BIOLOGY INTERNAL ASSESSMENT

An investigation into the correlation between percentage rate of deforestation and the incidence of malaria per 100,000 people in developing countries

Introduction

In light of the COVID-19 pandemic, I was interested in learning about global disease epidemiology and factors that affect disease transmission. While studying ecology in my Biology classes, I remember learning about climate change and was surprised to learn that climate change affects the transmission of diseases. I had until then never considered the direct effect of ecological changes on human health and this made me curious to see if this claim could be supported by a database investigation. Malaria is one such infectious disease and was associated with 219 million cases and 409,000 deaths globally in 2019 alone ("Malaria Fact Sheet"), and is a substantial socio-economic and health burden in many undeveloped and developing countries given the combination of the existing environmental, demographic and socio-economic factors in these countries. This problem hit especially close to home considering that in my country, India, infectious diseases account for almost half of the health burden (Dikid et al.), the second most common of which is malaria. Thus, I decided to devote this investigation to studying the correlates of malaria prevalence, in hopes of better understanding it in the larger picture.

In epidemiology, vectors are organisms that transmit infections by conveying pathogens from one host to another. Malaria is one of the most common vector-borne diseases globally, and is caused by the protozoan *Plasmodium* parasite that is transmitted to humans via infected female *Anopheles* mosquitoes. Mild symptoms consist of fever, tiredness, sweating and chills, and in severe cases it can cause anemia, breathing problems, low blood sugar and/or swelling of the brain (cerebral malaria) ("Malaria – Symptoms and Causes"). A female *Anopheles* mosquito gets infected with the *Plasmodium* parasite when it takes a blood meal from an already infected person. During a subsequent feeding, the infected mosquito injects the parasite into the bloodstream of the human host, from where the parasite first infects the liver. It then re-enters the bloodstream to grow successive broods inside red blood cells and then destroys them, releasing daughter parasites or merozoites that in turn, continue the cycle by invading other red blood cells ("CDC - Malaria - About Malaria – Biology"). Currently, no effective vaccine against malaria has been approved for use but certain prevention measures are taken to reduce risk of transmission, which are mainly focused on preventing contact between mosquitoes and humans for instance destroying adult mosquitoes by indoor residual spraying and insecticide-treated nets or killing mosquito larvae using larvicides (Wangdi et al.).

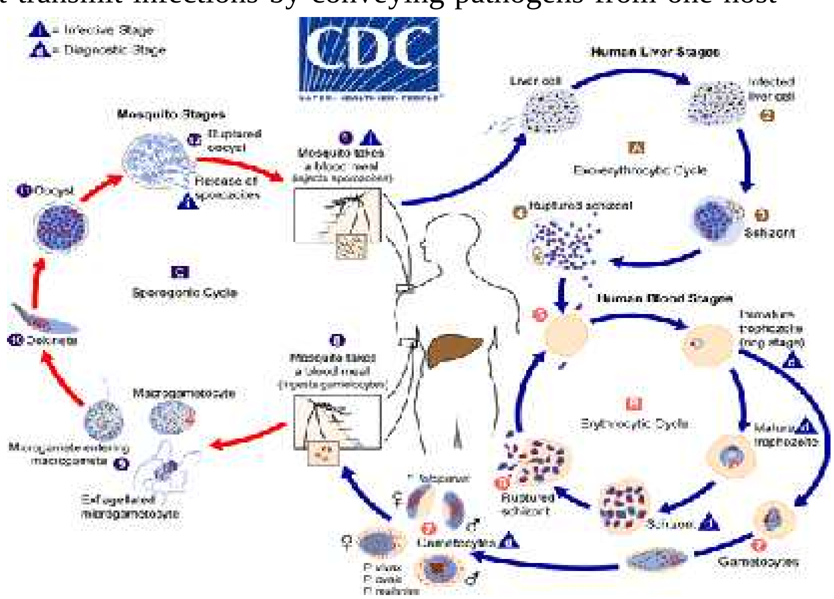

Figure 1: The life cycle of malaria (CDC)

The loss of natural forest cover, or deforestation, has been increasing at an unprecedented rate in recent years; studies show that about 46% of all trees have been cut since humans started deforestation activities (Crowther et al.). It is often followed by land use change for agricultural development, urbanization or mining activities and impacts every component of an ecosystem including its microclimate, soil and water conditions, and the ecology of local flora and fauna, including human disease vectors (Rejmánková et al.). Some studies show that deforestation can increase malaria risk factors in some settings (Kweka et al.), which implies that forest conservation could be a potential preventive measure for malaria. However, additional research is required to establish the definite impact of deforestation on malaria incidence in humans.

Understanding the factors that affect malaria transmission is essential to prevent further risk of infectious diseases and protect public health. I specifically chose to investigate this research question because as anthropogenic changes to the natural environment have only been increasing over time, it has become all the more important to conduct research investigating how human activities may be changing ecosystem conditions in ways that harm human life in the long run. Investigating this topic can have important implications for formulating policies regarding forest conservation and the prevention of infectious diseases; if deforestation amplifies risk of malaria for humans, forest conservation will then reduce the transmission of malaria. Therefore, this investigation aims to answer the following research question: **What is the correlation between percentage rate of deforestation and the incidence of malaria per 100,000 people in developing countries?**

Variables

Independent variable: Percentage rate of deforestation.

Dependent variable: Incidence of malaria per 100,000 people.

Controlled Variables

To ensure that the results of this investigation are as accurate as possible, certain controls were employed to reduce the effect of confounding variables.

Table 1: Variables that were controlled

Control variable	Reason for control	How it has been controlled
Population	Countries with relatively small populations may have higher variability in malaria incidence that less closely approximates the true conditions of the country. This may give results that are highly flawed and unrepresentative.	Only countries with a minimum population of 5 million were included in the sample.
Geographical location	Rates of the *Plasmodium* parasite and *Anopheles* mosquito life cycles, the frequency of mosquito blood meals and the rate at which parasites are acquired by mosquitoes have all shown to be influenced by changes in temperature (Siraj et al.; Beck-Johnson et al.).	Only countries that are located in the Torrid Zone (23°27' N–23°27' S) were selected for analysis as they are assumed to have similar climatic conditions. This region was chosen in particular as most malaria-endemic

	This means that the transmission of malaria likely differs between tropical and cooler regions.	countries happen to be located here.
Economic status	The incidence of malaria and the level of development of countries are linked, with malaria rates being the highest in the poorest nations (Stratton). Countries having a higher economic status will find it easier to prevent the spread of infectious diseases due to better healthcare systems, higher levels of education, better prevention measures etc. On the other hand, less-developed countries are less likely to have these facilities to the same extent; many third-world citizens also lack access to clean water or sanitation, increasing the risk of mosquito habitats that can spread malaria.	Only nations categorized as 'developing countries' by the International Monetary Fund (Word Economic Outlook) were included in the sample as malaria chiefly affects people in lower-income countries.
Area under forest cover	Meaningful inferences regarding deforestation can only be made in countries with significant area of forest land.	Countries that mostly consist of desert land (e.g. Libya) were excluded from the analysis.

A positive correlation can be predicted between the percentage rate of deforestation and the incidence of malaria per 100,000 people in developing countries. This is because, as shown by Vittor et al. (2006), deforestation alters the biotic and abiotic conditions of the surrounding ecosystem in a way that creates habitats conducive for *Anopheles* mosquitoes to breed, develop, and transmit disease.

Thus, the hypotheses formulated for statistical testing are as follows:

- **Null hypothesis (H_0):** There is no correlation between the percentage rate of deforestation and incidence of malaria per 100,000 people in developing countries.
- **Alternative hypothesis (H_1):** There is a positive correlation between the percentage rate of deforestation and the incidence of malaria per 100,000 people in developing countries.

As this is a database investigation, there are no significant ethical, safety or environmental considerations. However, a small detail to note is that all data was sourced ethically with due credit given and in accordance with the guidelines set by the sources.

Methodology

After controlling for the possible confounding variables presented in Table 1, 40 developing countries were available for analysis. A large sample size is essential for obtaining results with a small margin of error, thus random sampling was not conducted and all countries that fit the control criteria were selected; this will help give a clearer idea of the relationship between the two variables.

Data was sourced from the following online publicly available databases known for the credibility and reliability of their data:

1. World Health Organization database for the number of reported confirmed cases of Malaria per country (available at: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/malaria---number-of-reported-confirmed-cases)
2. World Bank database for total population per country (available at: https://data.worldbank.org/indicator/SP.POP.TOTL)
3. Food and Agriculture Organization database for forest land per country (available at: http://www.fao.org/faostat/en/#data/GF)

These sources in particular were chosen as they are run by reputable, well-known international organizations and provide rich data on the variables, and I assume that the information they have provided is reliable.

Initially, I planned on comparing the values for rates of deforestation and incidence of malaria of countries for one year, but this could produce erroneous results as the results of one year wouldn't be entirely representative of the true conditions of the country. Thus, to improve accuracy and get a sufficient understanding of the correlation between the two variables, data for 5 years between 2013 and 2017 were extracted and averaged. These years in particular were selected as they were the most recently available data for cases of malaria. Accordingly, the data for population and forest land from the same years were used. The raw data was extracted from the respective sources listed above and logged into the Microsoft Excel program for processing.

Table 2: Raw data for reported confirmed cases of malaria and the population of each country from 2013-2017

Country	**Cases of malaria**					**Population**				
	2017	**2016**	**2015**	**2014**	**2013**	**2017**	**2016**	**2015**	**2014**	**2013**
Angola	3874892	3794253	2769305	2298979	1999868	29,816,750	28842480	27,884,380	26,941,780	26,015,780
Bangladesh	4893	4787	6608	10216	3864	159670590	157970840	156256280	154,520,170	152764680
Benin	1573163	1324576	1268347	1044235	1078834	11,175,200	10872070	10,575,950	10,286,840	10,004,590
Bolivia	4572	5542	6874	7401	7301	11,192,850	11031810	10,896,730	10,706,520	10,542,380
Brazil	189503	124178	138229	139272	168862	207833830	206163060	204471770	202,763,730	201035900
Cameroon	1191257	1675264	1162784	n/a	26651	24,566,040	23926540	23,298,370	22,681,860	22,077,300
China	0	3	39	56	86	1386395000	1378665000	1371220000	1364270000	1357380000
Colombia	52805	82609	47616	40768	51696	48,909,840	48,175050	47,520,670	46,967,710	46,495,490
Dem. Rep. of the Congo	15176927	15330841	11627473	9968983	6715223	81,398,760	78789130	76,244,540	73,767,450	71,385,810
Ivory Coast	3274683	3471024	3375904	3712831	2506953	24,437,470	23822710	23,226,140	22,647,680	22,087,510
Dominican Republic	341	690	631	459	579	10,513,130	10397740	10,281,680	10,165,180	10,048,220
Ecuador	1275	1191	618	242	368	16,785,360	16491120	16,212,020	15,951,840	15,707,470
Ethiopia	1530739	1718504	1867059	2118815	2645454	106400020	103603500	100835460	98,094,250	95,385,790
Guatemala	3743	4853	6836	5685	6214	16,087,420	15827690	15,567,420	15,306,320	15,043,980
Guinea	1335323	992146	810979	660207	211257	12,067,540	11738440	11,432,090	11,150,980	10,892,810
Haiti	19135	21430	17583	17696	20957	10,982,370	10839970	10,695,540	10,549,010	10,400,670
Honduras	1277	4094	3575	3378	5428	9,429,010	9,270,800	9,112,920	8,955,590	8,798,520
India	844558	1087285	1169261	1102205	881730	1338658830	1324509590	1310152400	1295604180	1280846130
Indonesia	261617	218450	217025	252027	343527	264645890	261554230	258383260	255129000	251806400
Laos	9333	11223	36056	48071	38131	6,953,030	6,845,850	6,741,160	6,639,760	6,541,300
Madagascar	800661	475333	744103	377963	385598	25,570,540	24894380	24,234,090	23,589,890	22,961,250
Malawi	4901344	4827373	3661238	2905310	1280892	17,670,260	17205290	16,745,300	16,289,540	15,839,270
Malaysia	85	266	242	3147	2921	31,105,030	30684800	30,270,960	29,866,560	29,468,870

Myanmar	19619	110146	182768	205658	333871	53,382,580	53045230	52,680,730	52,280,810	51,852,450
Nepal	623	507	591	832	1974	27,617,120	27261130	27,015,030	26,906,930	26,917,910
Niger	2638580	4148167	2272000	1953309	2353422	21,602,470	20788840	20,001,660	19,240,160	18,504,260
Nigeria	11571958	9234387	6850782	7826954	n/a	190873310	185960290	181137450	176,404,900	171765770
Papua New Guinea	478340	478497	297787	281182	279994	8,438,030	8,271,760	8,107,770	7,946,730	7,788,380
Peru	55367	56623	61865	65252	48719	31444300	30926030	30,470,730	30,090,360	29,773,990
Philippines	3950	6680	8266	4903	6514	105173260	103663930	102113210	100,513,140	98,871,550
Sierra Leone	1651236	1775306	1483376	1374476	1701958	7488430	7328840	7171910	7,017,140	6,863,980
South Africa	22061	4323	555	11705	8645	57,000,450	56203650	55386370	54,545,990	53,689,240
Sudan	720879	575015	586827	1068506	592383	40813400	39847440	38902950	37,977,650	37,072,550
Tajikistan	0	0	0	2	3	8,880270	8,663,580	8,454,030	8,252,830	8,059,770
Tanzania	5354486	5193520	4241364	680442	1551923	54,663,910	53050790	51482630	49,959,820	48,482,270
Thailand	11440	11522	8022	37921	33302	69,209,860	68971330	68714510	68,438,730	68,155,500
Uganda	11667831	9385132	7137662	3631939	1502362	41,162,460	39647510	38225450	36,912,150	35,695,250
Venezuela	411586	240613	136402	90708	78643	29,390,410	29846180	30081830	30,045,130	29,783,570
Vietnam	4548	4161	9331	15752	17128	94,596,640	93638720	92677080	91,714,600	90,753,470
Zimbabwe	315624	279988	391651	535931	422633	14,236,750	14030390	13814630	13,586,680	13,350,360

Table 3: Raw data for forest land (in hectares) per year

Country	Forest area (in hectares)					
	2017	2016	2015	2014	2013	2012
Angola	68272566	68827628	69382690	69937752	70492814	71047876
Bangladesh	1885376.8	1883400	1883400	1884388.4	1883400	1886365.2
Benin	3285150	3335150	3385150	3435150	3485150	3535150
Bolivia	51549836	51788528	52027220	52238978	52450736	52662494
Brazil	500978720	502431760	503884800	505423980	506963160	508502340
Cameroon	20508480	20564480	20620480	20676480	20732480	20788480
China	214167822	212231036	210294250	208357476	206420702	204483928
Colombia	59737560	59936110	60134660	60269286	60403912	60538538
Dem. Rep. of the Congo	129459368	130560744	131662120	132763496	133864872	134966248
Ivory Coast	3175374	3288262	3401150	3514036	3626922	3739808
Dominican Rep.	2119830	2111740	2103650	2097544	2091438	2085332
Ecuador	12690640	12754910	12819180	12860982	12902784	12944586
Ethiopia	17287500	17360500	17433500	17506500	17579500	17652500
Guatemala	3562600	3574200	3585800	3613160	3640520	3667880
Guinea	6309000	6349000	6389000	6425000	6461000	6497000
Haiti	356630	359740	362850	365956	369062	372168
Honduras	6422002	6442916	6463830	6486158	6508486	6530814
India	71360800	71094400	70828000	70561600	70295200	70028800
Indonesia	93870020	94448960	95027900	95954160	96880420	97806680
Laos	16699000	16733500	16768000	16802500	16837000	16871500
Madagascar	12469458	12482674	12495890	12509108	12522326	12535544
Malawi	2367700	2409700	2451700	2493700	2535700	2577700
Malaysia	19324148	19394184	19464220	19360906	19257592	19154278
Myanmar	29413020	29702730	29992440	30282152	30571864	30861576
Nepal	5962030	5962030	5962030	5962030	5962030	5962030
Niger	1116960	1129380	1141800	1154220	1166640	1179060
Nigeria	22116850	22280150	22443450	22606756	22770062	22933368
Papua New Guinea	35956956.5	35990688.7	36024420.9	36055313.8	36086206.6	36117099.5
Peru	72848866	73021698	73194530	73365584	73536638	73707692
Philippines	7083926	7049038	7014150	6979264	6944378	6909492
Sierra Leone	2594070	2613800	2633530	2653256	2672982	2692708
South Africa	17159290	17195690	17232090	17268490	17304890	17341290

Sudan	18869778	19039854	19209930	19384178	19558426	19732674
Tajikistan	422600	422200	421800	419440	417720	414720
Tanzania	47152000	47621000	48090000	48462002	48834004	49206006
Thailand	19985800	20023400	20061000	20063400	20065800	20068200
Uganda	2461656	2502908	2544160	2585412	2626664	2667916
Venezuela	46502160	46592580	46683000	46847400	47011800	47176200
Vietnam	14294352	14178106	14061860	13927100	13792340	13657580
Zimbabwe	17582790	17628860	17674930	17721000	17767070	17813140

As the cases of malaria are relative to the population of the country, the data had to be standardized. This was done by calculating the incidence of malaria per 100,000 people for each year. The annual malaria cases of a country were divided by the population of the country in that given year, and the obtained value was finally multiplied by 100,000. For example, the incidence of malaria per 100,000 people in Angola for the year 2017 was calculated as follows:

$$Incidence\ of\ malaria\ per\ 100{,}000\ people\ in\ 2017 = \frac{Cases\ of\ malaria}{Total\ population} \times 100{,}000$$

$$Incidence\ of\ malaria\ per\ 100{,}000\ people\ in\ 2017 = \frac{3{,}874{,}892}{29{,}816{,}750} \times 100{,}000$$

$$\therefore Incidence\ of\ malaria\ per\ 100{,}000\ people\ in\ 2017 = 12995.69$$

Furthermore, in any given year, there may have been unknown factors that influenced the malaria incidence data in a given country (for example, errors in sampling or the introduction of new malaria prevention measures). To account for such random errors, the values across 5 sampling years were averaged to help get a clearer and more accurate idea of the incidence of malaria in each country.

Table 4: Raw data for incidence of malaria per 100,000 people for each year and the average incidence of malaria per 100,000 people from 2013-2017 for each country

Country	**Incidence of malaria per 100,000 people**					**Average incidence of malaria per 100,000 people**
	2017	**2016**	**2015**	**2014**	**2013**	
Angola	12995.69	13155.09	9931.38	8533.14	7687.13	10460.49
Bangladesh	3.06	3.03	4.23	6.61	2.53	3.89
Benin	14077.27	12183.29	11992.75	10151.17	10783.39	11837.58
Bolivia	40.85	50.24	63.08	69.13	69.25	58.51
Brazil	91.18	60.23	67.60	68.69	84.00	74.34
Cameroon	4849.20	7001.70	4990.84	n/a	120.72	3392.49
China	0.00000	0.000218	0.00284	0.00410	0.00634	0.0027
Colombia	107.96	171.48	100.20	86.80	111.18	115.53
Democratic Rep. of the Congo	18645.16	19458.07	15250.24	13514.07	9406.94	15254.89
Ivory Coast	13400.25	14570.23	14534.93	16393.87	11350.09	14049.88
Dominican Rep.	3.24	6.64	6.14	4.52	5.76	5.26
Ecuador	7.60	7.22	3.81	1.52	2.34	4.50
Ethiopia	1438.66	1658.73	1851.59	2159.98	2773.43	1976.48
Guatemala	23.27	30.66	43.91	37.14	41.31	35.26
Guinea	11065.41	8452.11	7093.88	5920.62	1939.42	6894.29
Haiti	174.23	197.69	164.40	167.75	201.50	181.11
Honduras	13.54	44.16	39.23	37.72	61.69	39.27
India	63.09	82.09	89.25	85.07	68.84	77.67

Indonesia	98.86	83.52	83.99	98.78	136.43	100.32
Laos	134.23	163.94	534.86	723.99	582.93	427.99
Madagascar	3131.19	1909.40	3070.48	1602.22	1679.34	2278.53
Malawi	27737.81	28057.49	21864.27	17835.43	8086.81	20716.37
Malaysia	0.27	0.87	0.80	10.54	9.91	4.48
Myanmar	36.75	207.65	346.94	393.37	643.89	325.72
Nepal	2.26	1.86	2.19	3.09	7.33	3.35
Niger	12214.25	19953.82	11359.06	10152.25	12718.27	13279.53
Nigeria	6062.64	4965.78	3782.09	4436.93	n/a	3849.48
Papua New Guinea	5668.86	5784.71	3672.86	3538.34	3595.02	4451.96
Peru	176.08	183.09	203.03	216.85	163.63	188.54
Philippines	3.76	6.44	8.09	4.88	6.59	5.95
Sierra Leone	22050.50	24223.56	20683.14	19587.41	24795.50	22268.02
South Africa	38.70	7.69	1.00	21.46	16.10	16.99
Sudan	1766.28	1443.04	1508.44	2813.51	1597.90	1825.83
Tajikistan	0.00	0.00	0.00	0.02	0.04	0.0123
Tanzania	9795.29	9789.71	8238.44	1361.98	3201.01	6477.29
Thailand	16.53	16.71	11.67	55.41	48.86	29.84
Uganda	28345.81	23671.43	18672.54	9839.41	4208.86	16947.61
Venezuela	1400.41	806.18	453.44	301.91	264.05	645.20
Vietnam	4.81	4.44	10.07	17.18	18.87	11.07
Zimbabwe	2216.97	1995.58	2835.05	3944.53	3165.70	2831.57

Note that data for the cases of malaria for Cameroon in the year 2014 and Nigeria in 2013 were not available, so the average incidence rate for these countries were calculated using the values of the remaining 4 years.

Accordingly, to calculate the annual rate of deforestation, the following formula recommended by the Food and Agriculture Organization (1995) was used:

$$Annual\ rate\ of\ deforestation = \left(\frac{A_1}{A_2}\right)^{1-(t_1/t_2)} - 1$$

Where A_1 and A_2 are the areas of forest cover in the country at reference years t_1 and t_2 respectively. For example, the rate of deforestation in Angola for the year 2017 was calculated as follows:

$$Annual\ rate\ of\ deforestation = \left(\frac{68{,}272{,}566}{68{,}827{,}628}\right)^{1-(2016/2017)} - 1$$

$$\therefore Annual\ rate\ of\ deforestation = 0.0000040145$$

As the values obtained are very small the deforestation rates were then converted into percentage form to make it easier to compare between countries.

$$Annual\ percentage\ rate\ of\ deforestation = \frac{0.0000040145}{1} \times 100$$

$$\therefore Annual\ percentage\ rate\ of\ deforestation = 0.00040145\%$$

The value can be either positive (representing a loss of forest cover or deforestation), negative (denoting a gain or afforestation) or zero, showing no change. These calculations were applied to all 40 countries across the 5 sampling years, and the data obtained was then averaged and rounded to 3 significant figures.

Table 5: Calculated deforestation percentage rates for each year and the average percentage rate of deforestation from 2013-2017 for each country

Country	Percentage rate of deforestation					Average deforestation rate (in %)
	2017	2016	2015	2014	2013	
Angola	0.000401%	0.000398%	0.000395%	0.000392%	0.000389%	0.000395%
Bangladesh	-0.0000520%	0.0000000%	0.0000260%	-0.0000260%	0.0000780%	0.000005%
Benin	0.000749%	0.000738%	0.000727%	0.000716%	0.000706%	0.000727%
Bolivia	0.000229%	0.000228%	0.000201%	0.000201%	0.000200%	0.000212%
Brazil	0.000144%	0.000143%	0.000151%	0.000151%	0.000150%	0.000148%
Cameroon	0.000135%	0.000135%	0.000134%	0.000134%	0.000134%	0.000134%
China	-0.000450%	-0.000455%	-0.000459%	-0.000463%	-0.000467%	-0.000459%
Colombia	0.000165%	0.000164%	0.000111%	0.000111%	0.000110%	0.000132%
Democratic Republic of the Congo	0.000420%	0.000416%	0.000413%	0.000410%	0.000406%	0.000413%
Ivory Coast	0.001732%	0.001674%	0.001619%	0.001568%	0.001520%	0.001622%
Dominican Republic	-0.000190%	-0.000190%	-0.000144%	-0.000145%	-0.000145%	-0.000163%
Ecuador	0.000250%	0.000249%	0.000161%	0.000161%	0.000160%	0.000196%
Ethiopia	0.000209%	0.000208%	0.000207%	0.000206%	0.000205%	0.000207%
Guatemala	0.000161%	0.000161%	0.000377%	0.000374%	0.000371%	0.000289%
Guinea	0.000313%	0.000311%	0.000279%	0.000277%	0.000275%	0.000291%
Haiti	0.000430%	0.000427%	0.000423%	0.000419%	0.000416%	0.000423%
Honduras	0.000161%	0.000161%	0.000171%	0.000170%	0.000170%	0.000167%
India	-0.000185%	-0.000186%	-0.000187%	-0.000188%	-0.000188%	-0.000187%
Indonesia	0.000305%	0.000303%	0.000481%	0.000476%	0.000472%	0.000407%
Laos	0.000102%	0.000102%	0.000102%	0.000102%	0.000101%	0.000102%
Madagascar	0.000053%	0.000052%	0.000052%	0.000052%	0.000052%	0.000052%
Malawi	0.000872%	0.000857%	0.000842%	0.000828%	0.000814%	0.000843%
Malaysia	0.000179%	0.000179%	-0.000264%	-0.000265%	-0.000267%	-0.000088%
Myanmar	0.000486%	0.000481%	0.000477%	0.000472%	0.000468%	0.000477%
Nepal	0.000000%	0.000000%	0. 000000%	0.000000%	0.000000%	0.000000%
Niger	0.000548%	0.000542%	0.000536%	0.000531%	0.000525%	0.000537%
Nigeria	0.000365%	0.000362%	0.000359%	0.000357%	0.000354%	0.000359%
Papua New Guinea	0.0000465%	0.0000464%	0.0000425%	0.0000425%	0.0000424%	0.000044%
Peru	0.000117%	0.000117%	0.000116%	0.000115%	0.000115%	0.000116%
Philippines	-0.000245%	-0.000246%	-0.000247%	-0.000248%	-0.000250%	-0.000247%
Sierra Leone	0.000376%	0.000373%	0.000370%	0.000367%	0.000365%	0.000370%
South Africa	0.000105%	0.000105%	0.000105%	0.000104%	0.000104%	0.000105%
Sudan	0.000445%	0.000441%	0.000448%	0.000444%	0.000440%	0.000443%
Tajikistan	-0.000047%	-0.000047%	-0.000278%	-0.000204%	-0.000357%	-0.000187%
Tanzania	0.000491%	0.000486%	0.000382%	0.000379%	0.000376%	0.000423%
Thailand	0.000093%	0.000093%	0.000006%	0.000006%	0.000006%	0.000041%
Uganda	0.000824%	0.000810%	0.000797%	0.000785%	0.000773%	0.000798%
Venezuela	0.0000963%	0.0000961%	0.000174%	0.000174%	0.000173%	0.000143%
Vietnam	-0.000405%	-0.000408%	-0.000477%	-0.000482%	-0.000487%	-0.000452%
Zimbabwe	0.000130%	0.000129%	0.000129%	0.000129%	0.000128%	0.000129%

Finally, the values for average rate of deforestation and incidence of malaria per 100,000 people for each country were plot on a scatter plot using Excel, with average rate of deforestation on the y-axis and incidence of malaria on the x-axis. The regression line was found along with the R^2 value using Excel. The values for average incidence of malaria in the dataset were very widespread, ranging from 0.0027 to 15038.86, and as a result produced a skewed graph; so the data points were instead plot on a logarithmic scale with base 10 on the x-axis to obtain a suitable scatter plot where each data point was easily visible. Note that this does not affect the R^2 value in any way.

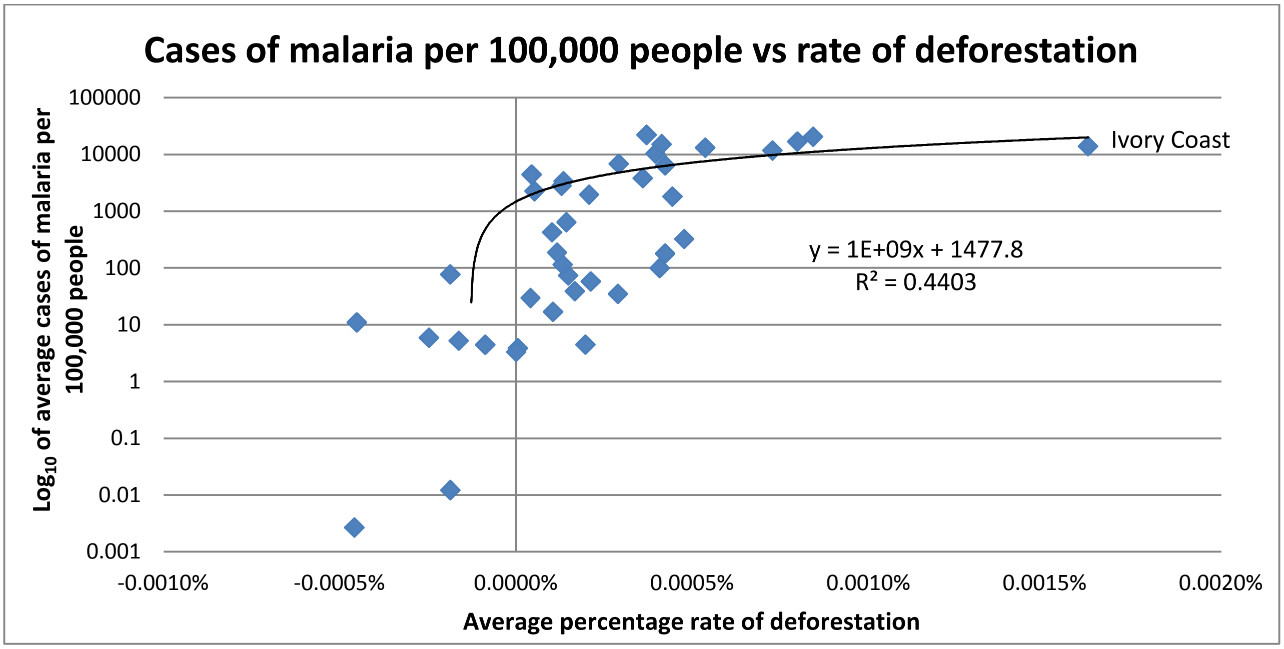

Figure 2: Correlation between $\log_{10}$ of incidence of malaria per 100,000 people & average percentage rate of deforestation from 2013-2017

Visually, there seems to be a weak positive linear correlation between the two variables. A slight increase in average incidence of malaria can be observed as rate of deforestation increases, as can also be seen via the regression line. There is high variation between data points, with very few values falling close to the regression line, while the majority is scattered below it. This is supported by the coefficient of determination R^2 produced (R^2= 0.4403) according to which only 44.03% of the variation in average cases of malaria per 100,000 people can be explained in a linear direction by the average percentage rate of deforestation, which means it is unlikely that the two variables are linearly correlated. The presence of an outlier belonging to Ivory Coast can be made out in the far right corner of the graph; this will be further discussed later on in the analysis.

A test for skewness was then conducted to determine the normality of the data in order to understand which statistical test would be the most suitable for further analyzing the data. Using Excel, the skewness coefficients were calculated for both variables; skewness coefficients for the incidence of malaria per 100,000 people and average rate of deforestation were found to be 1.66 and 1.28 respectively. As both values are higher than 1, they show positive skewness of both sets of data points (GoodData). Thus, the Spearman's correlation test is best suited for statistical testing as it is appropriate for testing non-incremental data that is not normally distributed. Additionally, this test is also less sensitive to outliers as compared to other forms of statistical testing and is not only restricted to testing linear relationships but monotonic relationships as well which is important because as shown by the R^2 value, a linear correlation is unlikely to be present between the two variables.

Using Excel, the Spearman's correlation coefficient, r_s, calculated is 0.742. As the value is in the range of 0.500 and 0.750, it shows a moderate positive correlation between average incidence of malaria per 100,000 people and average rate of deforestation. Furthermore, in order to determine whether this correlation is statistically significant or not, hypothesis testing will be conducted.

The critical value r_{crit} for 40 data points with a degree of freedom of 38 and testing at a significance level of 5% (α=0.05) is 0.271 (University of York). Comparing the Spearman's correlation coefficient to the r_{crit} value, we find that the correlation coefficient is higher (0.742 > 0.271), which means that the null hypothesis is rejected and we can conclude with the alternative hypothesis stating that there is a positive correlation between the percentage rate of deforestation and the incidence of malaria per 100,000 people in developing countries.

Discussion

In Figure 2, a single outlier belonging to Ivory Coast can be identified at the top right corner of the scatter plot. Ivory Coast has high average percent rates of deforestation (0.0016%) as compared to other countries with similar average values of malaria incidence (14049.88), which means that deforestation has not greatly influenced the malaria incidence in this country. Malaria has likely been controlled to some extent in this country, reducing the effect of deforestation on malaria incidence. As the geographical and ecological conditions of Ivory Coast such as altitude, rainfall and average yearly temperature are similar to those of other countries in Sub-Saharan Africa utilized in the sample, climatic conditions are not the main factor in this regard either. Although through online research I was not able to gather relevant information on the national health policies of Ivory Coast to support this assertion, it is likely that public health conditions (e.g. relatively higher immunity of the population, better sanitation and access to clean water etc.) and malaria intervention strategies are responsible for this difference.

A possible explanation of the positive correlation obtained (r_s= 0.742) is that deforestation alters the meteorological conditions of the surrounding ecosystem in a way that favors the growth of *Anopheles* mosquito and *Plasmodium* parasite populations. The loss of temperature regulation services previously provided by forest cover through cooling and precipitation (as a result of transpiration) results in an increase in temperature and a decrease in humidity of the area. Mosquitoes are cold-blooded organisms, because of which each life stage of theirs is especially sensitive to changes in climatic conditions ("Vector-Borne Disease"). Warmer and drier microclimates favor mosquito growth; these conditions enable them to feed and lay eggs more often and increase rates of mosquito development and reproduction (Rejmánková et al.). Higher temperatures also directly affect the life cycle of the *Plasmodium* parasite, reducing its incubation time in the gut of the mosquito and making mosquitoes infectious quicker (Gilles).

Furthermore, deforestation causes the loss of insectivores that help control mosquito populations and reduces diversity in "dead-end hosts"–wild warm-blooded animals that cannot develop high levels of the *Plasmodium* parasite in their bloodstream, and as a result do not pass the parasite on to other biting mosquitoes. Such animals thus provide an indirect method of malaria control by reducing the chances of infective bites in humans (Laporta et al.). Cleared forest land can also raise the availability of stagnant surface-water by causing depressions in the forest floor that hold water, causing an increase in new breeding sites for *Anopheles* mosquitoes. Land-use changes can increase risk of malaria due to the adaptation of vectors to newly created niches, for example when the vector adapts to arid land caused due to desertification. Additionally, when urbanization occurs on deforested land where there are still *Anopheles* mosquito breeding sites it can increase the exposure of people to the mosquito, increasing the chances of getting infected.

Evaluation

There are certainly some strengths of the methodology of this investigation, including a sufficient sample size consisting of countries with a minimum population of 5 million to give representative results and reducing error variability by using the average of data points of the two variables over 5 years. The use of Spearmann's correlation test helped ensure that the results weren't affected by errors that may have occurred due to averaging of values or anomalous data points and hypothesis testing helped determine whether the correlation obtained between the two variables occurred by chance or not. All three databases used (WHO, FAO and World Bank) contained relevant information that I needed for this investigation and were quite convenient to use as they allowed me to selectively view data for the countries and years I wanted.

However given the epidemiology of vector-borne diseases, a wide range of variables play a role in the spread of malaria and it is difficult to control for all these factors in a database investigation. Thus, these limitations have been summarized in the table below along with suggestions for improvement.

Table 6: Limitations of the present study

Limitation	How it affects the investigation	Suggestions for improvement
Lack of control for climatic factors	Although there was an attempt to control for climate by selecting countries in the same geographic zone, weather conditions still differ from country to country and more specifically, region to region even within countries due to factors such as altitude, presence of large water bodies etc. that can affect life cycle of *Plasmodium* parasite and *Anopheles* mosquitoes and thus the transmission of malaria. This may have caused variance in the results of the present study.	Larger malaria-endemic countries (e.g. India, Nigeria etc.) could be separated into their constituent states to account for climatic factors that differ even within countries and data for deforestation rates and incidence of malaria in these states could be collected and compared. This would also allow a more in-depth analysis.
Lack of control for sociocultural factors	Sociocultural factors such as education level, GDP per capita, rural population growth and public health conditions such as fertility rates and access to clean water have shown to be associated with malaria incidence (Austin et al.) and may have been responsible for the variance in the results.	Countries could be grouped based on sociocultural factors such as minimum schooling years, average number of births per woman, GDP per capita etc. to further standardize the sample, despite the fact that it may limit the sample size.

Impact of different land use changes not accounted for	Different land use changes after deforestation can have varying impacts on mosquito survival and are an important factor to consider. For instance, Paul et al. (2018) found that land use changes for agriculture, settlements and developmental projects all increase habitats for malaria infected mosquitoes.	The analysis can be divided on the basis of different land use changes (cultivation, grazing, urbanization etc.) and the impact of each on malaria incidence can be compared.

Conclusion

This investigation aimed to answer the following research question: **What is the correlation between percentage rate of deforestation and the incidence of malaria per 100,000 people in developing countries?** The results support the original hypothesis stating that there is a positive correlation between the percentage rate of deforestation and incidence of malaria per 100,000 people in developing countries. These results are consistent with the findings of Vittor et al. (2006), which found a close positive association between deforestation and breeding rates of potential malaria vectors. Although the present study was not investigating breeding rates, an increase in breeding rates of malaria vectors will surely increase the subsequent transmission of malaria and thus support the results of this study. The findings can be used as an important tool to help guide forest conservation policies in order to prevent the amplification of infectious diseases in developing countries. They also illustrate the importance of maintaining balance in our ecosystems, as change in any one factor of an environment can heavily impact the others in a process of ecological succession.

An alternative approach to this study can be examining the changes in deforestation rates and the incidence of malaria of a particular country over time through a longitudinal analysis. This would allow us to establish whether there is a clear direct correlation between the deforestation increasing or decreasing and incidence of malaria in a given country. It would also be easier to account for confounding variables and eliminate discrepancies. Additionally, I think an interesting extension would be to examine the effect of loss of biodiversity in particular on malaria incidence to test the 'Dilution-Effect Theory', which postulates that higher species-diversity results in a reduction in the prevalence of an infectious disease (Keesing et al.). Although deforestation somewhat accounts for loss of biodiversity as well, its direct impact is not observed in the present study. Since malaria is a vector-borne disease and its transmission also relies on surrounding ecology, it would be interesting to observe the effect of fluctuations in biodiversity on the impact of this transmission.

Bibliography

Ahmed, Toqeer et al. “Climatic Conditions: Conventional and Nanotechnology-Based Methods for the Control of *Mosquito* Vectors Causing Human Health Issues.” *International journal of environmental research and public health* vol. 16,17 3165. 30 Aug. 2019, doi:10.3390/ijerph16173165. Accessed 16 March 2021.

Austin, Kelly & Bellinger, Megan & Rana, Priyokti. (2017). *Anthropogenic forest loss and malaria prevalence: a comparative examination of the causes and disease consequences of deforestation in developing nations*. AIMS Environmental Science. 4. 217-231. 10.3934/environsci.2017.2.217.

Bauhoff, Sebastian, and Jonah Busch. “Does Deforestation Increase Malaria Prevalence? Evidence from Satellite Data and Health Surveys.” *World Development*, vol. 127, 2020, p. 104734., doi:10.1016/j.worlddev.2019.104734. Accessed 16 March 2021.

Beck-Johnson, Lindsay M et al. (2013). The effect of temperature on Anopheles mosquito population dynamics and the potential for malaria transmission. *PloS one* vol. 8,11 e79276. 14 Nov. 2013, doi:10.1371/journal.pone.0079276. Accessed 11 January 2021.

CDC - Malaria - About Malaria - Biology. Cdc.gov. 2020. Available at: https://www.cdc.gov/malaria/about/biology/index.html. Accessed 18 January 2021.

Crowther, T., Glick, H., Covey, K. *et al.* Mapping tree density at a global scale. *Nature* 525**,** 201–205 (2015). https://doi.org/10.1038/nature14967

Dikid, T et al. “Emerging & re-emerging infections in India: an overview.” *The Indian journal of medical research* vol. 138,1 (2013): 19-31. Accessed 16 March 2021.

FAO. (1995). Forest resources assessment 1990. *Global Synthesis*. FAO, Rome

F. Keesing, L.K. Belden, P. Daszak, A. Dobson, C.D. Harvell, R.D. Holt, P. Hudson, A. Jolles, K.E. Jones, C.E. Mitchell, S.S. Myers, et al. (2010). *Impacts of biodiversity on the emergence and transmission of infectious diseases***.** Nature, 468 (2010), pp. 647-652.

Gilles HM . 1999. Epidemiology. Oxford University Press.

GoodData Documentation. (n.d). *Normality Testing - Skewness and Kurtosis.* help.gooddata.com/doc/en/reporting-and-dashboards/maql-analytical-query-language/maql-expression-reference/aggregation-functions/statistical-functions/predictive-statistical-use-cases/normality-testing-skewness-and-kurtosis. Accessed 10 January 2021.

Kweka, Eliningaya J. et al. "Effect Of Deforestation And Land Use Changes On Mosquito Productivity And Development In Western Kenya Highlands: Implication For Malaria Risk". *Frontiers In Public Health*, vol 4, 2016. *Frontiers Media SA*, doi:10.3389/fpubh.2016.00238. Accessed 14 March 2021.

Laporta, Gabriel Zorello et al. "Biodiversity Can Help Prevent Malaria Outbreaks In Tropical Forests". *Plos Neglected Tropical Diseases*, vol 7, no. 3, 2013, p. e2139. *Public Library Of Science (Plos)*, doi:10.1371/journal.pntd.0002139. Accessed 16 March 2021.

Mayo Clinic. 2018. *Malaria - Symptoms And Causes*. Available at: https://www.mayoclinic.org/diseases-conditions/malaria/symptoms-causes/syc-20351184.Accessed 22 January 2021.

MacDonald, Andrew J., and Erin A. Mordecai. "Amazon Deforestation Drives Malaria Transmission, And Malaria Burden Reduces Forest Clearing". *Proceedings Of The National Academy Of Sciences*, vol 116, no. 44, 2019, pp. 22212-22218. *Proceedings Of The National Academy Of Sciences*, doi:10.1073/pnas.1905315116. Accessed 10 March 2021.

Oaks SC Jr., Mitchell VS, Pearson GW, et al. *Malaria: Obstacles and Opportunities.* Institute of Medicine (US) Committee for the Study on Malaria Prevention and Control, Washington (DC): National Academies Press (US); 1991. 12, *Social and Behavioral Aspects of Malaria*.

Paul, P., Kangalawe, R.Y.M. & Mboera, L.E.G. Land-use patterns and their implication on malaria transmission in Kilosa District, Tanzania. *Trop Dis Travel Med Vaccines* 4, 6 (2018). https://doi.org/10.1186/s40794-018-0066-4. Accessed 16 March 2021.

Rejmánková, E., Grieco, J., Achee, N., and Roberts, R. (2013). *Ecology of Larval Habitats, Anopheles mosquitoes - New insights into malaria vectors*, Sylvie Manguin, IntechOpen, DOI: 10.5772/55229. Accessed 1 April 2021.

Robbins, J., 2016. *How Forest Loss Is Leading To A Rise In Human Disease*. Yale E360. Accessed 12 January 2021. https://e360.yale.edu/features/how_forest_loss_is_leading_to_a_rise_in_human_disease_malaria_zika_climate_change.

Siraj AS, Santos-Vega M, Bouma MJ, Yadeta D, Ruiz Carrascal D, Pascual M. *Altitudinal changes in malaria incidence in highlands of Ethiopia and Colombia.* Science. 2014 Mar 7;343(6175):1154-8. doi: 10.1126/science.1244325. PMID: 24604201. Accessed 8 December 2021.

Stratton L, O'Neill MS, Kruk ME, Bell ML. The persistent problem of malaria: addressing the fundamental causes of a global killer. *Soc Sci Med*. 2008 Sep;67(5):854-62. doi:10.1016/j.socscimed.2008.05.013. Epub 2008 Jun 26. PMID: 18583009. Accessed 10 January 2021.

Tuno N, Wilberforce O, Minakawa N, Takagi M, Yan G. (2005). Survivorship of *Anopheles gambiae* sensu stricto (Diptera: Culicidae) larvae in western Kenya highland forest. *Journal of Medical Entomology.* 42: 270 - 277.

"Vector-Borne Disease". European Centre For Disease Prevention And Control. 2021. https://www.ecdc.europa.eu/en/climate-change/climate-change-europe/vector-borne-diseases#:~:text=Vector%2Dborne%20diseases%20are%20infections,especially%20sensitive%20to%20climatic%20factors. Accessed 5 March 2021.

Vittor, AY Gilman, RH Tielsch, J Glass, G Shields, TIM Lozano, WS Pinedo-Cancino, V Patz JA. 2006. The effect of deforestation on the human-biting rate of Anopheles darlingi, the primary vector of falciparum malaria in the Peruvian Amazon. American Journal of Tropical Medicine and Hygiene 74: 3–11.

Wangdi, K., Furuya-Kanamori, L., Clark, J. *et al.* Comparative effectiveness of malaria prevention measures: a systematic review and network meta-analysis. *Parasites Vectors* 11**,** 210 (2018). https://doi.org/10.1186/s13071-018-2783-y. Accessed 14 March 2021.

World Economic Outlook Database October 2018 -- WEO Groups and Aggregates Information. *International Monetary Fund*, www.imf.org/external/pubs/ft/weo/2018/02/weodata/groups.htm. Accessed 10 December 2020.

World Health Organization, *Malaria Fact Sheet.* World Health Organization, 30 November 2020, https://www.who.int/news-room/fact-sheets/detail/malaria. Accessed 6 January 2021.

World Health Organization, *Vector-Borne Diseases.* World Health Organization, 2 March 2020, www.who.int/news-room/fact-sheets/detail/vector-borne-diseases. Accessed 6 January 2021.

World Meteorological Organization & World Health Organization. Factsheet #2: Climate information for protecting human health (2009).

Yasuoka J Levins R. 2007. Impact of deforestation and agricultural development on anopheline ecology and malaria epidemiology. *American Journal of Tropical Medicine and Hygiene* 76: 450–460.

University of York, 2021, https://www.york.ac.uk/depts/maths/tables/spearman.pdf. Accessed 13 March 2021.