Clustering Community Risk Behaviors for Non-Communicable Diseases in Indonesia: Based on Indonesian Health Survey Data

Hanif Pandu Suhito[1]*, Mahalul Azam[2], Dina Nur Anggraini Ningrum[2], Sholikun[3]

[1]Doctoral of Public Health, Faculty of Medicine, Semarang State University, Semarang, Indonesia
[2]Department of Public Health, Faculty of Medicine, Semarang State University, Semarang, Indonesia
[3]Department of Health Resource, Semarang City Health Office, Semarang, Indonesia
*Authors Correspondence: pandu.suhito@gmail.com/+628116110101

ARTICLE INFO

ABSTRACT

Non-communicable Diseases (NCDs) such as hypertension, stroke, diabetes mellitus, and heart disease are major causes of morbidity and mortality in Indonesia. Community behaviors, including smoking, low physical activity, and unhealthy diets, significantly influence NCD incidence. To analyze associations between behavioral risk factors (daily smoking, physical inactivity, consumption of sugary foods and drinks, salty foods, fatty foods, soft drinks, and inadequate fruit intake) and the prevalence of NCDs, and to classify regions based on risk factors and disease prevalence. A cross-sectional study was conducted using secondary data from the 2023 Indonesian Health Survey (IHS), encompassing 345,000 households across 39 provinces. The Ministry of Health collected the data through standardized interviews and health examinations conducted by trained surveyors. Data analysis employed descriptive statistics, Pearson correlation, and K-Means Clustering. Daily smoking correlated positively with hypertension, stroke, and diabetes. Fatty food consumption showed strong positive associations with the same diseases (r = 0.45–0.48, p < 0.01). Soft drink consumption demonstrated a significant negative correlation, especially with stroke (r = -0.66, p < 0.001). Hypertension (7.2%) and stroke (7.1%) were more prevalent than diabetes (2.0%) and heart disease (0.76%). North Maluku showed the highest behavioral risks, Bali the lowest. DI Yogyakarta and DKI Jakarta had the highest NCD prevalence, while the Papua Mountains had the lowest. K-Means clustering grouped provinces into low, medium, and high-risk clusters. Community behaviors, particularly smoking and fatty food intake, significantly affect NCD prevalence. Public health programs should target smoking cessation, healthy diet promotion, and increased physical activity to reduce the NCD burden.

INTRODUCTION

Non-communicable Diseases (NCDs) are a significant public health challenge in many countries, including Indonesia. According to the World Health Organization (WHO), NCDs such as hypertension, stroke, diabetes mellitus and heart disease contribute to more than 70% of global deaths each year.[1] In Indonesia, the trend of NCD prevalence continues to increase, resulting in a huge economic burden to the health system as well as a decrease in the quality of life of affected individuals. The significant increase in NCD cases is expected to make the burden even greater on society and the government, as handling them is costly and requires high technology.[2]

Smoking, physical inactivity, and an unhealthy diet have been identified as major risk factors in numerous epidemiological studies.[3] Smoking, for example, is a major cause of various cardiovascular and respiratory complications, while consumption of a high-fat, low-fiber diet is associated with increased cholesterol levels and obesity risk. On the other hand, physical inactivity has also been linked to insulin resistance and high blood pressure.

Some national health survey-based studies, such as the 2023 Indonesian Health Survey (IHS), provide comprehensive data on people's behavior patterns and NCD incidence at the provincial level. These data-driven analyses provide insights into the distribution of risk factors across regions and assist in designing more effective public health interventions.[4]

This study aims to gain insight into patterns of NCD cases and people's unhealthy behaviors and explore the relationship between the latter and NCD prevalence in Indonesia. Furthermore, it attempts to identify patterns that can help in risk mitigation strategies for NCD cases. A machine-learning clustering approach was also used to classify provinces based on their level of risk factors and NCD cases. In turn, this can help formulate more focused health policies. By understanding these risk patterns, this study is expected to provide a scientific basis for the government and policymakers to develop better targeted NCD prevention and control strategies.

MATERIAL AND METHOD

Study Design

This study used a cross-sectional study design with the data it utilized being secondary data from the 2023 IHS managed by the Ministry of Health of the Republic of Indonesia. The IHS collected data from January to December 2023 through standardized interviews and health checks performed by trained health workers. The current study did not collect data directly from participants but analyzed the publicly available dataset. All 39 provinces in Indonesia were included to ensure national representativeness and comprehensive coverage of regional variations in behavioral risk factors and NCD prevalence.

Data Sources

The data were obtained from the official 2023 IHS, accessible at layanandata.kemkes.go.id, which the authors accessed on March 5, 2025. This study analyzed a subset of variables relevant to NCD risk factors and outcomes, including daily smoking, physical inactivity, consumption of sugary foods and drinks, salty foods, fatty foods, soft drinks, low fruit intake, and NCD cases (hypertension, stroke, diabetes mellitus, and heart disease). All variables were measured using standardized methods validated by the Indonesian Ministry of Health. As this study used secondary data, the authors did not collect data directly from respondents. Sex and gender information were not available in the dataset; therefore, no sex or gender-based analyses were conducted.

Sample Size and Sampling

The 2023 IHS dataset included 345,000 households, covering all 39 provinces of Indonesia. This study utilized the entire dataset to ensure national representativeness. Minimum sample size for the primary variable of interest, hypertension, was calculated based on a 95% confidence level, an expected prevalence of 7.28%, and a 5% margin of error, resulting in a minimum requirement of approximately 104 respondents. The IHS dataset far exceeded this number, ensuring sufficient statistical power for all analyses. The IHS employed a stratified sampling design at both census block and household

levels, with explicit stratification for urban and rural areas. Households with incomplete information on the primary variables of interest were excluded from the analysis. Inclusion criteria consisted of all individuals aged ≥10 years with complete data on NCD outcomes and behavioral risk factors.

Data Collection

Although SKI 2023 collected a wide range of indicators, this study specifically analyzed behavioral risk factors and NCD outcomes. Behavioral variables included daily smoking, physical inactivity, and the consumption of sweet foods and drinks, salty foods, fatty foods, soft drinks, and low intake of vegetables and fruits. Smoking activity was defined as daily use of tobacco products (yes/no). Physical inactivity was defined as engaging in less than 150 minutes of moderate-to-vigorous physical activity per week. Food and beverage consumption was measured based on daily frequency. NCD outcomes included hypertension, stroke, diabetes mellitus, and heart disease, defined as cases diagnosed by a physician and recorded in the IHS dataset.

Variables

The variables in this study consisted of independent and dependent variables. The former included risk behavior factors in the community, such as (i) smoking, (ii) physical inactivity, (iii) consumption of sweet foods, (iv) consumption of sweet drinks, (v) consumption of salty foods, (vi) consumption of fatty foods, (vii) consumption of soft drinks, and (ix) low fruit intake. Meanwhile, the latter was NCD cases, which covered such cases as hypertension, stroke, diabetes mellitus, and heart disease.

Data Analysis

Data were analyzed using IBM SPSS Statistics for Windows, version 26.0 (IBM Corp., Armonk, NY, USA). Univariate analysis summarized both dependent and independent variables. Associations between behavioral risk factors and NCD outcomes were assessed using pearson correlation analysis. K-Means clustering was used to categorize provinces based on risk factors and NCD prevalence, with the optimal number of clusters determined using the elbow method based on within-cluster sum of squares. A significance threshold of $p < 0.05$ was applied. The STROBE guidelines were followed for reporting observational study results, but they were not part of the data analysis process.[5]

Ethical Consideration

This study fully used secondary data that had been published to the public at layanandata.kemkes.go.id page in the data catalog and had received ethical approval number LB.02.01/I/KE/L/287/2023 from the Health Research Ethics Committee of Health Polytechnics Jakarta II of the Ministry of Health of the Republic of Indonesia.

RESULTS

The highest average prevalence of NCDs in Indonesia was observed in DKI Jakarta, with hypertension at 12.6%, stroke at 10.7%, and diabetes mellitus at 3.9%. DI Yogyakarta had the highest prevalence of heart disease at 1.67% and stroke at 11.4%. Central Papua and the Papua Mountains had the lowest prevalence of NCDs, including hypertension, stroke, and diabetes mellitus. North Maluku had the highest average behavioral risk factors among provinces. Bali had the lowest average behavioral risk factors.
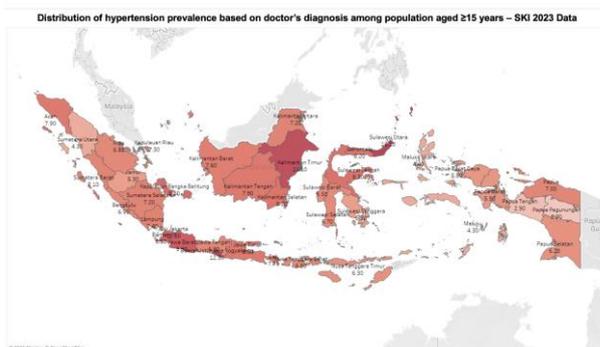
Source: Processed from IHS, Ministry of Health, 2023
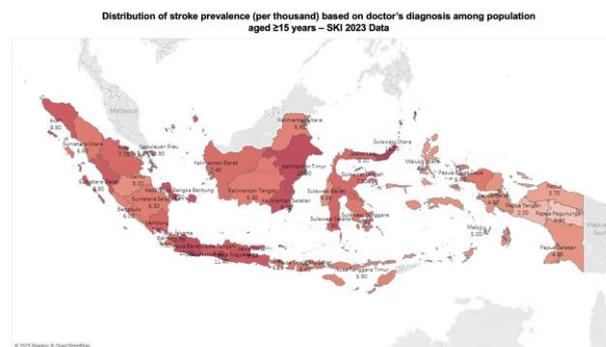
Source: Processed from IHS, Ministry of Health, 2023
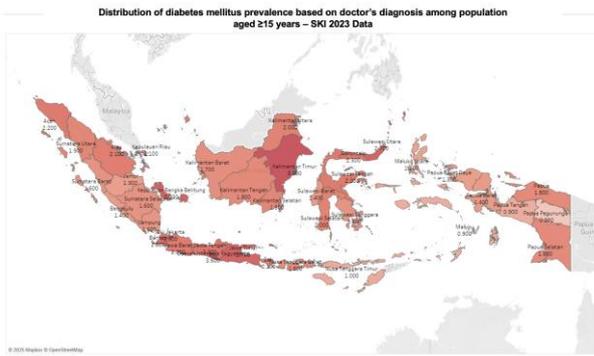
Source: Processed from IHS, Ministry of Health, 2023

Source: Processed from IHS, Ministry of Health, 2023

Figure 1. Spatial Analysis of NCDs Rates Based in Indonesia

Source: Processed from IHS, Ministry of Health, 2023

Source: Processed from IHS, Ministry of Health, 2023

Source: Processed from IHS, Ministry of Health, 2023

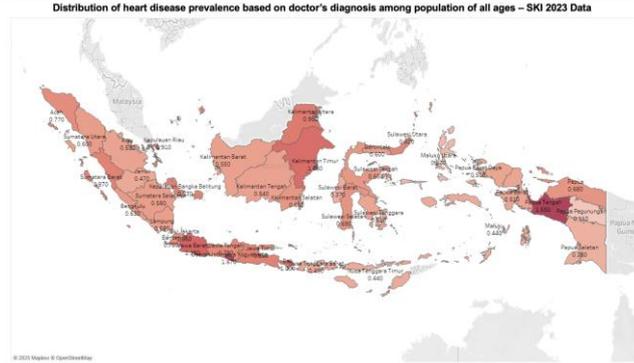Source: Processed from IHS, Ministry of Health, 2023

Source: Processed from IHS, Ministry of Health, 2023
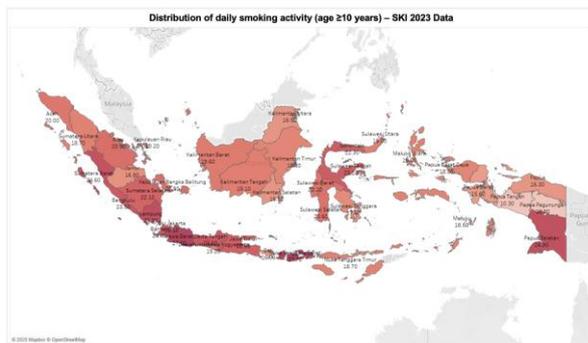

Source: Processed from IHS, Ministry of Health, 2023
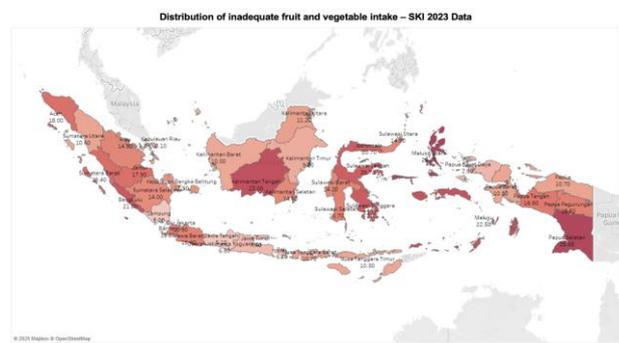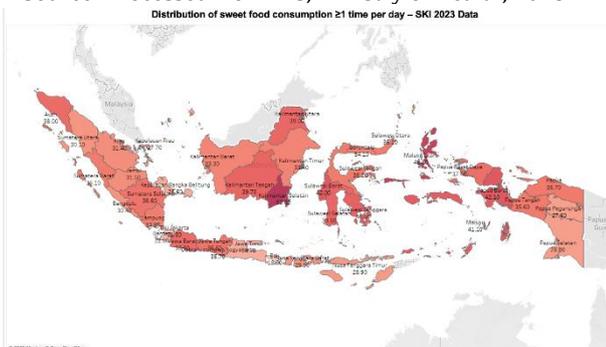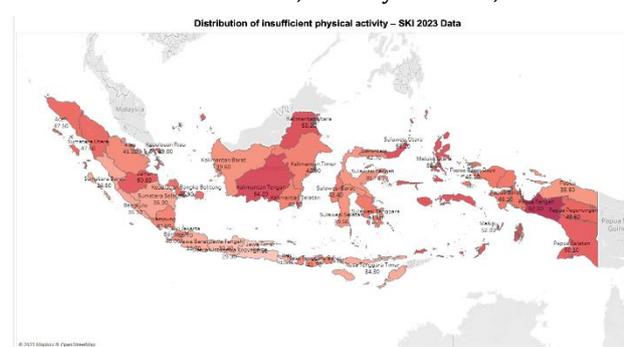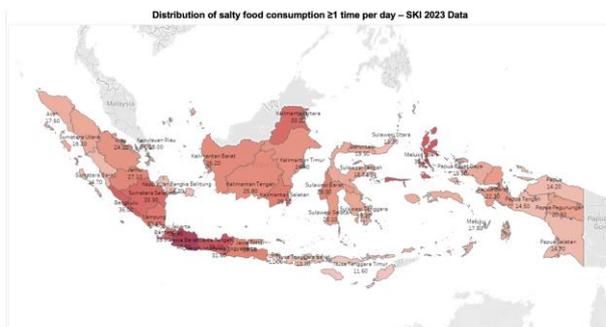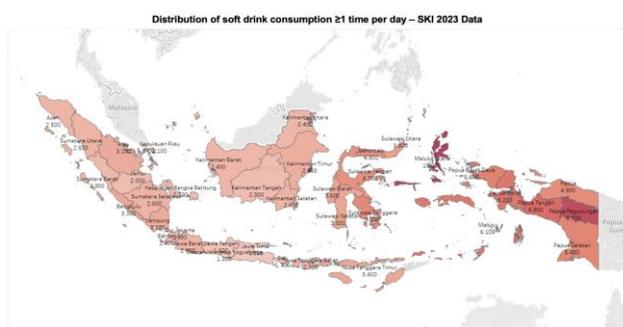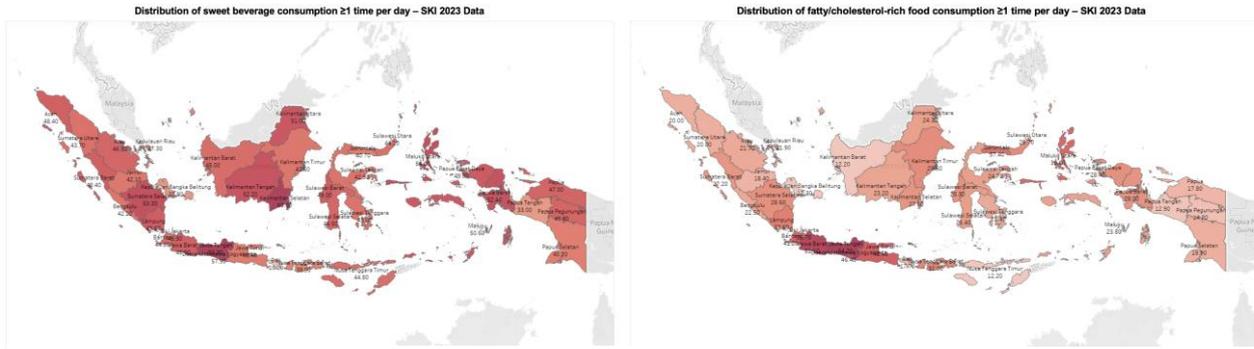
Source: Processed from IHS, Ministry of Health, 2023

Source: Processed from IHS, Ministry of Health, 2023

Figure 2. Spatial Analysis of Risk Behavior Rates in Indonesia

The comparison of several variables between risk behavior and NCD cases revealed that provinces with high risk factors did not always have the highest disease cases. North Maluku, for example, had the highest risk factors, yet it had relatively low disease cases. DKI Jakarta and DI Yogyakarta, on the other hand, had the highest cases of diseases even though their risk factors were not the highest. An interesting fact was that Papua Mountains had both low risk factors and low disease cases.

The people's behavior-related variables in this survey were found to contribute to the risk of NCDs. Frequent smoking activity on a daily basis in the population aged ≥10 years old reached an average of 19.92%. Physical inactivity showed an average value of 43.31%, meaning that almost half the population had low physical activity. The mean value of consumption of sugary foods ≥1 time per day was 34.41%. The average consumption of sugary drinks ≥1 time per day of 45.71% was the most dominant lifestyle behavior. For consumption of salty foods ≥1 time per day, the average was 23.61%, with variations being close to the median value. Consumption of fatty (cholesterol) foods ≥1 time per day was found, on average, at 27.04%. This indicated a significant fat intake. The average consumption of soft drinks ≥1 time per day at 3.58% was the lowest among other lifestyle behaviors. Finally, the low fruit and vegetable intakes showed an average value of 14.23% (Table 1).

Analysis of the NCD case variables showed that hypertension had an average prevalence of 7.28%, with a range from 2.2% to 12.6%. The 95% confidence interval of 6.55%-8.02% suggested that hypertension was quite common in most provinces. Stroke also had a fairly high prevalence, averaging 7.07% with a range of 0.9% to 11.4%, 95% confidence interval of 6.31% to 7.82%. Meanwhile, diabetes mellitus had a lower average prevalence of 1.92%, at a range of 0.2% to 3.9% and a 95% confidence interval of 1.69%-2.15%. Heart disease recorded the lowest prevalence among the four diseases, averaging 0.73% with a range of 0.11% to 1.67%, and a 95% confidence interval of 0.61% to 0.84% (Table 1).

Table 1. Descriptive Analysis of Risk Behaviors and NCD Cases

Community Behavior Variables	Mean	SD	95% CI
Daily smoking activity (Age ≥10 Years)	19.92	3.87	18.69, 21.15
Physical inactivity	43.31	8.06	40.74, 45.87
Consumption of sweets ≥1 time per day	34.41	6.65	32.3, 36.53
Consumption of sugary drinks ≥1 time per day	45.71	7.64	43.28, 48.14
Salty food consumption ≥1 times per day	23.61	8.64	20.86, 26.35
Fatty/cholesterol consumption ≥1 time per day	27.04	10.57	23.68, 30.4
Soft drink consumption ≥1 time per day	3.58	2.02	2.94, 4.23
Not eating fruits & vegetables	14.23	5.73	12.41, 16.05
Hypertension	7.28	2.31	6.55, 8.02
Stroke (Age ≥ 15)	7.07	2.38	6.31, 7.82
Diabetes Mellitus (DM)	1.92	0.73	1.69, 2.15
Disease	0.73	0.35	0.61, 0.84

Source: Processed from Indonesian Health Survey (IHS), Ministry of Health, 2023
Note: Mean values represent the percentage of respondents reporting each behavior or NCD outcome. SD = Standard deviation; 95% CI = 95% confidence interval.

Pearson correlation analysis between behavioral risk factors and NCD prevalence revealed the following: daily smoking was positively correlated with hypertension (r=0.38, p=0.019), stroke (r=0.48, p=0.002), and diabetes mellitus (r=0.33, p=0.046), but not significantly correlated with heart disease (p=0.724). Physical inactivity showed non-significant negative correlations with all NCDs (p > 0.05). Consumption of sugary foods and drinks had no significant associations with NCDs. Salty food consumption showed positive but non-significant correlations with hypertension (r=0.20, p=0.230) and stroke (r=0.19, p=0.261). Fatty food consumption was strongly correlated with hypertension (r=0.45, p=0.005), stroke (r=0.48, p=0.002), and DM (r=0.46, p=0.004), and moderately associated with heart disease (r=0.28, p=0.090). Soft drink consumption showed negative correlations with NCDs, strongest for stroke (r=-0.66, p<0.001) and DM (r=-0.62, p<0.001). Low fruit and vegetable intake had negative correlations with DM (r=-0.32, p=0.047) and heart disease (r=-0.35, p=0.033) (Table 2).

Using the Elbow method analysis, up to 5 clusters were tested to get the optimal inertia value. The inertia values for 1, 2, 3, 4, and 5 clusters were 15, 444.05 (k=1), 10,789.79 (k=2), 8,709.71 (k=3), 7,616.07 (k=4), and 6,953.00 (k=5) respectively. From the pattern of inertia values above, it could be seen that after getting to 3 clusters (k=3) the decrease in inertia started to slow down. On one hand, if we choose k = 2, we still lose the pattern of variation that clustering could capture. If we chose k = 4 or more, we started to get less significant results in inertia reduction. Thus, choosing 3 clusters (k = 3) was the wisest decision since it would allow us to get a balance between keeping the main patterns in the data without making the model too complicated (Table 3).

Based on the results of K-Means clustering, provinces were grouped into 3 clusters as per the patterns of risk behavior factors and NCD prevalence. The three clusters were cluster 0, 1, and 2. Cluster 0 (medium-risk one) was a mixture of healthy and risky patterns. Cluster 1 (low-risk one) was a healthy lifestyle and a low NCD prevalence pattern moreover cluster 2 (high-risk one) was a pattern of risky behavior and high NCD prevalence (Table 3).

In cluster 0, we had Aceh, North Sumatra, West Sumatra, Riau, Jambi, Bangka Belitung, Riau Islands, Bali, NTB, NTT, West Kalimantan, North Sulawesi, Central Sulawesi, Southeast Sulawesi, Papua, South Papua, Central Papua, and Papua Mountains. In cluster 1, we had South Sumatra, Bengkulu, Lampung, DKI Jakarta, Central Kalimantan, South Kalimantan, East Kalimantan, North Kalimantan, South Sulawesi, Gorontalo, West Sulawesi, Maluku, North Maluku, West Papua, and Southwest Papua. Finally, in cluster 2 we had West Java, Central Java, DI Yogyakarta, East Java, and Banten (Table 3).

Table 2. Analysis of The Relationship Between Risk Behavior Factors and NCD Cases

Behavioral Variables	Hypertension (r, r², p-value)	Stroke (r, r², p-value)	DM (r, r², p-value)	Heart Disease (r, r², p-value)
Smoking Every Day	0.38, 0.14, 0.019	0.48, 0.23, 0.002	0.33, 0.11, 0.046	-0.06, 0.00, 0.724
Less Physical Activity	-0.26, 0.07, 0.114	-0.32, 0.10, 0.053	-0.21, 0.04, 0.216	-0.04, 0.00, 0.830
Sugary Food Consumption ≥1/day	-0.11, 0.01, 0.494	-0.02, 0.00, 0.887	-0.21, 0.04, 0.206	-0.14, 0.02, 0.391
Sugary Drink Consumption ≥1/day	0.06, 0.00, 0.722	0.13, 0.02, 0.440	-0.03, 0.00, 0.837	-0.16, 0.03, 0.340
Salty Food Consumption ≥1/day	0.20, 0.04, 0.230	0.19, 0.03, 0.261	0.15, 0.02, 0.356	0.11, 0.01, 0.528
Fatty Food Consumption ≥1/day	0.45, 0.20, 0.005	0.48, 0.23, 0.002	0.46, 0.21, 0.004	0.28, 0.08, 0.090
Soft drink consumption ≥1/day	-0.61, 0.37, 0.000	-0.66, 0.44, 0.000	-0.62, 0.39, 0.000	-0.43, 0.18, 0.008
Not Eating Fruits & Vegetables	-0.28, 0.08, 0.091	-0.31, 0.09, 0.062	-0.32, 0.10, 0.047	-0.35, 0.12, 0.033

Source: Processed from Indonesian Health Survey (IHS), Ministry of Health, 2023

Table 3a. Provincial Clustering Results Based on Risk Behavior Factors and NCD Cases

Province	Daily Smoking (%)	Less Physical Activity (%)	Consumption of Sweets (%)	Consumption of Sugary Drinks (%)	Consumption of Salty Foods (%)	Consumption of Fatty Foods (%)	Soft Drink Consumption (%)	Not Eating Fruits & Vegetables (%)	Prevalence of Hypertension (%)	Stroke Prevalence (%)	Diabetes Prevalence (%)	Heart Disease Prevalence (%)	Cluster (%)
Aceh	20	47.5	38	48.4	17.6	20	2.5	18	7.9	8.8	2.2	0.77	0
North Sumatra	18.7	47.6	30.1	43.7	19.2	20	2.6	10.6	4.3	6.6	1.9	0.6	0
West Sumatra	24.6	34.6	35.1	48.4	14.7	27.2	2.3	18.4	7.1	8.8	1.6	0.87	0
Riau	20.9	41.2	31.4	46.5	24.1	21.7	3.1	14.9	6.8	7.7	2.1	0.53	0
Jambi	16.8	50.8	31.5	42.1	27.1	18.4	2	17.9	5.3	5	1.3	0.47	0
South Sumatra	22.1	35.3	36.6	53.3	33.3	28.6	2.6	14	7.2	6.3	1.6	0.58	1
Bengkulu	23.5	35.1	30.4	42.2	36	22.5	3.3	21.6	6.9	6	1.4	0.63	1
Lampung	25.3	42.6	34.8	49.6	30.4	33.6	3.2	5	7.4	7.9	1.6	0.58	1
Bangka Belitung	20.8	46.9	25.5	33.8	14.2	12.3	2.3	12.9	8.8	9.5	2.8	0.87	0
Riau Islands	20.2	49.8	27.7	37.3	18	21.9	2.1	8.1	7.3	8.9	2.1	0.91	0
DKI Jakarta	19.1	55.7	31.3	45.3	26.9	35.7	1.5	10.9	12.6	10.7	3.9	1.56	1
West Java	27	33.7	38.3	48.9	46.5	51.1	2.6	15.4	9.9	10	2.2	1.18	2
Central Java	23.6	30.4	36.5	60.3	39.1	54.2	1.6	5.1	6.9	8.4	2.3	0.79	2
DIY	19.2	29.3	35.7	57.9	31.6	46.4	1.3	6.8	12.3	11.4	3.6	1.67	2
East Java	22.8	27.8	24.7	41.4	28	42.6	1.3	7.4	8	9	2.7	0.88	2
Banten	24.4	40	22.6	44.3	33.9	41.2	2.4	15.8	9.5	7.3	2.5	0.78	2
Bali	16	32.4	18.8	19.2	11	17.7	1.7	6.2	7.7	6.2	2.1	1	0
NTB	27.7	41.7	29.9	38.9	13.3	32	2.7	10.7	6.8	6.6	1.8	0.49	0
NTT	18.7	34.3	28.9	44.8	11.6	12.2	3.4	10.3	6.3	5.9	1	0.44	0
West Kalimantan	19.6	39.6	33.3	45	26.2	12.2	2.4	10	7.6	7.4	1.7	0.55	0
Central Kalimantan	19.1	54.2	39.7	52.2	25.8	23.2	2	23	7.9	6.3	1.8	0.54	1
South Kalimantan	16.5	42.8	52	60	29.1	27.6	2.6	14.6	8.2	9.5	1.9	0.66	1
East Kalimantan	18.3	42.9	33.4	42.5	24.4	29.6	2.9	9.2	11.1	10	3.1	1.08	1
North Kalimantan	16.5	53.2	39	51	33.2	24.3	2.4	11.2	7.2	5.9	2	0.95	1
North Sulawesi	19	53.2	35.1	44.1	13.3	29.7	3.8	14.9	12.1	11.3	2.7	0.82	0

Table 3b. Provincial Clustering Results Based on Risk Behavior Factors and NCD Cases

Province	Daily Smoking (%)	Less Physical Activity (%)	Consumption of Sweets (%)	Consumption of Sugary Drinks (%)	Consumption of Salty Foods (%)	Consumption of Fatty Foods (%)	Soft Drink Consumption (%)	Not Eating Fruits & Vegetables (%)	Prevalence of Hypertension (%)	Stroke Prevalence (%)	Diabetes Prevalence (%)	Heart Disease Prevalence (%)	Cluster
Central Sulawesi	23.1	38.7	35.2	42.9	18.8	24.7	4.7	20.7	8.3	7.5	2.2	0.65	0
South Sulawesi	20.6	39.6	39.5	46.5	23.1	25.4	3.5	15.7	6.7	7.9	2	0.69	1
Southeast Sulawesi	17.5	44.6	40.4	43	19.2	21.3	5.2	21.1	5.3	6.6	1.6	0.51	0
Gorontalo	22.3	42.7	34.1	40.7	19.9	37.4	4.8	20.7	8.2	8	2.3	0.6	1
West Sulawesi	20.1	43.4	42	49	28	28.3	3.5	14.2	6.5	4.8	1.4	0.37	1
Maluku	18.6	52	41.1	50.6	17.8	23.8	6.1	22.5	4.3	5	0.9	0.44	1
North Maluku	19	52.6	48.5	54	37.2	39.6	10	26.4	4.4	4.5	1.2	0.27	1
West Papua	15.6	45.1	42.1	52.4	22.3	28	6.2	10.5	5.9	4.8	1.4	0.51	1
Southwest Papua	18.2	45.1	37.6	49.8	18.3	28.4	5.6	7.6	5.8	6.6	1.3	0.51	1
Papua	16.3	38.5	35.7	47	14.2	17.8	4.9	10.7	7	3.7	1.8	0.68	0
South Papua	24.9	50.1	28	40.2	14.7	19.9	5.4	25.5	6.2	4.8	1.8	0.38	0
Central Papua	10.3	62	35.6	33	14.5	12.9	6.8	16.8	2.9	2	0.9	1.65	0
Papua Pegunungan	10	48.6	27.6	46.8	20.5	14.2	8.9	15.5	2.2	0.9	0.2	0.11	0

Source : Processed from Indonesian Health Survey (IHS), Ministry of Health, 2023
Note : All behavioral and disease variables are presented as percentages of respondents or prevalence in the province.
Cluster indicates the group assigned by K-Means analysis (0 = Medium-risk, 1 = Low-risk, 2 = High-risk).

DISCUSSION

The various analyses showed that provinces with high smoking rates tended to have a higher prevalence of hypertension. This finding confirmed the research by Leone (2011) who found that smoking was a significant risk factor for hypertension.[6] Meanwhile, provinces with low physical activity had higher stroke cases. This was in line with Börjesson et al. (2016)[7] who suggested that consumption of sugary drinks and low physical activity were the most dominant behavioral characteristics in stroke patients. However, hypertensive disease and stroke had a higher prevalence than diabetes mellitus and heart disease in these provinces.

The results showed that risky behaviors had a significant correlation with the incidence of non-communicable diseases (NCDs). Daily smoking was significantly and positively correlated with hypertension, stroke, and diabetes (p-value < 0.05). This suggested that smoking could increase the risk of NCDs.[8] Consumption of fatty foods also showed a strong positive correlation with hypertension, stroke, and diabetes (r ranging from 0.45-0.48, p-value < 0.01), supporting previous findings regarding the negative impact of a high-fat diet on health especially NCDs.[9] Regarding previous studies on this matter, the association between a high-fat diet and hypertension or stroke has been confirmed in many epidemiological studies.[10]

In contrast, consumption of soft drinks had a highly significant negative correlation with all diseases, especially stroke (r = -0.66, p-value < 0.001). This finding contradicted previous studies which mentioned that consumption of soft drinks had a significant association with stroke for both men and women.[11] Therefore, further review with a more in-depth approach was needed to understand the underlying factors by removing possible bias or confounding factors. In addition, low fruit and vegetable intakes were significantly negatively correlated with diabetes and heart disease. This was in line with past research which found that adequate fruit and vegetable intakes were associated with reduced risk of cardiovascular disease, cancer, and all-cause mortality.[12] Less physical activity also showed a negative correlation with stroke (p-value = 0.053).

In terms of NCD risk behaviors, hypertension and stroke had strong correlations with daily smoking and consumption of fatty foods. Whereas diabetes and heart disease showed more complex relationships with food and beverage consumption factors.[13,14]

In addition to analyzing individual risk-behavior associations, we further explored clustering patterns to identify province-level groupings based on NCD risks and health behaviors. As already explained above, the provinces were grouped into 3 clusters (0, 1, 2). Those included in cluster 0 had a moderate level of smoking and consumption of unhealthy foods as well as moderate prevalence of hypertension, stroke, and diabetes, i.e., neither too high nor too low. Hence, it showed a mixed pattern between high and low risk behavior, but not on the extreme side. In this cluster, public health interventions needed to be directed towards secondary prevention such as educational campaigns and healthy food regulations. Meanwhile, those included in cluster 1 had the lowest rates of smoking and consumption of fatty foods. They also had lower prevalence of NCDs than other clusters.

This showed that they had a healthier lifestyle and higher physical activity compared to other provinces. Therefore, the risk of NCDs was more controlled. This cluster could serve as a good example of public health. The focus should be directed towards maintaining healthy habits through continuous education and promotion of healthy lifestyles. Those in cluster 2 had the highest levels of daily smoking, consumption of fatty foods and soft drinks, and low physical activity. This exacerbated NCD risk. Also, the prevalence of hypertension, stroke, and diabetes here was very high compared to the other two clusters. This cluster reflected the provinces with the highest NCD risk, thus requiring more attention in health policy through various public health programs such as more aggressive anti-smoking campaigns. It was also important to apply regulations strictly for consuming soft drinks and fatty foods here. Community-based sports programs could also be recommended to increase people's physical activity.[15]

CONCLUSION AND RECOMMENDATION

People's behaviors, particularly smoking and consumption of fatty foods, had a significant influence on the incidence of NCDs in Indonesia. Public health interventions needed to focus on

making people more aware of the dangers of smoking and educating them on the importance of a healthy diet and increased physical activity, especially in areas with health problems, to reduce the burden of NCDs in Indonesia.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: HPS, MA, DNAR and SL; data collection: HPS, MA, DNAR and SL; analysis and interpretation of results: HPS, MA, DNAR and SL; draft manuscript preparation: HPS, MA, DNAR and SL. All authors reviewed the results and approved the final version of the manuscript. HPS = Hanif Pandu Suhito; MA = Mahalul Azam; DNAR = Dina Nur Anggraini Ningrum; SL = Sholikun.

CONFLICTS OF INTEREST

There are no conflicts of interest.

REFERENCES

1. WHO. Noncommunicable Diseases 2023. 2025. https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases

2. Deonisia Arlinta. Rapor Merah Indonesia Dalam Pengendalian Tembakau. Jakarta: KOMPAS;2024. https://www.kompas.id/artikel/rapor-merah-indonesia-dalam-pengendalian-tembakau

3. Sakaria N, Indongo N. Socioeconomic and Behavioral Factors that Contribute to the Co-occurrence of Risk Factors for Noncommunicable Diseases. *BMC Public Health*. 2025;25(1):165. https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-024-20993-w

4. Kemenkes BKPK. Indonesian Health Survey (IHS)/*Survei Kesehatan Indonesia* (SKI); 2023. https://www.badankebijakan.kemkes.go.id/en/hasil-ski-2023/

5. STROBE. (n.d.). Strengthening the Reporting of Observational Studies in Epidemiology. 2025. https://www.strobe-statement.org/

6. Leone A. Smoking and Hypertension: Independent or Additive Effects to Determining Vascular Damage?. *Current Vascular Pharmacology*. 2011;9(5):585–593. https://doi.org/10.2174/157016111796642706

7. Borjesson, M., Onerup, A., Lundqvist, S., & Dahlöf, B. Physical Activity and Exercise Lower Blood Pressure in Individuals with Hypertension: Narrative Review of 27 RCTs. *British Journal of Sports Medicine.* 2016;50(6):356-361. https://bjsm.bmj.com/content/50/6/356.short

8. Adhikari K, Gupta N. Tobacco use: A Major Risk Factor for Non-communicable Diseases in Central Nepal. *Research Journal of Pharmaceutical, Biological and Chemical Sciences.* 2014;5(2):19-25. https://www.cabidigitallibrary.org/doi/full/10.5555/20143126253

9. Billingsley HE, Carbone S, Lavie CJ. Dietary Fats and Chronic Noncommunicable Diseases. *Nutrients.* 2018;10(10):1–16. https://doi.org/10.3390/nu10101385

10. Clemente-Suárez, V. J., Beltrán-Velasco, A. I., Redondo-Flórez, L., Martín-Rodríguez, A., & Tornero-Aguilera, J. F. Global Impacts of Western Diet and its Effects on Metabolism and Health: A Narrative Review. *Nutrients.* 2023;15(12);2749. https://doi.org/10.3390/nu15122749

11. Bernstein, A. M., de Koning, L., Flint, A. J., Rexrode, K. M., & Willett, W. C. Soda Consumption and the Risk of Stroke in Men and Women123. *The American Journal of Clinical Nutrition.* 2012;95(5):1190-1199. https://ajcn.nutrition.org/article/S0002-9165(23)02783-1/fulltext

12. Aune D, Giovannucci E, Boffetta P, Fadnes LT, et al. Fruit and Vegetable Intake and the Risk of Cardiovascular Disease, Total Cancer and All-cause Mortality: A Systematic Review and Dose-response Meta-analysis of Prospective Studies. *International Journal of Epidemiology are Provided here Courtesy of Oxford University Press.* 2022;46(3):1029

1056.
https://doi.org/10.1093/ije/dyw319

13. Howard B V., Wylie-Rosett J. Sugar and Cardiovascular Disease: A Statement for Healthcare Professionals from The Committee on Nutrition of the Council on Nutrition, Physical Activity, and Metabolism of the American Heart Association. *AHA Journals Switchboard.* 2002;106(4): 523-527.
https://doi.org/10.1161/01.CIR.00000195
52.77778.04

14. Arsyad DS, Westerink J, Cramer MJ, Ansar J, Wahiduddin, Visseren FLJ, et al. Modifiable Risk Factors in Adults with and without Prior Cardiovascular Disease: Findings from The Indonesian National Basic Health Research. *BMC Public Health*. 2022;22(1): 660.
https://doi.org/10.1186/s12889-022-13104-0

15. Heath GW. The Role of the Public Health Sector in Promoting Physical Activity: National, State, and Local Applications. *Journal of Physical Activity and Health*. 2009;6(2):159–167.
https://pubmed.ncbi.nlm.nih.gov/2012012
5/