Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi
Sayı 8, S. 96-112, 4, 2024

Araştırma Makalesi

International Journal of Advanced Natural Sciences and Engineering Researches
Volume 8, pp. 96-112, 4, 2024

Research Article

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

Statistical modelling of global communicable diseases in the aspect of demographic, economic, and environmental indicators using generalized linear mixed models with multi-random effects

Neslihan İYİT

Statistics Department, Faculty of Science, Selcuk University, Konya, Turkiye

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

(Received:10 April 2024, Accepted: 24 May 2024)

(3rd International Conference on Engineering, Natural and Social Sciences ICENSOS 2024, May 16-17, 2024)

ATIF/REFERENCE: İyit, N. (2024). Statistical modelling of global communicable diseases in the aspect of demographic, economic, and environmental indicators using generalized linear mixed models with multi-random effects. *International Journal of Advanced Natural Sciences and Engineering Researches*, 8(4), 96-112.

Abstract **–** Communicable diseases are infectious diseases that can spread from person to person every year, causing the death of hundreds of thousands of people and significantly threatening public health. Disability-adjusted life years (DALYs) is an important criterion measuring the years of life lost by a person due to a negative situation such as illness, injury or infectious diseases, and the quality of life. In this study, the relationships between DALYs from communicable diseases and countries' income levels, urbanization, net immigration rate, median age, forest area, and human development index (HDI) for 187 countries from six continents in 2019. Four generalized linear models and twelve generalized linear mixed models (both GLMs and GLMMs) having binomial distribution with different random effects such as countries, continents, and both of them under "logit", "probit", "cloglog", and "cauchit" link functions are used for modelling DALYs data in the global aspect of population and demographic change, and also economic, development, and environmental indicators. As a result of sixteen modelling, GLMM having binomial distribution with country and continent-random effects under "logit" link function is detected as the best fitted model according to information criteria as AIC with 100.766, AICc with 102.870, BIC with 142.770, and CAIC with 155.770. According to these statistical findings, it has been detected that increases in urbanization and net immigration rates have a positive effect on DALYs from communicable diseases, while increases in countries' income levels, median age, forest area, and HDI have a negative effect.

Keywords – Generalized Linear Model, Generalized Linear Mixed Model, Random Effect, Logit, Probit, Cloglog, Cauchit Link Function, Communicable Disease, Disability-Adjusted Life Years.

I. INTRODUCTION

Communicable diseases are infectious diseases that can cause the death of hundreds of thousands of people every year and pose a significant threat to public health [1-5]. The "tree map" in terms of visualization of the percentages of causes of deaths worldwide in 2019 are given in Figure 1 [6].

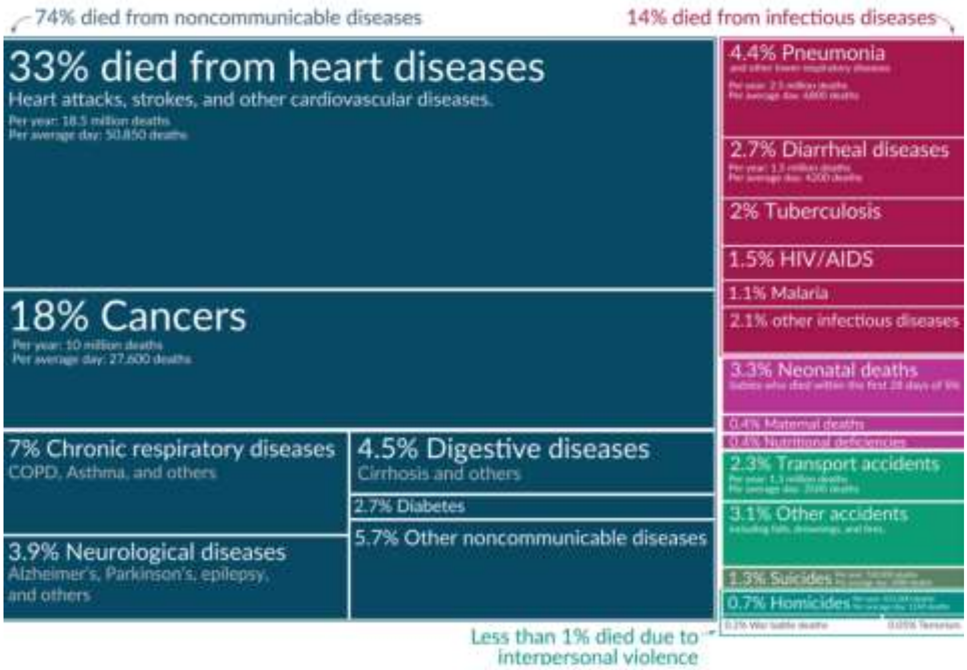

Figure 1. The treemap in terms of visualization of the percentages of causes of deaths worldwide in 2019 [6].
Source: https://ourworldindata.org/causes-of-death-treemap [6]

Approximately 7.7 million of the 55 million deaths in 2019 are caused by communicable diseases. As can be seen from Figure 1, deaths from communicable (infectious) diseases are accounted for 14% worldwide in 2019. These diseases come after non-communicable diseases among the causes of deaths ranking 2nd worldwide. From here it is understood how important these diseases are [6].

Disability-adjusted life years (DALYs) is an important criterion as the years of life lost by a person due to a negative situation such as illness, injury or infectious diseases, and the quality of life [7-11]. Since DALYs resulting from a particular disease or disability, this measurement is very important in determining the burden of that disease or disability [12-15].

Intercountry migration means people changing their place of residence from one country to another for any purpose, and urbanization means the mobility of a country's population from rural areas to cities, that is, migration within the country [16-20]. It is known that migration and urbanization as population and demographic changes indicators have a significant impact on the spread and treatment of communicable diseases as they will cause dense population [21-32]. Forest area is part of the ecosystem in a region or a country with its vegetation, trees and biodiversity [33-35]. Therefore, the decrease in the forest areas may lead to a decrease in access to clean water, which may significantly reduce the rate of decrease in the air pollution and may lead to an increase in the communicable diseases [36-39]. Human Development Index (HDI) is an index that measures a country's development with indicators such as life expectancy, income level and education level [40-42]. In countries with a high HDI, early diagnosis, treatment opportunities and prevention of infectious diseases will be more effective than in countries with low HDI [43-46].

In the literature, there are some studies investigating the relationships between communicable diseases and countries' income levels, urbanization, immigration, median age, forest area, and HDI. Wood et al. [47] investigated the effects of environmental, demographic and economic factors on DALYs for 60 countries using spatial analysis. Guégan et al. [48] examined the effect of clearing forest areas on the increased risk of infection. Napolioni and MacMurray [49] investigated the relationships between HDI and communicable (infectious) diseases with other indicators for 189 countries. Buheji et al. [50] used the Poisson regression model for modeling mortality rates and recovery from the COVID-19 pandemic with infection diseases data for 173 countries. Xu et al. [51] modelled global age-standardized DALYs data using the GLM approach. Gianino et al. [52] measured DALYs for sixteen Europe countries using clustered analysis and Kolmogorov-Smirnov (KS) test. Montagu et al. [53], Forrester et al. [54], Sharma [55], Daroudi [56], Pattanayak [57], Sun et al. [58], and İyit et al. [59] investigated the relationships between DALYs from communicable diseases and the different indicators mentioned above.

In this study, the relationships between DALYs from communicable diseases and countries' income levels, urbanization, net immigration rate, median age, forest area, and HDI are investigated for 187 countries from six continents in 2019. Four generalized linear models and twelve generalized linear mixed models (GLMs and GLMMs) having binomial distribution with different random effects such as countries, continents, and both of them under "logit", "probit", "cloglog", and "cauchit" link functions are used for modelling DALYs data in the global aspect of population and demographic change, and also economic, development, and environmental indicators.

II. MATERIALS AND METHOD

In this study, DALYs from communicable diseases in 2019 per 100.000 person is taken as the binary response variable with cut-off value as its median value 2106.86 [60]. Income levels, urbanization, net migration rate, median age, forest area, and HDI are taken as the explanatory variables in this study. According to the World Data Bank's income classifications in 2019, income levels of 187 countries from six continents consist of four income categories as "low-income", "low-middle-income", "upper-middle-income", and "high-income" [61]. Net migration rate is calculated by subtracting the number of people per 1,000 people migrating from a given country from the number of people moving to that country in the same year [62]. Urbanization is the share of the population living in urban areas calculated as the percentage of urban population relative to the total population [63]. Median age is the value that divides the age distribution of people living in a country in a year into two in terms of probability distribution [64]. Forest area is the share of the forest area of at least five meters in length, excluding agricultural production, to the total area in that country [65]. Lastly, HDI is an indicator measuring given country's education, life expectancy and income level and takes values between 0 and 1 [66]. All data set of the response and explanatory variables included in this study belongs to 187 countries from six continents in 2019. All variables used in this study are given in Table 1.

Table 1. Definitions of the response/explanatory variables taken in this study [60-66]

Names of response/explanatory variables	**Description**
DALYs from communicable diseases	Binary variable according to its median value at 2106.86 of DALYs from communicable diseases in 2019 per 100.000 person for 187 countries from six continents in 2019 [60]
Income	Income levels according to the World Data Bank's classification of 187 countries from six continents in 2019 [61]
Urbanization	The share of the population living in urban areas calculating as the percentage of urban population relative to the total population of 187 countries from 6 continents in 2019 [62]
Net migration	Subtracting the number of people per 1,000 people migrating from a given country from the number of people moving to that country in 2019 [63]
Median age	The value that divides the age distribution of people living in 187 countries from 6 continents in 2019 into two in terms of probability distribution [64]
Forest area	The share of the forest area of at least five meters in length, excluding agricultural production, to the total area in 187 countries from six continents in 2019 [65]
Human development index (HDI)	An indicator measuring a given country's education, life expectancy and income level and taking values between 0 and 1 for 187 countries from six continents in 2019 [66]

Descriptive statistics such as minimum, mean, 1st quantile, 2nd quantile (median), 3rd quantile, maximum, and standard deviation values of the response and explanatory variables are given in Table 2.

Table 2. Descriptive statistics such as minimum, mean, 1st quantile, 2nd quantile, 3rd quantile, and maximum values of response and explanatory variables

Names of response/ explanatory variables	**Min.**	**1st quantile**	**2nd quantile (median)**	**Mean**	**3rd quantile**	**Max.**	**Standard deviation (s.d.)**
DALYs from communicable diseases	415.7071	917.9943	2106.8596	5442.0608	6268.7229	39702.7906	7378.1840
Urbanization	13.2500	41.1890	60.0370	59.2667	78.0205	100.0000	22.8959
Net migration	-41.6310	-2.3080	-0.4290	-0.6788	2.0955	25.2100	7.9113
Median age	14.4000	21.0000	28.3000	28.8626	36.7500	47.6000	9.0817
Forest area	0.0000	11.5112	31.1231	32.3991	49.5799	97.4831	23.8184
HDI	0.3900	0.6080	0.7460	0.7460	0.8410	0.9610	0.1511

On the other hand, descriptive statistics such as minimum, mean, 1st quantile, 2nd quantile (median), 3rd quantile, maximum, and standard deviation values of DALYs from communicable diseases in 2019 per 100.000 person for 187 countries according to their income levels as “low-income”, “low-middle-income”, “upper-middle-income”, and “high-income”, and also six continents as Europe, North America, Asia, South America, Africa, Oceania and Antarctica are given in Table 3, respectively.

Table 3. Descriptive statistics of DALYs from communicable diseases in 2019 per 100.000 person for 187 countries according to their income levels, and also for six continents

Income levels / Continents	**Number of countries**	**Minimum**	**1st quantile**	**2nd quantile (median)**	**Mean**	**3rd quantile**	**Maximum**	**Standard deviation (s.d.)**
Low-income	27	882.2754	9317.7103	13553.1152	15370.9871	19908.0818	38072.8814	9220.9955
Low-middle-income	47	911.7114	2430.3096	5400.0980	7975.2617	11113.9608	39702.7906	7600.0662
Upper-middle-income	54	545.1941	1067.6574	2013.5665	3002.6990	2699.5184	18900.4823	3931.5094
High-income	59	415.7071	644.7708	836.0992	1112.9777	1150.8612	4875.9924	910.4690
Europe	42	415.7071	641.6667	733.7178	876.1479	971.7213	2247.9315	418.4459
North America	23	691.6689	1513.6234	1909.3746	2433.7560	2481.3156	12139.2751	2319.6562
Asia	45	435.8627	988.7096	1539.3566	2604.6223	3443.3476	8605.7535	2146.2253
South America	12	891.4329	1515.0315	2065.3543	2172.0689	2588.6001	4111.2836	923.8290
Africa	52	911.7114	8000.3846	13175.3572	13866.0718	18518.9646	39702.7906	9259.5041
Oceania and Antarctica	13	447.4426	2722.1080	3643.3579	4660.1701	5412.2458	13160.7536	3448.2189

According to the income levels, the top five countries with the highest DALYs from communicable disease per 100.000 person are Central African Republic with 38072.8814, Chad with 35389.1239, Niger with 28658.8973, Eritrea with 25857.5411, and Mozambique with 24379.3795 among “low-income” countries; Lesotho with 39702.7906, Eswatini with 25946.0406, Nigeria with 21631.7396, Cameroon with 18553.3984, and Zimbabwe with 18507.4866 among “low-middle-income” countries; South Africa with 18900.4823, Botswana with 18229.6411, Namibia with 14711.6622, Equatorial Guinea with 11732.7681, and Gabon with 6055.8858 among “upper-middle-income” countries; and finally Saint Kitts and Nevis with 4875.9924, Palau with 4714.5883, Nauru with 3516.1237, Seychelles with 2862.9414, and Bahamas with 2464.8324 among “high-income” countries, respectively.

In the GLM, the response variable belongs to the exponential distribution family, the systematic component consists of linear combinations of explanatory variables with their parameters, and the link function connects these two parts as follows [59,68-70];

$$\eta_{ij} = g\left(E\left[y_{ij}\right]\right) = \sum_{j=0}^{p} \beta_j x_{ij} \tag{1}$$

where $E\left[y_{ij}\right]$, x_{ij}, β_j, and g are expected value of the response variable, explanatory variables, parameters to be estimated, and the link function, respectively [71-73].

The GLMM as an extended version of the GLM is formed by adding random effect(s) to the systematic part of the GLM as follows [74-77];

$$\eta_{ij} = g\left(E\left[y_{ij}\right]\right) = \sum_{j=0}^{p} \beta_j x_{ij} + \sum_{k=1}^{q} u_k z_{ik} \tag{2}$$

where u_k and z_{ik} are random effects and their coefficients, respectively [78-80].
The logit, probit, cloglog, and cauchit link functions generally used for the GLM and GLMM having binomial distribution are given as follows, respectively [68,71,81-84];

$$g(p_i) = \text{logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right) \tag{3}$$

$$g(p_i) = \Phi^{-1}(p_i) \tag{4}$$

$$g(p_i) = 1 - \exp\{1 - \exp\{p_i\}\} \tag{5}$$

$$g(p_i) = \tan\left(\pi(p_i - 0.5)\right) \tag{6}$$

The formulas of information criteria (IC) as AIC, AICc, BIC, and CAIC used to statistically compare the performances of the GLMs and GLMMs are $-2l + 2k$, $-2l + (2kN)/(N-k-1)$, $-2l + k\ln(N)$, and $-2l + k\{\ln(N)+1\}$, respectively where l is the value of the log-likelihood function, k is number of parameters to be estimated, and finally N is the number of observations taken into the model [85-88].

In this study, iteratively reweighted least squares (IRLS) algorithm as a special case of Fisher-Scoring (FS) method and Laplace approximation method are used as the parameter estimation method for fitting the GLMs and GLMMs, respectively [75,77,88,90-94].

III. RESULTS AND DISCUSSION

In this study, four GLMs and twelve GLMMs having binomial distribution with different random effects such as countries (id), continents (cn), and both of them under the "logit", "probit", "cloglog", and "cauchit" link functions are used for modelling the DALYs data in the global aspect of urbanization, net migration rate, and median age as the population and demographic change indicators, income levels and HDIs of these countries as the economic and development indicators, and also forest area as the environmental indicator. Net migration rate belonging to 187 countries is taken as the categorical variable according to its quantiles. IRLS algorithm as special cases of the FS method and the Laplace approximation method are used as the parameter estimation method for fitting the GLMs and GLMMs having binomial distribution under "logit", "probit", "cloglog", and "cauchit" link functions, respectively. RStudio program is utilized in all statistical iterative approaches used in this study [95-97].

The results of the GLM having binomial distribution under "logit" link function using the IRLS algorithm for modelling DALYs data from 187 countries' communicable (infectious) diseases are given in Table 4.

Table 4. The results of the GLM having binomial distribution under "logit" link function using IRLS algorithm for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.(\hat{\beta})$	$p-$values	$\exp(\hat{\beta})$	Lower and upper bounds for $\exp(\beta)$
Intercept	-15.58534	4.53533	0.00059*	0.00000	(0.00000, 0.00124)
Income [1st level]	-1.80183	1.89581	0.341897	0.16500	(0.00402, 6.77941)
Income [2nd level]	0.35619	1.10631	0.747483	1.42788	(0.16330, 12.48487)
Income [3rd level]	0.67658	0.79169	0.392773	1.96714	(0.41682, 9.28373)
Urbanization	0.03197	0.01504	0.03353*	1.03249	(1.00250, 1.06338)
Net migration [2nd level]	-0.10390	0.75790	0.89096	0.90132	(0.20406, 3.98109)
Net migration [3rd level]	1.57912	0.77652	0.04199*	4.85069	(1.05883, 22.22174)
Net migration [4th level]	-0.29748	0.81525	0.71519	0.74269	(0.15027, 3.67069)
Median age	-0.25959	0.09031	0.00405*	0.77137	(0.64623, 0.92073)
Forest area	-0.03917	0.01303	0.00265*	0.96159	(0.93734, 0.98646)
HDI	-10.17223	7.01565	0.14708	0.00004	(0.00000, 35.80872)

*0.05 significance level

The results of the GLM having binomial distribution under "probit" link function using IRLS algorithm for modelling DALYs data from 187 countries' communicable (infectious) diseases are given in Table 5.

Table 5. The results of the GLM having binomial distribution under "probit" link function using IRLS algorithm for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.(\hat{\beta})$	$p-$values	$\exp(\hat{\beta})$	Lower and upper bounds for $\exp(\beta)$
Intercept	-8.74638	2.44078	0.00034*	0.00016	(0.00000, 0.01901)
Income [1st level]	-0.95786	1.05939	0.36591	0.38371	(0.04811, 3.06028)
Income [2nd level]	0.15201	0.62940	0.80915	1.16418	(0.33906, 3.99725)
Income [3rd level]	0.34991	0.45050	0.43733	1.41894	(0.58681, 3.43109)
Urbanization	0.01806	0.00850	0.03364*	1.01822	(1.00140, 1.03533)
Net migration [2nd level]	0.02506	0.42579	0.95307	1.02537	(0.44509, 2.36221)
Net migration [3rd level]	0.90581	0.43168	0.03588*	2.47394	(1.06155, 5.76551)
Net migration [4th level]	-0.15775	0.46737	0.73572	0.85406	(0.34171, 2.13461)
Median age	-0.13905	0.04692	0.00304*	0.87018	(0.79373, 0.95400)
Forest area	-0.02122	0.00723	0.00331*	0.97900	(0.96523, 0.99296)
HDI	-5.94869	3.80309	0.11778	0.00261	(0.00000, 4.50522)

*0.05 significance level

The results of the GLM having binomial distribution under "cloglog" link function using IRLS algorithm for modelling DALYs data from 187 countries' communicable (infectious) diseases are given in Table 6.

Table 6. The results of the GLM having binomial distribution under "cloglog" link function using IRLS algorithm for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.(\hat{\beta})$	$p-$values	$\exp(\hat{\beta})$	Lower and upper bounds for $\exp(\beta)$
Intercept	-11.36556	3.25749	0.00049*	0.00001	(0.00000, 0.00687)
Income [1st level]	-0.83291	1.48086	0.57381	0.43478	(0.02387, 7.92096)
Income [2nd level]	-0.07403	0.71558	0.91760	0.92865	(0.22843, 3.77530)
Income [3rd level]	0.50210	0.50144	0.31668	1.65218	(0.61835, 4.41452)
Urbanization	0.01971	0.01025	0.05448	1.01990	(0.99962, 1.04059)
Net migration [2nd level]	0.04690	0.52007	0.92815	1.04801	(0.37816, 2.90439)
Net migration [3rd level]	1.11993	0.51648	0.03013*	3.06464	(1.11366, 8.43348)
Net migration [4th level]	-0.25560	0.54688	0.64023	0.77445	(0.26515, 2.26204)

Median age	-0.13350	0.05270	0.01130*	0.87503	(0.78916, 0.97024)
Forest area	-0.02723	0.00883	0.00205*	0.97314	(0.95643, 0.99013)
HDI	-9.13696	5.03915	0.06980	0.00011	(0.00000, 2.09505)

*0.05 significance level

The results of the GLM having binomial distribution under “cauchit” link function using IRLS algorithm for modelling DALYs data from 187 countries’ communicable (infectious) diseases are given in Table 7.

Table 7. The results of the GLM having binomial distribution under “cauchit” link function using IRLS algorithm for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.(\hat{\beta})$	$p-$ values	$\exp(\hat{\beta})$	Lower and upper bounds for $\exp(\beta)$
Intercept	-25.0068	10.10181	0.01331*	0.00000	(0.00000, 0.00547)
Income [1st level]	-2.35282	4.91571	0.6322	0.09510	(0.00001, >1000)
Income [2nd level]	0.37474	1.84269	0.83885	1.45461	(0.03929, 53.85774)
Income [3rd level]	1.7369	1.38435	0.2096	5.67971	(0.37667, 85.64203)
Urbanization	0.04546	0.02555	0.0752	1.04651	(0.99539, 1.10025)
Net migration [2nd level]	-0.28915	1.17812	0.80612	0.74890	(0.07441, 7.53774)
Net migration [3rd level]	2.30643	1.25676	0.06647	10.03852	(0.85490, 117.87613)
Net migration [4th level]	-1.10893	1.21505	0.36142	0.32991	(0.03049, 3.56985)
Median age	-0.64754	0.25022	0.00966*	0.52333	(0.32047, 0.85460)
Forest area	-0.06881	0.02709	0.01108*	0.93350	(0.88523, 0.98441)
HDI	-8.02403	12.59788	0.52417	0.00033	(0.00000, >1000)

*0.05 significance level

The results of the GLMM having binomial distribution under “logit” link function using Laplace approximation method with “id” random effect for modelling DALYs data from 187 countries’ communicable (infectious) diseases are given in Table 8.

Table 8. The results of the GLMM having binomial distribution under “logit” link function using Laplace approximation with “id” random effect for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.(\hat{\beta})$	$p-$values	$\exp(\hat{\beta})$	Lower and upper bounds for $\exp(\beta)$
Intercept	-15.97061	6.97307	0.0220*	0.00000	(0.00000, 0.09989)
Income [1st level]	-1.65786	2.05802	0.42050	0.19055	(0.00337, 10.75946)
Income [2nd level]	0.21969	1.19666	0.85430	1.24569	(0.11935, 13.00197)
Income [3rd level]	0.64450	0.86888	0.45820	1.90503	(0.34699, 10.45911)
Urbanization	0.03450	0.02020	0.08760	1.03510	(0.99492, 1.07691)
Net migration [2nd level]	-0.07558	0.82400	0.92690	0.92721	(0.18441, 4.66193)
Net migration [3rd level]	1.69487	1.02698	0.09890	5.44594	(0.72762, 40.76051)
Net migration [4th level]	-0.28203	0.88240	0.74930	0.75425	(0.13379, 4.25222)
Median age	-0.28289	0.13012	0.0297*	0.75360	(0.58396, 0.97253)
Forest area	-0.04250	0.01981	0.0319*	0.95839	(0.92189, 0.99633)
HDI	-9.84772	8.11164	0.22470	0.00005	(0.00000, 424.45347)

*0.05 significance level

The results of the GLMM having binomial distribution under “probit” link function using Laplace approximation method with “id” random effect for modelling DALYs data from 187 countries’ communicable (infectious) diseases are given in Table 9.

Table 9. The results of the GLMM having binomial distribution under "probit" link function using Laplace approximation with "id" random effect for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.(\hat{\beta})$	$p-$values	$\exp(\hat{\beta})$	Lower and upper bounds for $\exp(\beta)$
Intercept	-8.87570	3.88638	0.0224*	0.00014	(0.00000, 0.28407)
Income [1st level]	-0.97160	1.13578	0.39230	0.37848	(0.04086, 3.50604)
Income [2nd level]	0.15389	0.63202	0.80760	1.16636	(0.33795, 4.02541)
Income [3rd level]	0.35486	0.47850	0.45830	1.42598	(0.55823, 3.64262)
Urbanization	0.01833	0.01264	0.14710	1.01850	(0.99358, 1.04405)
Net migration [2nd level]	0.02620	0.43052	0.95150	1.02655	(0.44149, 2.38692)
Net migration [3rd level]	-0.91932	0.55857	0.09980	0.39879	(0.13344, 1.19179)
Net migration [4th level]	-0.16003	0.48173	0.73970	0.85212	(0.33147, 2.19053)
Median age	0.14107	0.07216	0.05060	1.15151	(0.99964, 1.32644)
Forest area	-0.02153	0.01049	0.0402*	0.97870	(0.95878, 0.99903)
HDI	-6.03813	4.37961	0.16800	0.00239	(0.00000, 12.75276)

*0.05 significance level

The results of the GLMM having binomial distribution under "cloglog" link function using Laplace approximation method with "id" random effect for modelling DALYs data from 187 countries' communicable (infectious) diseases are given in Table 10.

Table 10. The results of the GLMM having binomial distribution under "cloglog" link function using Laplace approximation with "id" random effect for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.(\hat{\beta})$	$p-$values	$\exp(\hat{\beta})$	Lower and upper bounds for $\exp(\beta)$
Intercept	-11.36623	3.30925	0.000593*	0.00001	(0.00000, 0.00760)
Income [1st level]	-0.83334	1.46132	0.56850	0.43459	(0.02479, 7.62019)
Income [2nd level]	-0.07336	0.69908	0.91643	0.92927	(0.23609, 3.65761)
Income [3rd level]	0.50217	0.50135	0.31652	1.65230	(0.61850, 4.41405)
Urbanization	0.01971	0.01053	0.06123	1.01991	(0.99907, 1.04118)
Net migration [2nd level]	0.04692	0.54623	0.93155	1.04804	(0.35927, 3.05724)
Net migration [3rd level]	1.11958	0.54133	0.038623*	3.06355	(1.06033, 8.85133)
Net migration [4th level]	-0.25529	0.54914	0.64201	0.77469	(0.26406, 2.27278)
Median age	-0.13354	0.05321	0.012090*	0.87499	(0.78833, 0.97118)
Forest area	-0.02723	0.00939	0.003720*	0.97314	(0.95541, 0.99121)
HDI	-9.13482	5.11526	0.07413	0.00011	(0.00000, 2.43733)

*0.05 significance level

The results of the GLMM having binomial distribution under "cauchit" link function using Laplace approximation method with "id" random effect for modelling DALYs data from 187 countries' communicable (infectious) diseases are given in Table 11.

Table 11. The results of the GLMM having binomial distribution under "cauchit" link function using Laplace approximation with "id" random effect for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.\left(\hat{\beta}\right)$	$p-$values	$\exp\left(\hat{\beta}\right)$	Lower and upper bounds for $\exp(\beta)$
Intercept	-25.00409	9.74996	0.0103*	0.00000	(0.00000, 0.00275)
iIncome [1st level]	-2.35044	4.64216	0.61260	0.09533	(0.00001, 852.37492)
Income [2nd level]	0.37344	2.14806	0.86200	1.45272	(0.02157, 97.86219)
Income [3rd level]	1.73413	1.87287	0.35450	5.66400	(0.14419, 222.49125)
Urbanization	0.04544	0.02578	0.07800	1.04649	(0.99493, 1.10072)
Net migration [2nd level]	-0.29154	1.63871	0.85880	0.74711	(0.03010, 18.54634)
Net migration [3rd level]	2.30798	1.45786	0.11340	10.05409	(0.57731, 175.09578)
Net migration [4th level]	-1.11004	1.44707	0.44300	0.32955	(0.01933, 5.61906)
Median age	-0.64716	0.29753	0.0296*	0.52353	(0.29220, 0.93799)
Forest area	-0.06879	0.02869	0.0165*	0.93352	(0.88248, 0.98752)
HDI	-8.03992	12.91671	0.53370	0.00032	(0.00000, >1000)

*0.05 significance level

The results of the GLMM having binomial distribution under "logit" link function using Laplace approximation method with "cn" random effect for modelling DALYs data from 187 countries' communicable (infectious) diseases are given in Table 12.

Table 12. The results of the GLMM having binomial distribution under "logit" link function using Laplace approximation with "cn" random effect for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.\left(\hat{\beta}\right)$	$p-$values	$\exp\left(\hat{\beta}\right)$	Lower and upper bounds for $\exp(\beta)$
Intercept	-15.59486	4.74118	0.00100*	0.00000	(0.00000, 0.00183)
Income [1st level]	-1.64440	1.98339	0.40705	0.19313	(0.00396, 9.42130)
Income [2nd level]	0.38239	1.15294	0.74014	1.46578	(0.15300, 14.04282)
Income [3rd level]	0.66395	0.83843	0.42842	1.94245	(0.37556, 10.04668)
Urbanization	0.03602	0.01661	0.03016*	1.03668	(1.00347, 1.07098)
Net migration [2nd level]	0.02556	0.80057	0.97453	1.02589	(0.21363, 4.92660)
Net migration [3rd level]	-1.54214	0.80594	0.05569	0.21392	(0.04408, 1.03818)
Net migration [4th level]	-0.26558	0.84448	0.75315	0.76676	(0.14650, 4.01313)
Median age	-0.26240	0.09292	0.00474*	0.76920	(0.64113, 0.92286)
Forest area	-0.04311	0.01486	0.00371*	0.95781	(0.93031, 0.98611)
HDI	-9.75022	7.31677	0.18267	0.00006	(0.00000, 98.53244)

*0.05 significance level

The results of the GLMM having binomial distribution under "probit" link function using Laplace approximation method with "cn" random effect for modelling DALYs data from 187 countries' communicable (infectious) diseases are given in Table 13.

Table 13. The results of the GLMM having binomial distribution under "probit" link function using Laplace approximation with "cn" random effect for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.\left(\hat{\beta}\right)$	$p-$values	$\exp\left(\hat{\beta}\right)$	Lower and upper bounds for $\exp(\beta)$
Intercept	-8.73529	2.55198	0.000619*	0.00016	(0.00000, 0.02391)
Income [1st level]	-0.91667	1.09203	0.40124	0.39985	(0.04703, 3.39964)
Income [2nd level]	0.15233	0.64722	0.81393	1.16454	(0.32752, 4.14065)
Income [3rd level]	0.35901	0.46700	0.44204	1.43191	(0.57333, 3.57625)
Urbanization	0.01991	0.00887	0.024845*	1.02011	(1.00252, 1.03801)
Net migration [2nd level]	0.07342	0.44193	0.86805	1.07618	(0.45260, 2.55894)
Net migration [3rd level]	0.88546	0.44422	0.046230*	2.42410	(1.01491, 5.78991)
Net migration [4th level]	-0.14699	0.47673	0.75783	0.86330	(0.33913, 2.19763)
Median age	-0.14117	0.04900	0.003962*	0.86834	(0.78884, 0.95587)
Forest area	-0.02392	0.00777	0.002084*	0.97637	(0.96161, 0.99135)
HDI	5.71739	3.93752	0.14649	304.11016	(0.13533, >1000)

*0.05 significance level

The results of the GLMM having binomial distribution under "cloglog" link function using Laplace approximation method with "cn" random effect for modelling DALYs data from 187 countries' communicable (infectious) diseases are given in Table 14.

Table 14. The results of the GLMM having binomial distribution under "cloglog" link function using Laplace approximation with "cn" random effect for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.\left(\hat{\beta}\right)$	$p-$values	$\exp\left(\hat{\beta}\right)$	Lower and upper bounds for $\exp(\beta)$
Intercept	-11.36526	3.30878	0.000593*	0.00001	(0.00000, 0.00760)
Income [1st level]	-0.83270	1.46106	0.56872	0.43487	(0.02481, 7.62108)
Income [2nd level]	0.07418	0.69905	0.91549	1.07700	(0.27364, 4.23881)
Income [3rd level]	0.50206	0.50131	0.31659	1.65213	(0.61848, 4.41329)
Urbanization	0.01970	0.01053	0.06136	1.01990	(0.99906, 1.04116)
Net migration [2nd level]	0.04689	0.54619	0.93159	1.04800	(0.35929, 3.05687)
Net migration [3rd level]	1.12007	0.54135	0.038542*	3.06506	(1.06083, 8.85589)
Net migration [4th level]	-0.25575	0.54913	0.64140	0.77433	(0.26394, 2.27167)
Median age	-0.13348	0.05321	0.012130*	0.87505	(0.78838, 0.97124)
Forest area	-0.02723	0.00939	0.003713*	0.97314	(0.95540, 0.99120)
HDI	-9.13809	5.11461	0.07399	0.00011	(0.00000, 2.42626)

*0.05 significance level

The results of the GLMM having binomial distribution under "cauchit" link function using Laplace approximation method with "cn" random effect for modelling DALYs data from 187 countries' communicable (infectious) diseases are given in Table 15.

Table 15. The results of the GLMM having binomial distribution under "cauchit" link function using Laplace approximation with "cn" random effect for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.\left(\hat{\beta}\right)$	$p-$values	$\exp\left(\hat{\beta}\right)$	Lower and upper bounds for $\exp(\beta)$
Intercept	-24.19312	0.00778	<2e-16*	0.00000	-
Income [1st level]	-0.61724	0.00795	0.87200	0.53943	(0.53109, 0.54790)
Income [2nd level]	-0.26184	0.00795	<2e-16*	0.76964	(0.75773, 0.78173)
Income [3rd level]	-0.59940	0.00754	<2e-16*	0.54914	(0.54109, 0.55732)
Urbanization	0.05385	0.00522	<2e-16*	1.05533	(1.04459, 1.06617)
Net migration [2nd level]	1.07864	0.00822	<2e-16*	2.94067	(2.89365, 2.98844)

Net migration [3rd level]	2.93802	0.00781	<2e-16*	18.87839	(18.59158, 19.16963)
Net migration [4th level]	1.77148	0.00778	<2e-16*	5.87954	(5.79054, 5.96990)
Median age	-0.49792	0.00719	<2e-16*	0.60779	(0.59929, 0.61642)
Forest area	-0.06393	0.00620	<2e-16*	0.93807	(0.92675, 0.94953)
HDI	-13.39818	0.00778	<2e-16*	0.00000	-

*0.05 significance level

The results of the GLMM having binomial distribution under “logit” link function using Laplace approximation method with “id” and “cn” random effects for modelling DALYs data from 187 countries’ communicable (infectious) diseases are given in Table 16.

Table 16. The results of the GLMM having binomial distribution under “logit” link function using Laplace approximation with “id” and “cn” random effects for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.\left(\hat{\beta}\right)$	$p-$values	$\exp\left(\hat{\beta}\right)$	Lower and upper bounds for $\exp(\beta)$
Intercept	-15.84617	0.00778	<2e-16*	0.00000	-
Income [1st level]	-1.67807	0.00795	<2e-16*	0.18673	(0.18385, 0.18967)
Income [2nd level]	-0.37173	0.00795	<2e-16*	0.68954	(0.67888, 0.70037)
Income [3rd level]	-0.67647	0.00754	<2e-16*	0.50841	(0.50095, 0.51598)
Urbanization	0.03618	0.00522	4.03e-12*	1.03684	(1.02629, 1.04749)
Net migration [2nd level]	0.02277	0.00822	0.00563*	1.02303	(1.00667, 1.03965)
Net migration [3rd level]	1.58073	0.00781	<2e-16*	4.85851	(4.78469, 4.93346)
Net migration [4th level]	0.24757	0.00778	<2e-16*	1.28090	(1.26152, 1.30059)
Median age	-0.26576	0.00719	<2e-16*	0.76662	(0.75589, 0.77750)
Forest area	-0.04393	0.00620	1.34e-12*	0.95702	(0.94547, 0.96871)
HDI	-9.99171	0.00778	<2e-16*	0.00005	(0.00005, 0.00005)

*0.05 significance level

The results of the GLMM having binomial distribution under “probit” link function using Laplace approximation method with “id” and “cn” random effects for modelling DALYs data from 187 countries’ communicable (infectious) diseases are given in Table 17.

Table 17. The results of GLMM having binomial distribution under “probit” link function using Laplace approximation with “id” and “cn” random effects for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.\left(\hat{\beta}\right)$	$p-$values	$\exp\left(\hat{\beta}\right)$	Lower and upper bounds for $\exp(\beta)$
Intercept	-8.80462	2.57141	0.00062*	0.00015	(0.00000, 0.02317)
Income [1st level]	0.91181	1.09877	0.40662	2.48883	(0.28888, 21.44203)
Income [2nd level]	-0.15362	0.65028	0.81324	0.85760	(0.23975, 3.06760)
Income [3rd level]	-0.36150	0.46934	0.44116	0.69663	(0.27765, 1.74784)
Urbanization	0.02031	0.00894	0.02230*	1.02052	(1.00281, 1.03855)
Net migration [2nd level]	-0.08317	0.44466	0.85163	0.92020	(0.38493, 2.19977)
Net migration [3rd level]	0.89774	0.44705	0.04463*	2.45405	(1.02177, 5.89405)
Net migration [4th level]	0.14268	0.47896	0.76579	1.15336	(0.45110, 2.94888)
Median age	-0.14394	0.04951	0.00365*	0.86594	(0.78586, 0.95418)
Forest area	-0.02407	0.00783	0.00210*	0.97621	(0.96135, 0.99131)
HDI	-5.67037	3.96359	0.15254	0.00345	(0.00000, 8.15093)

*0.05 significance level

The results of the GLMM having binomial distribution under “cloglog” link function using Laplace approximation method with “id” and “cn” random effects for modelling DALYs data from 187 countries’ communicable (infectious) diseases are given in Table 18.

Table 18. The results of the GLMM having binomial distribution under "cloglog" link function using Laplace approximation with "id" and "cn" random effects for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.(\hat{\beta})$	$p-$values	$\exp(\hat{\beta})$	Lower and upper bounds for $\exp(\beta)$
Intercept	-11.36619	3.30911	0.00059*	0.00001	(0.00000, 0.00759)
Income [1st level]	0.83300	1.46118	0.56862	2.30022	(0.13122, 40.32087)
Income [2nd level]	-0.07386	0.69906	0.91585	0.92880	(0.23598, 3.65561)
Income [3rd level]	-0.50207	0.50133	0.31660	0.60528	(0.22658, 1.61691)
Urbanization	0.01971	0.01053	0.06128	1.01990	(0.99907, 1.04117)
Net migration [2nd level]	-0.04698	0.54620	0.93145	0.95410	(0.32709, 2.78308)
Net migration [3rd level]	1.11978	0.54133	0.03859*	3.06417	(1.06054, 8.85314)
Net migration [4th level]	0.25557	0.54913	0.64164	1.29120	(0.44012, 3.78805)
Median age	-0.13349	0.05321	0.01212*	0.87504	(0.78837, 0.97123)
Forest area	-0.02723	0.00939	0.00372*	0.97314	(0.95540, 0.99121)
HDI	-9.13766	5.11511	0.07403	0.00011	(0.00000, 2.42973)

*0.05 significance level

The results of the GLMM having binomial distribution under "cauchit" link function using Laplace approximation method with "id" and "cn" random effects for modelling DALYs data from 187 countries' communicable (infectious) diseases are given in Table 19.

Table 19. The results of the GLMM having binomial distribution under "cauchit" link function using Laplace approximation with "id" and "cn" random effects for modelling DALYs data from communicable (infectious) diseases

Variables	$\hat{\beta}$	$s.e.(\hat{\beta})$	$p-$values	$\exp(\hat{\beta})$	Lower and upper bounds for $\exp(\beta)$
Intercept	-25.89646	12.18007	0.0335*	0.00000	(0.00000, 0.13213)
Income [1st level]	-0.94812	6.77910	0.88880	0.38747	(0.00000, >1000)
Income [2nd level]	0.17000	2.12836	0.93630	1.18530	(0.01829, 76.82338)
Income [3rd level]	-0.65227	1.80098	0.71720	0.52086	(0.01527, 17.77132)
Urbanization	0.06098	0.03141	0.05220	1.06288	(0.99942, 1.13037)
Net migration [2nd level]	-1.22425	1.48509	0.40970	0.29398	(0.01600, 5.40040)
Net migration [3rd level]	3.08029	1.87492	0.10040	21.76471	(0.55185, 858.39615)
Net migration [4th level]	-1.83127	1.53029	0.23140	0.16021	(0.00798, 3.21569)
Median age	-0.59394	0.27510	0.0308*	0.55215	(0.32203, 0.94672)
Forest area	-0.07210	0.03125	0.0210*	0.93044	(0.87516, 0.98921)
HDI	-11.93799	18.15342	0.51080	0.00001	(0.00000, >1000)

*0.05 significance level

To statistically compare the performances of four GLMs and twelve GLMMs having binomial distribution with different random effects such as countries, continents, and both of them under "logit", "probit", "cloglog", and "cauchit" link functions, IC values as AIC, AICc, BIC, and CAIC of these GLMs and GLMMs are given in Table 20.

Table 20. IC values of the GLMs and GLMMs having binomial distribution with different random effects such as countries, continents, and both of them under "logit", "probit", "cloglog", and "cauchit" link functions

Models	Link Functions	Log-likelihood	AIC	AICc	BIC	CAIC
GLMs	logit	-45.260	114.520	116.028	150.062	161.062
	probit	-45.885	113.770	115.279	149.313	160.313
	cloglog	-47.328	116.656	118.154	152.198	163.198
	cauchit	-49.088	120.175	121.684	155.717	166.717
GLMMs with id-random effect	logit	-45.112	114.223	116.016	152.997	164.996
	probit	-45.884	115.768	117.561	154.541	166.541
	cloglog	-45.589	115.178	116.971	153.951	165.951
	cauchit	-49.088	122.175	123.968	160.949	172.949
GLMMs with cn-random effect	logit	-45.998	115.996	117.789	154.769	166.769
	probit	-45.594	115.188	116.981	153.961	165.961
	cloglog	-45.589	115.178	116.971	153.951	165.951
	cauchit	-48.367	120.733	122.526	159.506	171.506
GLMMs with id and cn-random effects	logit	-38.036*	102.073*	104.177*	144.077*	157.077*
	probit	-45.590	117.180	119.284	159.184	172.184
	cloglog	-45.589	117.178	119.282	159.183	172.183
	cauchit	-48.256	122.513	124.617	164.517	177.517

IV. CONCLUSION

In this study, as a final summary, four GLMs and twelve GLMMs having binomial distribution with different random effects such "id", "cn", and also both "id" and "cn" under "logit", "probit", "cloglog", and "cauchit" link functions are used for modelling DALYs data from 187 countries' communicable (infectious) diseases in the global aspect of urbanization, net migration rate, median age, income level, HDI, and forest area. As statistical conclusions from these GLMs and GLMMs with multi-random effects, the GLMM having binomial distribution under "logit" link function using Laplace approximation with "id" and "cn" random effects is statistically determined as the best fitted model according to the largest value of the log-likelihood with -38.036 and the smallest IC values as AIC with 102.073, AICc with 104.177, BIC with 144.077, and CAIC with 157.077.

The model equation of the GLMM having binomial distribution under "logit" link function using Laplace approximation with "id" and "cn" random effects as statistically the best fitted model according to Table 16 is given as follows;

$$\hat{\eta}_i = \log\left(\frac{p_i}{1-p_i}\right) = \left\{\begin{array}{l} -15.84617 - 1.67807 * Income[1] - 0.37173 * Income[2] - 0.67647 * Income[3] \\ +0.03618 * Urban + 0.02277 * Net\,Mig.[2] + 1.58073 * Net\,Mig.[3] + 0.24757 * Net\,Mig.[4] \\ -0.26576 * Med.Age - 0.04393 * Forest.Area - 9.99171 * HDI \end{array}\right\} \quad (7)$$

$$i = 1, 2, \ldots, 187$$

As the main important statistical inferences and conclusions from this study, according to Table 16 and Eq.(7);

When 4th level ("high-income") is taken as the reference category, for the 1st level ("low-income"), 2nd level ("low-middle-income"), and 3rd level ("upper-middle-income") of 187 world countries, the odds of being above rather than being below of the median value (2106.86) of DALYs from communicable diseases in 2019 per 100.000 person decrease by a factor level of $e^{-1.67807}=0.18673$, $e^{-0.37173}=0.68954$, and $e^{-0.67647}=0.50841$, respectively.

One percentage increase in the share of the population living in urban areas calculated as the percentage of urban population increases the odds of being above rather than being below of median value (2106.86) of DALYs from communicable diseases in 2019 per 100.000 person for 187 countries from six continents $e^{0.03618}=1.03684$ times.

When 1st level is taken as the reference category, for 2nd level, 3rd level, and 4th level of net migration, the odds of being above rather than being below of median value (2106.86) of DALYs from communicable diseases in 2019 per 100.000 person for 187 countries from six continents increase by a factor level of $e^{0.02277}=1.02303$, $e^{0.58073}=4.85851$, and $e^{0.24757}=1.28090$, respectively.

One year increase in the median age of 187 countries from six continents decreases the odds of being above rather than being below of median value (2106.86) of DALYs from communicable diseases in 2019 per 100.000 person $e^{-0.26576}=0.76662$ times.

0.1 unit increase in the HDI decreases the odds of being above rather than being below of median value (2106.86) of DALYs from communicable diseases in 2019 per 100.000 person for 187 countries from six continents $e^{-9.99171*0.1}=0.36819$ times.

REFERENCES

[1] Carballo, M. (2009). Communicable diseases. Health and migration in European Union: better health for all in an, 2006, 11-28.

[2] Connolly, M. A., Gayer, M., Ryan, M. J., Salama, P., Spiegel, P., & Heymann, D. L. (2004). Communicable diseases in complex emergencies: impact and challenges. The Lancet, 364(9449), 1974-1983.

[3] Wörmann, T., & Krämer, A. (2011). Communicable diseases. In Migration and health in the European Union (pp. 121-138). Maidenhead: Open University Press.

[4] Waring, S. C., & Brown, B. J. (2005). The threat of communicable diseases following natural disasters: a public health response. Disaster Management & Response, 3(2), 41-47.

[5] Tulchinsky, T. H., & Varavikova, E. A. (2014). Communicable diseases. The New Public Health, 149.

[6] Our world in data. Causes of death globally: what do people die from?. Available from: https://ourworldindata.org/causes-of-death-treemap [Accessed 20 Feb 2024].

[7] Arnesen, T., & Nord, E. (1999). The value of DALY life: problems with ethics and validity of disability adjusted life years. Bmj, 319(7222), 1423-1425.

[8] Devleesschauwer, B., Havelaar, A. H., Maertens de Noordhout, C., Haagsma, J. A., Praet, N., Dorny, P., ... & Speybroeck, N. (2014). Calculating disability-adjusted life years to quantify burden of disease. International journal of public health, 59, 565-569.

[9] Murray, C. J. (1994). Quantifying the burden of disease: the technical basis for disability-adjusted life years. Bulletin of the World health Organization, 72(3), 429.

[10] Grosse, S. D., Lollar, D. J., Campbell, V. A., & Chamie, M. (2009). Disability and disability-adjusted life years: not the same. Public Health Reports, 124(2), 197-202.

[11] Rushby, J. F., & Hanson, K. (2001). Calculating and presenting disability adjusted life years (DALYs) in cost-effectiveness analysis. Health policy and planning, 16(3), 326-331.

[12] Disability-adjusted life years (DALYs). World Health Organization. Available from: https://www.who.int/data/gho/indicator-metadata-registry/imr-details/158.

[13] Sassi, F. (2006). Calculating QALYs, comparing QALY and DALY calculations. Health policy and planning, 21(5), 402-408.

[14] Murray, C. J., & Acharya, A. K. (1997). Understanding dalys. Journal of health economics, 16(6), 703-730.

[15] Cuschieri, S., Pallari, E., Terzic, N., Alkerwi, A. A., & Sigurðardóttir, Á. K. (2021). Mapping the burden of diabetes in five small countries in Europe and setting the agenda for health policy and strategic action. Health Research Policy and Systems, 19, 1-10.

[16] Williamson, J. G. (1988). Migration and urbanization. Handbook of development economics, 1, 425-465.
[17] Yan, T. (2016). Migration and urbanization. In Critical Issues in Contemporary China (pp. 174-192). Routledge.
[18] Zhang, K. H., & Shunfeng, S. (2003). Rural–urban migration and urbanization in China: Evidence from time-series and cross-section analyses. China economic review, 14(4), 386-400.
[19] Adger, W. N., Crépin, A. S., Folke, C., Ospina, D., Chapin, F. S., Segerson, K., ... & Wilen, J. (2020). Urbanization, migration, and adaptation to climate change. One Earth, 3(4), 396-399.
[20] Collyer, M. (2020). Movement of Peoples: Urbanization and Migration. Mental Health and Illness of Children and Adolescents, 1-7.
[21] Moore, M., Gould, P., & Keary, B. S. (2003). Global urbanization and impact on health. International journal of hygiene and environmental health, 206(4-5), 269-278.
[22] Utzinger, J., & Keiser, J. (2006). Urbanization and tropical health—then and now. Annals of Tropical Medicine & Parasitology, 100(5-6), 517-533.
[23] Nauman, E., VanLandingham, M., & Anglewicz, P. (2016). Migration, urbanization and health. International handbook of migration and population distribution, 451-463.
[24] Zhang, P., & Atkinson, P. M. (2008). Modelling the effect of urbanization on the transmission of an infectious disease. Mathematical biosciences, 211(1), 166-185.
[25] Alirol, E., Getaz, L., Stoll, B., Chappuis, F., & Loutan, L. (2011). Urbanisation and infectious diseases in a globalised world. The Lancet infectious diseases, 11(2), 131-141.
[26] Connolly, C., Keil, R., & Ali, S. H. (2021). Extended urbanisation and the spatialities of infectious disease: Demographic change, infrastructure and governance. Urban studies, 58(2), 245-263.
[27] Neiderud, C. J. (2015). How urbanization affects the epidemiology of emerging infectious diseases. Infection ecology & epidemiology, 5(1), 27060.
[28] McMichael, C. (2015). Climate change-related migration and infectious disease. Virulence, 6(6), 548-553.
[29] Quinn, T. C., Bartlett, J. G., Vlahov, D., Boufford, J. I., Pearson, C., & Norris, L. (2010). Global infectious diseases and urbanization. Urban Health: Global Perspectives, 18, 105.
[30] Castelli, F., & Sulis, G. (2017). Migration and infectious diseases. Clinical Microbiology and Infection, 23(5), 283-289.
[31] Gushulak, B. D., & MacPherson, D. W. (2004). Globalization of infectious diseases: the impact of migration. Clinical infectious diseases, 38(12), 1742-1748.
[32] Alirol, E., Getaz, L., Stoll, B., Chappuis, F., & Loutan, L. (2011). Urbanisation and infectious diseases in a globalised world. The Lancet infectious diseases, 11(2), 131-141.
[33] Kurz, W. A. (2010). An ecosystem context for global gross forest cover loss estimates. Proceedings of the National Academy of Sciences, 107(20), 9025-9026.
[34] Sedjo, R. A. (1993). The carbon cycle and global forest ecosystem. Water, Air, and Soil Pollution, 70, 295-307.
[35] Lund, H. G. (2002). When is a forest not a forest?. Journal of Forestry, 100(8), 21-28.
[36] McCallum, H. (2008). Landscape structure, disturbance, and disease dynamics. Infectious disease ecology: effects of ecosystems on disease and of disease on ecosystems, 100, 122.
[37] Saha, K., Ghatak, D., & S. Muralee, N. S. (2022). Impact of Plantation Induced Forest Degradation on the Outbreak of Emerging Infectious Diseases—Wayanad District, Kerala, India. International Journal of Environmental Research and Public Health, 19(12), 7036.
[38] Muleme, M. (2012). Drivers of Infectious Disease Outbreaks: How Climate, Environment and Disease Control Programs Influence Occurence of Infectious Disease Outbreaks.
[39] Lee Cruz, L., Degenne, P., Cappelle, J., Caron, A., De Nys, H., Schloesing, E., ... & Roger, F. (2022). Spatial modeling of Ebola virus transmission in a changing forest landscape. ESA.
[40] Anand, S., & Sen, A. (1994). Human Development Index: Methodology and Measurement.
[41] Sagar, A. D., & Najam, A. (1998). The human development index: a critical review. Ecological economics, 25(3), 249-264.
[42] Anand, S., & Sen, A. (2000). The income component of the human development index. Journal of human development, 1(1), 83-106.
[43] Boutayeb, A. (2010). The impact of infectious diseases on the development of Africa. Handbook of disease burdens and quality of life measures, 1171.
[44] Khodabakhshi, A. (2011). Relationship between GDP and human development indices in India. In Society of Interdisciplinary Business Research (SIBR) 2011 Conference on Interdisciplinary Business Research.
[45] Franco-Herrera, D., González-Ocampo, D., Restrepo-Montoya, V., Gómez-Guevara, J. E., Alvear-Villacorte, N., & Rodríguez-Morales, A. J. (2018). Relationship between malaria epidemiology and the human development index in Colombia and Latin America. Infez Med, 26(3), 255-262.
[46] Boutayeb, A. (2010). The impact of infectious diseases on the development of Africa. Handbook of disease burdens and quality of life measures, 1171.
[47] Wood, C. L., McInturff, A., Young, H. S., Kim, D., & Lafferty, K. D. (2017). Human infectious disease burdens decrease with urbanization but not with biodiversity. Philosophical Transactions of the Royal Society B: Biological Sciences, 372(1722), 20160122.

[48] Guégan, J. F., Ayouba, A., Cappelle, J., & De Thoisy, B. (2020). Forests and emerging infectious diseases: unleashing the beast within. Environmental Research Letters, 15(8), 083007.

[49] Napolioni, V., & MacMurray, J. (2016). Infectious diseases, IL6− 174G> C polymorphism, and human development. Brain, Behavior, and Immunity, 51, 196-203.

[50] Buheji, M., AlDerazi, A., Ahmed, D., Bragazzi, N. L., Jahrami, H., Hamadeh, R. R., & BaHammam, A. S. (2022). The association between the initial outcomes of COVID-19 and the human development index: An ecological study. Human Systems Management, 41(2), 303-313.

[51] Xu, Y. Y., Xie, J., Yin, H., Yang, F. F., Ma, C. M., Yang, B. Y., ... & Li, S. L. (2022). The Global Burden of Disease attributable to low physical activity and its trends from 1990 to 2019: An analysis of the Global Burden of Disease study. Frontiers in public health, 10, 1018866.

[52] Gianino, M. M., Savatteri, A., Politano, G., Nurchis, M. C., Pascucci, D., & Damiani, G. (2021). Burden of COVID-19: disability-adjusted life years (DALYs) across 16 European countries. European Review for Medical and Pharmacological Sciences, 25(17), 5529-5541.

[53] Montagu, D., Ngamkitpaiboon, L., Duvall, S., & Ratcliffe, A. (2013). Applying the disability-adjusted life year to track health impact of social franchise programs in low-and middle-income countries. BMC public health, 13, 1-12.

[54] Forrester, J. D., Cai, L. Z., Mbanje, C., Rinderknecht, T. N., & Wren, S. M. (2017). Clostridium difficile infection in low-and middle-human development index countries: a systematic review. Tropical Medicine & International Health, 22(10), 1223-1232.

[55] Sharma, R. (2019). The burden of prostate cancer is associated with human development index: evidence from 87 countries, 1990–2016. EPMA Journal, 10, 137-152.

[56] Daroudi, R., Akbari Sari, A., Nahvijou, A., & Faramarzi, A. (2021). Cost per DALY averted in low, middle-and high-income countries: evidence from the global burden of disease study to estimate the cost-effectiveness thresholds. Cost Effectiveness and Resource Allocation, 19, 1-9.

[57] Pattanayak, S., Dickinson, K., Corey, C., Murray, B., Sills, E., & Kramer, R. (2006). Deforestation, malaria, and poverty: a call for transdisciplinary research to support the design of cross-sectoral policies. Sustainability: Science, Practice and Policy, 2(2), 45-56.

[58] Sun, H. Z., Zhao, J., Liu, X., Qiu, M., Shen, H., Guillas, S., ... & Archibald, A. T. (2023). Antagonism between ambient ozone increase and urbanization-oriented population migration on Chinese cardiopulmonary mortality. The Innovation, 4(6).

[59] 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).

[60] Our world in data. Burden of disease by cause, World, 2019. Available from: https://ourworldindata.org/grapher/burden-of-disease-by-cause [Accessed 20 Feb 2024].

[61] Our world in data. World Bank income groups. Available from: https://ourworldindata.org/grapher/world-bank-income-groups?tab=table [Accessed 20 Feb 2024].

[62] Our world in data. Share of the population living in urban areas, 2021. Available from: https://ourworldindata.org/grapher/share-of-population-urban [Accessed 20 Feb 2024].

[63] Our world in data. Net İmigration rate, 1950 to 2021. Available from: https://ourworldindata.org/explorers/population-and-demography [Accessed 20 Feb 2024].

[64] Our world in data. Median age. Available from: https://ourworldindata.org/grapher/median-age [Accessed 20 Feb 2024].

[65] Our world in data. Forest area, 2020. Available from: https://ourworldindata.org/grapher/forest-area-km [Accessed 20 Feb 2024].

[66] Our world in data. Human Development Index. Available from: https://ourworldindata.org/grapher/human-development-index [Accessed 20 Feb 2024].

[67] Our world in data. Continents according to Our World in Data. https://ourworldindata.org/grapher/continents-according-to-our-world-in-data [Accessed 20 Feb 2024].

[68] Agresti, A. (2015). Foundations of linear and generalized linear models. John Wiley & Sons.

[69] İyit, N., Yonar, H., & Genç, A. (2016). Generalized linear models for European Union countries energy data. Acta Physica Polonica A, 130(1), 397-400.

[70] İyit, N., Yonar, H., & Yonar, A., 2021, An application of generalized linear model approach on econometric studies, In: Research & Reviews in Science and Mathematics-II, Eds, p. 181-194.

[71] Fox, J. (2015). Applied regression analysis and generalized linear models. Sage publications.

[72] Dobson, A. J., & Barnett, A. G. (2018). An introduction to generalized linear models. Chapman and Hall/CRC.

[73] Hardin, J. W., & Hilbe, J. M. (2007). Generalized linear models and extensions. Stata press.

[74] Stroup, W. W. (2012). Generalized linear mixed models: modern concepts, methods and applications. CRC press.

[75] 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.

[76] McCulloch, C. E., Searle, S. R., & Neuhaus, J. M. (2001). Generalized, linear, and mixed models (Vol. 325). New York: John Wiley & Sons.

[77] İyit, N., & Sevim, F. (2023). A novel statistical modeling of air pollution and the COVID-19 pandemic mortality data by Poisson, geometric, and negative binomial regression models with fixed and random effects. Open Chemistry, 21(1), 20230364.
[78] İyit, N., & Genc, A. (2009). Constitution of linear mixed models (LMMs) in the analysis of correlated data: random intercept model (RIM) for repeated measurements data. Journal of Modern Mathematics and Statistics, 3(3), 60-68.
[79] Faraway, J. J. (2016). Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models. Chapman and Hall/CRC.
[80] Joe, H. (2008). Accuracy of Laplace approximation for discrete response mixed models. Computational Statistics & Data Analysis, 52(12), 5066-5074.
[81] Morgan, B. J., & Smith, D. M. (1992). A note on Wadley's problem with overdispersion. Journal of the Royal Statistical Society: Series C (Applied Statistics), 41(2), 349-354.
[82] İyit, 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.
[83] Özaltın, Ö., & İyit, N. (2018). Modelling the US diabetes mortality rates via generalized linear model with the Tweedie distribution. Int. J. Sci. Res, 7(2), 1326-1334.
[84] Koenker, R., & Yoon, J. (2009). Parametric links for binary choice models: A Fisherian–Bayesian colloquy. Journal of Econometrics, 152(2), 120-130.
[85] Schwarz, G. (1978). Estimating the dimension of a model. The annals of statistics, 461-464.
[86] Akaike, H. (1974). A new look at the statistical model identification. IEEE transactions on automatic control, 19(6), 716-723.
[87] Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika, 52(3), 345-370.
[88] Cavanaugh, J. E. (1997). Unifying the derivations for the Akaike and corrected Akaike information criteria. Statistics & Probability Letters, 33(2), 201-208.
[89] Jordan, M. I., & Jacobs, R. A. (1994). Hierarchical mixtures of experts and the EM algorithm. Neural computation, 6(2), 181-214.
[90] Müller, M. (2012). Generalized linear models. Handbook of Computational Statistics: Concepts and Methods, 681-709.
[91] İyit, N. (2024). The global effects of digestive system cancers worldwide on the COVID-19 pandemic by negative binomial (mixed) regression models. Journal of Radiation Research and Applied Sciences, 17(3), 100944.
[92] Güney, E., Aydemir, A. F., İyit, N., & Alkan, Ö. (2024). Gender differences in psychological help-seeking attitudes: a case in Türkiye. Frontiers in Psychology, 15, 1289435.
[93] Tekin, K. U., Mestav, B., & İyit, N. (2021). Robust logistic modelling for datasets with unusual points. Journal of New Theory, (36), 49-63.
[94] İyit, N., Al Mashhadani, A. A. (2017). An application of generalized linear model (GLM) to child mortality data in Iraq based on socio-economic indicators. Turkish World Socio-Economic Strategies. Ed. Ugur, A., Tozak, K., & Yatbaz, A., LAP Lambert Academic Publishing, 195-203, Beau Bassin, Mauritius.
[95] Team, R. C. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing. (No Title).
[96] Team, R. C., Team, M. R. C., Suggests, M. A. S. S., & Matrix, S. (2018). Package stats. The R Stats Package.
[97] Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., ... & Bolker, M. B. (2015). Package 'lme4'. convergence, 12(1), 2.