Regression-based Model Deciphers Progression on Health Management for Two Long-standing and Unresolved Infectious Diseases in Indonesia:

Integration of Outpatient Data by the National Universal Health Care (BPJS) and Fluctuation of Registered HIV and Malaria Cases

*Alam A.*Hidayat[1], and *Ika* Nurlaila[2,3*]

[1]Mathematics Departement, School of Computer Science, Bina Nusantara University, Jakarta Barat-Indonesia
[2]Information System Department, BINUS Online Learning (BOL), Bina Nusantara University, Jakarta Barat-Indonesia
[3]National Research and Innovation Agency (Badan Riset dan Inovasi Nasional, BRIN), Tangerang Selatan-Indonesia
[*]Corresponding author : ika.nurlaila@brin.go.id

Abstract. Although the COVID-19 outbreak was recently declared overcome COVID-19 is not the sole life-threatening infectious disease for which we direct our all resources. HIV and malaria are, on the other hand, two long-standing national health issues that are yet to find a proper solution. We query how these two infectious diseases load volume on National Burden as represented by the claims National Health Resilience and how far our efforts paid thus far affect the progression on optimally putting the cases in a break. We approach these mathematically as we hardly found integrated reports on this matter. Hence, we employed regression linear to model the Health Resilience of Indonesia based on the dynamics of outpatients and inpatients across facility categories (class) provided by National Universal Health Care (BPJS) in the context of HIV and malaria infection cases respectively. The estimation of the two regression parameters was done via the ordinary least square (OLS) method. We used pandas 1.3.5 for performing the data analysis, seaborn 0.11.2 and matplotlib 3.5.1 for the data visualization, and scipy 1.7.3 and statmats 0.31.2 for the statistical analysis. Our results show that the number of outpatients declines as the number of HIV and malaria cases increase. Furthermore, we obtained significant associations between the increased HIV rate and decreased number of outpatients in class 2 (P=0.030), class 3 (P=0.002), and the total outpatients (P=0.0019). These patterns are not observed for malaria cases. Meanwhile, the increase in HIV cases was found to be associated with the decreased number of registered BPJS in class 1 (P=4e-4), class 3 (P=1.5e-6), and total participants in all classes (P=3.6e-6). Less strong associations were found between the malaria cases with decreased number of participants in BPJS class 1 (P=0.010), class 3 (P=0.042), and the total registered to BPJS (P=0.045). Our data suggest that the higher the HIV incidence rate the lesser likelihood of the affected patients being treated as outpatients using facility classes 2 and 3 since this may lose the transmission control intervention. Intriguingly, in the sense of the BPJS facility category where an increased HIV rate is seen to be strongly associated with a decreased registration number for class 1 and class 3, the monthly dues and the adequateness of expected facility items may be the most abductive reasoning. Despite our understanding that these require validation cohort, taken altogether our data hint the threat of HIV is steady and demands a new approach to strategizing for the sake of limiting the transmission as well as maintaining the quality of life of the affected patients.

Keywords : Regression, Model, outpatients, inpatients, HIV, Malaria, Health Resilience, BPJS

1 Introduction

The pandemic status of COVID-19 has recently been revoked and turned to endemic as the President of the Republic of Indonesia, Joko Widodo, announced it through some social media platforms (1). Consequently, previous restriction regulations are loosened, and any new COVID-19 case will be handled as per respective hospitals' regulations. COVID-19 for years has been the main focus of health management and economic aspects, despite its dilemmatic prioritization (2). It caused other long-standing infectious diseases less prioritized. Overcoming the COVID-19 outbreak does not suggest a status of national burden-free (3). We have yet to show satisfactory progression in fighting unresolved infectious diseases being that malaria dan HIV (4,5). In the era of the COVID-19 pandemic, where orientation was almost solely put on the mitigation of the disease, the case numbers for malaria and HIV rose (6,7). This is a big alert that requires more attention cross-sectionally. Here, we stress information system intervention which we expect to be more integrated and carry adequate capacity in providing a broad scope of health. As reviewed in (8) medical information system has been established in many countries globally with Indonesia as no exception. The benefit it offers is increased service availability and reliability to the patients as well as diminished expenditure.

Despite the plausible offers that might also be lucrative, an undeniable challenge lies in the data gathering as shown in a published report which highlighted multiple non-interoperable health information systems (HIS) in Iran (9). In addition, managing and incorporating health information and statistical reports in multiple HISs was also an issue (9). Unfortunately, this is also the case with the HIS in Indonesia. The Ministry of Health (MOH) only regulates the health referral systems of individual health services as stated in the Law on the National Social Security System (Law No. 40/2004) which assures the inception of national health insurance or universal health coverage (10,11). For this purpose, Indonesia developed a health referral system that defines a health service arrangement that is technically organized by the Social Security Agency for Health (BPJS-K) (11). This is a great movement but is yet to be sufficient to address health-related issues in one go. To illustrate, BPJS-K since 2014 had disseminated annual reports which covered claims and frequency of use of each facility level. The reports also included ranks of health issues being treated. However, in the context of major infectious diseases, it missed incorporating these ranks into other aspects reported in the same reports. We see this as crucial to do as we do not only need to have a global picture of what we have done and how successful they were but also need to have some projected figures on how these infectious diseases burden our national strength. This integrated information, we believe, will benefit people across social strata and the relevant stakeholders who are mandated to guard health resilience at the national level. For this purpose, we developed a regression model to see mathematical interactions of major infectious diseases in Indonesia and rates of outpatients as well as inpatients recorded by BPJS-K for the last five years. We expect to produce models with statistical significance. This is a preliminary plotting where we sought to find the fittest models of tandem variables. Overall, our models provide ideas on how infectious diseases, whenever interventions are not well performed, impact our national health resilience.

2 Materials and Methods

2.1 Data Extraction

We combined data that were produced by different sources. Malaria and HIV data were extracted from https://statista.com from 2010-2022 with respective keywords (malaria, HIV). In the case that data was missing in particular years, we removed that and only proceeded with years that were complete data for both malaria and HIV. Outpatients and inpatients data were derived from annual reports by BPJS-K termed Statistic Book JKN 2015-2019 (Buku Statistik JKN 2015-2019) which were downloadable at https://bpjs-kesehatan.go.id.

2.2 Statistical Analysis

A descriptive statistic via data visualization was performed to find the general overview of the data and their related trend used in this study. A number of inpatients and outpatients (regardless of the diseases) were chosen to represent the yearly hospital burden nationally. On the other hand, the number of citizens registered to the national universal health care (BPJS) with three-class premiums (classes 1, 2, and 3 are categorized by rates) provides a rough approximation for studying the national medical care economic burden.

By aggregating the incidence of diseases such as HIV and malaria, the analysis sought a relationship to describe how a disease occurrence in the population acts on the health burden. Therefore, to determine the statistical relationship between health-burden variables with variables related to various disease incidences, we performed a linear regression analysis to measure the strength of the relationship. The following equation is imposed to determine the relationship between a response variable as a dependent variable (DV) and a predictor, which is an independent variable (IV):

$$\boldsymbol{DV = \alpha + \beta * IV} \qquad (1)$$

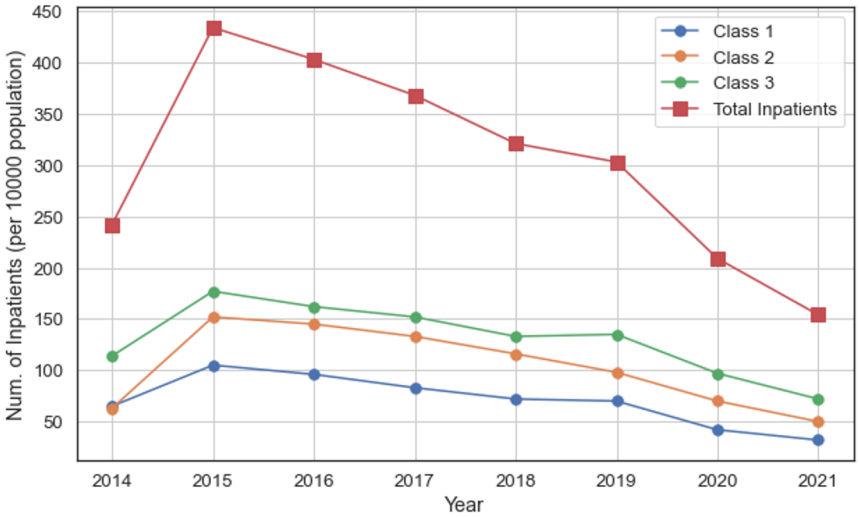
where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are regression parameters. More specifically, $\boldsymbol{\beta}$ coefficient quantifies how a unit change in the predictor affects the average change in the response variable. The estimation of the two parameters was performed using the ordinary least square (OLS) method. The significance of statistical association for $\boldsymbol{\beta}$ was measured using a P-value, for which the P-value < 0.05 indicates that the null hypothesis ($\boldsymbol{\beta}$ is not different from zero) can be rejected and hence the alternative hypothesis is accepted. In this work, Python 3.7 was used as a framework for performing analysis. Several Python libraries used in this work: pandas 1.3.5 for performing the data analysis, seaborn 0.11.2 and matplotlib 3.5.1 for the data visualization, and scipy 1.7.3 and statmats 0.31.2 for the statistical analysis.

3 Results and Discussion

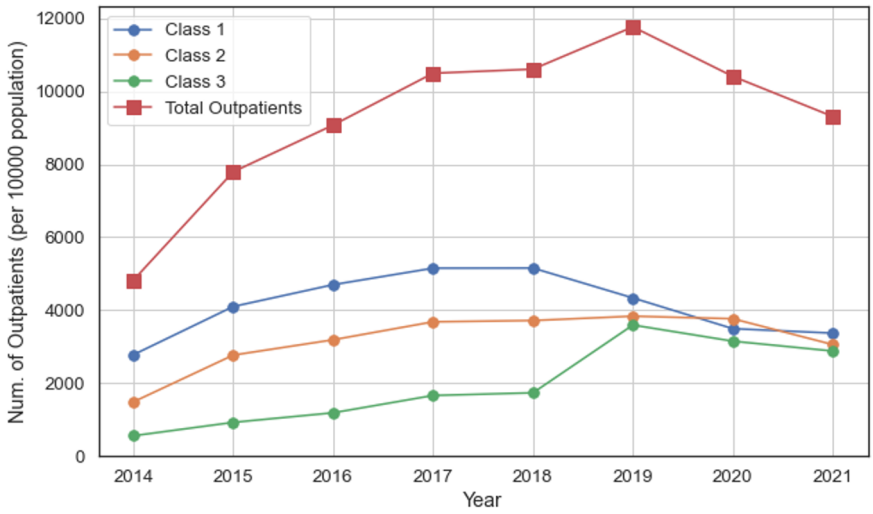
Prior to plotting either HIV or malaria to total inpatients or outpatients, we started with 4 major infectious diseases being that malaria, HIV, dengue, and tuberculosis. However, as we searched for exact data (shown as a total number of total cases in a particular portion of the population, instead of range or estimation) we hardly found complete information regarding dengue and tuberculosis. We, hence, proceeded with malaria and HIV case only.

3. 1 The General Trend of Inpatients, Outpatients, and BPJS Participants

We plotted inpatients who were administered in three different facility levels (class 1, class 2, and class 3, respectively) within 8 years commenced since the first active year of BPJS-K (2014) and found out that throughout 8 years (except 2014-2015) the numbers of inpatients tend to fall. A slight rise was detected in 2019 but this is compensated by a steep reduction in the following year.



(a) Inpatients



(b) Outpatients

Fig. 1. The yearly trend of the number of inpatients (a) and outpatients (b) in Indonesia

On the contrary, the outpatient number increased and culminates in 2019 prior to having a slight reduction in two subsequent years. Interestingly, in regard to facility class being used class 1 outnumbers the two counterparts up to 2018 then all head off toward similar points in 2020 and 2021. This differs from the inpatients trend where class 3 outnumbers others, followed by class 2 then class 1 as the least used for 8 years. Abductive reasoning might hint severity of the diseases being the reason for these dynamics. Diseases that can be handled without strict and direct observation from clinicians do not require to be administered to hospitals or nearby public health centers. They rather are treated as outpatients but with a class 1 facility. We, next, question if the people participating in each class will also demonstrate a specific pattern or rather reflecting either inpatient or outpatient model.

The number of people registered in BPJS class 1 was almost similar to that of class 2, which is around 1000 people per 10000 population, and the trend was relatively stagnant over the period 2014-2021. In contrast, the number of citizens in BPJS class 3 dominates the total number of BPJS participants as the number in class 3 was around four times to six times larger than the number in class 1 and class 2. Besides, a steady increase of people registered to class 3 is noticeable from Figure 2, starting from around 3500 people in 2014 to 6000 in 2021 per 10000 population.

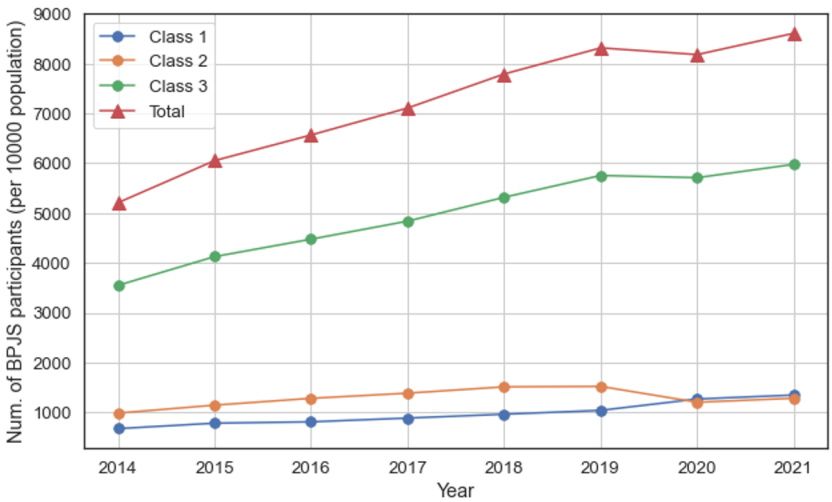


Fig. 2. The yearly trend of the number of people registered to BPJS-K

As seen in Figure 2. the yearly trend of the number of people registered to BPJS is similar to that seen for outpatients where in 2020 and 2021 there is no marked difference between class 1 and class 2 registration but, what makes the pattern distinguished from outpatients is the registration for class 3 which makes up to four folds higher than class 1 and class 2 and steadily rising afterward. This probably reflects the affordability of the people to the set monthly dues.

3.2 Regression Analysis

We, further, regress malaria and HIV to each class 1, 2, and 3, respectively to assess the fitness of the models. Using a regression in the form of Equation 1, we performed a statistical analysis to study how the incidence rate of some common diseases, represented by the yearly number of HIV and malaria cases in our case, as the predictors to 12 dependent variables. The regression results were summarized in **Table 1**.

Table 1. Linear Regression Results

Disease (IV)	Variables (DV)	Coefficient (β)	STD of β	P-value
HIV	Inpatients 1	72.00	32.86	0.071
	Inpatients 2	65.95	64.11	0.343
	Inpatients 3	96.93	50.43	0.136
	Inpatients Total	224.88	146.45	0.176
	Outpatients 1	-455.32	1527.81	0.776
	Outpatients 2	-2591.35	918.58	0.030*
	Outpatients 3	-4408.13	821.74	0.002*
	Outpatients Total	-7455.23	2343.91	0.019*
	BPJS 1	-972.69	135.96	4e-4*
	BPJS 2	-473.71	260.14	0.118
	BPJS 3	-3746.28	200.53	1.5e-6*
	BPJS_Total	-5192.67	321.85	3.6e-6*
Malaria	Inpatients 1	8.12	3.73	0.072
	Inpatients 2	9.57	6.83	0.211
	Inpatients 3	11.46	5.18	0.069
	Inpatients Tot	29.15	15.52	0.109
	Outpatients 1	95.39	169.75	0.595
	Outpatients 2	-96.87	153.55	0.551
	Outpatients 3	-311.60	184.21	0.142
	Outpatients Tot	-313.08	415.49	0.480
	BPJS 1	-97.18	26.11	0.010*
	BPJS 2	-18.83	35.87	0.618
	BPJS 3	-310.10	120.17	0.042*
	BPJS Total	-426.11	169.15	0.045*

*Significant at $\alpha = 0.05$

The graphical representation of each regression analysis is shown in Figures 3, 4, 5, and 7. A visual comparison between the regression trend of inpatients and outpatients demonstrates that the number of patients in hospital care increased over the years with respect to the increased number of reported HIV and malaria cases (Figure 3 and 4). These results, however, were not supported statistically by the linear regression analysis reported in Table 1, since no significant association was found between the trend of the inpatients and the incidence of the two diseases ($P>0.05$).

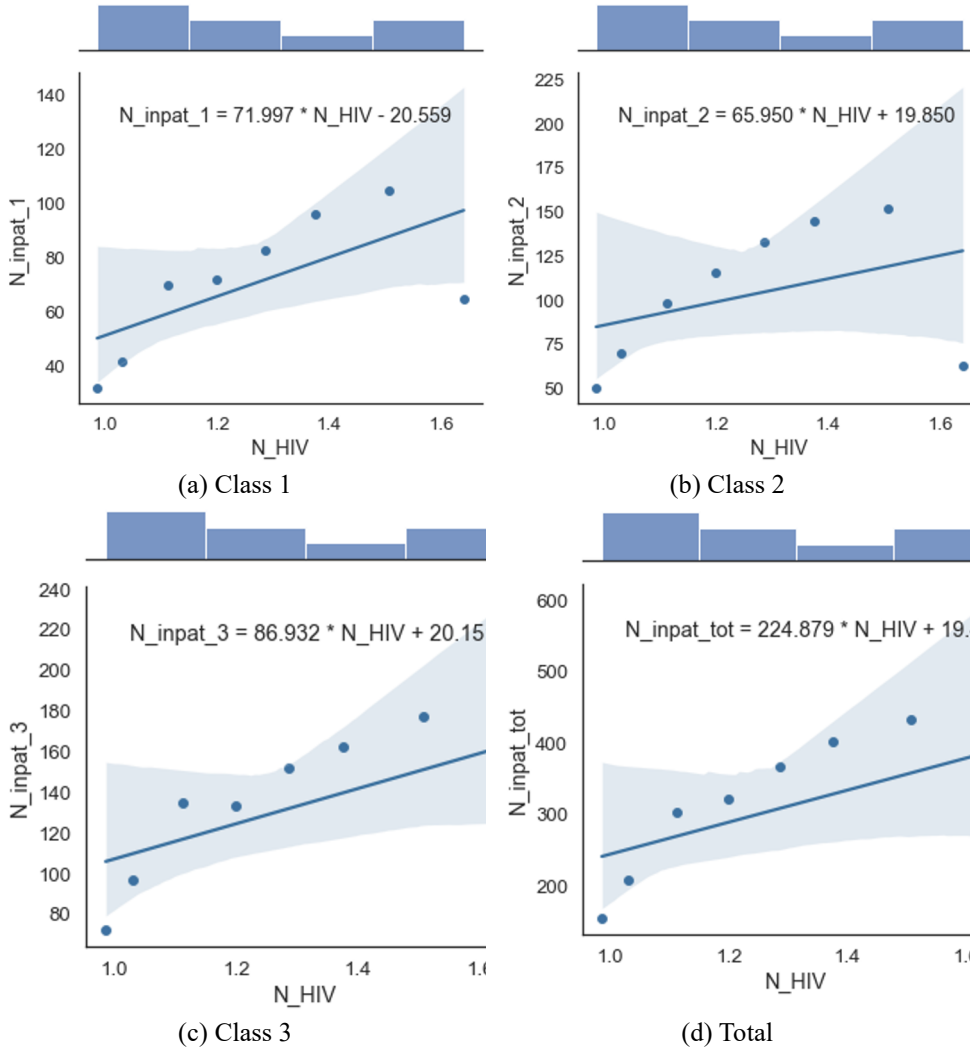
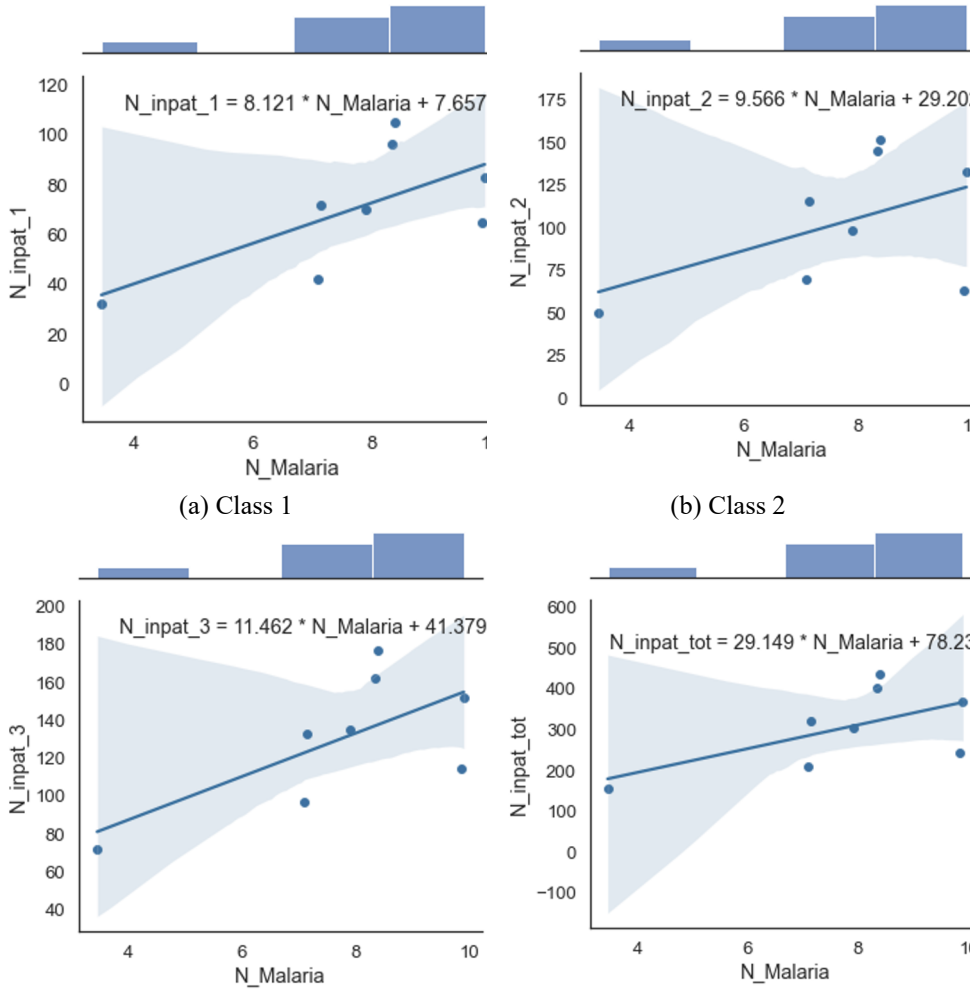


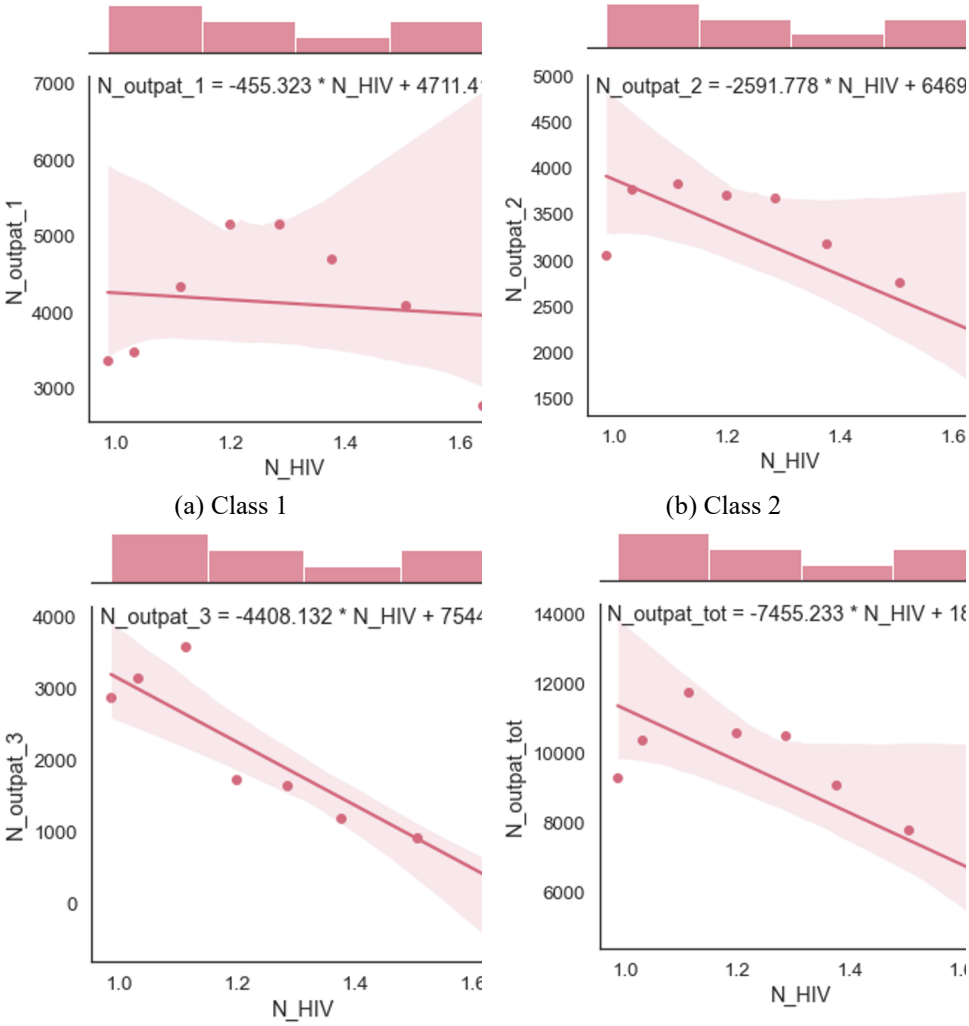
Fig. 3. Regression analysis between the number of outpatients with the number of HIV incidents



(a) Class 1 (b) Class 2

(c) Class 3 (d) Total

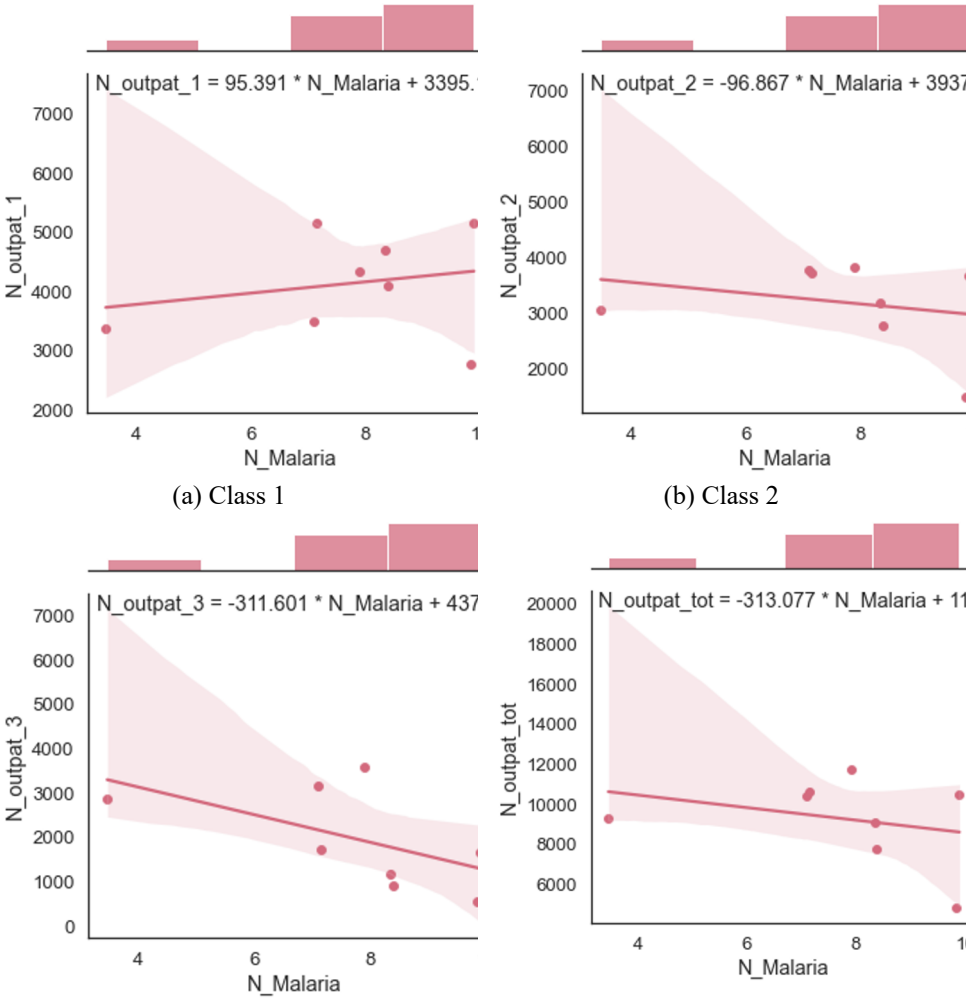
Fig. 4. Regression analysis between the number of inpatients with the number of malaria incidents



(a) Class 1 (b) Class 2

(c) Class 3 (d) Total

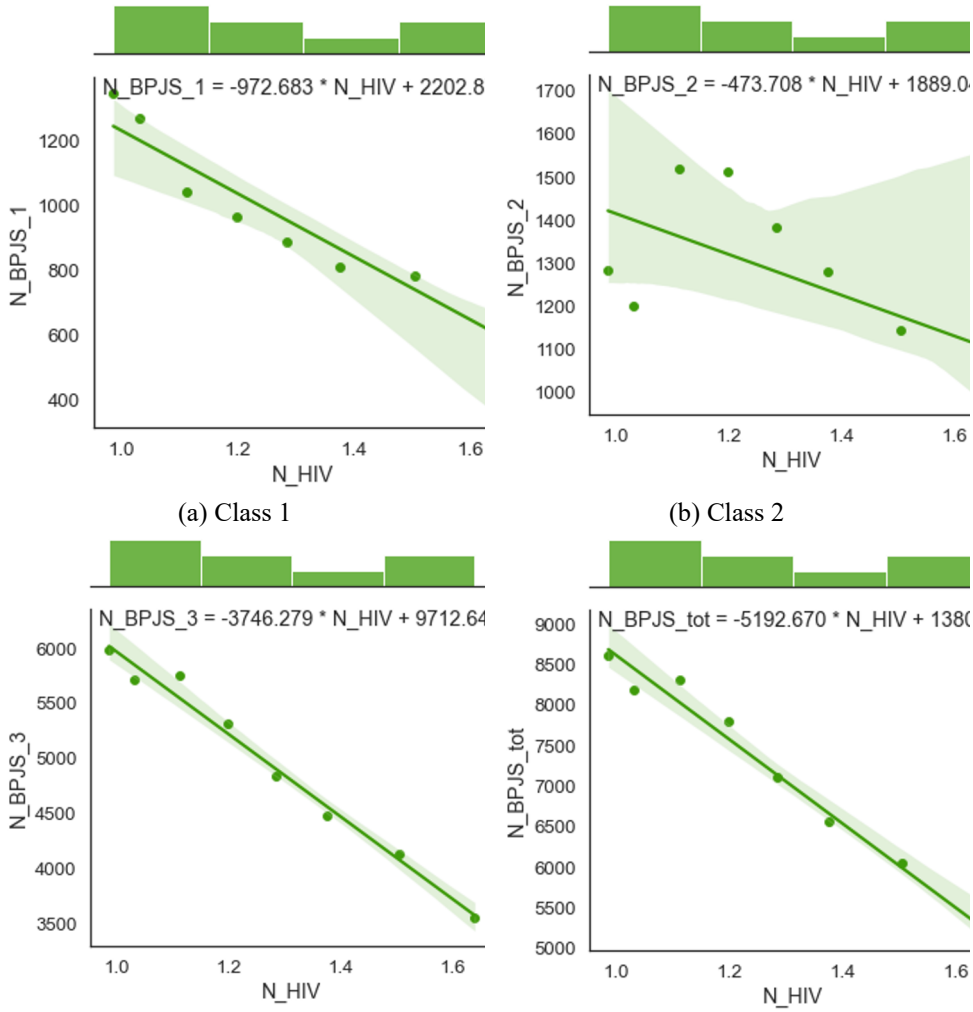
Fig. 5. Regression analysis between the number of outpatients with the number of HIV incidents



(a) Class 1 (b) Class 2

(c) Class 3 (d) Total

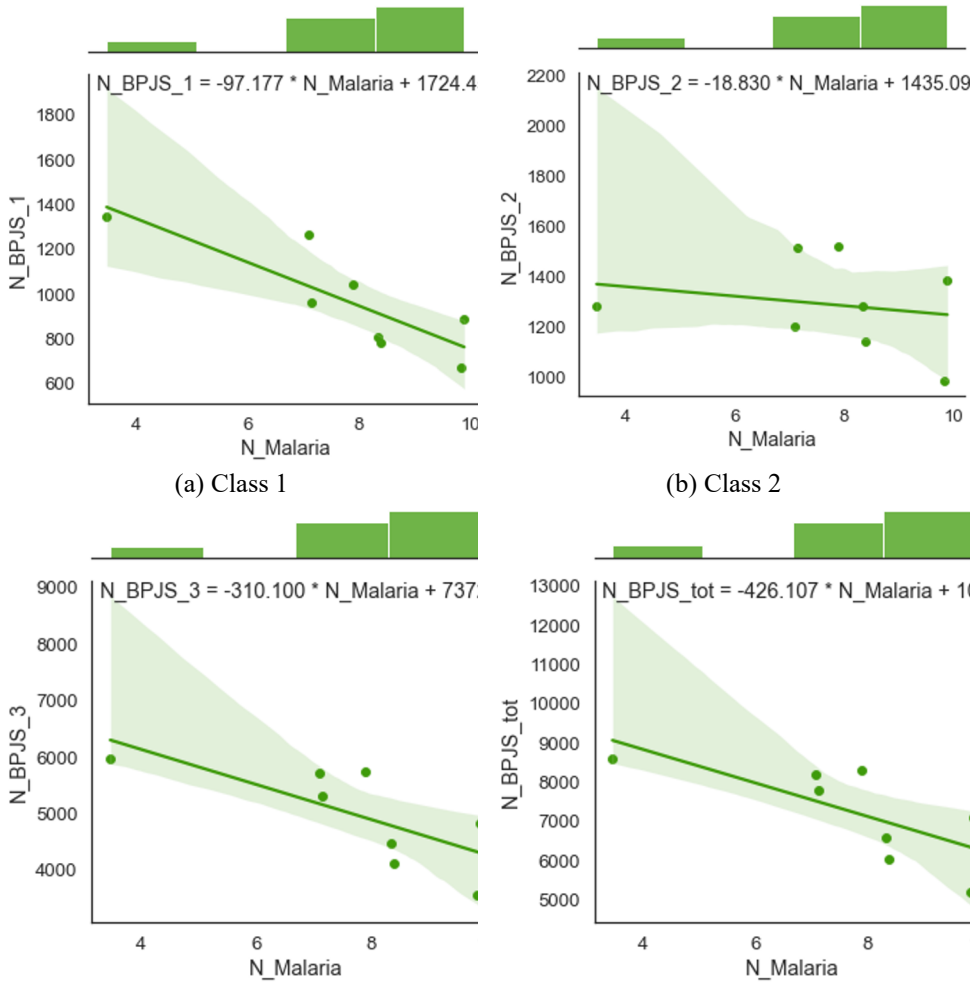
Fig. 6. Regression analysis between the number of outpatients with the number of malaria incidents



(a) Class 1 (b) Class 2

(c) Class 3 (d) Total

Fig. 7. Regression analysis between the number of BPJS participants and the number of HIV incidents



(a) Class 1 (b) Class 2

(c) Class 3 (d) Total

Fig. 6. Regression analysis between the number of BPJS participants and the number of malaria incidents

We can see that the number of outpatients declined as the number of HIV and malaria cases increased during the eight years. Referring to the HIV cases, we obtained significant associations of the increase in HIV incidence to the decreased number of outpatients in class 2 (P=0.030), class 3 (P=0.002), and the total outpatients (P=0.0019). However, all malaria cases did not statistically associate with the trend of the outpatients (P>0.05).
Meanwhile, the increase in HIV cases was found to be significantly associated with the decreased number of registered BPJS in class 1 (P=4e-4), class 3 (P=1.5e-6), and total participants in all classes (P=3.6e-6). Relatively weak statistical associations were found between the malaria cases with a reduced number of participants in BPJS class 1 (P=0.010), class 3 (P=0.042), and the total registered to BPJS (P=0.045).

Apart from the significance values we have obtained from our regression models, all these show that regression is a plausible instrument to map how strong the incidence rate of infectious diseases will burden our national health resilience. The number of published

reports implementing regression for health-related variables is limited. It was used to draw an association between health surveillance personnel's performance in the early detection of pre-eclampsia with work motivation where all datasets were collected through questionnaires (12). Other studies showed that regression modeling could be a useful tool for studying the simultaneous effects of multiple variables in a single dependent variable where clinical data, imaging data dan laboratory parameters were integrated to generate a diagnosis although this study was yet to have validation (13). In Indonesia, the data alone is not well-organized. Each organization has its strategy to manage its data which in most cases turns out to be challenging just to assure that the data is not overlapping. For example, BPJS-K issued reports on outpatients, inpatients, and progression on health-related facilities for each stratum of public health centers but we could not have an idea if some fractions in the inpatient or outpatient list have been recorded by other organizations which released reports on infectious disease mitigation. This might confound the measurement. Hence, we put our effort to tailor secondary data from BPJS-K and Statista to see if there were specific patterns, we could gain to help us understand where we are at right now in the constellation of national health-resilience. This is critical for a nation to redefine its intervention to serve its people for their highest good. Particularly in Indonesia, it is mandated by the 1945 State Constitution of the Republic of Indonesia. We acknowledge that the data points used for these regressions are limited. Data availability is of importance in regression. The more data we have, the stronger regression we can produce. However, at this point, we have been able to provide models with statistical significance. This suggests that health management progression toward specific diseases can be modeled using regression. Not only does the model show the amplitude of the burden (here, the burden is the infectious diseases be that malaria and HIV) but it also provides future propensity.

4 Conclusion

Our regression model is able to project the interaction of either outpatient or inpatient numbers and malaria and HIV, as the two most common infectious diseases in Indonesia. Although significances are detected in almost all combinatorial arrangements, regression with HIV as an independent variable emerges as stronger model. More in-depth, it is shown that increased HIV incidence is significantly associated with decreased number of outpatients who opted class 2 facility as well as to the total outpatients. Increased HIV case rate is also associated with decreased number of class 1 and class 3 registration. Taking all these together, our data suggest that HIV and malaria should be prioritized as their dynamics impact, either directly or indirectly, national burden.

Acknowledgement

We are thankful for the support given by Bina Nusantara University. We also extend our gratitude to Statista that providing us with detailed yet free data for our exploration in this study.

References

1. Indonesia's COVID-19 Task Force. Government Decides Indonesia Has Entered

Endemic Period [Internet]. covid19.go.id. 2023 [cited 2023 Jun 30]. Available from: https://covid19.go.id/artikel/2023/06/21/pemerintah-putuskan-indonesia-masuki-masa-endemi

2. Setiati S, Azwar MK. Dilemma of Prioritising Health and the Economy During COVID-19 Pandemic in Indonesia. Acta Med Indones. 2020;52(3):196–8.
3. Fekadu G, Bekele F, Tolossa T, Fetensa G, Turi E, Getachew M, et al. Impact of COVID-19 pandemic on chronic diseases care follow-up and current perspectives in low resource settings: a narrative review. Int J Physiol Pathophysiol Pharmacol. 2021;13(3):86–93.
4. Magnani RJ, Wirawan DN, Sawitri AAS, Mahendra IGAA, Susanti D, Utami Ds NKAD, et al. The short-term effects of COVID-19 on HIV and AIDS control efforts among female sex workers in Indonesia. BMC Womens Health. 2022 Jan;22(1):21.
5. UNAIDS. Indonesia [Internet]. unaids.org. 2023. Available from: https://www.unaids.org/en/keywords/indonesia
6. Asmarawati TP, Martani OS, Bramantono B, Arfijanto MV. Prolonged fever and exaggerated hypercoagulopathy in malaria vivax relapse and COVID-19 co-infection: a case report. Malar J. 2022 Jun;21(1):199.
7. Merati TP, Karyana M, Tjitra E, Kosasih H, Aman AT, Alisjahbana B, et al. Prevalence of HIV Infection and Resistance Mutations in Patients Hospitalized for Febrile Illness in Indonesia. Am J Trop Med Hyg. 2021 Aug;105(4):960–5.
8. Limanto S, Andre. Information system to enhance medical services quality in Indonesia. Int J Electr Comput Eng. 2019;9(3):2049–56.
9. Ebnehoseini Z, Ebrahimipour H, Koohjani Z, Adel A, Aval SB, Hoseini SJ, et al. Challenges of Health Information Systems (HISs) During COVID-19 Pandemics: Lessons Learned from Teaching Hospitals in a Developing Country. Shiraz E Med J. 2022;23(2).
10. National Team for the Acceleration of Poverty Reduction. The Road to National Health Insurance (JKN). JKN [Internet]. 2015;108. Available from: http://www.tnp2k.go.id/images/uploads/downloads/FINAL_JKN_road to national health insurance.pdf
11. Handayani PW, Pinem AA, Azzahro F, Hidayanto AN, Ayuningtyas D. The Information System/Information Technology (IS/IT) practices in the Indonesia health referral system. Informatics Med Unlocked [Internet]. 2019;17(October):100263. Available from: https://doi.org/10.1016/j.imu.2019.100263
12. Cahyandaru Z, Tamtomo D, Soemanto RB. Linear Regression Analysis on Factors Affecting Health Surveillance Personnel Performance in Early Detection of Preeclampsia in Semarang, Central Java. J Heal Policy Manag [Internet]. 2019 Jan 21;4(1 SE-Articles):13–22. Available from: https://thejhpm.com/index.php/thejhpm/article/view/78
13. Mukhopadhyay K, Singh R, Dalai CK. Understanding the application of multiple linear regression model in health care research using simulation with computer generated data . Introduction : Methods : J West Bengal Univ Heal Sci. 2020;1(2):44–52.