ОЦЕНКА ФАКТОРОВ РИСКА ДЕФИЦИТА ВИТАМИНА D У ПАЦИЕНТА С ХРОНИЧЕСКИМ ГЕПАТИТОМ В С ПРИМЕНЕНИЕМ АЛГОРИТМА ИЗУЧЕНИЯ ДЕРЕВА РЕШЕНИЙ

Фрэштэх Османи

Исследовательский центр инфекционных заболеваний, Бирджандский университет медицинских наук, Бирджанд, Иран

ASSESSMENT OF THE RISK FACTORS FOR VITAMIN D DEﬁCIENCY IN CHRONIC HEPATITIS B PATIENT USING THE DECISION TREE LEARNING ALGORITHM

Freshteh Osmani,

Dentistry Clinical Research Development Center, Birjand University of Medical Sciences, Birjand, Iran

Резюме.

История вопроса: инфекция ВГВ является серьезной проблемой для здоровья, которая может быть опасной для жизни. Витамин D (ВД) участвует в различных патофизиологических механизмах при различных заболеваниях. Также есть большая необходимость в прогнозировании его степени тяжести при помощи различных методов. Наше исследование было направлено на оценку эффективности DT как одной из моделей машинного обучения для прогнозирования степени дефицита витамина D.

Методы. Всего было обследовано 292 пациента с ХГВ, у которых были оценены уровни ВД в сыворотке. Для прогнозирования дефицита ВД использовались независимые характеристики, такие как возраст, пол, вес, рост, цинк, ИМТ, жировые отложения, воздействие солнечного света и потребление молока. 60% из них были внесены в обучающий набор данных случайным образом. Для оценки производительности дерева решений оставшиеся 40% были использованы в качестве набора данных тестирования. Валидация модели оценивалась кривой ROC.

Результаты. Распространенность дефицита ВД среди пациентов была высокой (63,0%). Окончательные результаты экспериментов показали, что классификатор DT обеспечивает лучшую точность 96% и превосходит по производительности при обучении и тестировании набора данных VD. Кроме того, площади под кривой ROC AUC составила (0,78) при применении алгоритма DT со значимыми переменными путем перекрестной проверки, с получением значения AUC = 0,78 и точности 85,3%.

Заключение. Мы пришли к выводу, что уровень цинка в сыворотке крови является важным сопутствующим фактором риска для выявления случаев дефицита витамина D. Кроме того, риск дефицита ВД можно предсказать с высокой точностью с использованием алгоритма обучения

дерева решений, который можно применять для противовирусной терапии у пациентов с ХГВ.

Ключевые слова: дефицит витамина D, дерево решений, машинное обучение, вирус гепатита B, витамин D, кривая ROC.

Abstract.

Background: HBV infection is a major health problem which may be life-threatening. vitamin D (VD) is involved in various pathophysiological mechanisms in a plethora of diseases. And also, there is a strong demand for the prediction of its severity using different methods. The study aims to evaluate performance of DT as one of the machine learning models in the prediction of severity in vitamin D deficiency.

Methods: In total, data containing serum VD levels were collected from 292 CHB patients. The independent characteristics such as: age, sex, weight, height, zinc, BMI, body fat, sunlight exposure, and milk consumption were used for prediction of VD deficiency. 60% of them were allocated to a training dataset randomly. To evaluate the performance of decision-tree the remaining 40% were used as the testing dataset. The validation of the model was evaluated by ROC curve.

Results: The prevalence of VD deficiency was high among patients (63.0%). The final experimentation results showed that DT classifier achieves better accuracy of 96 % and outperforms well on training and testing of VD dataset. Also, the areas under the ROC curve AUC is (0.78), when we applied DT algorithm with the significant variables by cross validation, the values of AUC= 0.78 and 85.3% accuracy were obtained

Conclusion: We concluded that the serum level of Zn is an important associated risk factor for identifying cases with vitamin D deficiency. Also, the risk of VD deficiency could be predicted with high accuracy using decision tree learning algorithm that could be used for antiviral therapy in CHB patients.

Key words: Vitamin D deficiency, Decision tree, Machine Learning, Hepatitis B virus, Vitamin D, ROC curve

Introduction

The liver is a basic place for vitamin D_3synthesis, where 25-hydroxylation occurs and a large portion vitamin D_3binding protein is manufactured(1). Vitamin D_3 plays an emerging role in inflammatory and metabolic liver diseases. There is strong evidence about the association between vitamin $D_3$and various chronic liver diseases in different stages (2). About 240 million individuals are infected with HBV chronically in the whole world (3) It has been shown that vitamin D_3has very important biologic effects. Vitamin D_3 levels may affect the immune system and host response to viral infections, like HBV infection(4).

Recent researches have shown a relationship between hematological factors and vitamin D. As an example, a study had shown that hematocrit (HCT) and hemoglobin (HB) levels are increased by taking vitamin D significantly for a period of 4 months in chronic kidney patients undergoing hemodialysis(5). Different researches indicated that low levels of vitaminD_3 are associated with high levels of HBV replication in CHB infection recently. Although, a study found a positive correlation between HBs-Ag seroclearance and vitamin D_3 levels. The other reported a significant correlation between low levels of serum vitamin D_3 and higher levels of HBV replication (5). Recently, data mining techniques such as Classification trees have been used as prediction, classification, and diagnosis tool(6). Alternating decision tree (ADT) technique combines the simplicity and effectiveness of boosting together in a single decision tree(7).

Therefore, in this study, at combining the serum biomarkers and clinical information, our basic purpose was to determine the associated potential risk factors related with vitamin D deficiency by using decision tree algorithms, in an Iranian CHB patients in compare to healthy population.

Method

This study was undertaken in Khorasan Jonoobi province in 2019 in the one of outpatient clinic of infection diseases.

Patients were selected randomly according to consent to patriciate since the beginning of the study. Finally, 292 patients with CHB (Hbs-Ag positive, anti-HBs negative), were included in the study according to calculated sample size by the following formula with power of 90%. In addition, 330 natural immunized persons (HBsAg negative, anti HBs has normal liver enzymes who have not received antiviral treatment were included. healthy group were selected from collected samples of the master plan of the province(8).

$$n = \left(\frac{z_{1-\frac{\alpha}{2}} + z_{1-\beta}}{0.5 \ln\left(\frac{1+r}{1-r}\right)} \right)^2 + 3$$

Written consent was obtained from the all of patients. Patients with any auto-immune diseases, cancer, severe renal disease, cardiovascular disorders, pregnancy, diabetic disease, thyroid disorders, other viral hepatitis (HCV, HDV, HIV) and other causes of liver disease, vitamin D, calcium supplement use or injection and hormone therapy in the last six months were excluded.

The inclusion criteria for CHB patients were: patients who were admitted to the infectious disease's outpatient clinic with the diagnosis of CHB with the approval of the infectious specialist according to clinical and serological signs, age ≥18 years and willingness to participate in the study. Finally, a total of 292 patients were met the inclusion criteria.

In addition, 330 healthy subjects without a history of hepatitis B disease with any auto-immune diseases, cancer, metabolic bone disease, cardiovascular disorders, malabsorption or thyroid, adrenal diseases were selected as healthy control group. The laboratory tests were performed with 10 cc of venous blood was taken from patients and healthy controls (14 h overnight fast). The blood serum was separated by centrifuging blood samples (Hettich model D-78532) and stored at 80 $^{\circ}$C for future analysis. The serum levels of vitamin D_3 were measured using a COBAS e411 analyzer, manufactured by Mannheim Roch diagnostic Gmbh

 ISSN 2220-7619 (Print)
ISSN 2313-7398 (Online)

in Germany, with the Elecsys kit (REF 0589413). Complete blood count (CBC) was measured in whole blood samples.

The data were including demographic characteristics such as age, sex, and also clinical factors including: body mass index (BMI), grade and the activity of fibrosis, and routine laboratory tests, like, alanine aminotransferase (ALT), aspartate aminotransferase (AST), alpha-fetoprotein (AFP), hemoglobin (Hb), levels of total cholesterol, LDL, HDL, FBS platelet count, glycated hemoglobin (HbA_1C), creatinine, serology finding, glucose and platelet count. Total vitamin D_3levels were measured in the serum samples. Based on the WHO, a level of 30 ng/mlor above is considered as vitamin D_3 sufficiency(9). Then, vitamin D_3 status was classified as normal (≥30 ng/ml), insufficient (20–29.9 ng/ml), and deficient (<20 ng/ml)(10).

Decision Tree (DT)

Data mining technique is favorite for all researchers as a method to extract unknown patterns. Decision tree algorithm is one of most popular procedure which creates a tree-structured model(11). By using this method, cases divide into separate groups or according to values of predictor variables, values of a target variable predicts (12). This algorithm has three types of nodes, including, the root, internal, and end nodes. By using all predictor variables, root node, including the all dataset, is split into subgroups. This method makes the nodes several times, up to form the homogenous subsets according to target variable(13). Variables with the best rate of splitting criterion remain in the model. Gini index is used as a splitting criterion(14). Decision nodes determine a collection of features. By this procedure, progressively all leaves are created. To confirm the reproducibility of the results, both internal and external validation was used to validate the generality of the DT algorithm.

Internal validation was performed by 10- fold cross-validation which is used to assess the performance of a model and a technique for evaluating how the results

ISSN 2220-7619 (Print)
ISSN 2313-7398 (Online)

of a statistical analysis will generalize.

External validation was performed using samples collected from a collaborating study by 150 patients which was independent from our dataset completely. With no significant differences in the variables of these two groups(15).

Statistical Analysis

The analyses were done by R statistical software, ver (3.4.1). to construct the decision tree, 60% of the data were chosen as training dataset. The remaining 40% were used to assess the performance of the model as testing dataset. For the multivariable analysis for factors associated with vitamin D deficiency, logistic regression model was used. The significance in all of these tests was two-tailed with a 5% significant level. The ROC curve sensitivity, specificity were measured for comparison. Several types of decision tree learning techniques (CART) [24], C4.5(16), were implemented on the datasets.

This study was approved by the ethics board committee of Birjand University of Medical Sciences, reference number: IR.BUMS.REC.1398.324.

Results

Out of the all subjects who participated in the study, 48.6% were male in the case group; with mean age 29 ± 5.3; and 52.2% female with mean age (31.5 ± 7.8); also, of the 330 healthy subjects, 22.1% male; mean age 23.1±7.7; and 77.8% female with mean age29.1±11.7. So; distribution of gender was similar in patients' group.

Data expressed as mean ± SD unless otherwise was stated. The data were divided into a training and testing dataset (60% vs 40%) respectively. A decision tree was built on the training dataset. The testing dataset were used to assess the model. Gini index was used for selecting the variables in the algorithm to achieve final tree. Variables such as age, sex, zinc, FBG, HDL, RBC, MCV, MCHC, and HCT were entered to this model. These variables were different between two groups significantly. The training and test datasets were similar to each other roughly. The results identified age, BMI and AST as independent predictors of vitamin D deficiency, with significant correlation (P.value<0.001) (Table.1). Therefore, these variables were used in model 1(DT for healthy subjects).

Decision tree was learned for the training dataset by using variables with significant correlation with vitamin D deficiency (p.value<0.001) including variables such as age, BMI, AST, ALT, and AFP.

Figure.1 shows the decision tree diagram of fitted model 1. In this figure, vitamin D deficiency is the main variable as shown with YES while insufficiency or sufficient vitamin D were shown as No in both groups. The patient and control groups' vitamin D deficiency were scored using summing all of the prediction nodes. If it is more than or equal to zero, then the case is high risked to have vit D deficiency and conversely. Positive figure of prediction nodes raise the probability of vit D deficiency or a negative one decreases it. Age group, sex, serum levels of ALT, ALP, AFP, histological stage of fibrosis were associated by the standard univariate and multivariate analysis (Table 2).

Figure.3 showed the accuracy, ROC curve, sensitivity, specificity values for predicting vit D deficiency in training set. The areas under the ROC curves is (0.78), When we applied DT algorithm with the significant variables by cross validation, the values of 0.78 ROC and 85.3% accuracy were obtained, which is similar to the obtained results of applying training and test sets one by one.

The extracted rules from both DT algorithm for CHB patients and healthy groups were shown in Table 3 and Table 4 respectively.

Discussion

This study was conducted for the first time in this region in East of Iran regarding vitamin D_3pattern in patients with CHB and healthy group to investigate factors associated with vitaminD_3 deficiency by using DT algorithm technique.

In a period (1990-2010), the prevalence of vitamin D_3deficiency was studied in Iranian society and according to the results, in all regions; both sexes had moderate and significant vitamin D_3deficiency(17, 18). Nghiem.et al showed that vitamin D_3deficiency existed in many CHB patients and this deficiency had a relationship with the complications and outcome of the disease. Decreased liver function due to HBV-induced injuries to liver cells can be one of the causes of vitamin D3 deficiency in CHB(19, 20).

A substantial key finding of this study was exploring a new usage of the DT model for analyzing the predicators associated with vitamin D deficiency, which has not been studied formerly. In addition, the other major strong of this study was that it is the first model that determine vitamin D deficiency by using blood count parameters along with other variables as a component.

One of the interesting finding of this study was the pretense of serum zinc as a substantial factor for vitamin D deficiency which is in line with other researches (21).

A previous study has shown a significant correlation between serum level of vit D and low serum levels of zinc among Iranian population aged 10-18 years old(22). In addition, another study had reported a statistically association between serum level of vit D and serum levels of zinc among Iranian pregnant women. Their findings showed 37% vitamin D deficiency in pregnant women and 23% of them had zinc deficiency(23, 24).

One other interesting obtained result, was the presence of FBG after serum zinc. If FBG was <98 (mg/dl), then, hematological factors specify the other layers of tree. In these persons MCV and MCHC were the other effective agents for vitamin D deficiency. It has been made clear that deficiency of vitamin D can weaken insulin secretion which is a major agent in the pathogenesis of T2DM(25). Despite of vitamin D receptor, it was shown that the synthesis and secretion of insulin can be induced by 25OHD in islet beta cells(26). Another important finding of this model was that for the individuals had FBG<98 (mg/dl), MCHC and MCV were the indicators of having or not having vitamin D deficiency. These findings are in line with other studies reported a relationship between 25OHD and hematological factors(27).

Data mining analysis has been used to investigate hidden patterns in datasets , so it can be applied to build prediction models(28). In the present study, the relationship between AFP levels and chronic HBV patients was investigated. This finding that AFP was obtained as a significant factor related with vitamin D deficiency in both DT and the multivariate logistic models was interesting.

Antiviral therapy, the concomitant decrease in AFP and ALT levels indicated that inflammation presence and hepatocellular injury were the most important reasons of high AFP in these patients. This model have the potential ability to select patients with Vitamin D deficiency based on the possibility of response against a various factor. Moreover, it may provide a rationale to improve the efficacy of therapy. Similarly, CART analysis recognized several variables which were not associated with response by standard statistical model significantly(29). The fitted DT model could identify few demographic characteristics such as age and sex as significant factors associated with Vitamin D deficiency.

BMI higher than normal is considered to be an effective factor in the level of vitamin D3stores. In this study, significant relationship was found between serum levels of vitamin D3and BMI. This result was in line with other studies.

In the published studies on the prevalence of vitamin D3deficiency showed that vitamin D3deficiency prevalence was significantly different based on geographical regions in Iranian population(30, 31). In this study, however, vitamin D3deficiency was not correlated with liver function parameters significantly, probably due to that vitamin D3serum levels are affected by multiple factors (1, 32, 33), our results confirm a correlation between HbA1C loads and vitamin D3levels in patients' group.

ISSN 2220-7619 (Print)
ISSN 2313-7398 (Online)

Some limitations of this study were: influencing of several factors on Serum vitamin D_3 levels; such as: seasonal variation, diet and geographical habitation. However, any information about these affecting factors for study subjects was not available. The authors aim to develop more sensitivity and specificity prediction models, which be able to specify having vitamin D deficiency more exactly in their future studies.

Conclusion

In this study, we used biochemical and hematological factors as input variables in both CHB and healthy Iranian subjects and the findings specified the associated factors to vitamin D deficiency. Current study acquires a convenient way to apply classification rules for classifying potential factors related to vitamin D deficiency.

Acknowledgement

We would like to thank all the study subjects for their participation. Dentistry Clinical Research Development Center, Birjand University of Medical Sciences, Birjand, Iran for consulting**.**

Conflict of interest

None

FIGURES

Fig. 1. Decision tree with training dataset in healthy group (model.1)

Рис. 1. Дерево решений с обучающим набором данных в здоровой группе (модель 1)

ISSN 2220-7619 (Print)
ISSN 2313-7398 (Online)

Fig. 2. Decision tree with training dataset in CHB group (model.2)

Рис. 2. Дерево решений с обучающим набором данных в группе ХГВ (модель 2)

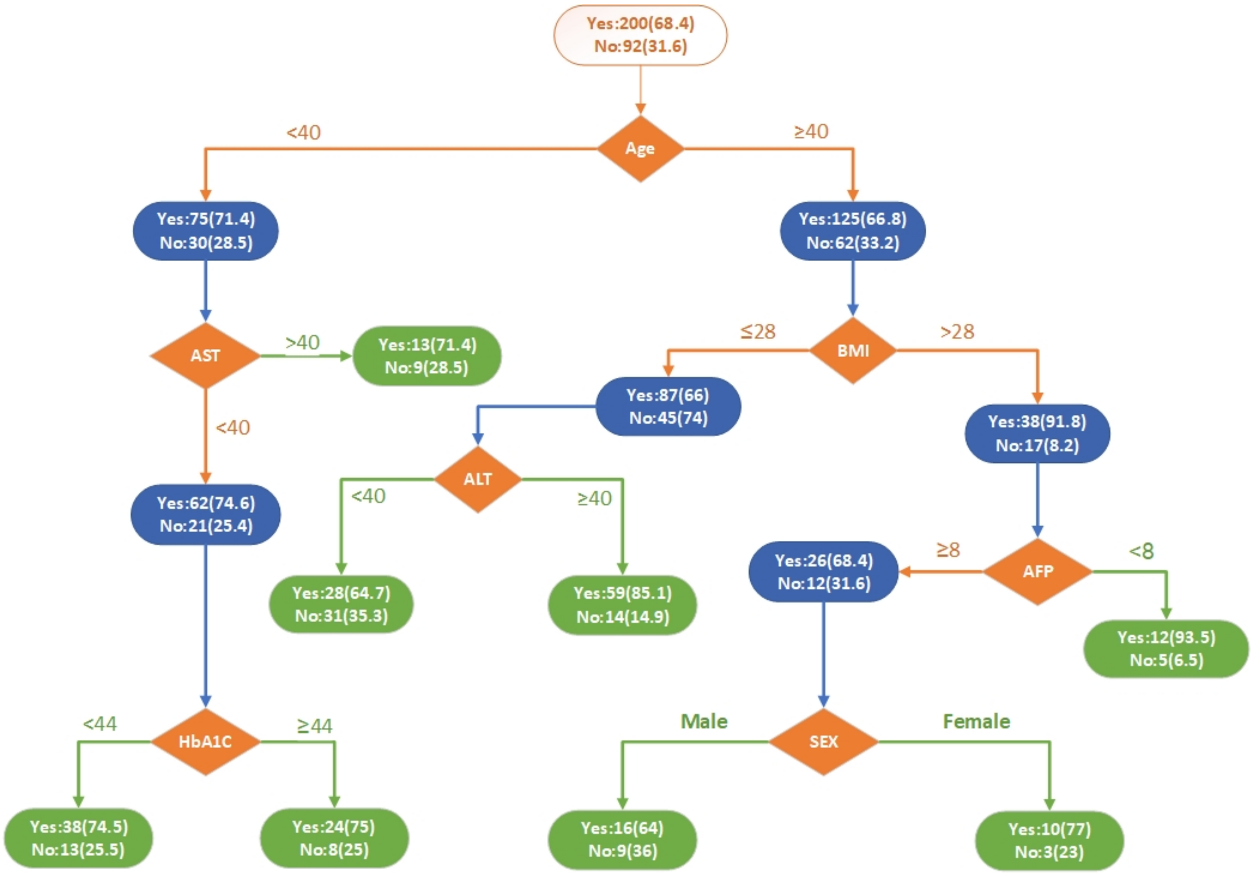
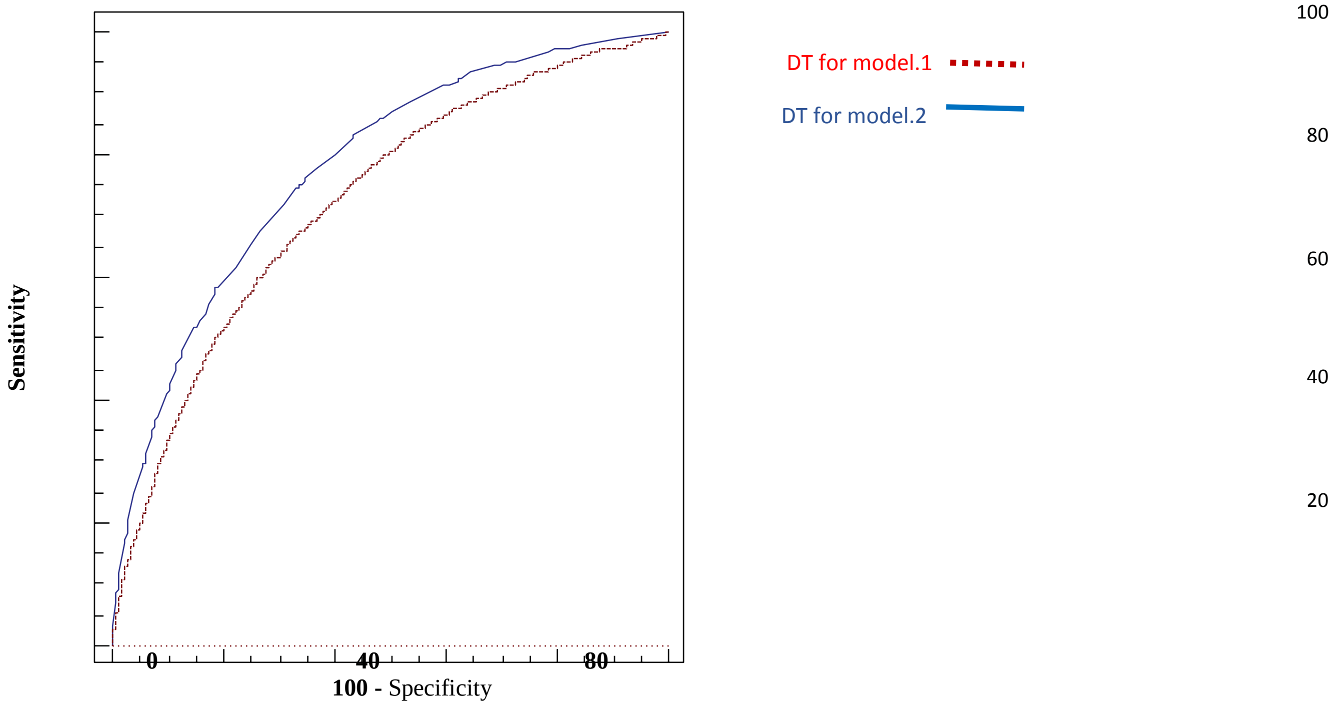


Fig 3. Roc curve of both decision tree models.

Рис. 3. Roc-кривая обеих моделей дерева решений.



TABLES

Table 1: Characteristics of variables

Таблица 1: Характеристики переменных

Variables Переменные		**Training dataset Набор данных для обучения**	**Validation dataset Набор данных для валидации**	**Pearson correlation coefficients Коэффициенты корреляции Пирсона**	***P* value *P* значение**
Age Возраст		36 ± 12	35 ± 11	0.32	<0.01
Sex Пол	Male	(72.7%)	(73.1%)		
	Female	(27.3%)	(26.9%)	-0.03	-0.03 0.008
BMI ИМТ		24.20 ± 4.21	25.32 ± 2.84	0.10	<0.001
AFP (U/L) альфа-фетопротеин (ед/л)		6.56 ± 16.51	6.69 ± 21.49	0.10	<0.001
AST (U/L) АСТ (ед/л)		54.17 ± 23.73	55.78 ± 31.62	0.12	<0.001
ALT (U/L) АЛТ (ед/л)		61.84 ± 36.89	61.84 ± 38.19	0.06	0.008
Hemoglobin (Hb) Гемоглобин (Hb)		14.03 ± 1.47	14.03 ± 1.62	−0.02	0.005

ISSN 2220-7619 (Print)
ISSN 2313-7398 (Online)

Albumin (g/dL) Альбумин (г / дл)		3.39 ± 0.47	3.40 ± 0.53	0.05	0.064
Platelet count (∗109/L) Количество тромбоцитов (* 109 / л)		216.48 ±53.64	214.55 ± 55.5	0.07	<0.12

Table 2: Multivariate logistic regression analysis for factors associated with vitamin D deficiency in CHB patients

Таблица 2: Многомерный логистический регрессионный анализ факторов, связанных с дефицитом витамина D у пациентов с ХГВ

Parameter **Параметр**	**Odds ratio** **Соотношение шансов**	**SE** **среднеквадратическая ошибка**	**95% CI** **95% ДИ**	***P***
Age > **40** **Возраст > 40**	1.0	0.13	0.8-1.3	**0.86**
AFP > 10 ng/ml **нг-фетопротеин > 10 ng/мл**	1.1	0.01	1.0-1.1	**< 0.01**
Male gender **Мужской пол**	0.67	0.08	0.53-0.85	**0.001**
Hb<14	0.47	0.06	0.37-0.59	**0.001**
ALP (> 290 IU/L) **ЩФ (> 290 МЕ/л)**	1.7	0.21	1.3-2.2	**< 0.01**
ALT (> 40 IU/L) **АЛТ (>40 МЕ/л)**	1.0	0.0	0.99-0.99	**0.01**
AST (> 40 IU/l) **АСТ (>40 МЕ/л)**	1.0	0.0	0.99-0.99	**0.01**

Table 3. The extracted rules through the DT model for healthy subjects (model.1)

R1: IF Zn < 88 (mg/dl), THEN class: subject with deficiency (89.2%)
R2: IF Zn ≥ 88 (mg/dl) and FBG ≤ 98, THEN class: subject with deficiency (91.8%)
R3: IF Zn ≥88 (mg/dl) and FBG ≤ 98and MCV< 92, THEN class: subject without deficiency (6.5%)
R4: IF Zn ≥88 (mg/dl), FBG ≤ 98 and MCV<92, and MCHC<32 THEN class: subject with deficiency (75%)
R5: IF Zn ≥88 and FBG ≤ 98, MCV<92 and MCHC≥32, THEN class: subject with deficiency (93.4%)
R6: IF Zn ≥88 and FBG≥98 and Hb<14, THEN class: subject with deficiency (64.7%)
R7: IF Zn ≥88 and FBG≥98 and Hb≥14, THEN class: person with deficiency (85.1%)
R8: IF Zn ≥88 (mg/dl), FBG≥98 and Hb≥14 and HCT ≤41, THEN class: person without deficiency (8.5%)
R9: IF Zn ≥88 (mg/dl), FBG≥98 and Hb≥14 and HCT >41, THEN class: person with deficiency (44.5%)

Таблица 3. Правила, извлеченные с помощью модели DT для здоровых субъектов (модель 1)

R1: ЕСЛИ Zn <88 (мг / дл), ТО класс: подвержен дефициту (89,2%)
R2: ЕСЛИ Zn ≥ 88 (мг / дл) и ВБР ≤ 98, ТО класс: подвержен дефициту (91.8%)
R3: ЕСЛИ Zn ≥88 (мг / дл) и FBG ≤ 98 и MCV <92, ТО класс: не подвержен дефициту (6.5%)
R4: ЕСЛИ Zn ≥88 (мг / дл), FBG ≤ 98 и MCV <92 и MCHC <32ТО класс: подвержен дефициту (75%)
R5: ЕСЛИ Zn ≥88 и FBG ≤ 98, MCV <92 и MCHC≥32, ТО класс: подвержен дефициту (93.4%)
R6: ЕСЛИ Zn ≥88 и FBG≥98 и Hb <14, ТО класс: подвержен дефициту (64.7%)
R7: ЕСЛИ Zn ≥88 и FBG≥98 и Hb≥14, ТО класс: подвержен дефициту (85.1%)
R8: ЕСЛИ Zn ≥88 (мг / дл), FBG≥98 и Hb≥14 и HCT ≤41, ТО класс: не подвержен дефициту (8.5%)

 10.15789/2220-7619-AOT-1642

R9: ЕСЛИ Zn ≥88 (мг / дл), FBG≥98 и Hb≥14 и HCT> 41, ТО класс: не подвержен дефициту (44.5%)

 ISSN 2220-7619 (Print)
ISSN 2313-7398 (Online)

Table 4. The extracted rules through the DT model for CHB patients (model.2)

R1: IF Age≥40, THEN class: subject with deficiency (66.8%)
R2: IF Age≥40and BMI>28, THEN class: subject with deficiency (91.8%)
R3: IF Age≥40, BMI>28 and AFP <8 THEN class: subject without deficiency (6.5%)
R4: IF Age≥40, BMI>28 and AFP ≥8 THEN class: subject with deficiency (75%)
R5: IF IF Age≥40 AND BMI≤28, THEN class: subject with deficiency (66%)
R6: IF Age<40 and AST>40 THEN class: subject with deficiency (71.4%)
R7: IF Age<40, THEN class: person with deficiency (85.1%)
R8: IF Age<40 and AST≤40, THEN class: person without deficiency (25.4%)
R9: IF Age<40, AST≤40, and HbA1C<44 THEN class: person with deficiency (74.5%)
R10: IF Age<40, AST≤40 and HbA_1C≥44 THEN class: person with deficiency (75%)

Таблица 4. Правила, извлеченные с помощью модели DT для пациентов с ХГВ (модель 2)

R1: ЕСЛИ Возраст ≥40, ТО класс: подвержен дефициту (66.8%)
R2: ЕСЛИ Возраст ≥40 и ИМТ> 28, ТО класс: подвержен дефициту (91.8%)
R3: ЕСЛИ Возраст ≥40, ИМТ> 28 и АФП <8ТО класс: не подвержен дефициту (6.5%)
R4: ЕСЛИ Возраст ≥40, ИМТ> 28 и АФП ≥8ТО класс: подвержен дефициту (75%)
R5: ЕСЛИ ЕСЛИ Возраст ≥ 40 И ИМТ ≤ 28, ТО класс: подвержен дефициту (66%)
R6: ЕСЛИ Возраст <40 и АСТ> 40 ТО класс: подвержен дефициту (71.4%)
R7: ЕСЛИ Возраст <40, ТО класс: человек с дефицитом (85.1%)
R8: ЕСЛИ Возраст <40 и АСТ≤40, ТО класс: человек без дефицита (25.4%)
R9: ЕСЛИ Возраст <40, АСТ≤40 и HbA1C <44ТО класс: человек с дефицитом (74.5%)

R10: ЕСЛИ Возраст <40, AST≤40 и HbA1C≥44 ТО класс: человек с дефицитом (75%)

TITLE PAGE_METADATA

Assessment of the Risk Factors for vitamin D Deficiency in Chronic Hepatitis B Patient using the Decision Tree Learning Algorithm

Freshteh Osmani ,

Dentistry Clinical Research Development Center, Birjand University of Medical Sciences, Birjand, Iran

ОЦЕНКА ФАКТОРОВ РИСКА ДЕФИЦИТА ВИТАМИНА D У ПАЦИЕНТА С ХРОНИЧЕСКИМ ГЕПАТИТОМ В С ПРИМЕНЕНИЕМ АЛГОРИТМА ИЗУЧЕНИЯ ДЕРЕВА РЕШЕНИЙ

Фрэштэх Османи

Исследовательский центр инфекционных заболеваний, Бирджандский университет медицинских наук, Бирджанд, Иран

Corresponding Author: Freshteh Osmani, Dentistry Clinical Research Development Center, Birjand University of Medical Sciences, Birjand, Iran Tel: 09151637653, E-mail: dr.osmani68 @gmail.com ORCID: https://orcid.org/0000-0002-6112-7131

Краткое название: Факторы риска дефицита витамина D у больных ХГВ
Running title: Risk Factors for vitamin D Deficiency in CHB Patients

Ключевые слова: дефицит витамина D, дерево решений, машинное обучение, вирус гепатита B, витамин D, кривая ROC.

Keywords: Vitamin D deficiency, Decision tree, Chronic HBV infection

REFERENCES

1	Coussens AK, Wilkinson RJ, Hanifa Y, Nikolayevskyy V, Elkington PT, Islam K, et al. Vitamin D accelerates resolution of inflammatory responses during tuberculosis treatment. Proceedings of the National Academy of Sciences. 2012;109(38):15449-54.
2.	Efe C, Kav T, Aydin C, Cengiz M, Imga NN, Purnak T, et al. Low serum vitamin D levels are associated with severe histological features and poor response to therapy in patients with autoimmune hepatitis. Digestive diseases and sciences. 2014;59(12):3035-42.
3.	Tseng T-C, Kao J-H. Clinical utility of quantitative HBsAg in natural history and nucleos (t) ide analogue treatment of chronic hepatitis B: new trick of old dog. Journal of gastroenterology. 2013;48(1):13-21.
4.	Chen E-Q, Shi Y, Tang H. New insight of vitamin D in chronic liver diseases. Hepatobiliary & Pancreatic Diseases International. 2014;13(6):580-5.
5.	Hauge Matzen L, Christensen J, Hintze H, Schou S, Wenzel A. Diagnostic accuracy of panoramic radiography, stereo-scanography and cone beam CT for assessment of mandibular third molars before surgery. Acta Odontologica Scandinavica. 2013;71(6):1391-8.
6.	Heshmat R, Mohammad K, Majdzadeh S, Forouzanfar M, Bahrami A, Ranjbar Omrani G, et al. Vitamin D deficiency in Iran: A multi-center study among different urban areas. Iran J Public Health. 2008;37(1):72-8.
7.	Ghaziasadi A, Ziaee M, Norouzi M, Malekzadeh R, Alavian SM, Saberfar E, et al. The prevalence of hepatitis B virus surface antigen (HBsAg) variations and

	correlation with the clinical and serologic pictures in chronic carriers from Khorasan Province, North-East of Iran. 2012.
8.	Trépo E, Ouziel R, Pradat P, Momozawa Y, Quertinmont E, Gervy C, et al. Marked 25-hydroxyvitamin D deficiency is associated with poor prognosis in patients with alcoholic liver disease. Journal of hepatology. 2013;59(2):344-50.
9.	Torresi J, Tran BM, Christiansen D, Earnest-Silveira L, Schwab RHM, Vincan E. HBV-related hepatocarcinogenesis: the role of signalling pathways and innovative ex vivo research models. BMC cancer. 2019;19(1):707.
10.	Wong GL-H, Chan HL-Y, Chan H-Y, Tse C-H, Chim AM-L, Lo AO-S, et al. Adverse effects of vitamin D deficiency on outcomes of patients with chronic hepatitis B. Clinical Gastroenterology and Hepatology. 2015;13(4):783-90. e1.
11.	Hoan NX, Van Tong H, Le Huu Song CGM, Velavan TP. Vitamin D deficiency and hepatitis viruses-associated liver diseases: A literature review. World journal of gastroenterology. 2018;24(4):445.
12.	Osmani F, Ziaee M. Assessment of the risk factors for vitamin D3 deficiency in chronic hepatitis B patient using the decision tree learning algorithm in Birjand. Informatics in Medicine Unlocked. 2021;23:100519.
13.	Parvaie P, Majd HS, Ziaee M, Sharifzadeh G, Osmani F. Evaluation of gum health status in hemophilia patients in Birjand (a case-control study). American Journal of Blood Research. 2020;10(3):54.

14.	Tayefi M, Saberi-Karimian M, Esmaeili H, Zadeh AA, Ebrahimi M, Mohebati M, et al. Evaluating of associated risk factors of metabolic syndrome by using decision tree. Comparative clinical pathology. 2018;27(1):215-23.
15.	Plum LA, DeLuca HF. Vitamin D, disease and therapeutic opportunities. Nature reviews Drug discovery. 2010;9(12):941-55.
16.	Schillie S, Xing J, Murphy T, Hu D. Prevalence of hepatitis B virus infection among persons with diagnosed diabetes mellitus in the United States, 1999–2010. Journal of viral hepatitis. 2012;19(9):674-6.
17.	Tabrizi R, Moosazadeh M, Akbari M, Dabbaghmanesh MH, Mohamadkhani M, Asemi Z, et al. High prevalence of vitamin D deficiency among Iranian population: a systematic review and meta-analysis. Iranian journal of medical sciences. 2018;43(2):125.
18.	Manco M, Ciampalini P, Nobili V. Low levels of 25-hydroxyvitamin D3 in children with biopsy-proven nonalcoholic fatty liver disease. Hepatology. 2010;51(6):2229-.
19.	Mahamid M, Nseir W, Elhija OA, Shteingart S, Mahamid A, Smamra M, et al. Normal vitamin D levels are associated with spontaneous hepatitis B surface antigen seroclearance. World journal of hepatology. 2013;5(6):328.
20.	Ziaei S, Norrozi M, Faghihzadeh S, Jafarbegloo E. A randomised placebo-controlled trial to determine the effect of iron supplementation on pregnancy outcome in pregnant women with haemoglobin≥ 13.2 g/dl. BJOG: An International Journal of Obstetrics & Gynaecology. 2007;114(6):684-8.
21.	Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. Hepatology. 1996;24(2):289-93.

22.	Wang D, Wang Q, Shan F, Liu B, Lu C. Identification of the risk for liver fibrosis on CHB patients using an artificial neural network based on routine and serum markers. BMC infectious diseases. 2010;10(1):251.
23.	Shoaei SD, Sali S, Karamipour M, Riahi E. Non-invasive histologic markers of liver disease in patients with chronic hepatitis B. Hepatitis monthly. 2014;14(2).
24.	Mohamadkhani A, Katoonizadeh A, Poustchi H. Immune-regulatory events in the clearance of HBsAg in chronic hepatitis B: focuses on HLA-DP. Middle East journal of digestive diseases. 2015;7(1):5.
25.	Farnik H, Bojunga J, Berger A, Allwinn R, Waidmann O, Kronenberger B, et al. Low vitamin D serum concentration is associated with high levels of hepatitis B virus replication in chronically infected patients. Hepatology. 2013;58(4):1270-6.
26.	Han J, Pei J, Kamber M. Data mining: concepts and techniques: Elsevier; 2011.
27.	Han J, Kamber M, Pei J. Data mining concepts and techniques third edition. The Morgan Kaufmann Series in Data Management Systems. 2011:83-124.
28.	Osmani F, Hajizadeh E, Rasekhi A, Akbari ME. Analyzing Relationship Between Local and Metastasis Relapses with Survival of Patients with Breast Cancer: A Study Using Joint Frailty Model. International Journal of Cancer Management. 2018;11(12).
29.	Azarkar G, Doosti Z, Osmani F, Ziaee M. Analysis Of Risk Factors For Nonalcoholic Fatty-Liver Disease In Hepatitis B Virus Infection: A Case–Control Study. Hepatic medicine: evidence and research. 2019;11:153.

30.	Osmani F. Effect evaluation of vitamin D level amongst patients with chronic hepatitis B. 2019.
31.	Osmani F. Problems with reporting accuracy in COVIDE-19 statistics in Iran. Gastroenterology and Hepatology from Bed to Bench.
32.	Hewison M. Vitamin D and immune function: autocrine, paracrine or endocrine? Scandinavian Journal of Clinical and Laboratory Investigation. 2012;72(sup243):92-102.
33.	Kitson MT, Roberts SK. D-livering the message: the importance of vitamin D status in chronic liver disease. Journal of hepatology. 2012;57(4):897-909.