DOI: https://doi.org/10.56712/latam.v5i6.3111

Estimation of the risk of death from COVID -19 in the cancer population

Estimación el riesgo de defunción por COVID - 19 en población con cáncer

Juan Bacilio Guerrero Escamilla
juan_guerrero9464@uaeh.edu.mx
https://orcid.org/0000-00002-0971-7564
Universidad Autónoma del Estado de Hidalgo
México

Sócrates López Pérez
sonia_bass10401@uaeh.edu.mx
https://orcid.org/0000-0002-1532-3254
Universidad Autónoma del Estado de Hidalgo
México

Sonia Bass Zavala
lopezs@uaeh.edu.mx
https://orcid.org/0000-0001-9261-9430
Universidad Autónoma del Estado de Hidalgo
México

Artículo recibido: 23 de noviembre de 2024. Aceptado para publicación: 07 de diciembre de 2024.
Conflictos de Interés: Ninguno que declarar.

Resumen

Uno de los principales problemas de salud que afronta México en los últimos años es el crecimiento acelerado del cáncer en sus distintas modalidades. De acuerdos con las estadísticas de la Secretaria de Salud, en el año 2022 se registraron más de 67 mil casos en un rango de 20 a 75 años. Con la presencia del covid, el virus generó una incertidumbre para la población con cáncer, pues al tener un sistema inmunitario debilitado, se corre con el mayor riesgo de fallecer, por tal motivo, es fundamental predecir el riesgo que conlleva el covid respecto al cáncer, por tanto, es esencial pronosticar sus efectos. En el presente trabajo de investigación se predice el riesgo de defunción que tiene la población de México con cáncer, a partir del contagio de covid, se toma como referencia la aplicación de vacunas para contrarrestar el virus, la edad, el peso, y el sexo de los pacientes. El sustento de esta investigación se base en la construcción de un modelo de matemático, el cual toma como referencia la metodología de la Investigación de Operaciones.

Palabras clave: covid-19, cáncer, vacunas, riesgo, modelamiento

Abstract

One of the principal health troubles that Mexico has faced in the last years is the accelerating growth of cancer in its diverse modalities. According to the statistics from the Secretary of Health, there were more than 67 thousand cases from 20 to 75 years in 2022. Since the occurrence of COVID-19, the virus produced uncertainty for the population with cancer. Thus, they have a weak immunological system, which leads to greater chances of dying. For this reason, it is fundamental to foresee the risk that causes COVID-19 concerning cancer, resulting in forecasting its effects. In the present research work, the risk of dying can be anticipated in the Mexican population who suffer from cancer, and the spread

of COVID-19 has referred to the administration of vaccines to counter the virus, age, weight, and sex of the patients. The justification of this research is based on constructing a mathematical model, which refers to an Operational Research Methodology.

Keywords: covid-19, cancer, vaccines, risks, modeling

Cómo citar: Guerrero Escamilla, J. B., López Pérez, S., & Bass Zavala, S. (2024). Estimation of the risk of death from COVID -19 in the cancer population. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades 5 (6), 1626 – 1647. https://doi.org/10.56712/latam.v5i6.3111*

INTRODUCTION

Cancer is a disease that has multiple cellular characteristics of growing and replication without control, and it can invade other parts of the body, different from the ones that originate them (Martinez, 2021). According to the National Institute of Cancer, the most common types of cancer and the ones that are more frequent diagnostics are the following (NIH, 2022):

Table 1

Types of cancer

Colorectal cancer	Pancreatic cancer	Bladder cancer
Endometria cancer	Prostate cancer	Thyroid cancer
Liver cancer	Lung cancer	Melanoma
Leukemia	Kidney cancer	Breast cancer
	Non-hodgkin's lymphoma	

Source: National Institute of Cancer, 2022.

Based on the statistics taken from the National Institute of Cancer in Mexico in 2022, the most recurrent are prostate, breast, and lung cancer, with more than 35 thousand cases, statistically, that means the following (NIH, 2022):

Table 2

Scores by type of cancer

Type of cancer	M	F	Type of cancer	M	F
Colorectal cancer	6,503		Pancreatic cancer	2,722	
Endometria cancer	2,814		Prostate cancer	12,253	
Liver cancer	1,751		Lung cancer	10,130	
Leukemia	2,533		Kidney cancer	3,477	
Non-hodgkin's lymphoma	3,423		Vaginal cancer	3,497	
Breast cancer	12,353	304	Thyroid cancer	1,858	
Melanoma	4,148				

Source: National Institute of Cancer, 2022.

COVID-19 is part of a family of viruses present in human beings and some species of animals. It is an infectious disease provoked by SARS–CoV–2. Moreover, it is transmittable from person to person, leading to a breathing disease pandemic (WHO,2020).

The first case of COVID-19 in Mexico was detected on February 27th, 2020, and March 11th, 2020. From this moment, the World Health Organization (WHO) declares a pandemic due to increasing cases from China. It was not until March 18th, 2020, that the first death was registered in Mexico. As a result, the Secretary of Health announced the National Day of Healthy Distance, which consisted of implementing actions referring to sanitary measures and social distancing, whose purpose was to decrease the spread (Inai, 2022).

On May 9th, 2023, the government of Mexico ended the COVID health emergency, with a total of 7,633,355 cases, from which 334,336 were deaths; having a higher incidence in the population from 30 to 40 years, with a greater degree of hospitalization from 65 to 69 years, people who had comorbidity, hypertension, diabetes, obesity or any genre of cancer (57.67%). Furthermore, most of the victims were

indigenous people and migrants with a higher economic degree of margination (Secretary of Health, 2023).

COVID-19 has affected the daily life of the population of Mexico, especially those with hypertension, diabetes, obesity, or any category of cancer. Considering this, the general objective of the present work is :

To estimate the risk of dying of a patient with cancer when developing COVID-19 and being vaccinated through the construction of a mathematical model taken from the information shared by the Secretary of Health during the pandemic period from February 19th to December 31st, 2022.

Based on the previously mentioned, the following questions are raised:

What type of cancer is more affected by COVID-19?

What is the risk of dying if a patient who has cancer develops COVID-19?

What is the vaccination that impacts the risk of dying of a patient with COVID-19 and cancer?

What is the risk level of dying from a patient with cancer who is vaccinated by COVID-19?

The scope of this research is to have a prognosis about the effects of COVID-19 and the vaccines administered to patients who have cancer. This research provides a diagnosis for diverse categories of cancer.

On the other hand, the limitation of this work is that some variables are not reflected, for instance, hypertension, diabetes, and social and economic conditions, among others. Hence, these are considered for a future job.

METHODOLOGY

The current research work has a methodological basis, which is Operational Research, which focuses on the construction and development of mathematical models that must fulfill the following stages (Hillier & Lieberman, 2010):

Definition of the problem: It raised a linear relationship between the dependent and independent variables. Afterwards, data collection is carried out.

Formulation of the model: it describes the mathematical technique to predict the behavior of the phenomenon under study.

Estimation of the parameters: by applying the technique of maximum verisimilar and computational systems, it implements the calculation of the model parameters.

Model validation: This stage involves confirming the assumptions of the inference model and assessing the fit of the mathematical equation.

Application of the model: it resides in interpreting the calculated parameters besides the possible sceneries of the phenomenon.

It uses RStudio software to construct the mathematical model, so it is in a free environment of integrated development by the programming language R, which focuses on statistical and graphic computing with more efficient calculations and prices obtained than any other commercial software(the programming code is in the annexes).

Definition of the problem

A sample of 1,184 patients was used to create the model, as we can observe in Table 3:

Table 3

Data basis

No.	Y	XC	XS	XE	XR	XV
1	REC	0	W	45	C2	V1
2	TRE	1	M	19	C3	V8
3	REC	0	M	27	C9	V3
4	DEA	1	M	13	C10	V1
---	---	---	---	---	---	---
---	---	---	---	---	---	---
---	---	---	---	---	---	---
---	---	---	---	---	---	---
---	---	---	---	---	---	---
1181	TRE	1	M	11	C7	V6
1182	REC	0	W	18	C6	NO
1183	REC	0	W	43	C3	V1
1184	REC	0	W	44	C3	V8

Source: The Secretary of Health, 2022.

Where:

Y = Health condition (REC = recovery = 1, TRE = Treatment = 2 , & DEA = death = 3).

XC = Covid-19 (1 = positive & 0 = negative).

XS = Sex (M = woman & H = men).

XE = Age (in years).

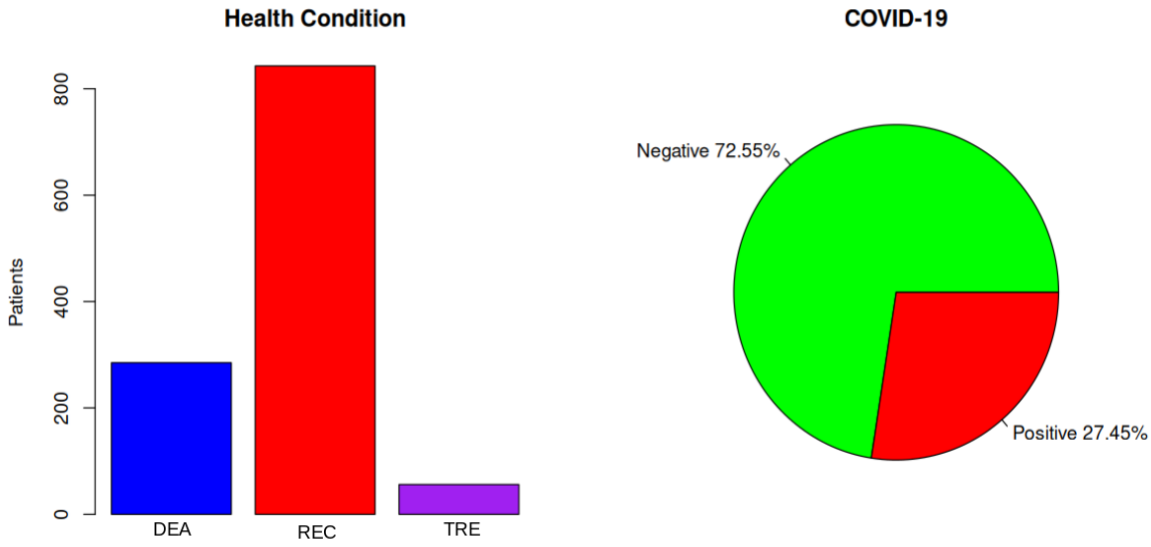
XR = Types of cancer (C1 = Prostate, C2 = Breast, C3 = Colorectal, C4 =Lung, C5 = Endometrial, C6 =Liver, C7 = Leukemia, C8 = Pancreas, C9 = Melanoma, C10 = Lymphoma, C11 = Kidney, C12 = Vaginal, C13 = Thyroid).

XV = Vaccines (V1 = Astrazeneca, V2 = Cansino, V3 = Pfizer, V4 = Gambelaya, V5 = Sinovac, V6 = Sinopharm, V7 = Janssen, V8 = Moderna, NO = Not vaccinated).

Statistically, the information has the following characteristics:

Graphic 1

Health condition and COVID-19

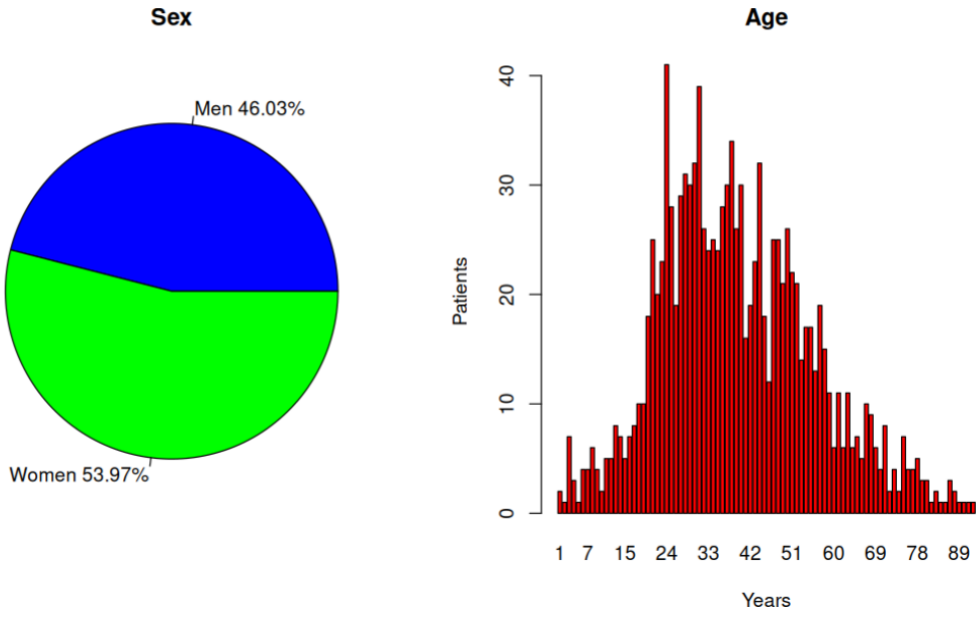


Source: own authorship.

Of the 1,184 patients, 46% are men and 54% are women. The age range is 1 to 90 years, predominantly those between 20 to 56 years, and the average age is 39 years (Graphic 2).}

Graphic 2

Sex & Age

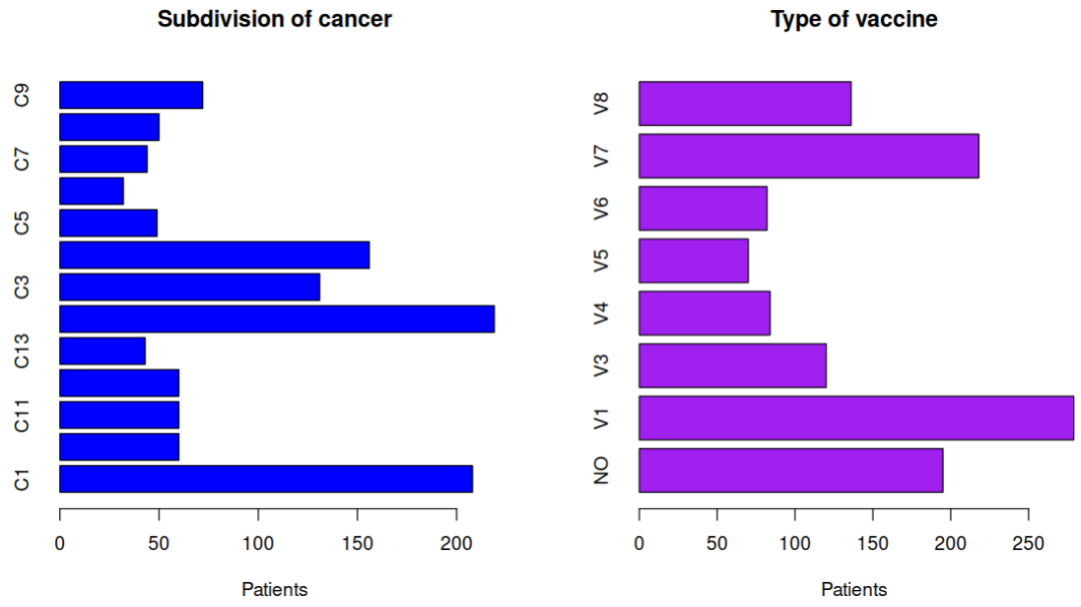


Source: own authorship.

100% of the patients have any subdivision of cancer, prevailing prostate (C1), breast (C2), and lung (C4). A dose to prevent COVID-19 was administered to most of them, being the most recurrent Astrazeneca (V1) & Janssen (V7). Nonetheless, 200 of them did not get vaccinated.

Graphic 3

Subdivision of Cancer and Type of Vaccine



Source: own authorship.

Model formulation

Health condition (Y) of the 1,1184 patients was in function of COVID-19(XC), sex(XC), age(XE), type of cancer(XR), and the variety of vaccination (XV):

$$Y = f(XC, XS, XE, XR, XV) \quad (1)$$

Whether Y has three events (DEA = death = 1, TRE = treatment = 2, & REC = recovery = 3), in terms of likelihood of linear relationship is:

$$Pr(Y) = \hat{B}_0 + \hat{B}_1 XC + \hat{B}_2 XS + \hat{B}_3 XE + \hat{B}_4 XR + \hat{B}_5 XV + ei \quad (2)$$

Where:

Pr(Y) is the probability of occurrence of any event or health condition.

$\hat{B}_i$; such that $i = 0, 1, 2, \ldots, 5$. They are the parameters to estimate the maximum plausibility.

ei is the error margin that does not explain the linear relationship of the phenomenon under study.

0 < Pr(Y) < 1, such that Pr(Y) = Pr(REC) + Pr(TRE) + Pr(DEA)

If Pr(Y) has events referring to the multi-logistic model, that means whether Y has more than two possible events with " n" independent rehearsals that allow k independent E1, E2,..., EK whose possibilities are the following (Montoya & Correa, 2017):

$$\boldsymbol{P(E_1) + P(E_2) + \ldots + P(E_k) = \sum_{i=1}^{k} P(E_i) = 1} \quad (3)$$

Where $n_1, n_2, \ldots, n_k$ is the number of occurrences of the events E1, E2,..., EK respectively in "n" rehearsals (Dekking, Lopuhaä, Kraaikamp & Meester, 2005):

$$\boldsymbol{n_1 + n_2 + \ldots + n_k = \sum_{i=1}^{k} n_i = n} \quad (4)$$

The function of likeliness of $n_1, n_2, \ldots, n_k$ events is proposed by (Evans & Rosenthal, 2010):

$$\boldsymbol{f(n_1, n_2, \ldots, n_k) = \frac{n!}{n_1!\, n_2! \ldots n_k!} P(E_1)^{n_1} P(E_2)^{n_2} \ldots P(E_k)^{n_k}} \quad (5)$$

For $\mathrm{n_i} = 0, 1, 2, \ldots, n$ subject to $\sum_{\mathrm{i=1}}^{\mathrm{k}} \mathrm{n_i} = \mathrm{n}$, where the average and variance of the multinomial distribution are provided by (Evans & Rosenthal, 2010):

$$\boldsymbol{E[n_i] = nP(E_i)} \quad (6) \qquad \boldsymbol{Var[n_i] = nP(E_i)\{1 - P(E_i)\}} \quad (7)$$

Whether Y has more than two possible events E1, E2,..., EK respectively in "n" trials, its most suitable model is the analysis of multi-logistic regression, in which, it is assumed that (Bocco & Herrero, 2009):

$$\boldsymbol{ln\left[\frac{Y}{1 - Y}\right] = \beta_0 + \sum_{i=1}^{n} \beta_i X_i;\ i = 1, 2, 3, 4, \ldots, k} \quad (8)$$

Where:

$\mathrm{X_i}$ are the dependent variables of the model.

β_i are the parameters to estimate.

The aim of this type of modeling is the estimation of the likelihood that any of the events occur taken from the independent variable, such that (Bocco & Herrero, 2009):

$$\boldsymbol{P[Y = i, X] = P(E_i) = \frac{e^{X_i \beta_i}}{1 + \sum_{j=1}^{g-1} e^{X_i \beta_i}}\ i = 1, \ldots, g - 1} \quad (9)$$

Where the estimation of the parameters is obtained through the maximum plausibility, which is given by the following algebraic expression (Pando, Martín, 2004):

$$\boldsymbol{L = \prod_{i=1}^{n} [P(E_{1i})^{Y_{1i}} * P(E_{2i})^{Y_{2i}} * \ldots * P(E_{ni})^{1 - (\Sigma_{i=1}^{n} Y_{ij})}]} \quad (10)$$

Therefore, the feasibility and adjustment of the model will depend on three elements (McCullagh & Nelder, 1983):

AIC (Akaike Information Criteria), evaluates the adjustment of the model to the data, as well as the complexity of the model:

$$\boldsymbol{AIC = 2p - 2Lm} \quad (11)$$

Where: 2p is the number of estimated parameters. Lm is the plausibility of the current model.

AIC is utilized to compare models when the size is smaller, the better the adjustment.

The final historical value of its interactions must be in the half of the deviance residuals.

The degree of incidence of the independent variables (Xi, i =1, 2, 3,....., k) about Y, to fulfill this, el P - Value must be higher in a level of significance.

The level of adjustment that a model has, in other words, the level of variability of the data that keeps the model, is done through deviance $(0 < \mathrm{D}^2 \leq 1)$.

$$\boldsymbol{D^2 = \frac{ND - Dr}{ND}} \qquad (12)$$

Where: ND is the null deviance and Dr is the deviance of the residuals.

The multi-logistic model assumes that the data of the phenomenon are specific to the case, where each independent variable has a unique value per each case. The null hypothesis of this type of modeling is that there is no relationship between the independent and dependent variables. In other words, the values of the dependent variable are anticipated from a multi-logistic equation, which is not closer to the actual values of the dependent variable (Bocco & Herrero, 2009).

Estimation of the parameters

Predicting the risk of health condition (0PY1)depends on the selection of the better adjustment model, and has a level of confidence of 0.95 and a significance of 0.05.

Nominal (A) $$\mathrm{E[Pr(Yn)]} = \widehat{\mathrm{B}}_0 + \widehat{\mathrm{B}}_1\mathrm{XC} + \widehat{\mathrm{B}}_2\mathrm{XS} + \widehat{\mathrm{B}}_3\mathrm{XE} + \widehat{\mathrm{B}}_4\mathrm{XR} + \widehat{\mathrm{B}}_5\mathrm{XV} + \mathrm{ei} \qquad (13)$$

Ordinal (B) $$\mathrm{E[Pr(Yo)]} = \widehat{\mathrm{B}}_0 + \widehat{\mathrm{B}}_1\mathrm{XC} + \widehat{\mathrm{B}}_2\mathrm{XS} + \widehat{\mathrm{B}}_3\mathrm{XE} + \widehat{\mathrm{B}}_4\mathrm{XR} + \widehat{\mathrm{B}}_5\mathrm{XV} + \mathrm{ei} \qquad (14)$$

Where:

$\mathrm{E[Pr(Yn)]}$ is the expected value of the plausibility of health condition (for the A model , dependent variable (Yn) is nominal (the A model: REC = Recovery, TRE = Treatment, DEA = Death) and ordinal (the B model: 1 = Recovery, 2 = Treatment, 3 = Death).

The estimation of the parameters is based on the maximum plausibility, consisting of adjusting the model with appropriate estimators through the calculation of parameters that might be optimal for the likelihood of the occurrence of the phenomenon.

Table 4

Model Run

The A Model (Nominal)				The B Model (Ordinal)		
	REC	**TRE**			**2**	**3**
Intercept	3.264	-16.326		Intercept	-20.005	-3.221
XC	-4.585	15.257		XC	20.255	4.542
XRC10	-0.217	0.687		XRC10	0.904	0.216
XRC11	0.392	0.554		XRC11	0.175	-0.379
XRC12	-0.615	0.481		XRC12	1.091	0.609
XRC13	0.934	1.751		XRC13	0.862	-0.881
XRC2	-0.277	0.863		XRC2	1.151	0.291
XRC3	-0.597	0.687		XRC3	1.192	0.506
XRC4	0.031	0.951		XRC4	0.923	-0.028

XRC5	-0.034	-0.209		XRC5	-0.173	0.039
XRC6	1.811	1.011		XRC6	-0.786	-1.796
XRC7	0.282	0.469		XRC7	0.188	-0.283
XRC8	-0.764	-1.103		XRC8	-0.341	0.762
XRC9	0.864	-0.242		XRC9	-1.088	-0.844
XSM	-0.249	-0.281		XSM	0.031	0.257
XE	0.006	-0.021		XE	-0.021	0.009
XVV1	-0.726	-0.269		XVV1	0.451	0.722
XVV3	-0.152	0.469		XVV3	0.621	0.152
XVV4	-0.256	-0.436		XVV4	-0.314	0.094
XVV5	-0.099	-0.339		XVV5	-0.236	0.104
XVV6	0.249	0.282		XVV6	0.0356	-0.246
XVV7	-0.385	0.005		XVV7	0.391	0.386
XVV8	-0.391	0.209		XVV8	0.599	0.391
AIC		1020		AIC		1024

Source: own authorship.

With the Akaike Information Criteria (AIC), it was observed in Table 4 that the A model (nominal) has a better adjustment respecting the B model (ordinal). It has a minor AIC predicting better the risk of health conditions (Yn). The results obtained from the A model are the following (See Table 5):

Table 5

Interactions of the A model

Weights: 84 (54 variable)					
Initial	Value	1300.756	Iter 50	Value	455.857
Iter 10	Value	482.855	Iter 60	Value	455.841
Iter 20	Value	458.565	Iter 70	Value	455.840
Iter 30	Value	456.551	Final	Value	455.840
Iter 40	Value	455.949			

Source: own authorship.

Table 6

Parameters run of the A model

	REC	**TRE**		**REC**	**TRE**
Intercept	3,264	-16,326	XRC8	-0,764	-1,103
XC	-4,585	15,257	XRC9	0,864	-0,242
XRC10	-0,217	0,687	XSM	-0,249	-0,281
XRC11	0,392	0.554	XE	0.006	-0.021
XRC12	-0,615	0,481	XVV1	-0,726	-0,269
XRC13	0,934	1,751	XVV3	-0,152	0,469
XRC2	-0,277	0,863	XVV4	-0,256	-0,436
XRC3	-0,597	0,687	XVV5	-0,099	-0,339
XRC4	0,031	0,951	XVV6	0,249	0,282
XRC5	-0,034	-0,209	XVV7	-0,385	0,005
XRC6	1,811	1,011	XVV8	-0,391	0,209
XRC7	0,282	0,469			

AIC	1020		Residual Deviance	911,681

Source: own authorship.

The collected information shows 70 interactions, which has a final value of 455.840 (Table 5). When this value is multiplied by two, the adjustment is suitable since it is representative of the residuals of deviance, resulting in the construction of sequenced models (See Table 6).

Table 7

Degree of significance of the parameters of the A model

P - value (Parameters)						
	REC	**TRE**			**REC**	**TRE**
Intercept	0.000	0.000		XRC8	0.196	0.353
XC	0.000	0.000		XRC9	0.118	0.795
XRC10	0.724	0.413		XSM	0.295	0.546
XRC11	0.536	0.551		XE	0.937	0.109
XRC12	0.292	0.555		XVV1	0.039	0.613
XRC13	0.198	0.036		XVV3	0.735	0.428
XRC2	0.525	0.185		XVV4	0.602	0.568
XRC3	0.193	0.313		XVV5	0.848	0.666
XRC4	0.948	0.177		XVV6	0.627	0.699
XRC5	0.959	0.866		XVV7	0.308	0.991
XRC6	0.034	0.441		XVV8	0.361	0.723
XRC7	0.683	0.647				

Source: own authorship.

Knowing that most of the independent variables are ordinal and nominal and that only age (XE) is numerical. As a result, the multi-logistical model to create is semi-parametric. It is observed in Table 7 that the variables of age (XE) and sex (XS) do not come into play in the dynamics of health conditions (Yn). For this reason, the significance level is higher to 0.05, and those variables do not impact the phenomenon under study, and because of that, they are eliminated from the model.

Table 8

Interactions of the A.1 Model

Weights: 60 (42 variables)					
Initial	Value	1300.756	Iter 50	Value	459.725
Iter 10	Value	514.160	Iter 60	Value	459.722
Iter 20	Value	466.886	Iter 60	Value	459.722
Iter 30	Value	460.151	Final	Value	459.722
Iter 40	Value	459.793			

Source: own authorship.

In the second model, there are 60 interactions with a final value of 459.722. When multiplying this value by two, we can observe that the adjustment is adequate since it is representative of the residuals of deviance of the A.1 Model (See Table 9).

Table 9

Parameter run of the A.1 Model

	REC	TRE		REC	TRE
Intercept	3.255	-17.054	XRC7	0.099	0.805
XC	-4.546	15.428	XRC8	-0.931	-0.691
XRC10	-0.354	0.881	XRC9	0.687	-0.005
XRC11	0.217	1.000	XVV1	-0.755	-0.196
XRC12	-7.193	0.824	XVV3	-0.164	0.451
XRC13	0.692	1.806	XVV4	-0.305	-0.425
XRC2	-0.534	0.727	XVV5	-0.093	-0.338
XRC3	-0.755	0.858	XVV6	0.247	0.741
XRC4	-0.119	1.215	XVV7	-0.407	-0.008
XRC5	-0.201	-0.052	XVV8	-0.427	0.147
XRC6	1.646	1.204			
AIC	1003.446		Residual Deviance		919.444

Source: own authorship.

Table 10

Significance of parameters of the A.1 Model

P - value (Parameters)						
	REC	TRE			REC	TRE
Intercept	0.000	0.000		XRC7	0.875	0.385
XC	0.000	0.000		XRC8	0.076	0.539
XRC10	0.518	0.219		XRC9	0.167	0.994
XRC11	0.705	0.207		XVV1	0.031	0.708
XRC12	0.171	0.255		XVV3	0.712	0.439
XRC13	0.301	0.013		XVV4	0.531	0.575
XRC2	0.136	0.188		XVV5	0.858	0.661
XRC3	0.058	0.163		XVV6	0.631	0.704
XRC4	0.773	0.038		XVV7	0.279	0.987
XRC5	0.739	0.946		XVV8	0.316	0.797
XRC6	0.044	0.338				

Source: own authorship.

With the implementation of Table 10 and with a significance level of 0.05, the model has a linear equation with significant variables that make the comparison of the category of recovery for death:

$$ln\left[\frac{Pr(REC)}{Pr(DEA)}\right] = 3.255 - 4.546XC + 1.646XRC6 - 0.755XVV1 \quad (16)$$

The same as in the first equation, and based on the significant variables, the second linear equation compares the category of treatment concerning death:

$$ln\left[\frac{Pr(TRE)}{Pr(DEA)}\right] = -17.054 + 15.428XC + 1.215XRC4 + 1.806XRC13 \quad (17)$$

Calculating deviance (D2):

$$D^2 = \frac{ND - Dr}{ND} = \frac{919.444 - 121.183}{919.44} = 0.8682 \quad (18)$$

Where:

ND is the null deviance (the A.1 model with a value of 919.444).

Dr are the deviation residuals(the A.1 model with a value of 121.183).

Both equations 16 and 17 have deviance (D2) of 0.8682, that means with a level of confidence of 9-95 and with a level of significance of 0.05, so equations 16 & 17 are viable to predict the likelihood of the events of health condition (Yn) of the 1,184 patients since the current model keeps 86.82% the variability of the data.

RESULTS AND DISCUSSION

Considering equations 16 & 17 the regression of both is:

Recovery

$$\left[\frac{\Pr(REC)}{\Pr(DEA)}\right] = 25.919e^{1.646RC6 - 4.546XC - 0.755XVV1} \quad (19)$$

Treatment

$$\left[\frac{\Pr(TRE)}{\Pr(DEA)}\right] = \frac{e^{15.428XC + 1.215XRC4 + 1.806XRC13}}{e^{17.054}} \quad (20)$$

In terms of the probability of each event that is part of the health condition (Yn), it can be expressed in the following way:

Likelihood of recovery

$$\Pr(REC) = \frac{\left[\frac{\Pr(REC)}{\Pr(DEA)}\right]}{1 + \left[\frac{\Pr(REC)}{\Pr(DEA)}\right] + \left[\frac{\Pr(TRE)}{\Pr(DEA)}\right]} \quad (21)$$

Plausibility of the treatment

$$\Pr(TRE) = \frac{\left[\frac{\Pr(TRE)}{\Pr(DEA)}\right]}{1 + \left[\frac{\Pr(REC)}{\Pr(DEA)}\right] + \left[\frac{\Pr(TRE)}{\Pr(DEA)}\right]} \quad (22)$$

$$1 = \Pr(TRE) + \Pr(REC) + \Pr(DEA)$$

Probability of death (23)

Such that:

$$\Pr(DEA) = 1 - [\Pr(TRE) + \Pr(DEA)]$$

Starting from the event of recovery of a patient, and using the equations 21,22 & 23, we can obtain the following:

Table 11

Scenario Probabilities

	I	II	III	IV	V	VI	VII	VIII	IX
Pr(REC)	0.8373	0.1868	0.0970	0.5436	0.3588	0.1418	0.1001	0.0719	0.0550
Pr(TRE)	0.1303	0.1338	0.1486	0.0751	0.0950	0.3420	0.4357	0.3699	0.5162
Pr(DEA)	0.0324	0.6794	0.7544	0.3813	0.5462	0.5162	0.4642	0.5582	0.4288

Source: own authorship.

Where:

I: Any subdivision of cancer without having COVID-19.

II: Any subdivision of cancer with having COVID-19.

III: Any subdivision of cancer with COVID-19 and vaccine (AstraZeneca).

IV: Liver cancer with COVID-19 and without getting vaccinated.

V: Liver cancer with COVID-19 and get vaccinated (AstraZeneca).

VI: With COVID-19, lung cancer without vaccine (AstraZeneca)in recovered patients.

VII: With COVID-19, thyroid cancer without vaccine (AstraZeneca) in recovered patients.

VIII: With COVID-19, lung cancer with the vaccine (AstraZeneca) in recovered patients.

IX: With COVID-19, thyroid cancer with the vaccine (AstraZeneca)in recovered patients.

In Table 11, we can observe that the plausibility of recovery of a patient who suffers from any subdivision of cancer and does not get infected by COVID-19 is 0.8373 (I scenery). Notwithstanding, if the patient develops COVID-19, the probability of recovery is 0.1868 (II scenery), which means the risk of dying [RDEA]will increase more than 20 times.

$$R_{DEA} = \frac{Pr(DEA)_{II}}{Pr(DEA)_{I}} = \frac{0.6794}{0.0324} = 20.969 \quad (24)$$

If a patient has any subdivision of cancer and develops COVID-19, he receives the Astrazeneca vaccination (III scenery), and the risk [RDEA] of dying will rise 1.11 times higher than scenery I.

$$R_{DEA} = 1 - \frac{Pr(DEA)_{III}}{Pr(DEA)_{II}} = \frac{0.7544}{0.6794} = 1.110 \quad (25)$$

Under a significance level of 0.05, the category XVV1 (AstraZeneca vaccine) of the variable XV affects the event of recovery (REC) of the patient. In other words, the significance level of each category of XV is found above 0.05. Therefore, it does not come into play with the events of the health condition of the patients(Yn), except for the application of the AstraZeneca vaccine.

AstraZeneca vaccine (XVV1) is the most significant about the event of recovery (REC) of the variable (Yn) and having a negative slope, it means the decrease of probability of recovery (REC) and the increase of plausibility of the event of death(DEA).In other words, applying the AstraZeneca vaccine to a patient who has any subdivision of cancer and who developed COVID-19, his likelihood of dying rises to 11%.

Whether a patient has liver cancer, and is positive to COVID-19 ,and does not get vaccinated by Astasenca,(IVscenary) has a likeliness of recovery of 0.5436, , being superior of any other subdivision of cancer, for this reason the risk of dying [RDEA] declines to 43.87%.

$$\boldsymbol{R_{DEA} = 1 - \frac{Pr(DEA)_{IV}}{Pr(DEA)_{II}} = 1 - \frac{0.3813}{0.6794} = 0.4387} \quad (26)$$

Nevertheless, if a patient gets vaccinated with AstraZeneca (V scenery). The risk of death [RDEA]will rise to 43.46%. As a consequence, recovery and treatment will be 0.3588 & 0.0950. Even though a patient with liver cancer and COVID-19, once the AstraZeneca vaccine is applied, his health condition will be reverted.

$$\boldsymbol{R_{DEA} = \frac{Pr(DEA)_{V}}{Pr(DEA)_{IV}} = \frac{0.5462}{0.3813} = 1.4325} \quad (27)$$

Whether a patient has lung or thyroid cancer and is positive for COVID-19(V1& VII sceneries), the likelihood of recovery will be minor to (0.1418 & 0.1001) or to any other subdivision of cancer. Because of this, it is recommended that they have a treatment (0.3420 & 0.4357). Hence, by doing so, it reduces the risk of death [RDEA] by 24.11% (lung) and 31.67%(thyroid), concerning any other subdivision of cancer.

Lung cancer – covid - 19

$$R_{DEA} = 1 - \frac{Pr(DEA)_{VI}}{Pr(DEA)_{II}} = 1 - \frac{0.5156}{0.6794} = 0.2411 \quad (28)$$

Thyroid cancer – covid - 19

$$R_{DEA} = 1 - \frac{Pr(DEA)_{VII}}{Pr(DEA)_{II}} = 1 - \frac{0.4642}{0.6794} = 0.3167 \quad (29)$$

Nonetheless, if a patient has lung or thyroid cancer, and he is convinced of developing COVID-19 and gets vaccinated with AstraZeneca (VIII & IX sceneries), his likelihood of recovery will relapse to 0.0719 & en 0.0550, having as a result the following effects:

The risk of death [RDEA] caused by lung cancer and COVID-19 will increase to 8.26%.

The risk of death [RDEA] caused by thyroid cancer and COVID-19 will decrease to 7.63% due to the increased probability of having a treatment (0.5162).

Lung cancer – covid -19 (astrazeneca) vaccine

$$R_{DEA} = \frac{Pr(DEA)_{VIII}}{Pr(DEA)_{VI}} = \frac{0.5582}{0.5156} = 1.0826 \quad (30)$$

Thyroid cancer – covid - 19 (astrazeneca) vaccine

$$R_{DEA} = 1 - \frac{Pr(DEA)_{VII}}{Pr(DEA)_{VII}} = 1 - \frac{0.4288}{0.4642} = 0.0763 \quad (31)$$

The development and construction of the current mathematical model provide elements to identify the factors that lead patients with cancer to develop COVID-19. Moreover, the possible effects that they obtain when vaccinated with AstraZeneca, with this mathematical model, adequate and irrefutable decisions could be considered to minimize the risk of death of the patients.

CONCLUSIONS

To conclude, there are two contexts that this research work considers. Firstly, the importance of mathematical models.Secondly, the environment that exists among cancer, COVID-19 and vaccines.

Mathematical models seek the simplifying representation of reality that a phenomenon occurs since thorough equations and sceneries of objects under study can be practically manipulated. Within public health, the behavior or the dynamics of diseases are predicted, and the objective is to direct appropriately the public policies of prevention. For this reason, it is essential to arrange the scenery with the evaluation of possible risks that the population faces.

The environment that exists among cancer, COVID-19, and vaccines from this present mathematical model, we can summarize the following aspects:

Whether a cancer patient develops COVID-19, his probability of recovery will diminish. Therefore, the risk of death will expand 20 times for the person who does not get infected. Furthermore, the risk will intensify to 11% due to the patient getting vaccinated with AstraZeneca.

If the patient has liver cancer and develops COVID-19, his risk of death will be 43.87% smaller than any other subdivision of cancer. However, his risk of death will magnify if he gets vaccinated with AstraZeneca.

The 13 subdivisions of cancer, lung, and thyroid can be contemplated as the most aggressive ones for patients who develop COVID-19. It is strongly recommended to administer any treatment that allows patients to minimize the risk of death, implementing the present mathematical model.

Whether a patient with lung or thyroid cancer is positive for COVID-19 and gets vaccinated with AstraZeneca, the risk of death in the first environment(lung cancer)will enlarge to 8.26%.In consequence, the second environment, thyroid cancer, will lessen the risk of death to 7.63%.

In general terms, if there is a linear relationship between cancer and COVID-19, the risk of dying of the patients will rise. This behavior will extend even more when the AstraZeneca vaccine is applied. That is why the rest vaccine doses do not provoke any reaction in the metabolism of patients with cancer and COVID-19. On the other side, liver cancer can be considered “the less aggressive one”.When a patient has lung or thyroid cancer and develops COVID-19, it is fundamental to be treated to reduce the risk of dying. Notwithstanding, if the AstraZeneca vaccine is implemented, the risk of death from lung cancer will intensify, on the contrary, thyroid cancer will be reduced.

REFERENCES

Bocco, M. and Herrero, V. (2009). "Multilogistic model to identify the determinants of joint labor participation modalities in Argentina." Argentina. Population Studies Association of Argentina.

Dekking, F., Lopuhaä, H., Kraaikamp, C., and Meester, L. (2005). "A Modern Introduction to Probability and Statistics". United States. Publisher: Springer.

Evans, M. and Rosenthal, J. (2010). "Probability and Statistics". United States. Publisher: University of Toronto.

Health Secretariat. (2023). "Informe Integral de Covid - 19 en México". Available online: https://epidemiologia.salud.gob.mx/gobmx/salud/documentos/covid19/Info-04-23-Int_COVID-19.pdf.

Hillier, F., and Lieberman, G. (2010). "Introduction to Operations Research". United States. Publisher: Mc Graw Hill.

Martínez, M. (2021). "Current status of breast cancer in Mexico: main types and risk factors". Mexico. Mexican journal of oncology.

Montoya, Y. and Correa, J. (2017). "Elicitation of multinomial distribution from several experts". Colombia. Communications in Statistics.

McCullagh, P. and Nelder, J. (1983). "Generalized Linear Models". United States. Chapman and Hall.

National Institute of Transparency, Access to Information and Protection of Personal Data (INAP). (2022). "Linea del Tiempo COVID - 19 en México". Available online: https://micrositios.inai.org.mx/conferenciascovid-19tp/?page_id=8432.

National Cancer Institute. (2022). Common types of cancer. Available online: https://www.cancer.gov/espanol/tipos/comunes.

Pando, V. and Martín, R. (2004). "Multinomial Logistic Regression". Spain. University of Valladolid.

World Health Organization. (2020). More information on the COVID - 19 pandemic. Available online: https://www.who.int/es/health-topics/coronavirus#tab=tab_1.

ACKNOWLEDGMENTS

The present research work is the result of a series of projects implemented within the Territorial Analysis, Environment, and Data Science Laboratory From the Institute of Social Sciences and Humanities (ICSHU) and from the Autonomous University of the Hidalgo State (UAEH), which resulting from a research project financed by National Council for the Humanities, Science, and Technology (CONACYT). The projects are the following:

The climate change agenda of the Pachuca municipal area.

Municipal intervention plans in Hidalgo in the face of climate change.

Estimation of water supply in the urban area of Pachuca.

Crime in Mexico and Social Cohesion.

Logistic Model to Predict the Contagion of COVID-19 in Mexico.

Estimation of the Risk of Death from Covid in a Cancer Population.

This laboratory has been the link between the Autonomous University of Hidalgo State and the public and private sectors since their research works have been relevant to decision-making.

ANNEXES

```
#####---------MULTI-LOGISTIC MODEL ---------#####

##--Libraries

library(tidyverse)

library(caret)

library(nnet)

library(ggplot2)

library(reshape2)

library(lessR)

##--Description of variables

#- Y = HEALTH CONDITION (REC = recovery, TRE = treatment, DEA = death)

#- XC = COVID (0 = negative & 1 = positive)

#- XS = SEX (W = woman  & M = man)

#- XE = AGE (years)

#- XT = TOBACCO (NO & YES)

#- XA = ALCOHOL (NO & YES)

#- XR = CANCER (C1 = PROSTATE, C2 = BREAST, C3 = LUNG, C4 = OTHERS)
```

```
#- XV = VACCINES (V1 = ASTRAZENECA, V2 = CANSINO, V3 = PFIZAR, V4 = GAMBELAYA, V5 = SINOVAC, V6 = SINOPHARM, V7 = JANSSEN, V8 = MODERNA, NO)

##--Descriptive Statistics

basis

##==Description of the information==##

par(mfrow=c(1,2))


###################

barplot(table(base$Y), main = "HEALTH CONDITION", &lab="Patients",

      col= c("blue","red","purple"))

percentages <- as.numeric(round(((prop.table(table(basis$XC)))*100),2))

percentages 1

labels1 <- c("Negative", "Positive")

labels1

labels1 <- paste(labels1, percentages1)

labels1

labels1 <- paste(labels1, "%", sep = "")

labels1

pie(percentages1, labels1,main = "COVID-19",col=c("green","red"))

###################


###################

percentages2 <- as.numeric(round(((prop.table(table(basis$XS)))*100),2))

percentages2

labels2 <- c("Men", "Women")

labels2

labels2 <- paste(labels2, percentages2)

labels2

labels2 <- paste(labels2, "%", sep = "")

labels2
```

```
pie(percentages2, labels2,main = "SEX",col=c("blue","green"))
barplot(table(base$XE),main = "AGE", col= c("red"), xlab="Years", &lab ="Patients")
##################


#################
par(mfrow=c(1,1))
PieChart(XN, data=basis,main ="PNEUMONIA",color = "black",lwd = 2, lty = 1)
PieChart(XO, data=basis, main ="OBESITY",color = "red",lwd = 2, lty = 1)
################


###############
par(mfrow=c(1,2))
percentages3 <- as.numeric(round(((prop.table(table(basis$XT)))*100),2))
labels3 <- c("Positive", "Negative")
labelss3 <- paste(labels3, percentages3)
labels3 <- paste(labels3, "%", sep = "")
pie(percentages3, labels3, main = "TOBACCO",col=c("red","green"))
percentages4 <- as.numeric(round(((prop.table(table(basis$XA)))*100),2))
labels4 <- c("Negative", "Positive")
labels4 <- paste(labels4, percentages4)
labels4 <- paste(labels4, "%", sep = "")
pie(percentages4, labels4, main = "ALCOHOL",col=c("red","green"))
#################


#################
par(mfrow=c(1,2))
barplot(table(basis$XR), main = "CANCER", horiz = TRUE, xlab="Patients", col= c("blue"))
barplot(table(basis$XV), main = "VACCINES",horiz = TRUE, xlab="Patients", col= c("purple"))
```

```
#################################################################
##---MODEL RUN
# ---- Prob(Y) = f(XC,XS,XE,XN,XO,XT,XA,XR,XV)
1model = multinom(Y~XC+XR+XS+XE+XV+XN+XO+XT+XA, data = basis)
summary(1model)
A.1model= multinom(Y1~XC+XR+XS+XE+XV+XN+XO+XT+XA, data = basis)
summary(A.1model)
###--Starting from the1model testA,the predictions will be the following
predicted.classes = model1%>% predict(basis)
head(predicted.class1)
# (Degree of adjustment)
mean(predicted.classes1 == basis$Y)
#Significance of the parameters
z1 = summary(1model)$coefficients/summary(1model)$standard.errors
# P-Value(Significance of the parameters)
p1 = (1 - pnorm(abs(z1), 0, 1)) * 2
p1
#####---Second model (eliminating XT)
#####---Fifth model (eliminating XS)
5model= multinom(Y~XC+XR+XV, data = basis)
summary(5model)
###--Starting from 1model testA, the predictions will be the following
predicted.classes5 = model5%>% predict(basis)
head(predicted.classes5)
# (Degree of adjustment)
mean(predicted.classes5 == basis$Y)
#Significance of the parameters
z5 = summary(5model)$coefficients/summary(5model)$standard.errors
# P-Valor (Significance of the parameters)
```

```
p5 = (1 - pnorm(abs(z5), 0, 1)) * 2

p5

# Predictions

pp = round(fitted(5model),3)

pp

par(mfrow=c(1,1))

plot(pp)
```