Multivariate Adaptive Regression Spline for Prediction of Hypertension Cases the Measurement of Blood Pressure in Indonesia

Suroto[1*], Bambang Widjanarko Otok[2], Suharto[3], Arief Wibowo[3]

[1] Health Polytechnic Banjarmasin, Ministry of Health & Ph.D Student Faculty of Public Health, Airlanga University, Surabaya
[2] Department of Biostatistics and Demography, Faculty of Public Health, Airlanga University, Surabaya,
[3] Laboratory of Environmental and Health Statistic, 'Sepuluh November' Institute of Technology (ITS), Surabaya

ABSTRACT

Hypertension is one of the disease is not contagious diseases which is a public health problem. Uncontrolled Hypertension can trigger degenerative diseases, such as congestive heart failure, renal failure and vascular disease. Hypertension is called the silent killer because his nature the condition is asymptomatic and can cause a fatal stroke. With the increasing prevalence of cases of degenerative diseases, one only hypertension, then the researchers want to predict the variables very big role as one of the risk factors of Genesis hypertension. With clearly know the risk factors that play against genesis hypertension is expected to be used as a reference for the prevention and control so that they can reduce the prevalence of hypertension and prevent deaths from degenerative diseases, especially hypertension. The results of the study showed that the approach of MARS through the criteria the smallest GCV and $(R-O)^2$ largest obtained factor that affect the percentage of hypertension measurement of blood pressure is the percentage of smoking rooms, the percentage of alcohol consumption 1 Months, the percentage of food consumption innards, the percentage of food consumption is sweet, percentage of alcohol consumption 12 Last Month and the percentage of less physical activity. The percentage of hypertension blood pressure measurements of a seat can be decreased by 3.863 percent if the percentage of alcohol consumption 1 Months in a region under the 4.4 percent, and fell 3.437 if the percentage of alcohol consumption 12 last month in a region below the 4 percent.

KEYWORDS*: hypertension, alcohol consumption, smoking, MARS, GCV

INTRODUCTION

Hypertension is the main cause of death and cardiovascular disorder [1]. The main problem in hypertension is that more than ninety percent of hypertension including the essential which are not or not yet known why, 75 percent including mild hypertension (105-115 diastolic mmHg) numbered more than ninety percent of the patients [2].

Hypertension is also often called as 'silent killer' because the cause of complications on the heart, brain and kidneys [3]. But unfortunately, around 50% patients with hypertension is not aware of the existence of hypertension, so that the patients can be treated in the meaning of hypertension controlled with good, only around 10%-12%. Hypertension is defined as blood pressure higher than 160/90 mmHg is a disease that many people suffered. The results of field research shows that the frequency of hypertension is approximately ten percent of the population of the Indonesian population adults and this is the same with the country surrounding [4][5].

Hypertension based on the cause is divided into 2 the namely hypertension essential or Primary hypertension of unknown causes or hypoparathyroidism and Secondary hypertension or also called renal hypertension. Various Factors associated with essential hypertension, but there was not a description surely can explain the causes. The behavior is the factors that affect the degree of community health because healthy or not people who health environment individuals, families and communities is very depending on the behavior of the man himself as the attitude and lifestyle [6].

Hypertension is one of the disease is not contagious diseases which is a public health problem [7]. Uncontrolled Hypertension can trigger a degenerative diseases such as congestive heart failure, renal failure and vascular disease. Hypertension is called the silent killer because his nature the condition is asymptomatic and can cause a fatal stroke [8]. Although not treated, prevention and managements can decrease the occurrence of hypertension and the accompanying disease [9]. Hypertension is the cause of the death of number 3 after stroke and tuberculosis, namely 6.7 % from the population of death in all age in Indonesia [10].

---

*Corresponding Author: Suroto, Health Polytechnic Banjarmasin, Ministry of Health & Ph.D Student Faculty of Public Health, Airlanga University, Surabaya. E-mail: surotojahrani@yahoo.co.id

Based on the Basic Health Research (*Riskesdas*) 2013 [5], the prevalence of hypertension in Indonesia that acquired through the measurement at the age $\geq$ 18 years of 25.8 %, the highest in Bangka Belitung securities have totaled 30.9 %, followed South Kalimantan 30,8 %, East Kalimantan 29.6 % and West Java 29.4 %. The prevalence of hypertension in Indonesia that is obtained through the questionnaire diagnosed health workers of 9.4 %, diagnosed health workers or is drinking drugs of 9.5 %. So there are 0.1 % that medication itself. Respondents who have a normal blood but is drinking Hypertension medications 0.7 %. So the prevalence of hypertension in Indonesia by 26,5 % [4][2].

With the increasing prevalence of hypertension in a region, then the researchers want to predict the variables very big role as one of the risk factors of Genesis hypertension. With clearly know the risk factors that play against genesis hypertension with the approach of MARS [11] is expected to be used as a reference for the prevention and control so that they can reduce the prevalence of hypertension and prevent deaths from degenerative diseases in Indonesia, especially hypertension.

LITERATURE RIVIEW

The method of MARS [12] can be considered as a multivariate generalization. MARS implement stepwise selection strategy forwad/ customer whether they. Forward selection only started with the basic function of the constant $B_0(\mathbf{x})=1$ in this model. On each iteration consider adding the two terms on the model [13][14]:

$$B_j(x - t)_+ \textbf{ dan } B_j(t - x)_+$$

Where B_j is one of the basic function that has been selected, x is one of the variables that is not shown on predictors B_j and t is the location of the knot on the variable. Both the term this form that cause the biggest square number of errors on this model. Forward selection process continues until the many basic function that is included in the effort to adjust the data. "pruning procedure" customer whether they, namely linier stepwise regression, used with the basic functions shows stock "variables." Selected the best adjustment model with adjustment of measured with general cross-validation criteria, GCV [12][13][14].

The general model of MARS equation can be written as follows [12][13][14]

$$f(x) = \alpha_0 + \sum_{m=1}^{M} \alpha_m \prod_{k=1}^{K_m} \left[s_{km} \cdot \left(x_{v(k,m)-} t_{km} \right) \right]_+ + \varepsilon_i$$
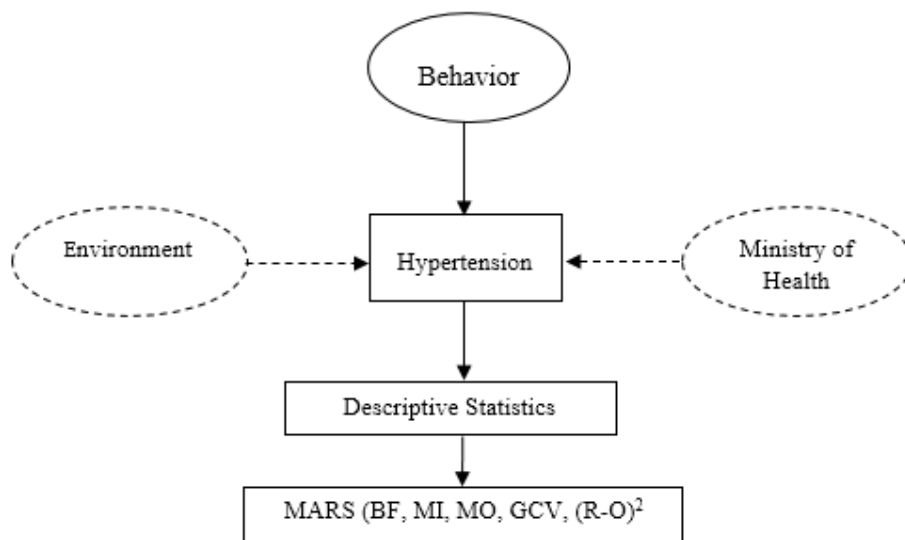
Where,

α_0	= constant basis function,
α_m	= coefficient from the basis function to-*m*,
M	= maximum basis function
K_m	= degrees of interaction,
s_{km}	= worth 1 if data is located to the right of the *knot point*, or worth -1 if data is located at the left of the *knot point*,
$x_{v(k,m)}$	= variables predictors,
t_{km}	= the value of the *knot* from the variables predictors $x_{v(k,m)}$.

METHODOLOGY

The data is hypertension secondary data obtained from Basic Health Research 2007 Indonesia [2], unit with the analysis of the province. Hypertension Blood Pressure measurement as a response variable, while the behavior as predictors variables which consists of 13 Indicators, namely less eat the fruit and vegetables (X1), alcohol consumption 12 Last Month (X2), alcohol consumption 1 Months (X3), sweet food consumption (X4), the consumption of salty food (X5), Fatty food consumption (X6), food consumption innards (X7), roasted food consumption (X8), Preserved food consumption (X9), the consumption of the food with caffeine (X10), the consumption of the Food Seasonings (X11), less physical activity (X12), Smoking (X13) and Obesity (X14) [4][15].

Hypertension based on the measurement result/blood pressure examination/blood pressure, assigned using measuring instruments digital tensimeter. Digital Tensimeter validated by using the basic standard measurement of blood pressure.The measurement of the blood pressure is done on the respondents age 15 years. Each of the respondents measured tensinya minimum 2 times, if the measurement result to two different more than 10 mmHg compared to the first measurement, then done measurement to three. Two measurement data with the smallest difference calculated reratanya as a result of the measurement of blood pressure. Hypertension Criteria used on the announcement of the case refers to the criteria for diagnosis of JNC VII 2003, namely systolic blood pressure measurement results 140 mmHg or diastolic blood pressure 90 mmHg [2][4].

Based on the results of the analysis of the previous library, hypertension served on the following conceptual framework [4][16].

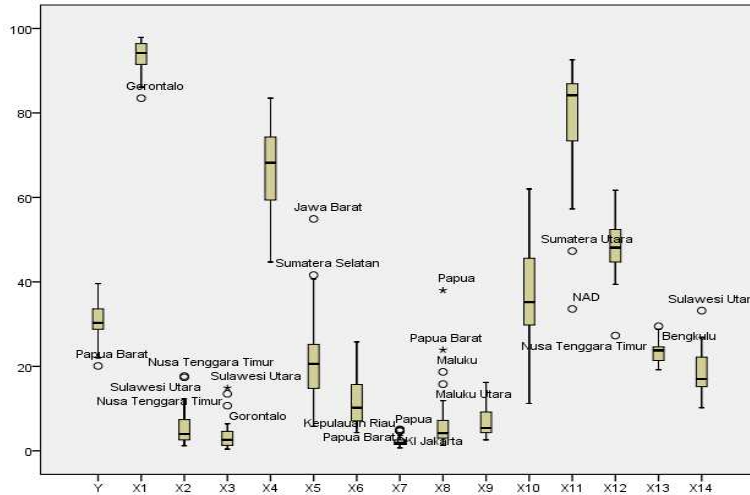**Figure 1. Modeling flow MARS on the case of hypertension**

The first step, done analysis of the descriptive statistics of the variables veriabel predictor. Second, to the formation of the model of MARS include: determine BF; determine MI; determine MO, the value of the smallest GCV.

RESULTS AND DISCUSSION

Description of this research includes the mean and standard deviation from each of the research variables. Now in detail is presented in the following table.

Table 1. The value of the Mean and Standard deviation of Variables Hypertension

	Minimum	Maximum	Mean	Std. Deviation
The percentage of Hypertension Measurement of Blood Pressure (Y)	20.1	39.6	30.71	4.40
The percentage of less eat the fruit and vegetables (X1)	83.5	97.9	93.49	3.40
The percentage of alcohol consumption 12 Last Month (X2)	1.2	17.7	5.50	4.14
The percentage of alcohol consumption 1 Months (X3)	0.4	14.9	3.68	3.47
The percentage of food consumption sweet (X4)	44.7	83.5	66.89	10.32
The percentage of food consumption of salt (X5)	5.8	54.9	21.66	10.27
The percentage of Fatty food consumption (X6)	4.3	25.8	12.02	6.09
The percentage of food consumption innards (X7)	0.7	5.1	2.22	1.19
The percentage of food consumption roasted (X8)	1.3	38.0	6.98	7.53
The percentage of food consumption was embalmed (X9)	2.6	16.2	6.65	3.23
The percentage of food consumption with caffeine (X10)	11.2	62.0	37.68	12.12
The percentage of food consumption seasonings (X11)	33.6	92.6	78.99	12.94
The percentage of less physical activity (X12)	27.3	61.7	48.31	6.53
The percentage of smoking (X13)	19.2	29.5	23.42	2.62
The percentage of obesity (X14)	10.2	33.2	18.78	4.92

Figure 2. Boxplot Research Variables Hypertension in Indonesia

Table 1 and Figure 2 shows that based measurement results of blood pressure, an average of the prevalence of hypertension in Indonesia by 30.706%, the prevalence of hypertension lowest in West Papua Province (20.1%) and highest in South Kalimantan Province (39.6%). The average population less eat the fruit and vegetables (X1) of 93.49%, lowest in Gorontalo Province (83.5%) and highest Riau (97.9%). The average alcohol consumption 12 Last Month (X2) by 5.5%, the lowest in South Kalimantan Province (39.6%) and the highest in the province of East Nusa Tenggara (17.7%). The average alcohol consumption 1 Months (X3) of 3.68 percent, the lowest in the province of Nanggroe Aceh Darussalam (0.4%) and the highest in the province of North Sulawesi (14.9%).

The average food consumption sweet (X4) of 66.8%, the lowest in the province of Bali (44.7%) and the highest in South Kalimantan Province (83.5%). The average consumption of salty food (X5) of 21.6%, lowest in Central Sulawesi province (5.8%) and the highest in the province of West Java (54.9%). The average Fatty food consumption (X6) of 12.02%, the lowest in the province of East Nusa Tenggara (4.3%) and the highest in Gorontalo Province (25.8%). The average food consumption innards (X7) of 2.22 percent, the lowest in the province of Central (0.7%) and the highest in West Papua province (5.1%). The average food consumption roasted (X8) of 6.98 percent, the lowest in the province of West Kalimantan (1.3%) and the highest in the province of Papua (38.0%). The average food consumption was embalmed (X9) of 6.65%, lowest in Gorontalo Province (2.6%) and the highest in DKI Jakarta Province (16.2%). The average food consumption with caffeine (X10) of 37.68%, the lowest in the province in Yogyakarta (11.2%) and the highest in the province of Bali (62.0%). The average food consumption seasonings (X11), USD 78.99%, the lowest in the province of Nanggroe Aceh Darussalam (33.6%) and the highest in the province of Central Kalimantan (92.6%). The average less physical activity (X12) of 48.31%, the lowest in the province of East Nusa Tenggara (27.3%) and the highest in East Kalimantan province (61.7%). The average Smoking (X13) of 23.42%, the lowest in the Maluku province (19.2%) and the highest in Bengkulu province (29.5%). The average Obesity (X14) of shed 18.78%, the lowest in the province of East Nusa Tenggara (10.2%) and the highest in the province of North Sulawesi (33.2%).

Furthermore done modeling the prevalence of hypertension by the method of MARS. The election of the best model obtained with how to compare the value of the smallest GCV and values $(R-O)^2$ greatest. The comparison is done on the various number of basis functions (BF), maximum interaction (MI) and minimum observation (MO). The best model obtained on GCV = **8.059** and $(R-O)2 = $ **0.999** as follows:

$$Y = 31.433 + 1.546\mathbf{BF1} + 2.456\mathbf{BF2} - 0.560\mathbf{BF3} - 1.854\mathbf{BF4} + $$
$$+ 2.508\mathbf{BF6} - 3.863\mathbf{BF7} - 3.437\mathbf{BF9} + 0.389\mathbf{BF10} + 2.701\mathbf{BF12} \qquad (1)$$

Where:

BF1 = max(0, X7 - 0.7);	**BF2 = max(0, X13 - 24.6);**	**BF3 = max(0, 24.6 - X13)**
BF4 = max(0, X4 - 79.3);	**BF6 = max(0, X3 – 4.4);**	**BF7 = max(0, 4.4 - X3)**
BF9 = max(0, 4.0 - X2);	**BF10 = max(0, X12 - 47.3);**	**BF12 = max(0, X2 – 8.8)**

Equation (1) obtained in interpret based on the basis of the function as follows:

- **BF1 = max(0, X7 - 0.7)**
  The basis function 1 loading food consumption innards variables (X7). It can be said that the hypertension in a region affected by food consumption innards (X7) of 1.546, if the percentage of food consumption innards (X7) in a region above 0.7 then will increase hypertension of 1.546, and when in aggregate and hypertension a predictable area of 32.979%.
- **BF2 = max(0, X13 - 24.6) and BF3 = max(0, 24.6 - X13)**
  The basis function 2 and 3 loading variables smoking (X13). It can be said that the hypertension in a region influenced by the percentage of smoking (X13) of which totaled 2.456, if the percentage of smoking (X13) in a region above 24.6 then will increase hypertension of which totaled 2.456, and when in aggregate and hypertension a predictable area of 33.889%. But if the percentage of smoking (X13) in a region under 24.6 then will reduce hypertension by 0.56 and when in aggregate and hypertension a predictable area of 30.873%.
- **BF4 = max(0, X4 - 79.3)**
  The basis function 4 load sweet food consumption variables (X4). It can be said that the hypertension in a region influenced by the percentage of food consumption sweet (X4) of -1.854, if the percentage of food consumption sweet (X4) in a region under 79.3m per cent then will reduce hypertension by 1.854 and when in aggregate and hypertension a predictable area of 29.579%.
- **BF6 = max(0, X3 – 4.4) and BF7 = max(0, 4.4 - X3)**
  The basis function 6 and 7 loading variables alcohol consumption 1 Months (X3). It can be said that the hypertension in a region influenced by the percentage of alcohol consumption 1 Months (X3) of which totaled 2.456 and (-3.863), if the percentage of alcohol consumption 1 Months (X3) in a region above 4.4 percent then will increase hypertension of 2.508, and when in aggregate and hypertension a predictable area of 33.941%. But if the percentage of alcohol consumption 1 Months (X3) in a region under 4.4 percent then will reduce hypertension by 3.863 and when in aggregate and hypertension a region predicted by 27.58%.

- **BF9 = max(0, 4.0 - X2)**
  The basis function 9 loading variables alcohol consumption 12 last month (X2). It can be said that the hypertension in a region influenced by the percentage of alcohol consumption 12 Last Month (X2) of -3.437, if the percentage of alcohol consumption 12 Last Month (X2) in a region under 4 percent and will decrease hypertension of 3.437 and when in aggregate and hypertension a predictable area of 27.996%.
- **BF10 = max(0, X12 - 47.3)**
  The basis function 10 loading variables less physical activity (X12). It can be said that the hypertension in a region influenced by the percentage of less physical activity (X12) of 0.389, if the percentage of less physical activity (X12) in a region above 47.3 percent then will increase hypertension of 0.389 percent and when in aggregate and hypertension a predictable area of 31.822%.
- **BF12 = max(0, X2 – 8.8)**
  The basis function 12 loading variables alcohol consumption 12 last month (X2). It can be said that the hypertension in a region influenced by the percentage of alcohol consumption 12 Last Month (X2) of 2.701, if the percentage of alcohol consumption 12 Last Month (X2) in a region above 8.8 percent then will increase hypertension of 2.701 percent and when in aggregate and hypertension a predictable area of 34.134%.

The importance level predictors variables in the influence of hypertension in MARS used the value of relative variables predictors. Complete Finances on the following table.

Table 2. Relative Variable Importance

The variables Predictors	Importance	-gcv
The percentage of smoking (X13)	100.000	28.823
The percentage of alcohol consumption 1 Months (X3)	76.559	20.229
The percentage of food consumption innards (X7)	66.742	17.308
The percentage of food consumption sweet (X4)	65.515	16.971
The percentage of alcohol consumption 12 Last Month (X2)	58.831	15.245
The percentage of less physical activity (X12)	50.975	13.454

Table 2 shows that the model of the genesis of hypertension of a region influenced by the percentage of smoking (X13) (100%), the percentage of alcohol consumption 1 Months (X3) (76.559%), the percentage of food consumption innards (X7) (66.742%), the percentage of food consumption sweet (X4) (65.515%), the percentage of alcohol consumption 12 Last Month (X2) (58.831%), and the percentage of less physical activity (X12) (50.975%).

CONCLUSION

The percentage of the prevalence of hypertension a region can be revealed through the percentage of smoking in a region under 24.6 percent, percentage of sweet food consumption in a region under 79.3m per cent, percentage of alcohol consumption 1 Months in a region under 4.4 percent and alcohol consumption 12 last month in a region under 4 per cent. On the contrary, the percentage of hypertension a region will be increased if the percentage of food consumption innards in a region above 0.7 percent if the percentage of smoking (X13) in a region above 24.6 percent, if the percentage of alcohol consumption 1 Months in a region above 4.4 percent, percentage less physical activity in a region above 47.3 percent and the percentage of alcohol consumption 12 last month in a region above 8.8 percent.

The province that need to be focused on the prevalence of hypertension based on alcohol consumption in the last 1 month of 4.4 percent is Bali (4.6%), the Gorontalo (10.7%), West Kalimantan (4.8%), Maluku (5%), East Nusa Tenggara (tumbled 13.5%), West Papua (4.9%), Central Sulawesi (6.4%), Southeast Sulawesi (5.8%) and North Sulawesi (14.9%).

REFERENCES

[1]   Noor, N.N., (2008). *Epidemiologi*, PT Rineka Cipta, Jakarta.
[2]   Kementrian Kesehatan RI., (2007). *Riset Kesehatan Dasar (Riskesdas 2007),* Badan Penelitian dan Pengembangan Kesehatan, Jakarta.
[3]   Yogiantoro, M., (2006). *Hipertensi Esensial dalam Buku Ajar Ilmu Penyakit Dalam*. Jilid I, Edisi IV, Jakarta: FKUI.
[4]   Rahajeng, E. dan Sulistyowati, T., (2009). Prevalensi Hipertensi dan Determinannya di Indonesia. *Majalah Kedokteran Indonesia*, Volum: 59, Nomor: 12
[5]   Kementrian Kesehatan RI., (2013). *Riset Kesehatan Dasar (Riskesdas 2013),* Badan Penelitian dan Pengembangan Kesehatan, Jakarta
[6]   Susilo Y. dan Wulandari, A., (2011). *Cara Jitu Mengatasi Hipertensi, Bahaya, Diagnosis dan Mitos. Pencegahan dan Terapi Penyembuhan*. Andi, Yogyakarta.
[7]   Tudor H.J, Tom, F., (2010). *Tanya Jawab Seputar Tekanan Darah Tinggi*, Edisi 2, Arcan, Jakarta.
[8]   Arfiza, R., Umaporn, B., and Piyanuch, J., (2012). Development and Pilot Study of Group-Based Dietary Self-Management Program for Community Dwellers with Hypertension. *Nurse Media Journal of Nursing*, 2, 2, 2012, 383-396
[9]   Wahdah, N., (2011). *Menaklukan Hipertensi mendeteksi, Mencegah dan Mengobati dengan Cara Medis dan Herbal*, Multi Press, Yogyakarta.
[10] Sutanto, (2010), *Cekal (cegah dan tangkal) Penyakit Modern (Hipertensi, Stroke, Jantung, Kolesterol dan Diabetes)*, CV Andi Offset, Yogyakarta.
[11] Isnawati, Bambang, W.O, Suharto, Arief W., (2017). Model Prediction of Incident Dengue Fever Based on Climate Factors with Multivariate Adaptive Regression Spline Longitudinal. *J. Appl. Environ. Biol. Sci.*, 7(4)175-179, 2017 © 2017, TextRoad Publication.
[12] Friedman, J.H., (1991). Multivariate Adaptive Regression Splines. *The Annals of Statistics*, Vol. 19.
[13] Otok, B.W. (2008). Bootstrap Pada Pemodelan Multivariat Adaptive Regression Spline. *Disertasi* UGM. Yogyakarta. (tidak dipublikasikan)
[14] Otok, B.W. and Arleina, O.D. (2014). Bootstrap Aggregating Multivariate Adaptive Regression Splines (Bagging MARS) for Poor Households Classification in Region of Jombang. *Available at SSRN 2489898*
[15] Rizka, H. dan Noor, C., (2016). Merokok dan stres dengan kejadian hipertensi, *The 3rd Universty Research Colloquium*. STIKES Muhammadiyah Kudus, Jawa Tengah.
[16] Somayeh, K., Alireza, R. and Hossein, J., (2012). The relationship between self-concept and adherence to medication regimens in patients with hypertension in Urmia, Iran. *ARYA Atherosclerosis Journal* 2012, 8(Special Issue in National Hypertension Treatment): S174-S177.