T. M. Tieu, M. G. Kamel, M. E. Morra, V. T. Q. Huong, N. T. Huy, K. Hirayama, Plasma cell-free DNA: A potential biomarker for early prediction of severe dengue, Annals of Clinical Microbiology and Antimicrobials 18 (2019). doi:10.1186/s12941-019-0309-x.

[38] Q. Jing, M. Wang, Dengue epidemiology, Journal of Global Health 3 (2019) 37–45. doi:10.1016/j.glohj.2019.06.002.

[39] World Health Organization, Dengue Guidelines for Diagnosis, Treatment, Prevention and Control: new edition, Technical Report, 2009. URL: https://apps.who.int/iris/bitstream/handle/10665/44188/9789241547871_eng.pdf?sequence=1&isAllowed=y.

[40] World Health Organization, Soil-transmitted helminth infections, 2022. URL: https://www.who.int/news-room/fact-sheets/detail/soil-transmitted-helminth-infections.

[41] P. J. Hotez, M. Alvarado, M. G. Basáñez, I. Bolliger, R. Bourne, M. Boussinesq, S. J. Brooker, A. S. Brown, G. Buckle, C. M. Budke, H. Carabin, L. E. Coffeng, E. M. Fèvre, T. Fürst, Y. A. Halasa, R. Jasrasaria, N. E. Johns, J. Keiser, C. H. King, R. Lozano, M. E. Murdoch, S. O'Hanlon, S. D. Pion, R. L. Pullan, K. D. Ramaiah, T. Roberts, D. S. Shepard, J. L. Smith, W. A. Stolk, E. A. Undurraga, J. Utzinger, M. Wang, C. J. Murray, M. Naghavi, The Global Burden of Disease Study 2010: Interpretation and Implications for the Neglected Tropical Diseases, PLoS Neglected Tropical Diseases 8 (2014) e2865. doi:10.1371/journal.pntd.0002865.

[42] R. L. Pullan, J. L. Smith, R. Jasrasaria, S. J. Brooker, Global numbers of infection and disease burden of soil transmitted helminth infections in 2010, Parasites and Vectors 7 (2014) 1–19. doi:10.1186/1756-3305-7-37.

[43] A. Furey, K. Moriarty, C. Silva, K. Fiumara, A. Kachalia, A guide to taking Warfarin, Partners Health Care, Washington D.C., 2014.

[44] G. Chen, D. Zeng, M. R. Kosorok, Personalized Dose Finding Using Outcome Weighted Learning, Journal of the American Statistical Association 111 (2016) 1509–1521. doi:10.1080/01621459.2016.1148611.

[45] S. S. Md-Sani, J. Md-Noor, W. H. Han, S. P. Gan, N. S. Rani, H. L. Tan, K. Rathakrishnan, M. A. A-Shariffuddin, M. Abd-Rahman, Prediction of mortality in severe dengue cases, BMC Infectious Diseases 18 (2018). doi:10.1186/s12879-018-3141-6.

[46] D. V. John, Y. S. Lin, G. C. Perng, Biomarkers of severe dengue disease - A review, 2015. doi:10.1186/s12929-015-0191-6.

[47] R. C. Pinto, D. B. De Castro, B. C. De Albuquerque, V. D. S. Sampaio, R. A. D. Passos, C. F. Da Costa, M. Sadahiro, J. U. Braga, Mortality predictors in patients with severe dengue in the state of Amazonas, Brazil, PLoS ONE 11 (2016) e0161884. doi:10.1371/journal.pone.0161884.

[48] L. Prieto-Pérez, R. Pérez-Tanoira, A. Cabello-Úbeda, E. Petkova-Saiz, M. Górgolas-Hernández-Mora, Geohelminths, Enfermedades Infecciosas y Microbiologia Clinica 34 (2016) 384–389. doi:10.1016/j.eimc.

2016.02.002.

[49] A. Montresor, A. F. Gabrielli, A. Yajima, N. Lethanh, B. A. Biggs, G. J. Casey, T. T. Tinh, D. Engels, L. Savioli, Markov model to forecast the change in prevalence of soil-transmitted helminths during a control programme: A case study in Vietnam, Transactions of the Royal Society of Tropical Medicine and Hygiene 107 (2013) 313–318. doi:10.1093/trstmh/trt019.

[50] R. G. Scholte, N. Schur, M. E. Bavia, E. M. Carvalho, F. Chammartin, J. Utzinger, P. Vounatsou, Spatial analysis and risk mapping of soil-transmitted helminth infections in Brazil, using Bayesian geostatistical models, Geospatial Health 8 (2013) 97–110. doi:10.4081/gh.2013.58.

[51] A. Loukouri, A. Méité, O. K. Kouadio, N. N. Djè, G. Trayé-Bi, B. G. Koudou, E. K. N'Goran, Prevalence, Intensity of Soil-Transmitted Helminths, and Factors Associated with Infection: Importance in Control Program with Ivermectin and Albendazole in Eastern Côte d'Ivoire, Journal of Tropical Medicine 2019 (2019). doi:10.1155/2019/7658594.

[52] C. A. Yaro, E. Kogi, S. A. Luka, M. A. Nassan, J. Kabir, K. N. Opara, H. F. Hetta, G. E. S. Batiha, Edaphic and climatic factors influence on the distribution of soil transmitted helminths in Kogi East, Nigeria, Scientific Reports 11 (2021) 1–12. doi:10.1038/s41598-021-88020-1.

[53] S. V. Kovalchuk, G. D. Kopanitsa, I. V. Derevitskii, G. A. Matveev, D. A. Savitskaya, Three-stage intelligent support of clinical decision making for higher trust, validity, and explainability, Journal of Biomedical Informatics 127 (2022) 1–14. doi:10.1016/j.jbi.2022.104013. arXiv:2007.12870.

[54] F. Al Zoubi, R. Gold, S. Poitras, C. Kreviazuk, J. Brillinger, P. Fallavollita, P. E. Beaulé, Artificial intelligence-driven prescriptive model to optimize team efficiency in a high-volume primary arthroplasty practice, International Orthopaedics 47 (2023) 343–350. doi:10.1007/s00264-022-05475-1.

[55] S. Fernando, A. Wijewickrama, L. Gomes, C. T. Punchihewa, S. D. Madusanka, H. Dissanayake, C. Jeewandara, H. Peiris, G. S. Ogg, G. N. Malavige, Patterns and causes of liver involvement in acute dengue infection, BMC Infectious Diseases 16 (2016) 1–9. doi:10.1186/s12879-016-1656-2.

[56] D. C. Malone, D. S. Hutchins, H. Haupert, P. Hansten, B. Duncan, R. C. Van Bergen, S. L. Solomon, R. B. Lipton, Assessment of potential drug-drug interactions with a prescription claims database, American Journal of Health-System Pharmacy 62 (2005) 1983–1991. doi:10.2146/ajhp040567.

[57] A. Berbudi, N. Rahmadika, A. I. Tjahjadi, R. Ruslami, Type 2 Diabetes and its Impact on the Immune System, Current Diabetes Reviews 16 (2019) 442–449. doi:10.2174/1573399815666191024085838.

[58] L. Kisling, J. Das, Prevention Strategies, 2023. URL: https://www.ncbi.nlm.nih.gov/books/NBK537222/.