



# Estimation of Parameters of Uterine Disease Risk in Medan Using Local Maximum Likelihood Method

Rina Widyasari

Program Studi Matematika, Universitas Islam Negeri Sumatera Utara Medan

---

## ARTICLE INFO

### Article history:

Received Sep 05, 2022  
Revised Sep 12, 2022  
Accepted Oct 03, 2022

### Keywords:

Estimation  
Uterine disease  
Logistic regression  
Local maximum likelihood

---

## ABSTRACT

The general public, especially women, still do not know about uterine disease, how to prevent and treat it. In fact, most of the Indonesian population still believes in traditional and occult medicine. Based on the problems that have been described, the researcher proposes a study related to knowledge of uterine disease. The researcher tried to estimate the multinomial logistic regression parameters on the risk of uterine disease (cancer stage) in the city of Medan with the local maximum likelihood method and its relationship with age, parity, age at marriage, birth status, use of family planning. The results showed that middle-aged women were susceptible to cervical cancer and uterine cancer and birth status affect uterine disease significantly.

*This is an open access article under the [CC BY-NC](https://creativecommons.org/licenses/by-nc/4.0/) license.*



---

### Corresponding Author:

Rina Widyasari,  
Fakultas Sains dan Teknologi/Program Studi Matematika,  
Universitas Islam Negeri Sumatera Utara Medan,  
Jalan Lapangan Golf Durin Jangak Tuntungan, Kabupaten Deli Serdang, Sumatera Utara, 20353,  
Indonesia.  
Email: [rinawidyasari@uinsu.ac.id](mailto:rinawidyasari@uinsu.ac.id)

---

## 1. INTRODUCTION

The uterus is an empty, pear-like organ located between the bladder and rectum in a woman's body. The uterus is the place where the fetus grows and develops during pregnancy. The uterine wall is called the endometrium. At the bottom of the uterus there is an organ that connects to the vagina, namely the cervix (Upahita, 2019). There are several types of uterine diseases including, excessive menstruation (menorrhagia), uterine fibroids (myoma), endometriosis, cervical cancer, descent, cysts, and others. However, what will be discussed in this study are fibroids, cysts, endometriosis, and cervical cancer.

Globocan data states that in 2018 there were 18.1 million new cases with a death rate of 9.6 million deaths in which 1 in 5 men and 1 in 6 women in the world had cancer. The incidence of cancer in Indonesia (136.2/100,000 population) ranks 8th in Southeast Asia. The highest incidence of cancer in women is breast cancer, which is 42.1 per 100,000 population with an average death rate of 17 per 100,000 population, followed by cervical cancer at 23.4 per 100,000 population with an average death rate of 13.9 per 100,000 population (depkes.go.id, 2019).

The government has made various efforts to prevent and control cancer in Indonesia, especially the two most common types of cancer in Indonesia, namely breast cancer and

cervical cancer using the Clinical Breast Examination (SADANIS) method and Visual Inspection with Acetic Acid (IVA) for the cervix. The general public, especially women, still do not know about uterine disease, how to prevent and treat it. In fact, most of the Indonesian population still believes in traditional and occult medicine.

Based on the problems that have been described, the researcher proposes a study related to knowledge of uterine disease. The researcher tried to estimate the multinomial logistic regression parameters on the risk of uterine disease (cancer stage) in the city of Medan with the local maximum likelihood method and its relationship with age, parity, age at marriage, birth status, use of family planning.

## 2. RESEARCH METHOD

The research method is basically a scientific way to obtain data with a specific purpose and use. The research methodology used in this research is quantitative research. The type of quantitative research used is descriptive research which is a fact finding using the right interpretation. In this research study about the problems that exist in society and also the procedures used in society and in certain situations. Descriptive research is a type of method that describes an object and subject being studied without any engineering. This includes the relationship between activities, views, attitudes and processes that influence a phenomenon that occurs (Sugiyono, 2018). In addition, researchers also used the causal associative method to determine the relationship between two or more variables that can explain risk factors for uterine disease on age, age at marriage, parity, childbirth, and use of family planning with data obtained from H. Adam Malik Hospital Medan.

### 2.1 Research Procedure

#### a. Formulation of the problem

This stage is the most basic stage in this research, namely formulating the problem of how much the parameter value for uterine disease risk in the city of Medan and its relationship to age, age at marriage, number of parity, childbirth process, and contraception.

#### b. Collect the data

The population of this study were women who were recorded at H. Adam Malik Hospital in Medan who had a history of uterine diseases, such as cervical cancer, uterine cancer. This study did not carry out sampling activities, the research data was obtained based on medical records of cervical cancer and uterine cancer inpatients at H. Adam Malik Hospital, Medan.

#### c. Variable of research

Table 1. Variable

Variable	Meaning	Value
Y	Cancer diagnose stage code C53.9	1 = Stadium IA 2 = Stadium IB 3 = Stadium IIA 4 = Stadium IIB 5 = Stadium IIIA 6 = Stadium IIIB 7 = Stadium IVA 8 = Stadium IVB
X <sub>1</sub>	Age	1 = <40, 2 = 40-51, 3 = 52-63, 4 = 64-75 (year)
X <sub>2</sub>	Age at marriage	1 = < 15, 2 = 16 – 25, 3 = 26 – 35, 4 = > 35
X <sub>3</sub>	Number of parity	1 = <= 2 children, 2 = > 2 children
X <sub>4</sub>	Childbirth Process	1 = Normal Parturition, 2 = Sectio Caesaria, 3 = Induction
X <sub>5</sub>	Contraceptions	1 = Injection, 2 = Pill, 3 = Implant, 4 = IUD, 5 = condom

#### d. Data Analysis

Research data obtained from RSUP H. Adam Malik Medan were analyzed using multinomial logistic linear regression. Multinomial logistic regression analysis is a logistic regression with nominal and ordinal scales used to explain the relationship between response variables and predictor variables that are polychotomous or multinomial. The multinomial logistic regression model according to Hosmer and Lemeshow (in Alan Agresti, 1980) states the model formula used as follows:

$$\pi(x) = \frac{e^{g(x)}}{1+e^{g(x)}} = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)} \quad (1)$$

The response variable used in this research uses 8 categories, so the general formation is as follows:

$$g_j(x) = \beta_{j0} + \beta_{j1}x_1 + \beta_{j2}x_2 + \dots + \beta_{jn}x_n \quad (2)$$

Cumulative logit models are used to compare opportunities that are less or equal to the  $j$ th response category on  $n$  predictor variables which are then presented in vector form  $x_i P(Y \leq j | x_i)$ , with probability greater than  $j$ th response category,  $P(Y > j | x_i)$ ). Formula of cumulative logit models as follows:

$$P(Y \leq j | x_i) = \log \left( \frac{P(Y \leq j | x_i)}{P(Y > j | x_i)} \right) \quad (3)$$

If the dependent variable has 8 categories, then the equations that will be formed are 7 logit equations. So that the multinomial logistic regression equation that will be formed due to cumulative logit models is as follows

$$g_1(x) = \log \left( \frac{P(Y=2|x)}{P(Y=1|x)} \right) = \log \left( \frac{\pi_2(x)}{\pi_1(x)} \right) = \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 + \dots + \beta_{1n}x_n \quad (4)$$

$$g_8(x) = \log \left( \frac{P(Y=8|x)}{P(Y=1|x)} \right) = \log \left( \frac{\pi_8(x)}{\pi_1(x)} \right) = \beta_{80} + \beta_{81}x_1 + \beta_{82}x_2 + \dots + \beta_{8n}x_n \quad (5)$$

Based on the logit equations (4) and (5) obtained, the probability of each response category is as follows

$$P(Y = 1|x) = \pi_1(x) = \frac{1}{1 + \exp g_1(x) + \exp g_2(x) + \dots + \exp g_8(x)}$$

$$\vdots$$

$$P(Y = 8|x) = \pi_8(x) = \frac{\exp g_8(x)}{1 + \exp g_1(x) + \exp g_2(x) + \dots + \exp g_8(x)} ; j = 1, 2, \dots, 8 \quad (6)$$

The data analysis of this research begins by changing the data in categorical form by labeling 1, 2, ..., 8 or 1, 2, 3, 4 according to their respective variables. Then tabulated the data to obtain information on the frequency of data on the variables. Furthermore, the data were analyzed by multinomial logistic regression method with local maximum likelihood estimator. Analysis of this data using the help of IBM SPSS software.

### 3. RESULTS AND DISCUSSIONS

Table 2. Distribution of the Number of Uterine Cancer Patients by Age

Age Group	Frequency	%
<40	28	17,4
40-51	70	43,5
52-63	50	31,0
64-75	13	8,1
Total	161	100

Based on Table 2, it was found that based on the patient's age as much as 17,4 percent suffered from uterine cancer in the age range below 40 years. This indicates that there are abnormalities in the uterus in this age group. Meanwhile, the majority were in the 40-51 year age group, namely 70 people (43.5%), followed by the 52-63 year age group as many as 50 people (31.0%). This shows that uterine cancer patients are dominated by middle adults.

Age is an important risk factor in the development of uterine cancer. Uterine cancer is also common in women with the age of the fourth decade and more. This happens because when the HPV infection starts to become invasive cancer takes about 10-20 years. In general, high-grade dysplasia can be detected 5-10 years before the occurrence of cancer.

Table 3. Distribution of Number of Patients Based on Uterine Cancer Stage

Stadium	Frequency	%
IA1	1	0,6
IB1	8	4,96
IB2	13	9,1
IIA1	6	3,7
IIA2	8	4,96
IIB	38	23,6
IIIA	5	3,1
IIIB	77	47,8
IVA	3	1,9
IVB	2	1,2
Total	161	100 %

In Table 3 can be seen that patients based on cervical cancer stage IIIB were 77 people (47.8%), followed by 38 patients based on uterine cancer stage IIB (23.6%). The results of this study are in line with the research of Sirait, et al (2003) which stated that uterine cancer mostly occurs at stage III B. This may be due to the low awareness of patients to

seek treatment or check themselves in health facilities. This study found only 1 person at stage IA. This may be due to the fact that the symptoms of uterine cancer are difficult to recognize and the lack of knowledge about early detection of uterine cancer (Sirait, et al., 2003).

Multinomial Logistics Regression Analysis Results with Maximum Local Likelihood Estimation. The following shows how the results of the analysis of cervical cancer patients with stages IA, IB, IIA, IIB, IIIA, IIIB, IVA, and IVB based on age, age at first marriage, parity, number of children born, and contraception used.

Table 4. Model Fitting Information

Model	Model Fitting Criteria		Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.	
Intercept Only	421,479				
Final	390,105	31,374	30	,397	

Based on Table 4. the final -2loglikelihood value is 390,105 with a significance of  $0.397 > 0.05$ , which means that  $H_0$  is accepted. This shows that the model after entering the FIT independent variable on the data.

Table 5. Parameter Estimation

Cancer Stadium a		B	Std. Error	Wald	df	Sig.	Upper Bound
I B	Intercept	47,909	3105,491	,000	1	,988	
	Childbirth Process	10,873	,641	288,128	1	<,001	185064,669
	Contraception	-12,369	621,096	,000	1	,984	. <sup>b</sup>
	Age	,061	,087	,498	1	,480	1,261
	Parity	-,460	,356	1,668	1	,197	1,269
	Age at Marriage	,195	,291	,450	1	,503	2,150
II A	Intercept	51,560	3105,492	,000	1	,987	
	Childbirth Process	12,387	,731	286,867	1	<,001	1004893,058
	Contraception	-12,330	621,096	,000	1	,984	. <sup>b</sup>
	Age	,014	,093	,022	1	,881	1,216
	Parity	-,114	,371	,095	1	,758	1,847
	Age at Marriage	-,112	,330	,116	1	,733	1,706
II B	Intercept	47,733	3105,490	,000	1	,988	
	Childbirth Process	10,613	,835	161,368	1	<,001	209128,102
	Contraception	-12,470	621,096	,000	1	,984	. <sup>b</sup>
	Age	,041	,087	,225	1	,635	1,235
	Parity	-,263	,344	,584	1	,445	1,509
	Age at Marriage	,261	,290	,813	1	,367	2,293
III A	Intercept	53,709	4160,610	,000	1	,990	
	Childbirth Process	-1,507	2768,858	,000	1	1,000	. <sup>b</sup>
	Contraception	-13,255	621,097	,000	1	,983	. <sup>b</sup>
	Age	,014	,131	,011	1	,916	1,310
	Parity	,328	,668	,241	1	,623	5,135
	Age at Marriage	,454	,394	1,324	1	,250	3,408
III B	Intercept	49,910	3105,490	,000	1	,987	
	Childbirth Process	11,742	,000	.	1	.	125761,276
	Contraception	-12,579	621,096	,000	1	,984	. <sup>b</sup>
	Age	,026	,085	,094	1	,759	1,213
	Parity	-,195	,328	,354	1	,552	1,566
	Age at Marriage	,182	,288	,398	1	,528	2,107
IV A	Intercept	64,671	3670,211	,000	1	,986	
	Childbirth Process	-1,482	1956,111	,000	1	,999	. <sup>b</sup>
	Contraception	-12,557	621,096	,000	1	,984	. <sup>b</sup>
	Age	-,073	,108	,455	1	,500	1,149
	Parity	-,559	,505	1,227	1	,268	1,538
	Age at Marriage	,256	,329	,606	1	,436	2,459

Based on Table 5., each variable, namely age, age at marriage, parity and contraception, had an insignificant effect on cervical cancer stage at a significant level of

5%. But for Stadium IA, IB and IIB effected significantly by childbirth process. This needs to be reviewed why only childbirth process which having a significant effect.

Table 6. Pseudo R-Square

Cox and Snell	,177
Nagelkerke	,191
McFadden	,074

Based on Table 6, to determine the ability of the independent variable to explain the dependent variable, the Cox and Snell, Nagelkerke and Mc Fadden values are used which are worth 0.177 which means that 17.1 percent of the independent variables can explain the dependent variable while 82.9 percent is explained by the factors in outside the model.

#### 4. CONCLUSION

Based on the characteristics of the patient's age to the stage of cancer, it was concluded that women with the age of >40 years were susceptible to uterine cancer. Therefore, it is necessary to take care of the health of the uterine organs such as doing pap smears, or routine examinations so that it can be known as early as possible about the disease. The estimation using the local maximum likelihood method results that only the childbirth process variable has a significant effect at the level of 5%. It is necessary to investigate what is the cause of the significance, whether there are other factors that influence uterine cancer.

#### ACKNOWLEDGEMENTS

I would like to give my appreciation to H. Adam Malik Hospital that give data and patient information about uterine disease, and big appreciation to Mathematic Study Program and LP2M Universitas Islam Negeri Sumatera Utara that give the researcher a chance to have a study about uterine disease.

#### REFERENCES

- Burhan, S. Jaya, A., K. (2018). Penaksir Parameter Regresi Linier Logistik dengan Metode Maksimum Likelihood Lokal pada Resiko Kanker Payudara di Makassar. *Jurnal Matematika, Statistika, dan Komputasi*. Vol. 14 No. 2 159-165.
- Depkes.go.id. (2019). Hari Kanker Sedunia. *Bulletin kanker.pdf*
- Nurjanah, A., L. Hajarisman, N. Yanti, T., S. (2016). Masalah Overdispersi dalam Model Regresi Logistik Multinomial. *Statistika*. Vol. 16 No. 1, 29-39.
- Rahmadeni, Safitri, E. (2016). Pemodelan Pasien Kanker Payudara Menggunakan Regresi Logistik Biner (Studi Kasus: Pasien Kanker Payudara di Rumah Sakit Umum Daerah Arifin Ahmad Pekanbaru). *Jurnal Sains, Teknologi dan Industri*. Vol. 13, No. 2, Juni 2016 168-173
- Sugiyono. (2018). *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, dan R & D*. Bandung : Alfabeta.
- Sutopo. (2006). *Metodologi Penelitian Kualitatif*. Surakarta: UNS.
- Titis, S., B. (2016). Estimasi Parameter Model Regresi Logistik dengan Metode Maksimum Likelihood. *Skrripsi Universitas Muhammadiyah Ponorogo*
- Upahita, dr. 2019. Halodokter.com