Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi
Sayı 7, S. 400-407, 4, 2023

Araştırma Makalesi

IJANSER
https://as-proceeding.com/index.php/ijanser
ISSN: 2980-0811

International Journal of Advanced Natural Sciences and Engineering Researches
Volume 7, pp. 400-407, 4, 2023

Research Article

Modeling COVID-19 Binary Data in the Aspect of Neoplasms as a Potential Indicator of Cancer by Logit and Probit Regression Models

Neslihan İyit*, Esra Sarı [2] and Ferhat Sevim [3]

[1] *Selcuk University, Statistics Department, Konya-TÜRKİYE*
[2] *Selcuk University, Statistics Department, Konya-TÜRKİYE*
[3] *Selcuk University, Statistics Department, Konya-TÜRKİYE*

(niyit@selcuk.edu.tr) Email of the corresponding author

(Received: 10 May 2023, Accepted: 26 May 2023)

(DOI: 10.59287/ijanser.754)

(1st International Conference on Contemporary Academic Research ICCAR 2023, May 17-19, 2023)

ATIF/REFERENCE: İyit, N., Sarı, E. & Sevim, F. (2023). Modeling COVID-19 Binary Data in the Aspect of Neoplasms as a Potential Indicator of Cancer by Logit and Probit Regression Models**.** *International Journal of Advanced Natural Sciences and Engineering Researches*, 7(4), 400- 407.

Abstract **–** In this study, the effects of disability-adjusted life years (DALYs) from neoplasms and concomitant non-communicable diseases (NCDs) on total deaths from the COVID-19 pandemic until 21 July 2021 are examined globally for 179 countries. For this purpose, the explanatory variables are taken as DALYs as a measure of total burden of diseases in life lost years and lived with a disability years from neoplasm and NCDs. In this study, the total number of deaths caused by the COVID-19 pandemic has been made categorical with the help of the indicator variable and then taken as the response variable. Thus, in this study, the effects of neoplasms and concomitant NCDs on the COVID-19 pandemic are investigated by using binary logit and binary probit regression models in the family of generalized linear models (GLMs) as statistical methods. Specific to this study, the superiority of the probit model which is based on the assumption that the errors have a normal distribution in the statistical sense over the logit model which is based on the assumption that the errors have a logistic distribution is emphasized. As principle results and major conclusion from this study, neoplasms, cirrhosis and other chronic liver diseases, cardiovascular diseases, skin and subcutaneous diseases and other non-communicable diseases have been found to have statistically significant effects on deaths due to the COVID-19 pandemic.

Keywords – COVID-19 Pandemic, Non-Communicable Diseases, Generalized Linear Model, Probit Model, Logit Model.

I. INTRODUCTION

The COVID-19 pandemic, officially declared on 11 March 2020, is an epidemic of disease affecting large numbers of people and occurring on a transnational scale [1]-[5]. Disability-adjusted life years (DALYs) is a metric measuring "burden of disease" which is the sum of morbility and mortality. One DALY is considered a loss of a healthy year [6],[7]. With DALYs, the burden of disease caused by premature deaths due to various diseases and injuries, and disease states that do not result in death but cause long-term disability and loss of function, are summarized with a single criterion [8]. Non-communicable diseases (NCDs) are basically the general expression of diseases such as cancers, chronic liver diseases, cardiovascular

diseases, diabetes, and others which are not caused by an acute infection and can lead to health problems taking a long time to treat [9]-[11].

In this study, the effects of DALYs from neoplasms and concomitant NCDs on the COVID-19 pandemic are investigated. In this context, Arsang-Jang et al. [12] examined the statistical relationship between COVID-19 mortality rate and total burden of NCDs by using a multilevel generalized linear model according to the income status of the countries. Azarpazhooh [13] investigated the relationships between DALYs from NCDs and COVID-19 pandemic data as cases and deaths using correlation tests. Azadnajafabad et al. [14] investigated the rate of various NCDs among deaths caused by COVID-19 pandemic using some clinical data in Iran. Sousa et al. [15] examined risk factors for death with various NCDs and other indicators in hospitalized children with COVID-19 in Brazil. Youn et al. [16] examined the spreading impact of the COVID-19 pandemic on people with some of the major NCDs. Nicoletti-Rojas et al. [17] modeled survey data with participants with NCDs during the COVID-19 pandemic in Chile with beta regression. Gaur et al. [18] investigated association between cases and deaths from the COVID-19 pandemic and various NCDs using multilevel regression in India. Pécout et al. [19] compared descriptive statistics according to various criteria on a given number of participants with NCDs in two waves during the COVID-19 pandemic in the USA and Europe. Further, there are more studies in literature for investigating the effects of neoplasms and concomitant NCDs on COVID-19 pandemic data [20]-[25].

In this study, the effects of DALYs from neoplasms and concomitant NCDs on the COVID-19 pandemic data by using binary logit and binary probit regression models in the family of GLMs.

II. Materials and Method

In this study, the response variable is taken as binary variable by categorizing the "total deaths attributed to COVID-19 pandemic" per 1.000.000 people until 21 July 2021 according to median value [26]. As it can be seen in Figure 1, when the daily total number of deaths for 179 countries is examined from 03.01.2020 to 16.03.2023, the date 21.07.2021 is determined at most number of deaths for these countries with 19.991.

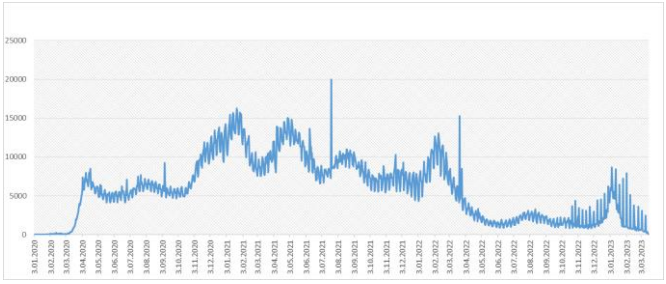

Fig. 1 Daily number of deaths attributed to COVID-19 pandemic for 179 countries

The explanatory variables are taken as DALYs per 100.000 individuals from "cirrhosis and other chronic liver diseases", "cardiovascular diseases", "skin and subcutaneous diseases", "neoplasms" and "other non-communicable diseases" belonging to 2019, respectively [27]. All of explanatory variables are transformed to the categorical variables, and the four possible levels for the these variables are taken as the first 25% percentage coded as 1, the second 25% percentage (median) coded as 2, the third 25% percentage coded as 3, the fourth 25% percentage (max) coded as 4.

Data from 179 world countries are taken as observations for this study. Descriptive statistics of all variables in this study are given in Table 1.

Table 1. Descriptive statistics of all variables given in this study

Variables	Min.	Median	Mean ± Sd	Max
Total Deaths attributed to the COVID-19 Pandemic	0.321	315.475	710.119 ± 880.738	5734.078
Disease burden from cirrhosis and other chronic liver diseases	110.510	546.124	596.279 ± 360.521	2616.745
Disease burden from cardiovascular diseases	1367.302	4166.910	4968.395 ± 3060.903	18936.288
Disease burden from skin and subcutaneous diseases	298.863	511.050	545.937 ± 128.148	1042.712
Disease burden from neoplasms	977.179	2718.874	3395.620 ± 1927.765	11701.152
Disease burden from other non-communicable diseases	1184.196	1911.310	2176.991 ± 907.094	7650.528

Generalized linear models (GLMs) are the expansion of the assumption that the response variable has a normal distribution in linear regression models with an exponential distribution family;

$$f\left(y_i;\theta,\phi\right)=\exp\left\{\frac{y_i\theta_i-b\left(\theta_i\right)}{a_i\left(\phi\right)}+c\left(y_i;\phi\right)\right\} \quad (1)$$

where $a(\square)$, $b(\square)$ and $c(\square)$ are known functions specific to the family of exponential distributions, θ and ϕ are the location and dispersion parameters, respectively [28]-[33]. For the binomial distribution, a member of the exponential distribution family, the functions of $a(\square)$, $b(\square)$ and $c(\square)$ are given in Table 2 [34]-[36].

Table 2. Known functions belonging to binomial distribution in Eq. (1).

$a(\phi)$	$b(\theta)$	$c(y;\phi)$
1	$n\log\left(1+e^{\theta}\right)$	$\log\binom{n}{y}$

In GLMs, the model is constructed by transforming the mean of the response variable to linear estimators $\eta_i=\beta_0+\sum_{j=1}^{p-1}x_{ij}\beta_j$ with a link function $g(\square)$. The logit and probit link function are given as:

$$g\left(\pi_i\right)=\text{logit}\left(\pi_i\right)=\log\frac{\pi_i}{1-\pi_i} \quad (2)$$

$$\eta_i=g\left(\pi_i\right)=\Phi^{-1}\left(\pi_i\right) \quad (3)$$

where $0\le\pi_i\le1$, $-\infty<\text{logit}\left(\pi_i\right)<\infty$ and $\Phi^{-1}(\square)$ is inverse standard normal cumulate distribution function [37]-[38]. So, binary logistic and binary probit model are given as following;

$$\eta_i=\log\frac{\pi_i}{1-\pi_i}=\beta_0+\sum_{j=1}^{p-1}x_{ij}\beta_j \quad (4)$$

$$\eta_i=\Phi^{-1}\left(\pi_i\right)=\beta_0+\sum_{j=1}^{p-1}x_{ij}\beta_j \quad (5)$$

respectively, where $i=1,2,...,n$ is size of observation, β_j is parameter estimation, x_{ij} is ij^{th} unit of explanatory variables, p are number of model parameters [39]-[40].

In this study, iteratively reweighted least squares (IRLS) method is used for fitting the binary logit and binary probit models [41].

Several goodness of test statistics for comparing among GLMs are Akaike's information criterion (AIC), Akaike's information corrected criterion (AICc), Bayesian information criterion (BIC) and consistent Akaike's information criterion (CAIC) given in Table 3 where L is log-likelihood statistics, L_{full} is log-likelihood for observed response and N is number of observation in the model [42]-[45].

Table 3. Various information criteria used in this study

Information Criteria	**Formula**
Deviance	$-2\left\{L_{full}-L\right\}$
AIC	$-2L+2p$
AICc	$-2L+\frac{2pN}{N-p-1}$
BIC	$-2L+p\ln\left(N\right)$
CAIC	$-2L+p\left(\ln\left(N\right)+1\right)$

III. RESULTS AND DISCUSSION

In this study, the effects of disability-adjusted life years (DALYs) from neoplasms and concomitant non-communicable diseases (NCDs) on total deaths from the COVID-19 pandemic are examined globally for 179 countries by using binary logit and binary probit regression models in the family of generalized linear models (GLMs) as statistical methods under IRLS method.

In accordance with this purpose, the response variable is categorically taken as "total deaths attributed to COVID-19 pandemic" per 1.000.000 until 21 July 2021 according to the median value of the related variable, and the explanatory variables are categorically taken as DALYs per 100.000 individuals from "cirrhosis and other chronic liver diseases", "cardiovascular diseases", "skin and subcutaneous diseases", "neoplasms" and "other non-communicable diseases" belonging to 2019

according to four possible levels of the related variables, respectively.

In this study, Rstudio statistical computing is used for all statistical modelling and inference [46].

Firstly, the result of binary logit model by using IRLS method is given in Table 4.

Table 4. The results of the binay logit regression model by using IRLS method

Explanatory Variables	$\hat{\beta}$	$s.e(\hat{\beta})$	**Wald Chi-Square**	$p-$ **value**	$\exp(\hat{\beta})$	95% **Wald C.I. for** $\exp(\hat{\beta})$	
						Lower	**Upper**
Intercept	2.6680	1.12439	5.6304	0.01765 *	14.4113	1.5907	130.5576
Cirrhosis and other chronic liver diseases [2nd level]	0.1002	0.89283	0.0126	0.91066	1.1054	0.1921	6.3606
Cirrhosis and other chronic liver diseases [3rd level]	2.1848	0.91701	5.6763	0.01720 *	8.8887	1.4732	53,6306
Cirrhosis and other chronic liver diseases [4th level]	0.0405	0.82506	0.0024	0.96086	1.0413	0.2067	5.2467
Cardiovascular diseases [2nd level]	2.6984	1.05428	6.5508	0.01048 *	14.8558	1.8814	117.3052
Cardiovascular diseases [3rd level]	1.8255	0.95361	3.6644	0.05559	6.2056	0.9573	40.2269
Cardiovascular diseases [4th level]	2.8170	1.31588	4.5829	0.03229 *	16.7268	1.2686	220,5523
Skin and subcutaneous diseases [2nd level]	2.7979	1.11037	6.3493	0.01174 *	16.4101	1.8619	144,6367
Skin and subcutaneous diseases [3rd level]	3.8879	1.15846	11.2631	0.00079 ***	48.8058	5.0393	472,6861
Skin and subcutaneous diseases [4th level]	3.5670	1.18812	9.0131	0.00268 **	35.4088	3.4496	363,4622
Neoplasms [2nd level]	3.7697	0.95914	15.4474	0.00008 ***	43.3684	6.6181	284,1904
Neoplasms [3rd level]	3.1494	1.03589	9.2431	0.00236 **	23.3211	3.0618	177,6303
Neoplasms [4th level]	22.0918	1347.79817	0.0003	0.98692	3929743525.4568	0.0000	Inf
Other non-communicable diseases [2nd level]	2.6862	0.96285	7.7835	0,00527 **	14.6764	2.2234	96,8754
Other non-communicable diseases [3rd level]	2.4436	0.89874	7.3926	0.00655 **	11.5145	1.9780	67,0300
Other non-communicable diseases [4th level]	0.0781	0.72107	0.0117	0.91372	1.0813	0.2631	4,4434

All parameter estimates obtained using the IRLS method are given in Table 4 and the model equation for the binary logit regression model is given in Eq. (6);

$$\eta_i = \log\left(\frac{\pi}{1-\pi}\right) = \begin{Bmatrix} 2.6680 + 2.1848\left(Cirrhosis[3]\right) + \\ 2.6984\left(Cardio[2]\right) + 2.8170\left(Cardio[4]\right) + \\ 2.7979\left(Skin[2]\right) + 3.8879\left(Skin[3]\right) + \\ 3.5670\left(Skin[4]\right) + 3.7697\left(Neoplasms[2]\right) + \\ 3.1494\left(Neoplasms[3]\right) + \\ 2.6862\left(Other[2]\right) + 2.4436\left(Other[3]\right) \end{Bmatrix} \quad (6)$$

Secondly, the result of binary probit model by using IRLS method is given in Table 5.

Table 5. The results of the binay probit regression model by using IRLS method

Explanatory Variables	$\hat{\beta}$	$s.e\left(\hat{\beta}\right)$	**Wald Chi-Square**	$p-$ **value**	$\exp\left(\hat{\beta}\right)$	95% **Wald C.I. for** $\exp\left(\hat{\beta}\right)$	
						Lower	**Upper**
Intercept	1.4943	0.61087	5.9837	0.01444 *	4.4562	1.3458	14.7552
Cirrhosis and other chronic liver diseases [2nd level]	0.0054	0.50250	0.0001	0.99137	1.0054	0.3755	2.6921
Cirrhosis and other chronic liver diseases [3rd level]	1.1897	0.50300	5.5942	0.01802 *	3.2861	1.2261	8.8072
Cirrhosis and other chronic liver diseases [4th level]	0.0685	0.47949	0.0204	0.88634	1.0709	0.4184	2.7410
Cardiovascular diseases [2nd level]	1.4649	0.56449	6.7344	0.00946 **	4.3271	1.4312	13.0826
Cardiovascular diseases [3rd level]	0.9615	0.51903	3.4319	0.06395	2.6157	0.9457	7.2342
Cardiovascular diseases [4th level]	1.6139	0.71629	5.0768	0.02425 *	5.0225	1.2337	20.4475
Skin and subcutaneous diseases [2nd level]	1.6776	0.61043	7.5525	0.00599 **	5.3526	1.6179	17.7080
Skin and subcutaneous diseases [3rd level]	2.2692	0.63705	12.6883	0.00037 ***	9.6716	2.7748	33.7103
Skin and subcutaneous diseases [4th level]	2.0471	0.65625	9.7308	0.00181 **	7.7457	2.1402	28.0333
Neoplasms [2nd level]	2.1224	0.52453	16.3720	5.2e-05 ***	8.3508	2.9871	23.3463
Neoplasms [3rd level]	1.7429	0.56852	9.3982	0.00217 **	5.7138	1.8750	17.4124
Neoplasms [4th level]	8.2294	282.49751	0.0008	0.97676	3749.6396	0.0000	Inf
Other non-communicable diseases [2nd level]	1.5779	0.51575	9.3604	0.00222 **	4.8448	1.7631	13.3135
Other non-communicable diseases [3rd level]	1.4391	0.49531	8.4419	0.00367 **	4.2170	1.5973	11.1334
Other non-communicable diseases [4th level]	0.0482	0.42344	0.0129	0.90942	1.0494	0.4576	2.4064

All parameter estimates obtained using the IRLS method are given in Table 4 and the model equation for the binary probit regression model is given in Eq. (7);

$$\eta_i = \Phi(\pi_i) = \left\{ \begin{array}{l} 1.4943 + 1.1897\left(Cirrhosis[3]\right) + 1.4649\left(Cardio[2]\right) + \\ 1.6139\left(Cardio[4]\right) + 1.6776\left(Skin[2]\right) + \\ 2.2692\left(Skin[3]\right) + 2.0471\left(Skin[4]\right) + \\ 2.1224\left(Neoplasms[2]\right) + 1.7429\left(Neoplasms[3]\right) + \\ 1.5779\left(Other[2]\right) + 1.4391\left(Other[3]\right) \end{array} \right\} \quad (7)$$

Goodness of test statistics for the binary logit and binary probit models are given in Table 6.

Table 6. Goodness-of-fit test statistics for the binay logit and binary probit regression models

Goodness-of-fit test statistics	Binary Logit Regression Model	Binary Probit Regression Model
Log-likelihood	-46.390	-46.099 *
Df.	16	16
Deviance	92.780	92.198 *
AIC	124.780	124.198 *
AICc	128.138	127.556 *
BIC	175.778	175.196 *
CAIC	191.778	191.196 *

IV. CONCLUSION

In this study, as a result of performances due to the binary logit and binary probit regression models, the best performancing model is determined the binary probit model according to log-likelihood, deviance statistic, AIC, AICc, BIC and CAIC with -46.099, 92.198, 124.198, 127.556, 175.196 and 191.196, respectively as can be seen Table 6.

Specific to this study, the superiority of the probit model which is based on the assumption that the errors have a normal distribution in the statistical sense over the logit model which is based on the assumption that the errors have a logistic distribution is emphasized.

According to the binary probit regression model given in Eq. (7);

It has been determined that being at 3rd level (546.124-717.272) of DALYs from cirrhosis and other chronic liver diseases when taking reference category at 1st level (110.510-337.989) of the related explanatory variable increase 1.1897 times the probability of being greater total deaths attributed to COVID-19 pandemic than the median value with 1206.606.

It has been determined that being at 2nd level (2917.532-4166.910) and 4th level (5614.920-18936.288) of DALYs from cardiovascular diseases when taking reference category at 1st level (1367.302-2917.532) of the related explanatory variable increase 1.4649 and 1.6139 times the probability of being greater total deaths attributed to COVID-19 pandemic than the median value with 1206.606, respectively.

It has been determined that being at 2nd level (444.101-511.050), 3rd level (511.050-640.589) and 4th level (640.589-1042.712) of DALYs from skin and subcutaneous diseases when taking reference category at 1st level (298.863-444.101) of the related explanatory variable increase 1.6776, 2.2692 and 2.0471 times the probability of being greater total deaths attributed to COVID-19 pandemic than the median value with 1206.606, respectively.

It has been determined that being at 2nd level (1872.697-2718.874) and 3rd level (2718.874-4842.845) of DALYs from neoplasm when taking reference category at 1st level (977.179-1872.697) of the related explanatory variable increase 2.1224 and 1.7429 times the probability of being greater total deaths attributed to COVID-19 pandemic than the median value with 1206.606, respectively.

It has been determined that being at 2nd level (1676.529-1911.310) and 3rd level (1911.310-2264.397) of DALYs from other non-communicable diseases when taking reference category at 1st level (1184.196-1676.529) of the related explanatory variable increase 1.5779 and 1.4391 times the probability of being greater total deaths attributed to COVID-19 pandemic than the median value with 1206.606, respectively.

ACKNOWLEDGMENT

This study is a part of Esra Sarı's M.Sc. Thesis entitled "Generalized Linear Models for Binomially Distributed Response Variable: An Application on Cancer and COVID-19 Pandemic Data" supervised by Assoc.Prof.Dr. Neslihan Iyit submitted by Statistics Department, Graduate School of Natural Sciences, Selcuk University, Konya, Türkiye.

REFERENCES

[1] Ciotti, M., Ciccozzi, M., Terrinoni, A., Jiang, W. C., Wang, C. B., & Bernardini, S. (2020). The COVID-19 pandemic. Critical reviews in clinical laboratory sciences, 57(6), 365-388.

[2] Lone, S. A., & Ahmad, A. (2020). COVID-19 pandemic–an African perspective. Emerging microbes & infections, 9(1), 1300-1308.

[3] Tuttle, K. R. (2020). Impact of the COVID-19 pandemic on clinical research. Nature Reviews Nephrology, 16(10), 562-564.
[4] Suryasa, I. W., Rodríguez-Gámez, M., & Koldoris, T. (2021). The COVID-19 pandemic. International Journal of Health Sciences, 5(2).
[5] Maeda, J. M., & Nkengasong, J. N. (2021). The puzzle of the COVID-19 pandemic in Africa. Science, 371(6524), 27-28.
[6] Anand, S., & Hanson, K. (1997). Disability-adjusted life years: a critical review. Journal of health economics, 16(6), 685-702.
[7] (2023) World Health Organization (WHO) website. [Online]. Available: https://www.who.int/data/gho/indicator-metadata-registry/imr-details/158
[8] Akman, M., & Civek, S. (2022). Dünyada ve Türkiye'de kardiyovasküler hastalıkların sıklığı ve riskin değerlendirilmesi. The Journal of Turkish Family Physician, 13(1), 21-28.
[9] Horton, R. (2013). Non-communicable diseases: 2015 to 2025. The Lancet, 381(9866), 509-510.
[10] Wagner, K. H., & Brath, H. (2012). A global view on the development of non communicable diseases. Preventive medicine, 54, S38-S41.
[11] Islam, S. M. S., Purnat, T. D., Phuong, N. T. A., Mwingira, U., Schacht, K., & Fröschl, G. (2014). Non-Communicable Diseases (NCDs) in developing countries: a symposium report. Globalization and health, 10(1), 1-8.
[12] Arsang-Jang, S., Belasi, M. T., Najafi, F., Darbandi, M., Raza, M. Z., Akhuanzada, H., ... & Azarpazhooh, M. R. (2021). The Association Between Healthcare Resources, Non-communicable Diseases, and Covid-19 Mortality: An Epidemiological Study of 139 Countries.
[13] Azarpazhooh, M. R., Morovatdar, N., Avan, A., Phan, T. G., Divani, A. A., Yassi, N., ... & Di Napoli, M. (2020). COVID-19 pandemic and burden of non-communicable diseases: an ecological study on data of 185 countries. Journal of Stroke and Cerebrovascular Diseases, 29(9), 105089.
[14] Azadnajafabad, S., Ghasemi, E., Moghaddam, S. S., Rezaei, N., & Farzadfar, F. (2021). Non-communicable diseases' contribution to the COVID-19 mortality: a global warning on the emerging syndemics. Archives of Iranian Medicine, 24(5), 445.
[15] Sousa, B. L. A., Brentani, A., Ribeiro, C. C. C., Dolhnikoff, M., Grisi, S. J. F. E., Ferrer, A. P. S., & Ferraro, A. A. (2021). Non-communicable diseases, sociodemographic vulnerability and the risk of mortality in hospitalised children and adolescents with COVID-19 in Brazil: a cross-sectional observational study. BMJ open, 11(9), e050724.
[16] Youn, H. M., Quan, J., Mak, I. L., Yu, E. Y. T., Lau, C. S., Ip, M. S. M., ... & Wan, E. Y. F. (2022). Long-term spill-over impact of COVID-19 on health and healthcare of people with non-communicable diseases: a study protocol for a population-based cohort and health economic study. BMJ open, 12(8), e063150.
[17] Nicoletti-Rojas, D., Retamal, R., Cerda-Rioseco, R., Rodríguez-Osiac, L., Fuentes-Alburquenque, M., & Araya-Bannout, M. (2022). Effects of sociodemographic and health factors on the self-management of non-communicable diseases among Chilean adults during the Covid-19 pandemic. PLOS Global Public Health, 2(7), e0000763.
[18] Gaur, K., Khedar, R. S., Mangal, K., Sharma, A. K., Dhamija, R. K., & Gupta, R. (2021). Macrolevel association of COVID-19 with non-communicable disease risk factors in India. Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 15(1), 343-350.
[19] Pécout, C., Pain, E., Chekroun, M., Champeix, C., Kulak, C., Prieto, R., et al. (2021). Impact of the COVID-19 Pandemic on Patients Affected by Non-Communicable Diseases in Europe and in the USA. International Journal of Environmental Research and Public Health, 18(13), 6697.
[20] Mistry, S. K., Ali, A. M., Yadav, U. N., Ghimire, S., Hossain, M. B., Das Shuvo, S., et al. (2021). Older adults with non-communicable chronic conditions and their health care access amid COVID-19 pandemic in Bangladesh: Findings from a cross-sectional study. PLoS One, 16(7), e0255534.
[21] Gutierrez, J. P., & Bertozzi, S. M. (2020). Non-communicable diseases and inequalities increase risk of death among COVID-19 patients in Mexico. PLoS One, 15(10), e0240394.
[22] Cobre, A. F., Surek, M., Vilhena, R. O., Böger, B., Fachi, M. M., Momade, D. R., et al. (2022). Influence of foods and nutrients on COVID-19 recovery: A multivariate analysis of data from 170 countries using a generalized linear model. Clinical Nutrition, 41(12), 3077-3084.
[23] Formenti, B., Gregori, N., Crosato, V., Marchese, V., Tomasoni, L. R., & Castelli, F. (2022). The impact of COVID-19 on communicable and non-communicable diseases in Africa: a narrative review. Le Infezioni in Medicina, 30(1), 30.
[24] Azzouzi, S., Stratton, C., Muñoz-Velasco, L. P., Wang, K., Fourtassi, M., Hong, B. Y., et al. (2022). The impact of the COVID-19 pandemic on healthy lifestyle behaviors and perceived mental and physical health of people living with non-communicable diseases: An international cross-sectional survey. International Journal of Environmental Research and Public Health, 19(13), 8023.
[25] Yadav, U. N., Rayamajhee, B., Mistry, S. K., Parsekar, S. S., & Mishra, S. K. (2020). A syndemic perspective on the management of non-communicable diseases amid the COVID-19 pandemic in low-and middle-income countries. Frontiers in public health, 8, 508.
[26] (2023) Our World in Data website. [Online]. Available: https://ourworldindata.org/coronavirus.
[27] (2023) Our World in Data website. [Online]. Available: https://ourworldindata.org/grapher/disease-burden-from-ncds.
[28] Özkaya, U., & Öztürk, Ş. (2022). Gaussian Regression Models for Day-Level Forecasting of COVID-19 in European Countries. Understanding COVID-19: The Role of Computational Intelligence, 339-356.
[29] İyit, N., Sevim, F., & Kahraman, Ü. M. (2023). Investigating the impact of CO2 emissions on the

COVID-19 pandemic by generalized linear mixed model approach with inverse Gaussian and gamma distributions. Open Chemistry, 21(1), 20220301.
[30] Yonar, H., & Neslihan, İ. Y. İ. T. (2018). Modeling the causality relationships between Gdp/Gni and electricity consumption according to income levels of countries by Generalized Estimating Equations. Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, (39), 191-200.
[31] Yonar, H., & İyit, N. (2021). Some Generalized Estimating Equations Models Based on Causality Tests for Investigation of The Economic Growth of The Country Groups. Foundations of Computing and Decision Sciences, 46(3), 297-315.
[32] Iyit, N., & Genc, A. (2011). Constitution of random intercept and slope model (RISM) as a special case of linear mixed models (LMMs) for repeated measurements data. Applied Mathematics and Computation, 218(3), 827-831.
[33] Tekin, K. U., Mestav, B., & Neslihan, İ. Y. İ. T. (2021). Robust Logistic Modelling for Datasets with Unusual Points. Journal of New Theory, (36), 49-63.
[34] Hilbe, J. M. (2011). Negative binomial regression. Cambridge University Press.
[35] İyıt, N., Yonar, H., & Genç, A. (2016). Generalized linear models for European Union countries energy data. Acta Physica Polonica A, 130(1), 397-400.
[36] İyit, N. (2021). An applicatıon of generalized lınear model approach on econometric studies. Research & Reviews in Science and Mathematics-II.
[37] Iyit N, Al Mashhadani AA. An application of generalized linear model (GLM) to child mortality data in Iraq based on socio-economic indicators. In: Ugur A, Tozak K, Yatbaz A, editors. Turkish World Socio Economic Strategies. Beau Bassin: LAP Lambert Academic Publishing; 2017. p. 195-203.
[38] Iyit, N. (2018). Modelling world energy security data from multinomial distribution by generalized linear model under different cumulative link functions. Open Chemistry, 16(1), 377-385.
[39] Fox, J. (2015). Applied regression analysis and generalized linear models. Sage Publications.
[40] Oznur, O., & Iyit, N. (2018). Modelling the US diabetes mortality rates via generalized linear model with the Tweedie distribution. Int. J. Sci. Res, 7(2), 1326-1334.
[41] Hardin, J. W., Hardin, J. W., Hilbe, J. M., & Hilbe, J. (2007). Generalized linear models and extensions. Stata press.
[42] Akaike, H. (1974). A new look at the statistical model identification. IEEE transactions on automatic control, 19(6), 716-723.
[43] Cavanaugh, J. E. (1997). Unifying the derivations for the Akaike and corrected Akaike information criteria. Statistics & Probability Letters, 33(2), 201-208.
[44] Schwarz, G. (1978). Estimating the dimension of a model. The annals of statistics, 461-464.
[45] Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika, 52(3), 345-370.
[46] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2021. Available from: https://www.R-project.org/.