International Journal of Innovative
Computing, Information and Control
Volume **18**, Number **5**, October **2022**

 ISSN 1349-4198
pp. 1463–1475

DEVELOPMENT OF A MODIFIED CERTAINTY FACTOR MODEL FOR PREDICTION OF METABOLIC SYNDROME

Sri Kusumadewi[1,*], Linda Rosita[2] and Elyza Gustri Wahyuni[1]

[1]Department of Informatics
[2]Faculty of Medicine
Universitas Islam Indonesia
Kaliurang Street km 14.5 Sleman, Yogyakarta 55584, Indonesia
{ linda.rosita; elyza }@uii.ac.id
*Corresponding author: sri.kusumadewi@uii.ac.id

Received March 2022; revised July 2022

Abstract. *Early detection of metabolic syndrome (MetS) is very important because it has the potential to cause several complications of the disease. In Indonesia, the criteria issued by the NCEP ATP III are often used to diagnose MetS. The diagnostic criteria from NCEP ATP III use indicators that cannot be tested on their own (require help from a clinical lab), like HDL and triglycerides. This is difficult to implement in clinics or with non-hospital health care providers. This study aims to develop a model of a decision support system for early detection of MetS using a modified Certainty Factor (CF). The model includes a knowledge base consisting of 50 IF-THEN rules. The model was tested on 198 data sets by implementing forward reasoning. The performance of the model is measured through indicators of accuracy, sensitivity, specificity, and precision of 100%. The same results were also shown in testing the data sets in Groups 1, 2, and 4. While the tests in Group 3 showed an accuracy of 87.7%. Overall, the proposed model is able to show an appropriate value. It can be seen that there is a good gradation of values from Group 1 to Group 4.*

Keywords: Certainty factor, Knowledge base, Rules, Indicators, Forward reasoning

1. **Introduction.** A metabolic syndrome (MetS) is a collection of health problems that occur at the same time. These conditions include high blood pressure, abdominal fat accumulation, and elevated blood sugar, cholesterol, and triglyceride levels [1,2]. In many cases, MetS has been shown to have an effect on disease complications in the future. Several researchers have analyzed the relationship between metabolic syndrome and the emergence of other health problems, including its association with chronic kidney disease [3] and coronary heart disease [4]. Some health organizations such as the World Health Organization (WHO), European Group for the Study of Insulin Resistance (EGIR), National Cholesterol Education Program Adult Treatment Panel III (NCEP ATP III), American Association of Clinical Endocrinologists (AACE), and International Diabetes Federation (IDF) have established several criteria for the correct clinical diagnosis of MetS [5].

In Indonesia, the diagnostic criteria for MetS that are often used are NCEP ATP III because this standard is simpler and most reliable to apply [6]. NCEP ATP III uses five indicators for the diagnosis of MetS, namely Waist Circumference (WC), High Density Lipoprotein (HDL), Triglycerides (TG), Fasting Blood Sugar (FBS), and Blood Pressure (BP) [7]. However, this approach remains dichotomous, with only two risk categories for MetS, namely high risk and non-high risk. This classifier could not reflect a gradation or range of dangerous levels increasing or decreasing. As a result, using a continuous

DOI: 10.24507/ijicic.18.05.1463

MetS risk score has greater advantages than using a dichotomous approach since it is more sensitive and has a smaller chance of inaccuracy [8]. The continuous approach is also more accommodating for solving problems that contain uncertainty. Uncertainty in the measurement of medical data is often encountered, especially clinical data. Therefore, a method is needed to measure the uncertainty of the measurement results [9], and in health care [10].

MetS must be detected early since it has the potential to cause serious problems. Early detection of MetS can be independently checked by considering the results of the required indicator test. The measurement of waist circumference as an indicator of obesity can be done on his own. However, the Body Mass Index (BMI) can also be used to assess obesity by looking at nutritional status. BMI considers nutritional status based on height and weight. Blood pressure can be tested independently by using a digital blood pressure meter. Although fasting blood sugar can be tested independently, people who want to take their blood sugar sometimes forget or do not have time to fast first. Random glucose is more reliable and can be done independently. Similarly, HDL and triglyceride indicators are still not widely used independently. So, total cholesterol is a better measure of the body's cholesterol levels than triglycerides.

This study aims to build a decision support system model using a modified certainty factor for early detection of MetS. In addition to using the indicators set by NCEP ATP III, this study also uses other indicators that are easier to do independently, such as weight, height, random glucose, and total cholesterol. The use of this modified CF will also solve the dichotomous prediction problem because CF is continuous. CF, as part of soft computing, can be used as a tool to solve the problems of day-to-day life. To deal with the uncertainty and risk of different processes, soft computing methods are better than hard computing methods [11-13].

MetS is diagnosed by considering a number of criteria. However, these criteria are still predominantly dichotomous, with only two risk categories for MetS, namely high risk and low risk. This binary classification could not indicate gradation or a spectrum of risk levels increasing or decreasing. Consequently, the use of a continuous MetS risk score has more advantages than the dichotomous approach because it is more sensitive and has a lower error risk [14,15]. We have conducted the same study for the prediction of MetS with a continuous approach using a fuzzy system, but the proposed model does not accommodate commonly used measurement indicators such as cholesterol, weight, height, and random glucose. The data sets used in the previous research still use simulation data. In this study, modified CF was used as a method for determining the certainty value. In order for the CF rule to be useful, it needs to be changed so that it does not just depend on the opinions of humans. It also needs to be based on objective data conditions.

The study focused on developing a knowledge base using the MetS diagnostic reference from the NCEP ATP III. Furthermore, each rule will be equipped with a CF value. Inference is performed on a number of data samples. Finally, the evaluation will be done by using a confusion matrix to measure how well the system works.

2. **MetS Diagnostic Criteria.** The classification of cholesterol according to ATP III is shown in Table 1. The table shows the classification of the severity of cholesterol measurement results. The classification of nutritional status based on BMI according to WHO is shown in Table 2. According to International Diabetes Federation criteria, normal Waist Circumference (WC) for South Asian ethnicities is lower than or equal to 90 cm for males, or lower than or equal to 80 cm for females [16].

Blood sugar levels in diagnosing diabetes according to the National Institute for Clinical Excellence (NICE) can be seen in Table 3. Samples of plasma glucose were taken in three

TABLE 1. ATP III classification of cholesterol [17]

Item (mg/dL)	Classification
LDL cholesterol	
< 100	Optimal
100-129	Near or above optimal
130-159	Borderline high
160-189	High
≥ 190	Very high
Total cholesterol	
< 200	Desirable
200-239	Borderline high
≥ 240	High
HDL cholesterol	
< 40	Low
≥ 60	High
Triglycerides	
< 150	Normal
150-199	Borderline high
200-499	High
≥ 500	Very high

TABLE 2. Nutritional status [18]

Body Mass Index (BMI)	Nutritional status
< 18.5	Under weight
18.5-24.9	Normal weight
25.0-29.9	Pre-obesity
30.0-34.9	Obesity class I
35.0-39.9	Obesity class II
≥ 40	Obesity class III

TABLE 3. Blood sugar levels in diagnosing diabetes according to NICE [19]

Plasma glucose test	Normal	Prediabetes	Diabetes
Random	< 200 mg/dL	N/A	≥ 200
Fasting	< 100 mg/dL	100-125	≥ 126
2 hour post-prandial	< 140 mg/dL	140-199	≥ 200

ways: while the person was fasting, at random, and two hours after eating. Table 4 shows the categorization of blood pressure according to the American Heart Association. The categories that need attention are high blood pressure and Hypertensive Crisis. According to NCEP ATP III, a proper clinical diagnosis of MetS is when at least three out of five criteria as shown in Table 5 are appropriate [21].

3. **Certainty Factor.** The Certainty Factor (CF) model is one of the models for the representation and manipulation of uncertain knowledge in a rule-based expert system. Certainty factor theory is an alternative to Bayesian reasoning. Bayesian formulas are quite complex and inadequate for human reasoning. Therefore, when reliable statistical

TABLE 4. Blood pressure categories according to American Heart Association [20]

Category	Systolic (mmHg)	And/or	Diastolic (mmHg)
Normal	< 120	and	< 80
Elevated	120-129	and	< 80
High Blood Pressure (Hypertension) Stage 1	130-139	or	80-89
High Blood Pressure (Hypertension) Stage 2	≥ 140	or	≥ 90
Hypertensive Crisis	> 180	and/or	> 120

TABLE 5. Diagnosis criteria of MetS according to NCEP ATP III

Criteria	Conditions
Central Obesity	WC $\geq$ 102 cm (male), WC $\geq$ 88 cm (female) In Asian: WC $\geq$ 90 cm (male), WC $\geq$ 80 cm (female)
Dyslipidemia Triglycerides (TG)	TG $\geq$ 150 mg/dL or medication
Dyslipidemia HDL	HDL-C $<$ 40 mg/dL (male), HDL-C $<$ 50 mg/dL (female), or medication
Blood Pressure (BP)	Systolic BP $\geq$ 130 mmHg or Diastolic BP $\geq$ 85 mmHg, or medication
Glucose	Fasting plasma glucose: $\geq$ 100 mg/dL or medication

information is not available or the independence of evidence cannot be assumed, a certainty factor can be used [22]. Certainty factors are generally based on the heuristics of human experts.

Given evidence E, the certainty value of hypothesis H is written as CF(H|E). A CF(H|E) between 0 and 1 means that a person's belief in H given E increases, while a CF(H|E) between -1 and 0 means that a person's belief decreases [23]. CF formula for a rule IF E THEN H is

$$\mathrm{CF}(\mathrm{H}, \mathrm{e}) = \mathrm{CF}(\mathrm{E}, \mathrm{e}) * \mathrm{CF}(\mathrm{H}, \mathrm{E}) \tag{1}$$

$\mathrm{CF}(\mathrm{H}, \mathrm{e})$ is the CF hypothesis H based on uncertain evidence e. $\mathrm{CF}(\mathrm{E}, \mathrm{e})$ is CF evidence E taken by uncertain evidence e. $\mathrm{CF}(\mathrm{H}, \mathrm{E})$ is the CF of hypothesis H with the assumption that the evidence is known certainly or $\mathrm{CF}(\mathrm{E}, \mathrm{e}) = 1$.

If the evidence is known, then $\mathrm{CF}(\mathrm{E}, \mathrm{e}) = 1$, so $\mathrm{CF}(\mathrm{H}, \mathrm{e}) = \mathrm{CF}(\mathrm{H}, \mathrm{E})$. This is called the attenuation factor. The attenuation factor is based on the assumption that the evidence is known absolutely. If the evidence is not known, then the rule must determine $\mathrm{CF}(\mathrm{E}, \mathrm{e})$. Then, if the model consists of several rules, it is necessary to calculate the CF combine of these rules as follows [24]:

$$\mathrm{CF}_{combine}(\mathrm{CF}_1, \mathrm{CF}_2) = \begin{cases} \mathrm{CF}_1 + \mathrm{CF}_2(1 - \mathrm{CF}_1) & \mathrm{CF}_1 \ \& \ \mathrm{CF}_2 > 0 \\ \dfrac{\mathrm{CF}_1 + \mathrm{CF}_2}{1 - \min(|\mathrm{CF}_1|, |\mathrm{CF}_2|)} & \mathrm{CF}_1 \ xor \ \mathrm{CF}_2 < 0 \\ \mathrm{CF}_1 + \mathrm{CF}_2(1 + \mathrm{CF}_1) & \mathrm{CF}_1 \ \& \ \mathrm{CF}_2 < 0 \end{cases} \tag{2}$$

So far, the value of the CF rule has been obtained based on the knowledge of experts. This allows the subjectivity of the data. In our study, the CF rule values were based on

the MetS diagnostic guidelines compiled by the NCEP ATP III. Thus, the objectivity of the data will be more guaranteed.

4. **Method.** This study was completed in stages, as shown in Figure 1.

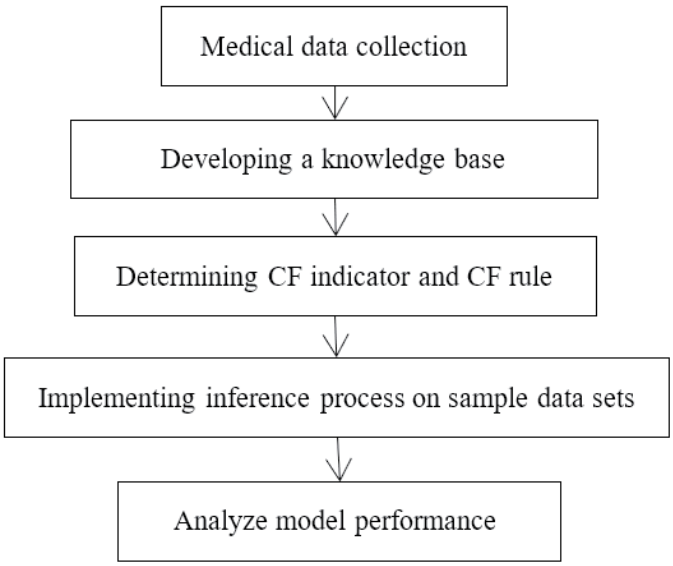

FIGURE 1. Research stages

Stage 1: medical data collection. Data collection aims to test the model and develop a knowledge base. The profile data sets for model testing have been described in the previous section. The data for developing the knowledge base is obtained from several references related to the MetS diagnostic criteria and the categorization of each indicator considered. Data from experts is needed to give the CF value for each rule.

Stage 2: developing a knowledge base. The knowledge base contains a collection of knowledge that is represented in the form of an IF-THEN rule. The knowledge base is grouped into four categories. Category 1 contains rules for the classification of indicators. This category will prove the truth of the hypothesis that the indicator meets the diagnostic criteria, for example, "obesity", "blood pressure is high", and "cholesterol is high". Category 2 contains the MetS diagnostic criteria rules adopted from the NCEP ATP III. This category will also generate rules involving three indicators in the antecedent section. In addition to using standard indicators set by NCEP ATP III, alternative indicators that are commonly used in independent tests are also used. Category 3 contains rules involving two indicators in the antecedent part. Category 4 contains rules that only involve one indicator in the antecedent part.

Stage 3: determining CF indicator and CF rule. CF is used as a method to determine the certainty value. There are two types of CF values needed as input, namely CF(E, e) or CF indicators, and CF(H, E) or CF rule. Generally, CF(E, e) is given by the user, but in this study we use the classification guide for each indicator as described previously. Similarly, CF(H, E) was given by human experts, but in this study we took the value of CF(H, E) based on the NCEP ATP III reference and approval from human experts.

Stage 4: implementing the inference process on sample data sets. After the knowledge base is formed, it is then implemented into the sample data sets. The inference process is directed by using forward reasoning. This reasoning concept starts from the antecedent part (the condition given after the IF) first to prove the truth of a hypothesis

(the part after THEN). Finding facts or knowledge in the knowledge base starts from Category 1 to Category 4. The combined CF value is calculated in each category. The final CF is calculated from the maximum CF value in each of the four categories.

Stage 5: analyze model performance. Model testing is done by matching the inference results on the data sample with the diagnosis results according to the NCEP ATP III criteria. There are two types of testing. First, test the performance of the model through indicators of accuracy, sensitivity, specificity, and precision. Second, evaluate the results of calculating the CF value by grouping the data set into four groups, namely groups that fulfill the criteria for three or more indicators (Group 1), groups that fulfill the criteria for two indicators (Group 2), groups that fulfill the criteria for only one indicator (Group 3), and groups that have no indicators that satisfy the criteria (Group 4). This evaluation aims to see whether the CF ranges in each group have been modeled properly.

5. **Material.** This study used 198 data sets. The data set includes ten indicators, namely Height (H), Weight (W), Waist Circumference (WC), HDL, Triglycerides (TG), total Cholesterol (Chol), Random Glucose (RG), Fasting Blood Sugar (FBS), Systolic blood pressure (Sys), and Diastolic blood pressure (Dias). Data sets are taken from medical records in hospitals and data collection by health care provider partners. The data sample required is at least 40 years old. This sample consisted of 69 males and 129 females. Each data set does not fully have these ten indicators. For example, there are 104 data sets that are not supported by the waist circumference indicator. The data sets profile is shown in Table 6. The data set is supported by at least two indicators and at most seven indicators. The highest 46% (91 data sets) is supported by seven indicators, and the smallest is supported by two indicators and only one data set.

Table 6. Data set profile

Not supported by indicators	Number of data sets
Height (H)	104
Weight (W)	104
Waist Circumference (WC)	104
HDL	94
Triglycerides (TG)	95
Cholesterol (Chol)	4
Fasting Blood Sugar (FBS)	125
Random Glucose (RG)	105
Systolic blood pressure (Sys)	11
Diastolic blood pressure (Dias)	11

6. **Results and Discussion.** Based on the data in Tables 1 to 4, CF can be set for each condition as shown in Table 7. For example, in Table 2, the categorization of BMI is divided into three groups. The obesity group (BMI ≥ 30) indicates that someone who has a BMI ≥ 30 is really obese (CF = 1). The pre-obesity group has a BMI in the range of 25.0-29.9, meaning that they are in moderately obese condition (CF = 0.5). The pre-obese and underweight groups have BMI < 25.0, meaning they are in a non-obese condition (CF = 0).

The determination of rules in Category 1 is based on the presence of indicators. Categories of obesity, high HDL, high triglycerides, high cholesterol, high blood sugar, and high blood pressure were obtained based on 14 rules in Category 1 as shown in Figure 2.

Table 7. CF of a condition in a certain class

Item	Unit	Range/condition	Classification	CF
Waist Circumference (WC)	cm	>90 (male) or >80 (female)	Large	1.00
		≤ 90 (male) or ≤ 80 (female)		0.00
Body Mass Index (BMI)	kg/cm^2	≥ 30	Obesity	1.00
		25.0-29.9		0.50
		<25.0		0.00
Total Cholesterol (Chol)	mg/dL	≥ 240	High	1.00
		200-239		0.50
		<200		0.00
HDL cholesterol (HDL)	mg/dL	>60 (male)	Low	0.00
		<40 (male)		1.00
		<50 (female)		1.00
Triglycerides (TG)	mg/dL	>150	High	1.00
		≤ 150		0.00
Fasting plasma glucose	mg/dL	≥ 100	High	1.00
		<100		0.00
Random Glucose	mg/dL	≥ 200	High	1.00
		<200		0.00
Systolic blood pressure	mmHg	>130	High	1.00
		120-130		0.50
		<120		0.00
Diastolic blood pressure	mmHg	>85	High	1.00
		80-85		0.50
		<80		0.00
In medication of HDL	–	Yes	Yes	1.00
		No		0.00
In medication of triglycerides	–	Yes	Yes	1.00
		No		0.00
In medication of diabetes	–	Yes	Yes	1.00
		No		0.00
In medication of hypertension	–	Yes	Yes	1.00
		No		0.00
In medication of cholesterol	–	Yes	Yes	1.00
		No		0.00

The CF for each rule is set at 1. Next, CF for "Obesity", "HDL is Low", "Triglycerides is High", "Cholesterol is High", "Blood Sugar is High", and "Blood Pressure is High" are calculated using the CF combine formula as shown in Equation (1). According to NCEP ATP III, a proper clinical diagnosis of MetS is when at least three out of five criteria are met. Suppose there are five conditions. If three of them are to appear in one rule, then ten combinations of rules like the one shown in Figure 3 will be made.

As previously explained, triglyceride checking is not commonly done independently (mostly it is still done in clinical laboratories). Therefore, in this study we used total cholesterol as another alternative to measure cholesterol levels. Thus, there will be some new rules that replace triglycerides with cholesterol. From the ten rules that have been

R1 : IF Waist Circumference is Large THEN Obesity (1.00)
R2 : IF BMI level is High THEN Obesity (1.00)
R3 : IF HDL level is Low THEN HDL is Low (1.00)
R4 : IF In medication of HDL THEN HDL is Low (1.00)
R5 : IF Triglycerides level is High THEN Triglycerides is High (1.00)
R6 : IF In medication of Triglycerides THEN Triglycerides is High (1.00)
R7 : IF Cholesterol level is High THEN Cholesterol is High (1.00)
R8 : IF In medication of Cholesterol THEN Cholesterol is High (1.00)
R9 : IF FBS level is High THEN Blood Sugar is High (1.00)
R10 : IF Random glucose level is High THEN Blood Sugar is High (1.00)
R11 : IF In medication of diabetes THEN Blood Sugar is High (1.00)
R12 : IF Systolic blood pressure level is High THEN Blood Pressure is High (1.00)
R13 : IF Diastolic blood pressure level is High THEN Blood Pressure is High (1.00)
R14 : IF In medication of hypertension THEN Blood Pressure is High (1.00)

FIGURE 2. Rules in Category 1

R15 : IF Obesity and HDL is Low and Triglycerides is High THEN MetS Risk is High (1.00)
R16 : IF Obesity and HDL is Low and Blood Sugar is High THEN MetS Risk is High (1.00)
R17 : IF Obesity and HDL is Low and Blood Pressure is High THEN MetS Risk is High (1.00)
R18 : IF Obesity and Triglycerides is High and Blood Sugar is High THEN MetS Risk is High (1.00)
R19 : IF Obesity and Triglycerides is High and Blood Pressure is High THEN MetS Risk is High (1.00)
R20 : IF Obesity and Blood Sugar is High and Blood Pressure is High THEN MetS Risk is High (1.00)
R21 : IF HDL is Low and Triglycerides and FBS is High THEN MetS Risk is High (1.00)
R22 : IF HDL is Low and Triglycerides is High and Blood Pressure is High THEN MetS Risk is High (1.00)
R23 : IF HDL is Low and Blood Sugar is High and Blood Pressure is High THEN MetS Risk is High (1.00)
R24 : IF Triglycerides is High and Blood Sugar is High and Blood Pressure is High THEN MetS Risk is High (1.00)

FIGURE 3. Combinations of rules according to NCEP ATP III (Category 2)

R25 : IF Obesity and HDL is Low and Cholesterol is High THEN MetS Risk is High (1.00)
R26 : IF Obesity and Cholesterol is High and Blood Sugar is High THEN MetS Risk is High (1.00)
R27 : IF Obesity and Cholesterol is High and Blood Pressure is High THEN MetS Risk is High (1.00)
R28 : IF HDL is Low and Cholesterol and FBS is High THEN MetS Risk is High (1.00)
R29 : IF HDL is Low and Cholesterol is High and Blood Pressure is High THEN MetS Risk is High (1.00)
R30 : IF Cholesterol is High and Blood Sugar is High and Blood Pressure is High THEN MetS Risk is High (1.00)

FIGURE 4. Rules made by replacing triglycerides with cholesterol (Category 2)

formed (R15-R24), there are six rules that contain triglycerides, so six new rules are generated as shown in Figure 4.

R15-R30 are rules belonging to Category 2. Each rule has three antecedent indicators, and the CF rule = 1.

In R15 to R30, the CF rule value for MetS risk equal to 1 will be given when the antecedent has three items (indicators). By using R15-R30, if a person fulfills two or one true condition, the resulting CF value is 0. This happens because the minimum CF value of the three items is 0. In fact, the MetS risk is still very possible. Therefore, it is necessary to add some rules with antecedents supported by two items. The CF value for these rules is given as 0.75. Here are the 14 rules that were generated in Category 3 as shown in Figure 5.

R31 : IF Obesity and HDL is Low THEN MetS Risk is High (0.75)
R32 : IF Obesity and Triglycerides is High THEN MetS Risk is High (0.75)
R33 : IF Obesity and Blood Sugar is High THEN MetS Risk is High (0.75)
R34 : IF Obesity and Blood Pressure is High THEN MetS Risk is High (0.75)
R35 : IF HDL is Low and Triglycerides is High THEN MetS Risk is High (0.75)
R36 : IF HDL is Low and Blood Sugar is High THEN MetS Risk is High (0.75)
R37 : IF HDL is Low and Blood Pressure is High THEN MetS Risk is High (0.75)
R38 : IF Triglycerides is High and Blood Pressure is High THEN MetS Risk is High (0.75)
R39 : IF Triglycerides is High and Blood Sugar is High THEN MetS Risk is High (0.75)
R40 : IF Blood Pressure is High and Blood Sugar is High THEN MetS Risk is High (0.75)
R41 : IF Obesity and Cholesterol is High THEN MetS Risk is High (0.75)
R42 : IF HDL is Low and Cholesterol is High THEN MetS Risk is High (0.75)
R43 : IF Cholesterol is High and Blood Pressure is High THEN MetS Risk is High (0.75)
R44 : IF Cholesterol is High and Blood Sugar is High THEN MetS Risk is High (0.75)

FIGURE 5. Rules in Category 3

R45 : IF Obesity THEN MetS Risk is High (0.50)
R46 : IF HDL is Low THEN MetS Risk is High (0.50)
R47 : IF Triglycerides is High THEN MetS Risk is High (0.50)
R48 : IF Cholesterol is High THEN MetS Risk is High (0.50)
R49 : IF Blood Sugar is High THEN MetS Risk is High (0.50)
R50 : IF Blood Pressure is High THEN MetS Risk is High (0.50)

FIGURE 6. Rules in Category 4

Actual value	Predicted value: +	Predicted value: -
+	TP = 41	FN = 0
-	FP = 0	TN = 157

FIGURE 7. Confusion matrix

We also added some rules with antecedents supported by only one item. The CF value for these rules is given as 0.5. Here are six rules that were generated in Category 4 as shown in Figure 6.

Furthermore, this model is implemented on 198 data samples. First, forward reasoning is implemented on R1 to R14 in Category 1. The calculation of the combined CF value in this Category 1 (CFC1) is based on Equation (2). Next, forward reasoning is implemented on the knowledge base from Category 2 to Category 4 sequentially. The CF combinations for each category are called CFC2, CFC3 and CFC4. The final CF, as the CF value of the data set is obtained by taking the maximum values of CFC1, CFC2, CFC3, and CFC4.

After the inference was done on 198 data sets, the results were tested for correctness. The first testing process is calculating the values of accuracy, sensitivity, specificity, and precision. Based on the confusion matrix in Figure 7, the value for the four test indicators is 100%.

The second test is to evaluate the suitability of the CF calculation. According to the NCEP ATP III standard, the diagnostic criteria are established if there are at least 3 suitable indicators. Because the CF value is continuous, the more appropriate indicators, the closer to the diagnostic criteria. Ideally, the more appropriate indicators, the greater

Table 8. CF evaluation results

Item	Group 1	Group 2	Group 3	Group 4
Number of data sets	41	77	56	24
Certainty Factor (CF)	CF = 1	$0.75 \leq CF < 1$	$0.50 \leq CF < 0.75$	$0 \leq CF < 0.50$
True Positive (TP)	41	77	50	24
Accuracy	100%	100%	89.29%	100%

the CF value. The data sets are divided into 4 groups, namely: groups that fulfill the criteria for three or more indicators (Group 1), groups that fulfill the criteria for two indicators (Group 2), groups that fulfill the criteria for only one indicator (Group 3), and groups that have no indicators that satisfy the criteria (Group 4). Table 8 summarizes the findings.

Group 1 is a group that contains data sets that are appropriate to the diagnostic criteria for MetS according to the NCEP ATP III standard. The CF in this model is set at 1, which means that it has a full certainty value. In Group 1, it shows that the accuracy is 100%, meaning that the inference results in all data sets (41 data sets) are the same as the results of the NCEP ATP III diagnosis, at risk of MetS. The same situation was also shown in Group 2 and Group 4. Both groups had an accuracy value of 100%. In Group 2, all data sets (77 data sets) with two indicators appropriate to the criteria have a CF in the range of $0.75 \leq CF < 1$. In Group 4, all data sets (24 data sets) that do not have indicators that are appropriate to the criteria have a CF in the range of $0 \leq CF < 0.5$.

Different results were given by Group 3. A total of 56 data sets with only one indicator are appropriate to the criteria, and 50 data sets had CF in the range of $0.50 \leq CF < 0.75$. As shown in Table 9, there are six data sets that have a $CF \geq 0.75$.

Table 9. Incompatibility of data sets in Group 3

No	Sex	Height	Weight	WC	HDL	TG	Chol	FBS	RG	Sys	Dias	CF
1	F	–	–	–	60.46	209.00	282	–	–	114	67	0.750
2	M	–	–	–	59.95	157.00	296	–	–	108	59	0.750
3	F	–	–	–	57.00	160.90	245	98	–	130	70	0.813
4	M	–	–	–	45.00	260.60	263	–	–	130	80	0.813
5	F	155	64	57	–	–	200	–	151	187	93	0.756
6	F	144	70	112	–	–	221	–	91	116	84	0.756

This incompatibility of data sets in Group 3 is caused by two factors: 1) High cholesterol and high triglycerides according to the Category 4 knowledge base will be counted twice, even though they both support high cholesterol; 2) If there are more than 2 indicators with $CF > 0$ (after being calculated based on the Category 1 knowledge base), then according to the Category 3 knowledge base, it will be calculated more than once.

Table 10 displays the comparison between the study's findings and those of the NCEP ATP III diagnosis. In the table, rules 1 through 10 for both methods yield the same result: risk value = 1. In rules 11 through 13, the value of CF is 1, whereas according to NCEP ATP III, the level of risk is still undetermined, depending on whether or not the obesity is central obesity. In rules number 17 to 44 in Figure 5, there is no MetS risk in NCEP ATP III, but the CF value is 0.75 or 0.5 in the proposed method. This demonstrates that CF allows for the gradation of risk values. In contrast to NCEP ATP III, which can only decide on two possible risks of MetS, namely risk (1) or no risk (0).

In the future, this decision support system can be integrated into the health care system by considering several factors related to the digitization of the health care system [25].

Table 10. Comparison between NCEP ATP III and study results

No	Rules	NCEP ATP III	CF
1	IF Obesity and HDL is Low and Triglycerides is High THEN MetS Risk is High	1.00	1.00
2	IF Obesity and HDL is Low and Blood Sugar is High THEN MetS Risk is High	1.00	1.00
3	IF Obesity and HDL is Low and Blood Pressure is High THEN MetS Risk is High	1.00	1.00
4	IF Obesity and Triglycerides is High and Blood Sugar is High THEN MetS Risk is High	1.00	1.00
5	IF Obesity and Triglycerides is High and Blood Pressure is High THEN MetS Risk is High	1.00	1.00
6	IF Obesity and Blood Sugar is High and Blood Pressure is High THEN MetS Risk is High	1.00	1.00
7	IF HDL is Low and Triglycerides and FBS is High THEN MetS Risk is High	1.00	1.00
8	IF HDL is Low and Triglycerides is High and Blood Pressure is High THEN MetS Risk is High	1.00	1.00
9	IF HDL is Low and Blood Sugar is High and Blood Pressure is High THEN MetS Risk is High	1.00	1.00
10	IF Triglycerides is High and Blood Sugar is High and Blood Pressure is High THEN MetS Risk is High	1.00	1.00
11	IF Obesity and HDL is Low and Cholesterol is High THEN MetS Risk is High	–	1.00
12	IF Obesity and Cholesterol is High and Blood Sugar is High THEN MetS Risk is High	–	1.00
13	IF Obesity and Cholesterol is High and Blood Pressure is High THEN MetS Risk is High	–	1.00
14	IF HDL is Low and Cholesterol and FBS is High THEN MetS Risk is High	0.00	1.00
15	IF HDL is Low and Cholesterol is High and Blood Pressure is High THEN MetS Risk is High	0.00	1.00
16	IF Cholesterol is High and Blood Sugar is High and Blood Pressure is High THEN MetS Risk is High	0.00	1.00
17	IF Obesity and HDL is Low THEN MetS Risk is High	0.00	0.75
18	IF Obesity and Triglycerides is High THEN MetS Risk is High	0.00	0.75
19	IF Obesity and Blood Sugar is High THEN MetS Risk is High	0.00	0.75
20	IF Obesity and Blood Pressure is High THEN MetS Risk is High	0.00	0.75
21	IF HDL is Low and Triglycerides is High THEN MetS Risk is High	0.00	0.75
22	IF HDL is Low and Blood Sugar is High THEN MetS Risk is High	0.00	0.75
23	IF HDL is Low and Blood Pressure is High THEN MetS Risk is High	0.00	0.75
24	IF Triglycerides is High and Blood Pressure is High THEN MetS Risk is High	0.00	0.75
25	IF Triglycerides is High and Blood Sugar is High THEN MetS Risk is High	0.00	0.75
26	IF Blood Pressure is High and Blood Sugar is High THEN MetS Risk is High	0.00	0.75
27	IF Obesity and Cholesterol is High THEN MetS Risk is High	0.00	0.75
28	IF HDL is Low and Cholesterol is High THEN MetS Risk is High	0.00	0.75
29	IF Cholesterol is High and Blood Pressure is High THEN MetS Risk is High	0.00	0.75
30	IF Cholesterol is High and Blood Sugar is High THEN MetS Risk is High	0.00	0.75
31	IF Obesity THEN MetS Risk is High	0.00	0.50
32	IF HDL is Low THEN MetS Risk is High	0.00	0.50
33	IF Triglycerides is High THEN MetS Risk is High	0.00	0.50
34	IF Cholesterol is High THEN MetS Risk is High	0.00	0.50
35	IF Blood Sugar is High THEN MetS Risk is High	0.00	0.50
36	IF Blood Pressure is High THEN MetS Risk is High	0.00	0.50

7. **Conclusions.** This modified CF model can prove the risk level of MetS well. The certainty value shows a consistent decreasing gradation from Group 1 to Group 4. The performance of the model is also shown by the values of accuracy, sensitivity, specificity, and precision of 100% according to the diagnostic criteria issued by NCEP ATP III. This model supports indicators that can be tested independently without having to go to a hospital or clinical laboratory.

This model does not yet accommodate the use of rules that involve two interchangeable indicators such as cholesterol and triglycerides. In future work, we will try to apply other soft computing methods that are able to handle the ambiguity to solve the problem.

Acknowledgment. We would like to thank the Directorate of Research and Community Service, Universitas Islam Indonesia, for the research grants and improvements in the writing of this paper.

REFERENCES

[1] S. M. Grundy et al., Diagnosis and management of the metabolic syndrome, *Circulation*, vol.112, no.17, pp.2735-2752, doi: 10.1161/CIRCULATIONAHA.105.169404, 2005.

[2] I. Lemieux and J. P. Després, Metabolic syndrome: Past, present and future, *Nutrients*, vol.12, no.11, doi: 10.3390/nu12113501, 2020.

[3] S. J. Lee et al., Metabolic syndrome status over 2 years predicts incident chronic kidney disease in mid-life adults: A 10-year prospective cohort study, *Sci. Rep.*, vol.8, no.1, doi: 10.1038/S41598-018-29958-7, 2018.

[4] C. Li et al., Effect of metabolic syndrome on coronary heart disease in rural minorities of Xinjiang: A retrospective cohort study, *BMC Public Health*, vol.20, no.1, pp.1-8, doi: 10.1186/S12889-020-08612-W/TABLES/6, 2020.

[5] S. Kusumadewi, L. Rosita and E. G. Wahyuni, *Model of Clinical Decision Support System for Metabolic Syndrome*, 1st Edition, UII Press, Yogyakarta, 2020.

[6] R. Christijan, Determination of diagnosis of metabolic syndrome based on assessment of metabolic syndrome and NCEP ATP III score in adolescents, *Penelitian Gizi dan Makanan*, vol.42, no.1, 2019.

[7] National Institute of Health, NCEP cholesterol guidelines, *[NCEP] National Cholesterol Education Program ATP III*, vol.329, no.3, 2001.

[8] H. Y. Ngai, K. K. S. Yuen, C. M. Ng, C. H. Cheng and S. K. P. Chu, Metabolic syndrome and benign prostatic hyperplasia: An update, *Asian Journal of Urology*, vol.4, no.3, doi: 10.1016/j.ajur.2017.05.001, 2017.

[9] N. Milinković, S. Jovičić and S. Ignjatović, Measurement uncertainty as a universal concept: Can it be universally applicable in routine laboratory practice?, *Crit. Rev. Clin. Lab. Sci.*, vol.58, no.2, pp.1-12, doi: 10.1080/10408363.2020.1784838, 2021.

[10] C. Dwan and C. Willig, Existential uncertainty in health care: A concept analysis, *J. Eval. Clin. Pract.*, vol.27, no.3, pp.562-570, doi: 10.1111/JEP.13536, 2021.

[11] N. Bharat, A. Kumar and P. S. C. Bose, A study on soft computing optimizing techniques, *Materials Today: Proceedings*, doi: 10.1016/J.MATPR.2021.08.068, 2021.

[12] P. Kumar, S. K. Singh and S. D. Choudhary, Reliability prediction analysis of aspect-oriented application using soft computing techniques, *Materials Today: Proceedings*, vol.45, pp.2660-2665, doi: 10.1016/J.MATPR.2020.11.518, 2021.

[13] S. Ramkumar et al., A bibliometric study on soft computing techniques and applications in top most populated Asian countries, *Materials Today: Proceedings*, vol.37, no.Part 2, pp.2938-2946, doi: 10.1016/J.MATPR.2020.08.702, 2021.

[14] A. Stabelini Neto, W. de Campos, G. C. dos Santos and O. Mazzardo Junior, Metabolic syndrome risk score and time expended in moderate to vigorous physical activity in adolescents, *BMC Pediatrics*, vol.14, no.1, doi: 10.1186/1471-2431-14-42, 2014.

[15] J. H. Huh, J. H. Lee, J. S. Moon, K. C. Sung, J. Y. Kim and D. R. Kang, Metabolic syndrome severity score in Korean adults: Analysis of the 2010-2015 Korea national health and nutrition examination survey, *Journal of Korean Medical Science*, vol.34, no.6, doi: 10.3346/JKMS.2019.34.E48, 2019.

[16] WHO, *Waist Circumference and Waist-Hip Ratio: Report of a WHO Expert Consultation*, www.who.int, Accessed on Dec. 13, 2021.

[17] Y. Lee and W. J. Siddiqui, Cholesterol levels, *StatPearls Publishing*, https://www.ncbi.nlm.nih.gov/books/NBK542294/, Accessed on Dec. 11, 2021.
[18] WHO, *Body Mass Index – BMI*, https://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi, 2021.
[19] NICE, *Normal and Diabetic Blood Sugar Level Ranges – Blood Sugar Levels for Diabetes*, https://www.diabetes.co.uk/diabetes_care/blood-sugar-level-ranges.html, Accessed on Dec. 13, 2021.
[20] AHA, *Understanding Blood Pressure Readings | American Heart Association*, https://www.heart.org/en/health-topics/high-blood-pressure/understanding-blood-pressure-readings, Accessed on Sep. 08, 2021.
[21] H.-Y. Ngai et al., Metabolic syndrome and benign prostatic hyperplasia: An update, *Asian Journal of Urology*, vol.4, no.3, pp.164-173, doi: 10.1016/J.AJUR.2017.05.001, 2017.
[22] E. Roventa and T. Spircu, Certainty factors theory, *Studies in Fuzziness and Soft Computing*, vol.227, pp.153-160, doi: 10.1007/978-3-540-77463-1_6, 2009.
[23] C. Jiang, W. Fan, N. Yu and E. Liu, Spatial modeling of gully head erosion on the Loess Plateau using a certainty factor and random forest model, *Science of the Total Environment*, vol.783, 147040, doi: 10.1016/J.SCITOTENV.2021.147040, 2021.
[24] A. Azareh et al., Modelling gully-erosion susceptibility in a semi-arid region, Iran: Investigation of applicability of certainty factor and maximum entropy models, *Science of the Total Environment*, vol.655, pp.684-696, doi: 10.1016/J.SCITOTENV.2018.11.235, 2019.
[25] N. Alharbe, Impact of digitization of healthcare system in Saudi Arabia, *ICIC Express Letters*, vol.15, no.3, pp.285-296, doi: 10.24507/icicel.15.03.285, 2021.

Author Biography

Sri Kusumadewi received the Doctor's degree in Computer Science from Gadjah Mada University, Yogyakarta, 2008, and the Master's degree in Electrical Engineering from Gadjah Mada University, Yogyakarta, 1997. She is currently working as lecturer in Informatics Department, Universitas Islam Indonesia (UII), Yogyakarta, Indonesia.

Her research interest lies in decision support systems, especially in the domain of medical informatics. With a Google Scholar H Index of 26, she has received over 10,737 citations. Many of her publications concentrated on clinical decision support systems, and various decision support system fields. She has also written dozens of books on the topic of fuzzy systems, medical informatics, and multi-attribute decision making.

Linda Rosita graduated from Gadjah Mada University in 2007 with a Specialist I degree in Clinical Pathology and in 2020 with a Specialist II degree in Hematology. She is currently working as lecturer in Faculty of Medicine, Universitas Islam Indonesia (UII), Yogyakarta, Indonesia.

Her main research interests include clinical pathology and laboratory. She has 119 citations, with Google Scholar at H Index at 6. She is a member of the European Hematology Association since 2021. Apart from being an academic, she is also active as a doctor at the "JIH" hospital and the Yogyakarta PDHI hospital.

Elyza Gustri Wahyuni received the Master's degree in Computer Science from Gadjah Mada University, Yogyakarta. She is currently working as lecturer in Informatics Department, Universitas Islam Indonesia (UII), Yogyakarta, Indonesia.

Her main research interests include medical informatics, decision support and intelligent systems. She has 185 citations, with Google Scholar at H Index at 5. She has published over 30 papers in journals, conferences, and books.