

Pregnancy Disease Diagnostic Expert System With Certainty Factor Method

Embun Fajar Wati¹, Elvi Sunita Perangin-Angin², Budi Sudrajat³
^{1,2,3}Universitas Bina Sarana Informatika, Indonesia
Email: ¹embun.efw@bsi.ac.id, ²elvi.evt@bsi.ac.id, ³budi.bst@bsi.ac.id

Abstract

Lack of knowledge and information about diseases in pregnancy can delay pregnant women from knowing there are diseases in their pregnancy. Diseases that attack a woman's womb need to be examined by an expert, while experts for this disease are still rare and require a lot of money. In order for the initial diagnosis to be carried out by pregnant women, a solution is proposed in the form of an expert system for diagnosing pregnant women's diseases using the Certainty Factor (CF) method based on the symptoms felt by pregnant women. The research stages used in this study used 4 steps, namely data collection consisting of disease data and symptom data, disease data and symptoms, as well as patient data, symptoms and weights, data analysis using the certainty factor method, validation and evaluation. Diagnostic results that are not in accordance with the CF calculation of around 37.5%, while the results in accordance with the CF calculation were 11 patients or 62.5%.

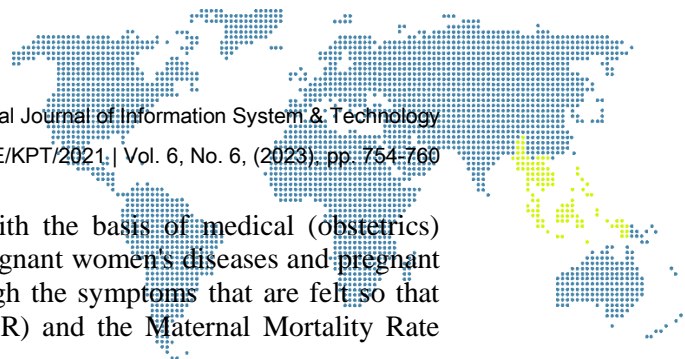
Keywords: pregnancy disease, certainty factor, expert system, patient, diagnostic

1. Introduction

The latest developments in the medical world use computers to help diagnose and prevent and treat a disease. Pregnancy is a very sensitive period in a woman's life, which is vulnerable to physical and mental disturbances and can affect health status which can lead to complications for the mother and fetus, so early pregnancy checks are needed, so that complications do not occur during pregnancy and childbirth [1]. Lack of knowledge and information about diseases in pregnancy can delay pregnant women from knowing there are diseases in their pregnancy [2]. Illness or disorders of pregnancy result in death, especially for pregnant women and babies. Diseases that attack a woman's womb need to be examined by an expert, while experts for this disease are still rare and require a lot of money [3]. So that the initial diagnosis can be carried out by pregnant women, a solution is proposed in the form of an expert system for diagnosing pregnant women's diseases using the Certainty Factor (CF) method based on the symptoms felt by pregnant women [4]. An expert system is a system that seeks to adopt expert knowledge into a computer so that it can solve problems as is usually done by experts [5].

An expert system for diagnosing diseases in pregnancy with the certainty factor method is an expert system designed as a tool for diagnosing diseases in pregnancy that are common in pregnant women with a dynamic knowledge base. The certainty factor method has characteristics that match the way of thinking of an expert, but uses mathematical calculations based on a measure of belief and a measure of disbelief, namely the concept of belief and the concept of uncertainty [6]. The certainty factor method is used to handle the level of certainty of the relationship between symptoms and disease diagnosis with certainty [7]. The CF method has performance that is able to run according to functional requirements, a good percentage of accuracy and can describe the level of confidence of an expert in the problem at hand [4].

From the facts in the form of symptoms from the patient, it can be concluded that the type of disease suffered by the mother during pregnancy [8]. The purpose of this research is to design and build an expert system whose function is to disseminate information to the public so that the handling of pregnancy disorders can be done quickly [9] It is hoped



that the expert system will use the CF method with the basis of medical (obstetrics) knowledge that an expert can help in diagnosing pregnant women's diseases and pregnant women can find out diseases that may occur through the symptoms that are felt so that they can reduce the Maternal Mortality Rate (MMR) and the Infant Mortality Rate (IMR) [10].

2. Research Methodology

The research stages used in this study used 4 steps, each of which also consisted of several parts, namely [11]:

- a) Data Collection: Disease Data and Symptoms Data, Disease Data and Symptoms, and Patient Data, Symptoms, and Weight
- b) Data Analysis with the Certainty Factor Method
- c) Validation and Evaluation

2.1. Data Collection

Data collection was carried out by taking data from previous research in the form of scientific articles that have been published in national journals. The data collected consists of: disease data, symptom data, disease data along with the symptoms, and patient data along with the symptoms they are suffering from.

a) Disease Data and Symptoms Data

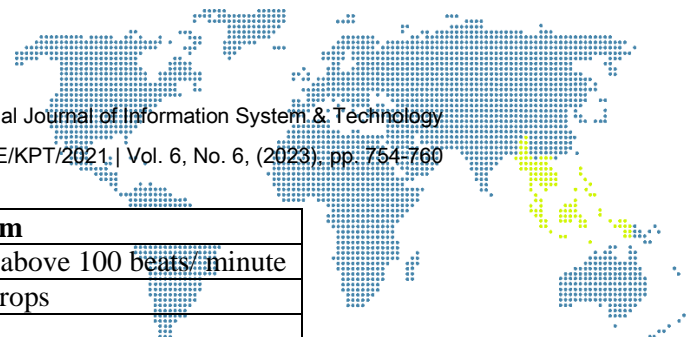
Disease and symptom data obtained from previous studies [11] and [12] which are summarized in the following tables. The code letter P indicates disease and the code letter G indicates symptoms.

Table 1. Disease Data

Code	Disease
P01	Hyperemesis third trimester
P02	Hyperemesis gravidarum in level 1
P03	Hyperemesis gravidarum in level 2
P04	Hyperemesis gravidarum in level 3
P05	Mild preeclampsia
P06	Preeclampsia
P07	Eclampsia

Table 2. Symptom Data

Code	Symptom
G01	Bleeding in young and old pregnant
G02	Amniotic water comes out before its time
G03	Excessive nausea or vomiting
G04	Upper abdominal pain
G05	Dry and dirty tongue
G06	Dehydration
G07	Decreased appetite
G08	Weight Loss
G09	Sunken eyes
G10	Increased pulse rate & lower blood pressure
G11	Pulse frequency about 100 beats/minute
G12	Looks weak and limp (not fit)
G13	Yellow eyes
G14	Difficulty defecating
G15	Decreased skin elasticity



Code	Symptom
G16	The faster the pulse frequency above 100 beats/ minute
G17	Small pulse as blood volume drops
G18	Increased body heat or fever
G19	Urine slightly until it does not come out urine
G20	Vomiting and mixed blood
G21	Decreased awareness
G22	Out of sight
G23	Headache or dizziness
G24	Excess weight or greater weight gain
G25	Swollen face or other parts of the body
G26	Blood pressure between 140/90 to 160/110
G27	Proteinuria +1
G28	Proteinuria +2
G29	Proteinuria +3 or more
G30	Tensions over 160/110
G31	Heart a pain
G32	Seizures

Explanation :
 P : Disease
 G : Symptom

b) Disease Data and Symptoms

Subsequent data is also from previous research [11] and [12] namely the disease data and symptoms that are felt and summarized in the following table.

Table 3. Disease dan Symptom Data

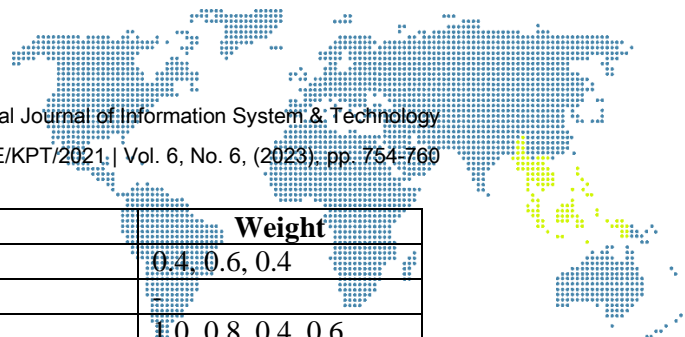
Disease	Symptom
P01	G01, G02
P02	G03, G04, G05, G06, G07, G08, G09, G10, G11, G12
P03	G03, G04, G05, G08, G09, G13, G14, G15, G16, G17, G18, G19, G20
P04	G03, G04, G05, G10, G13, G14, G17, G18, G19, G20, G21, G22
P05	G23, G24, G25, G26, G27
P06	G13, G23, G24, G25, G28, G30
P07	G13, G21, G22, G23, G24, G25, G29, G30, G31, G32

c) Patient Data, Symptoms, and Weight

Patient data, symptoms, weight, and disease estimates that have been diagnosed with an expert system application can be seen in the following table. These data come from previous research [11].

Table 4. Patient, Symptom, and Weight Data

Patient	Age	Symptom	Weight
A	26	G24	-
B	20	G06, G23, G24, G25, G26, G31	1.0, 0.8, 0.8, 0.8
C	30	G01, G02	1.0, 0.8
D	35	G23, G25, G26, G27	1.0, 1.0, 1.0, 0.6
E	37	G03, G04, G05, G06, G07, G08, G09, G13, G14, G15	0.4, 0.8, 1.0, 0.6, 0.8, 0.6, 0.8



Patient	Age	Symptom	Weight
F	28	G23, G24, G25	0.4, 0.6, 0.4
G	27	G04, G23	-
H	22	G23, G24, G28, G30	1.0, 0.8, 0.4, 0.6
I	27	G04, G23, G24, G25, G26	0.6, 0.6, 0.6, 0.6
J	37	G03, G04, G05, G07, G10, G11, G24, G26	1.0, 0.6, 0.4, 0.2, 0.4, 1.0
K	28	G04, G10, G13, G14, G15, G17, G20, G21, G22, G24	0.4, 0.6, 0.6, 0.6, 0.8, 0.4, 0.8, 0.4
L	24	G01, G03, G24, G26, G27	0.4, 0.6, 0.2
M	39	G24	-
N	21	G02, G04, G07, G14, G23, G24, G25, G26	1.0, 1.0, 1.0, 1.0
O	23	G01, G02, G03	1.0, 1.0
P	29	G24, G25, G28, G30	0.6, 1.0, 0.6, 0.2
Q	26	G25, G26, G27, G30	0.6, 0.8, 0.4
R	29	G03, G07, G13, G23, G24, G27, G28, G29	0.8, 0.8, 0.4, 0.6
S	36	G04	-
T	30	G23, G24, G26, G30	1.0, 0.6, 0.4

2.2. Data Analysis with the Certainty Factor Method

After going through the data collection stage, the next stage is carried out, namely data analysis using the certainty factor method. The certainty factor is a way of combining belief and disbelief in a single number. In certainty theory, qualitative data is represented as a degree of confidence.

In expressing the degree of confidence, a value called the Certainty Factor (CF) is used to assume the degree of confidence of an expert in a data. Certainty Factor Notation [7]:

$$CF[H,E] = MB[H,E] - MD[H,E] \quad (1)$$

Information :

CF = Certainty Factor in the H hypothesis which is influenced by the fact E.

MB = Measure of Belief is a measure of the increase in the confidence that hypothesis H is influenced by fact E.

MD = is the increase in distrust of hypothesis H influenced by fact E.

E = Evidence

H = Hypothesis

Certainty Factor for rules with single premise rules :

$$CF[H,E] = CF[E] * CF[rule] = CF[user] * CF[pakar] \quad (2)$$

Certainty Factor for rules with conclusions that similiary concluded rules :

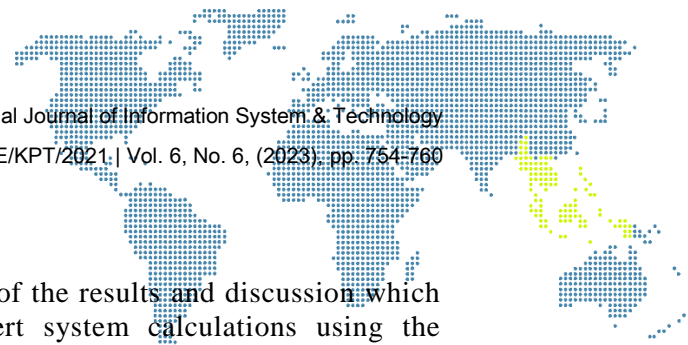
$$CF_{combine} CF[H,E]_{1,2} = CF[H,E]_1 + CF[H,E]_2 * [1 - CF[H,E]_1] \quad (3)$$

$$CF_{combine} CF[H,E]_{old,3} = CF[H,E]_{old} + CF[H,E]_3 * [1 - CF[H,E]_{old}] \quad (4)$$

The patient must choose an answer that each has the following value weights [13] :

- Absolutely Sure : 1.0
- Very Sure : 0.8
- Sure : 0.6
- Less Sure : 0.4
- Abstain : 0.2
- Not Sure : 0

The weights mentioned above will be used as values certainty factor (CF).



2.3. Validation and Evaluation

The validation stage is carried out at the end of the results and discussion which contains a comparison of the results of expert system calculations using the certainty factor method with the results of diagnoses by direct experts. While the evaluation stage is carried out by calculating the percentage of success by calculating the certainty factor method as well as the suitability of the results with experts. The percentage will be assessed with the following data [13] :

- 0% - 50% Possibility of not happening
- 51% - 75% Less likely to happen
- 76% - 84% Most likely to happen
- 86% - 100% Very confident it happened

3. Results and Discussion

From the results of matching the input of symptoms in patient L (table 4), there are 3 data on the same symptoms as mild preeclampsia, namely :

Table 5. Patient L Consultation Data

Gejala	CF
G24	0.4
G26	0.6
G27	0.2

The CF Combine calculation process is as follows :

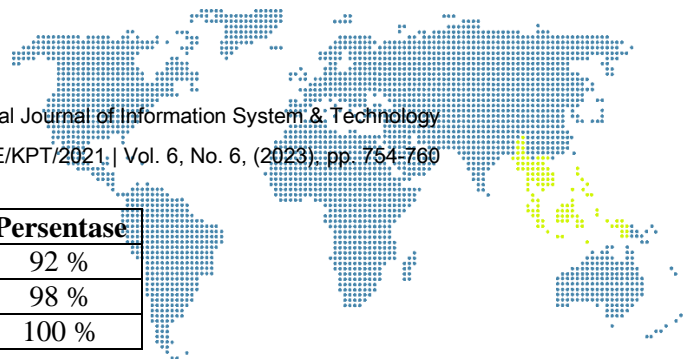
$$\begin{aligned}
 \text{CF(A)} &= \text{CF1} + (\text{CF2} * (1 - \text{CF1})) \\
 &= 0.4 + (0.6 * (1 - 0.4)) \\
 &= 0.76 \\
 \text{CF(B)} &= \text{CF3} + (\text{CF(A)} * (1 - \text{CF3})) \\
 &= 0.2 + (0.76 * (1 - 0.2)) \\
 &= 0.808
 \end{aligned}$$

From the results of the calculation above, the CF of the symptoms entered by patient L is likely to be 0.808 or 81% if rounded up, so the conclusion is that patient L is most likely to have mild preeclampsia.

After selecting and calculating the CF value of all patients, the data can be seen in table 6 below. For patients A, G, M, and S, data was omitted because they did not choose a weight.

Table 6. Percentage of CF Value

Patient	Disease	Persentase
B	P05	100 %
C	P01	100 %
D	P05	100 %
E	P02	99 %
F	P05	71 %
H	P06	100 %
I	P05	94 %
J	P02	100 %
K	P04	52 %
L	P05	81 %
N	P05	100 %
O	P01	100 %
P	P06	100 %



Patient	Disease	Persentase
Q	P05	92 %
R	P06	98 %
T	P05	100 %

Comparison of patient disease data between consultations with expert systems and consultations with experts who are experts in their fields can be seen in table 7 of expert validation [11] [12].

Based on the percentage data from the calculation of the CF value, the lowest score was 52% for only patient K and the highest score was 100% for 9 patients. From the calculation results, it can be concluded that the results of the accuracy of the CF value are less precise in diagnosing pregnancy disease because many percentages are worth more than 75% but the diagnosis results are not in accordance with experts as happened in patients H, I, O, and Q. Diagnostic results that are not in accordance with the CF calculation of around 37.5%, while the results in accordance with the CF calculation were 11 patients or 62.5%.

Table 7. Expert Validation

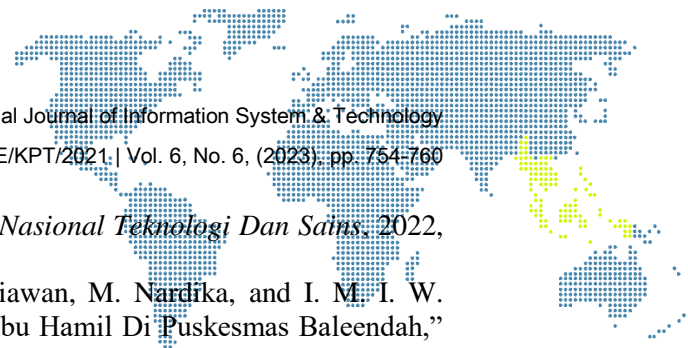
Patient	Certainty Factor	Expert
B	Mild preeclampsia	Mild preeclampsia
C	Hyperemesis third trimester	Hyperemesis third trimester
D	Mild preeclampsia	Mild preeclampsia
E	Hyperemesis gravidarum in level 1	Hyperemesis gravidarum in level 1
F	Mild preeclampsia	Preeclampsia
H	Preeclampsia	Eclampsia
I	Mild preeclampsia	Preeclampsia
J	Hyperemesis gravidarum in level 1	Hyperemesis gravidarum in level 1
K	Hyperemesis gravidarum in level 3	Hyperemesis gravidarum in level 1
L	Mild Preeklampsia	Mild Preeklampsia
N	Mild preeclampsia	Mild preeclampsia
O	Hyperemesis third trimester	Hyperemesis gravidarum in level 1
P	Preeclampsia	Preeclampsia
Q	Mild preeclampsia	Preeclampsia
R	Preeclampsia	Preeclampsia
T	Mild preeclampsia	Mild preeclampsia

4. Conclusion

Based on the percentage data from the calculation of the CF value, the lowest score was 52% for only patient K and the highest score was 100% for 9 patients. From the calculation results, it can be concluded that the results of the accuracy of the CF value are less precise in diagnosing pregnancy disease because many percentages are worth more than 75% but the diagnosis results are not in accordance with experts as happened in patients H, I, O, and Q. Diagnostic results that are not in accordance with the CF calculation of around 37.5%, while the results in accordance with the CF calculation were 11 patients or 62.5%. In future studies, it is expected to try the latest methods with the same or different systems so that they can be compared with previous studies

References

[1] M. H. Digdo, A. Wakhid, A. Wijayanti, A. A. Suryanto, and R. E. Putri, "Sistem Pakar Diagnosa Penyakit Pada Ibu Hamil Menggunakan Metode Forward



- Chaining Dan Certainty Factor,” in *Seminar Nasional Teknologi Dan Sains*, 2022, pp. 125–131.
- [2] R. Latuconsina, T. W. Purboyo, M. F. Ruriawan, M. Nardika, and I. M. I. W. Handi, “Aplikasi Pendeteksi Penyakit Pada Ibu Hamil Di Puskesmas Baleendah,” 2022.
- [3] D. Purnawan and J. E. H. S. R. M. Azmi, “Penerapan Metode Certainty Factor Dalam Mendiagnosa Penyakit Kehamilan Berbasis Web,” *Build. Informatics, Technol. Sci.*, vol. 4, no. 2, p. 614–623, 2022.
- [4] D. Nurdiansyah, R. R. El Akbar, and R. Gunawan, “Modifikasi Nilai Atribut Pada Sistem Pakar Dengan Metode Certainty Factor Untuk Meningkatkan Akurasi Diagnosa Penyakit Kehamilan,” *J. Siliwangi*, vol. 8, no. 1, pp. 27–33, 2022.
- [5] A. Hasbiyanor and Bahar, “Sistem Pakar Diagnosa Keluhan Selama Masa Kehamilan Menggunakan Metode Certainty Factor Berbasis Web,” *JUTISI*, vol. 6, no. 1, pp. 1345–1356, 2017.
- [6] A. Ramadhan, H. Mulyani, and K. Perdana, “Sistem Pakar Diagnosa Penyakit Pada Kehamilan Menggunakan Metode Certainty Factor,” *JURTIK*, vol. 8, no. 1, 2019.
- [7] Rawansyah, V. A. Lestari, and A. Sari, “Sistem Pakar Diagnosa Dini Preeklampsia Pada Ibu Hamil Menggunakan Metode Fuzzy Logic dan Certainty Factor,” in *Seminar Informatika Aplikatif Polinema (SIAP)*, 2020, pp. 221–225.
- [8] S. S. Sundari, Sarmidi, and A. F. Nuraropah, “Sistem Pakar Diagnosis Penyakit Gangguan Kehamilan Menggunakan Metode Forward Chaining Dan Certainty Factor Berbasis Web,” *IT J.*, vol. 9, no. 2, pp. 164–175, 2021.
- [9] N. Herawati, B. K. Simpony, and T. Alawiyah, “Implementasi Metode Certainty Factor Untuk Diagnosa Gangguan Kehamilan,” *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 4, no. 1, pp. 107–115, 2019.
- [10] A. H. Aji, M. T. Furqon, and A. W. Widodo, “Sistem Pakar Diagnosa Penyakit Ibu Hamil Menggunakan Metode Certainty Factor (CF),” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 5, pp. 2127–2134, 2018.
- [11] E. F. Wati and A. Puspitasari, “Expert System for Diagnosing Pregnancy Complaints by Forward Chaining,” *Sink. J. dan Penelit. Tek. Inform.*, vol. 5, no. 1, pp. 7–16, 2020.
- [12] E. F. Wati and B. Sudrajat, “Application of Naive Bayes Method For Diagnosis of Pregnancy Disease,” *Int. J. Inf. Syst. Technol.*, vol. 6, no. 1, pp. 93–100, 2022.
- [13] R. Avrizal, “Sistem Pakar Mendiagnosa Penyakit Flu Babi Menerapkan Metode Hybrid Case Based,” *J. Ris. Komput.*, vol. 6, no. 2, pp. 204–210, 2019.