LATE STAGE AT DIAGNOSIS OF COLORECTAL CANCER & ITS ASSOCIATED FACTORS AMONG PATIENTS IN KEDAH FROM 2007 TO 2011: A REVIEW FROM KEDAH CANCER REGISTRY

by

Dr Muhamad Fadhil Bin Mohamad Marzuki

Research Project Report submitted in partial fulfilment of the requirement for the degree of Master of Public Health

MAY 2016

ACKNOWLEDGEMENTS

Bismillahirrahmanirrahim.

In the name of ALLAH, the most Compassionate and the most Merciful. Salutations be upon His messenger Muhammad (peace be upon him), his family and his companions. With help and success granted by ALLAH, I have finished and completed this dissertation on time. I would like to express my deepest gratitude and thank to the following individuals who have contributed and supported me in so many ways in conducting this study:

1. This work would not have been possible without the continuous support and encouragement from my supervisor, Associate Professor Dr Nor Azwany Bt Yaacob, Senior Lecturer, Department of Community Medicine, School of Medical Sciences, Universiti Sains Malaysia. My utmost gratitude goes to her for making this research project possible, for her expertise, kindness, and most of all, for her patience.
2. My thanks and appreciations also go to my co-researcher, Dr Shahrul Bariyah Bt Ahmad, Public Health Physician, Principle Assistant Director of Non-Communicable Diseases Control Unit (NCD), Kedah State Health Department for her full cooperation in process of data collection, moral support, and willingness to offer help at any time that has made all the process in completing this research smooth sailing.
3. I also would like to express my thanks to Dato' Dr Norhizan Bin Ismail, Kedah State Health Director for allowing me to conduct this study in Kedah. Special

thanks also go to staff in NCD unit who involved in the data surveillance and to those who provide the data from all the respective hospitals.

4. I also would like to say thank you very much to all lectures and my fellow colleagues for all the guidance, assistance, and support all the way during the process of completing this research project.
5. And last but not least, a special thanks to my wife, Haslina and my children, Faiz, Hanis and Fawwaz for all the sacrification, understanding, full moral support and deepest affection that boost me further at all times.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

Adj.OR	Adjusted Odds Ratio
ASR	Age-Standardised Rate
CANREG	Cancer Registry
HPE	Histopathological Examination
ICD-O	International Classification of Diseases for Oncology
iFOBT	immunological Faecal Occult Blood Test
NCD	Non-Communicable Disease
NOS	Not Otherwise Specified
OR	Odds Ratio
ROC	Receiver Operating Characteristic
SD	Standard Deviation
SPSS	Statistical Package For The Social Science
TNM	TNM Classification of Malignant Tumours
USM	Universiti Sains Malaysia
WHO	World Health Organization

LIST OF SYMBOLS

>	More than
<	Less than
=	Equal to
≥	More than and equal to
≤	Less than and equal to
α	Alpha
β	Beta
%	Percentage
Δ	Precision / Delta
p	*p*-value

ABSTRAK

DIAGNOSA KANSER KOLOREKTAL DI PERINGKAT LEWAT & FAKTOR-FAKTOR YANG MEMPENGARUHINYA DI KALANGAN PESAKIT DI NEGERI KEDAH DARI 2007 KE 2011: TINJAUAN DARI DAFTAR KANSER KEDAH

Latar belakang: Kanser kolorektal adalah di antara kanser yang paling tinggi dilaporkan di Kedah. Ia adalah kanser kedua tertinggi di Kedah pada tahun 2007 hingga 2011 selepas kanser payudara. Pemerhatian klinikal juga menunjukkan sebahagian besar daripada kes kanser kolorektal di Kedah dikesan pada peringkat lewat yang selalunya memba wa kepada prognosis yang kurang baik dan mengurangkan peluang untuk terus hidup.

Objektif: Kajian ini dijalankan bertujuan untuk mengenalpasti ciri-ciri sosiodemografi dan klinikal pesakit kanser kolorektal dan peratusan pesakit yang dikesan pada peringkat lewat. Kajian ini juga bertujuan untuk mengenalpasti faktor yang mempengaruhi kejadian kanser kolorektal peringkat lewat di kalangan pesakit di negeri Kedah dari tahun 2007 hingga 2011.

Metodologi: Kajian secara keratan rentas telah dijalankan pada bulan Januari 2016 menggunakan data sekunder yang diperoleh daripada Daftar Kanser Kedah. Sebanyak 425 kes kanser kolorektal yang dipilih secara rawak telah dianalisa.

Keputusan: Min(Sisihan Piawai) umur kes adalah 60.30(13.40) tahun dengan majoriti adalah lelaki iaitu 250 kes (58.8%), berbangsa Melayu 233 kes (54.8%), tinggal di bandar 283 kes (66.6%), kanser berlaku di kolon 282 kes (66.4%) dan hampir semua daripadanya adalah dari jenis adenokarsinoma iaitu 397 kes (93.4%). Seratus tiga puluh satu (30.8%) daripada mereka telah didiagnos di tahap II diikuti oleh tahap III

iaitu 128 kes (30.1%), tahap IV 88 kes (20.7%) dan tahap I 78 kes (18.4%). Lebih daripada separuh adalah pada peringkat lewat iaitu 216 kes (50.8%). Analisis regresi logistik mendapati morfologi kanser merupakan satu-satunya faktor yang mempengaruhi kanser kolorektal peringkat lewat di mana lain-lain jenis kanser kolorektal seperti *mucinous adenocarcinoma, signet ring cell carcinoma,* dan *squamous cell carcinoma* mempunyai hampir 3 kali risiko yang lebih tinggi untuk dikesan pada peringkat lewat (Adj.OR 2.71, 95% CI: 1.17, 6.25; $p = 0.020$) berbanding *adenocarcinoma*.

Kesimpulan: Peratusan kes kanser kolorektal di Kedah yang dikesan di peringkat lewat adalah tinggi, terutama dari jenis selain daripada morphologi *adenocarcinoma.*

KATA KUNCI

Kanser kolorektal, diagnosa peringkat lewat, faktor yang mempengaruhi.

ABSTRACT

LATE STAGE AT DIAGNOSIS OF COLORECTAL CANCER & ITS ASSOCIATED FACTORS AMONG PATIENTS IN KEDAH FROM 2007 TO 2011: A REVIEW FROM KEDAH CANCER REGISTRY

Background: Colorectal cancer is among the highest number of cancer reported in Kedah. It is the second most common cancer in Kedah in year 2007 to 2011 after breast cancer. Clinical observation also suggested a high proportion of colorectal cancer cases in Kedah presented at late stage of diagnosis leading to poor outcome and reduce chances of survival.

Objectives: This study was conducted with the aim to describe the sociodemographic and clinical characteristics of the colorectal cancer patients and the percentage of late stage at diagnosis. It was also conducted to determine the associated factors of the late stage at diagnosis of colorectal cancer among patients in Kedah population from year 2007 to 2011.

Methodology: A cross sectional study was conducted in January 2016 using secondary data obtained from the Kedah Cancer Registry. Four hundred twenty five data was randomly selected for analysis.

Result: The mean(SD) age was 60.3(13.40) years old with majority were male 250 cases (58.8%), Malay race 233 cases (54.8%), urban patient 283 cases (66.6%), having colon cancer 282 cases (66.4%) and almost all of them are adenocarcinoma 397 cases (93.4%). One hundred thirty one of them (30.8%) are in stage II followed by stage III 128 cases (30.1%), stage IV 88 cases (20.7%) and stage I 78 cases (18.4%). More than half were in late stage 216 cases (50.8%). Logistic regression analysis found that only cancer morphology are associated with late stage at diagnosis of colorectal cancer in

which other type of colorectal cancer such as mucinous adenocarcinoma, signet ring cell carcinoma, and squamous cell carcinoma were nearly 3 times higher risk of having late stage at diagnosis of colorectal cancer (Adj.OR 2.71, 95% CI: 1.17, 6.25; p = 0.020) as compared to adenocarcinoma.

Conclusion: There are high proportion of colorectal cancer in Kedah presented with late stage at diagnosis, mainly if the morphology is other than adenocarcinoma.

KEY WORDS

Colorectal cancer, late stage at diagnosis, associated factor.

CHAPTER ONE

INTRODUCTION

1.1 The colorectal cancer

Colorectal cancer is a cancer that originates from colon and rectum. Colon and rectal, as part of large intestine are also been described as separate cancer when the disease raised from either part. Both are the part of large intestine. Most of the time, colorectal cancer starts with a growth called polyps in which may later change to cancer. Colorectal cancer is highly preventable and easily treatable. Removal of the polyps can prevent the occurrence of cancer and treatment at the early stage usually confer higher survival rate. Thus, early detection through colorectal cancer screening and prompt treatment can either prevent or ensure good prognosis and survival of the patient.

Risk factors for colorectal cancer include older age, lifestyle such as smoking, alcohol, and lack of physical activity and also inherited genetic disorder. Other factors such as family history of colon cancer or polyps, race, exposure to radiation and even other diseases such as diabetes and obesity are also shown to be associated with increased risk of colorectal cancer. A diet that high in red and processed meat and lack of fibre also confer a higher risk of colorectal cancer (American Cancer Society, 2014).

Based on the Kedah Cancer Registry Report 2007-2011, colorectal cancer is the second most common cancer in Kedah in year 2007 to 2011. The incidence of colorectal cancer was reported to be higher among male as compared to female with age-standardised rate (ASR) of 13.4 per 100,000 population and 9.7 per 100,000 population respectively. Chinese has the highest incidence with ASR 24.3 per 100,000 in male and

19.4 per 100,000 in female population, followed by Indian and Malays. Overall 5-year incidence an increasing trend with age.

1.2 Stage of colorectal cancer

Cancer staging is an important element in cancer management. It describes the severity of disease at the time of diagnosis, with consideration of the growth and size of the tumor and whether it has spread to the adjacent organs, lymph nodes, or distant organs (Edge *et al.*, 2010). Stage of a cancer is crucial for determining the most effective treatment and for survival prediction.

There are two commonly used staging system for colorectal cancer which are Dukes Classification and TNM system. Dukes classification is the earliest system used to described staging of colorectal cancer however it has been replaced by more advance TNM staging (Edge *et al.*, 2010). The TNM system is a system from American Joint Committee on Cancer (AACJ) and it is used widely to describe colorectal cancer. It describes the characteristics of cancer progression based on size and extend of the tumour (T), spread to regional lymph nodes (N), and presence of distant metastasis (M). The overall stage of cancer is assigned by combining the T, N, and M information for the tumour. As far as colorectal cancer is concern, it has four stages; stage I, II, III and IV. Generally, stage I and II are classified as early stage with better prognosis whereas stage III and IV are classified as late stage with poorer prognosis. Most of the cancer registry worldwide including National Cancer Registry in Malaysia has adopted the TNM System to standardise the data collection. It is also for the purpose to make the registration easier as this system is also used for staging of other types of cancer.

1.3 National Cancer Registry of Malaysia (NCRM) and Kedah Cancer Registry

World Health Organization (WHO) through International Agency for Research on Cancer has launched the Global Initiative for Cancer Registry Development (GICR) in order to coordinate multi-organization approach to deliver the required change particularly to make data cancer count. The effort to document cancer burden in Malaysia has begun in 1987 when the National Cancer Registry of Malaysia (NCRM) was established. It is a nationwide first population-based cancer registry with the aim to provide a reliable information on cancer incidence that would be an essential component for policy planning initiatives to meet the country's changing health needs. However due to various difficulties, it was held off in 1992. As an alternatives, the MOH decided to embark on manageable, smaller regional registries to ensure long term sustainability. Penang was chosen as a pilot project in 1994. Following the successful implementation of the pilot project, it was later extended to another 5 states namely Sarawak, Sabah, Kelantan, Pahang and Johor in 1994 before it is implemented in all states in 2007 (Noor Hashimah *et al.*, 2011).

Kedah Cancer Registry started formally in 2007. It registers all the cancer cases based on the patient's address as stated in the notification form. Only patient with the address within Kedah state will be registered in this registry. Therefore, it provides an epidemiological profile of cancer, cancer burden, trend and survival, and variation of cancer incidence for Kedah population. A patient will be registered once in the database based on the identification card number. All the information regarding cancer itself such as topography, morphology, staging, and treatment if available will be entered based on the primary lesion.

The cancer cases need to be notified through a notification form (Appendix A). The form has to be filled up by the hospital managing team and sent to the Non-Communicable Disease (NCD) Unit, Kedah State Health Department. There are three tertiary hospitals and six district hospitals that notify their cancer cases to the NCD unit in the whole Kedah state. NCD Unit also do receive notifications from local private hospitals, and several other hospital outside this state especially Penang where patients from Kedah seek treatment. Other than passive notification from hospitals, NCD unit also do a regular active case search using list of patient discharged from hospital medical record unit and through the list of histopathological examination (HPE) result from pathology department in tertiary hospital that has been requested by the managing doctors. Cases identified as not registered will be informed to the respective hospital for notification. The notification form received from various sources mentioned will then be verified by Public Health Specialist of the NCD unit and registered in the Kedah Cancer Registry database. This registry is an offline database using CanReg V4.0 software developed by International Agency for Research on Cancer (IARC). The data in this registry will be sent to the National Cancer Registry managed by National Cancer Institute on the three monthly basis (Figure 1.1).

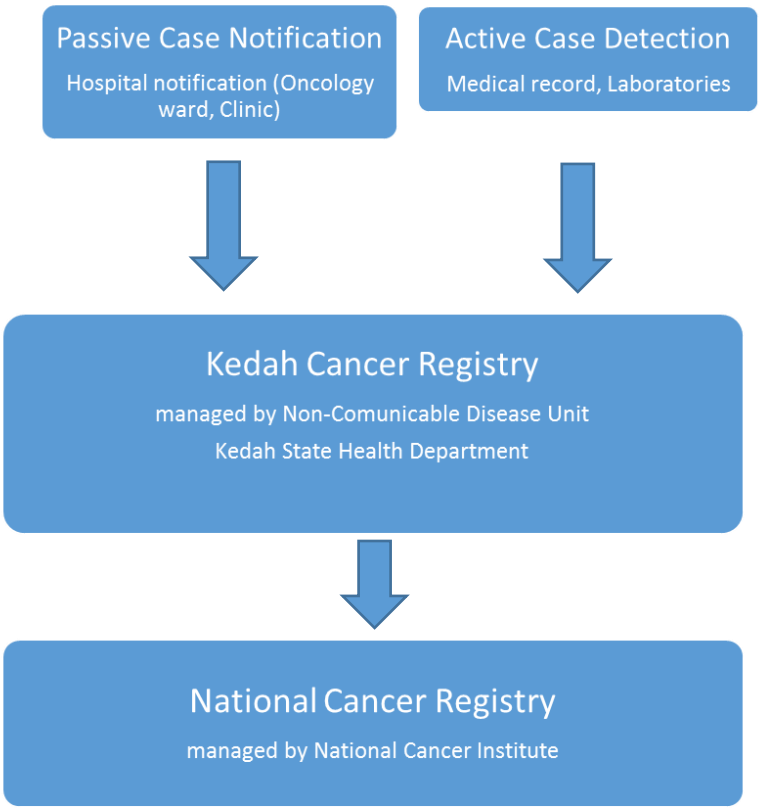

Figure 1.1: Flow chart of the surveillance registry in Kedah Cancer Registry

Until June 2015, 979 colorectal cancer with incidence date from 1st January 2007 until 31st December 2011 have been registered in the registry. This may not reflect the actual number of colorectal cancer patient in Kedah population as the cancer notification is not mandatory. However, the NCD unit of Kedah State Health Department do remind all the hospital director to make sure that they notify all cancer cases treated at their hospital. Active case detection helps to improve data quality. As a preparation for analysing and write up of the Kedah Cancer Report 2007-2011, the NCD unit has already completed the active case detection to maximise the notification of all cancer case including the colorectal cancer. To date, the notification of the cancer case for 2007-2011 is considered of a good quality coverage.

1.4 Problem statement

Based on the clinical observations on Kedah colorectal patients, more than half of patients presented at late stage. Late stage presentation is associated with complicated treatment, poor outcome and prognosis and reduced survival rate. Managing late stage colorectal cancer also requires more sophisticated treatment modality and increases burden to the health care system.

1.5 Rationale of study

Clinical observations suggested a high proportion of colorectal cancer patients in Kedah presented at late stage of diagnosis thus, end up with poor outcome of the management and reduce the risk for survival. Kedah State Health Department in the process of producing a cancer report in line with the National Cancer Report. However, those report only cover colorectal cancer in general without further analysis on the late stage presentation. Local studies done so far had only look into the associated factors of colorectal cancer in general and very few studies done at hospital level did mention about the late stage colorectal cancer. The associated factors might be different when we specifically analyse the late stage at diagnosis of colorectal cancer as compared to general colorectal cancer. Studies at different hospital setting and geographical area may revealed different factors for late presentation due to influence of socio-environment background.

Active case detection has been carried out recently by NCD unit for the preparation of Kedah Cancer Report 2007-2011. This study will be using registered data of patients who were diagnosed with colorectal cancer in Kedah in year 2007 to 2011 based on notification entry in Kedah Cancer Registry until 31st December 2015. Only data of patients diagnosed within the year 2007-2011 will be analysed for this study as

notification of cases within the range of these years are the most complete notification (cases diagnosed beyond 2011 are still underreported). The colorectal cancer screening programme started only in 2014, thus this study findings will provide the overall picture of late stage colorectal cancer in Kedah. This study also will highlight the late stage at diagnosis of colorectal cancer and its associated factors that may provide evidence to the need of reviewing the direction of colorectal cancer prevention campaign and screening programme in the future especially in Kedah.

1.6 Research questions

1. What is the proportion of late stage at diagnosis among colorectal cancer patient in Kedah population?
2. What factors are associated with the late stage at diagnosis of colorectal cancer patient in Kedah population?

1.7 Objectives

1.7.1 General

To study the proportion and associated factors of the late stage at diagnosis of colorectal cancer patient in Kedah population in year 2007 to 2011.

1.7.2 Specific

1. To describe the sociodemographic characteristics and clinical characteristics among colorectal cancer patient registered in Kedah among notified cases in year 2007-2011 Kedah Cancer Registry

2. To describe the proportion by stage at diagnosis among registered colorectal cancer in Kedah among notified cases in year 2007-2011 Kedah cancer registry
3. To determine the associated factors of late stage at diagnosis of colorectal cancer among notified cases in year 2007-2011 Kedah cancer registry.

1.8 Hypothesis

There are significant association between the sociodemographic characteristics (age group, sex, race, and geographical area) and clinical characteristics (cancer subsite and cancer morphology) with late stage at diagnosis among colorectal cancer patients in Kedah population in year 2007 to 2011.

CHAPTER TWO

LITERATURE REVIEW

2.1 Colorectal cancer burden

In year 2008 there were over 12 million new cases diagnosed, 7 million deaths from cancer and 25 million people live with cancer globally (Boyle and Levin, 2008). Colorectal cancer is the third most commonly diagnosed cancer among males and the second among females, with more than 1.2 million new cancer cases worldwide and 608,700 deaths estimated occurred in 2008 (Jemal *et al.*, 2011). The highest incidence rates are found in Australia and New Zealand, North America and Europe, whereas the lowest rates are found in Africa and South-Central Asia. Rates are substantially higher in males as compared to females. Colorectal cancer incidence rates increased rapidly in several areas historically at low risk, including Spain, and several other countries within Eastern Asia and Eastern Europe (Melissa *et al.*, 2009). Rates among males in the Czech Republic and Japan also have already exceeded the peak of incidence observed in the United States, Canada, and Australia, where rates are either declining or stabilizing.

Colorectal cancer was thought to be disease of the more developed and western countries before. GLOBOCAN 2002 reported that the Age Standardized Rate (ASR) in less developed countries was 134.0 per 100,000 male population whereas in more developed countries the ASR is 288.3 per 100,000 male population (Parkin *et al.*, 2005). However, the incidence of colorectal cancer are on the rise in some of the more developed and westernized Asian countries (Sung *et al.*, 2005). Data from the CancerBase of the International Agency for Research on Cancer (IARC) has shown that the incidence in

many wealthy Asian countries are similar to western countries (Bray *et al.*, 2013). The age-standardised rate (ASR) of colorectal cancer per 100 000 men is 49·3 in Japan, 24·7 in South Korea, and 35·1 in Singapore, compared to 44·4 in North America and 42·9 in western Europe (Yiu *et al.*, 2004).

According to National Cancer Registry Report 2008, colorectal cancer (CRC) was the second most common cancer among both male and female in Malaysia and represent 12.7% of all registered cancer case (*National Cancer Registry Report 2008*, 2012). It affects more males (ASR 12.7 per 100,000) as compared to females (ASR 10.1 in 100,000 population). The ASR was highest among Chinese male (20.2 per 100,000), in whom it was more than in Indian male (6.4 per 100,000) and Malay male (5.8 per 100,000). Chinese female also had highest ASR (17.8 per 100,000), which was more than in Indian female (6.5 per 100,000) and Malay female (4.9 per 100,000).

2.2 Clinical features, screening and detection of colorectal cancer

Symptoms of colorectal cancer usually related to the location of the cancer occurs. For example, left sided or distal tumour are more commonly in annular or encircling lesion that produced and 'apple-core' or 'napkin-ring' appearance (Macrae and Bendell, 2016). This will lead to obstructive symptoms such as abdominal distention, altered bowel habits, tenesmus, nausea and vomiting. Right sided or proximal tumour usually presented with haematochezia, abdominal pain and unexplained iron deficiency anaemia. (Saidi *et al.*, 2008). It also can present as emergency condition with intestinal obstruction, peritonitis or rarely acute gastrointestinal bleeding. Patient also may present with symptoms of metastasis. Local invasion can lead to formation of fistula into adjacent organ such as bladder causing pneumaturia or small bowel causing conditions mimicking diverticulitis.

Fever of unknown origin, abdominal wall, intraabdominal and peritoneal abscess can occur due to perforated colon cancer (Tsai *et al.*, 2007). Colorectal cancer can spread through lymphatic and haematogenous dissemination as well as by contiguous and transperitoneal routes (Macrae and Bendell, 2016).

Generally, early stage of colorectal cancer are asymptomatic. Most of the times, patient who are symptomatic at diagnosis already have more advanced disease with bad prognosis. Therefore it is very important to implement the colorectal screening to detect the asymptomatic patients. It is recommended for colorectal screening programme to be started at 50 years old and above using faecal occult blood test once a year, sigmoidoscopy every five years or colonoscopy every 10 years (U.S Preventive Service Task Force, 2014). In Kedah, colorectal screening programme has been offered in all health clinic. It is offered for all Malaysian aged 50 and above. Usually it is done as an opportunistic screening to a patient that comes for other health problems. Screening is done using immunochemical faecal occult blood test (iFOBT). Patient with positive faecal occult blood test will be referred for colonoscopy. However, in Malaysia the screening uptake for colorectal cancer is extremely low (0.7%) as compared to other countries like Korea (20.5%) and Finland (71.0%) (Yusoff *et al.*, 2012).

2.3 Stage at diagnosis of colorectal cancer and prognosis

Colorectal cancer is staged according to TNM system which divided the cancer stage into Stage I, II, III, and IV. This system is compatible with Duke's system but it adds greater precisions in identification in term of prognosis of every subgroups (Morris *et al.*, 2013). Prognosis of colorectal cancer decreased as the cancer stage progress (Table 2.2).

Table 2.1: Colorectal cancer staging and prognosis

Stage	5 years overall survival
I	80 – 95%
II	65 – 75%
III	25 – 60%
IV	<8%

Source: (Weitz *et al.*, 2005)

Survival rate for colorectal cancer is highly dependent on stage at diagnosis (Haggar and Boushey, 2009). In general, the earlier the stage at diagnosis, the higher the chances of survival. Colorectal cancer related morbidities can be avoided if diagnosed and treated early. As shown in the table 2.2, the prognosis drop from as high as 95.0% when patient presented at stage I to as low as less than 8.0% at stage IV.

Unfortunately proportion of patient presented with late stage colorectal cancer is still high. In the United States, 40.0% of colorectal cancer patients have localized disease (stage I and II), 36.0% are regional disease (stage III) and 20.0% have metastases at presentation (Jemal and Siegel, 2011). Clarke *et al.* (2014) found that almost half of cases in Ireland had relatively late stage at diagnosis (stage III/IV) and increased from 42.0% to 50.0% over a period of 15 years. Similarly, Fazio *et al.* (2005) found 40.0% of cases in North America presented at late stage. Malaysia reported a higher proportion of 64.0% of colorectal cancer presented at late stage (*National Cancer Registry Report 2008*, 2012).

2.4 Determinants of colorectal cancer

In general, colorectal cancer are more commonly associated with older age group, male, certain race, and urban residence (Haggar and Boushey, 2009). However, certain

associated factors have been shown to be associated with late stage at diagnosis of colorectal cancer. Many studies had showed that sociodemographic factors such as younger age, female, certain race and urban-rural area has an important association with late stage at diagnosis of colorectal cancer (Mandelblatt *et al.*, 1996; Wu *et al.*, 2004; Fazio *et al.*, 2005; Woods *et al.*, 2005; Paquette and Finlayson, 2007; Parra-Pérez *et al.*, 2015).

Younger patients were more likely to present with late stage colorectal cancer (p = 0.001) (Goldvaser *et al.*, 2016). Magaji *et al.* (2014) conducted a hospital based research looking at epidemiology of colorectal cancer in Universiti Malaya Medical Centre from 2001 to 2010 found the higher ratio of late stage colorectal cancer in younger age group as compared to older age group, 2.5:1.0 (p <0.007). A study by Chou *et al.* (2011) at Taipei Veterans General Hospital from 2001 to 2006 also found a significant difference between younger and older age group (82.6% vs 41.9%, p <0.001). Fazio *et al.* (2005) in Ontario had used the Ontario Familial Colon Cancer Registry (OFCCR) and found that younger age group are associated with late stage presentation where by case over 60 years old were 70% less likely to be diagnosed with late stage cancer than those under 45 years old (Adj.OR = 0.30, 95% CI 0.15,0.60). However, De Sousa *et al.* (2014) in their study in Brazil found no association between younger age group and late stage colorectal cancer (p = 0.338).

A study conducted in Ireland found that incidence of late stage of colorectal cancer was increasing in both sexes but the increment was more among female patient (1.6%, 95% CI 0.90, 2.30) as compared to male patient (1.3%, 95%CI 0.60, 2.10) (Clarke *et al.*, 2014). There was also a retrospective study in Australia reported women

had more advanced stage cancers than men (Koo and Leong, 2010). However, Fazio *et al.* (2005) and Magaji *et al.* (2014) found that there's no statistically significant association between sex and late stage at diagnosis of colorectal cancer.

Race also play as important association with late stage presentation at diagnosis of colorectal cancer. A study conducted in New York City revealed that although low socioeconomic status explained certain degree of the excessive rates of late stage at diagnosis among blacks, black race remained an independent predictor of late-stage diagnosis (OR=1.24, 95% CI 1.13, 1.36) (Mandelblatt *et al.*, 1996). This finding was also in line with finding by Fazio *et al.* (2005) in which the Non-white race was associated with late stage at diagnosis of colorectal cancer (Adj.OR=3.34, 95%CI 1.20, 9.26) . In study by Azmi *et al.* (2007) in Kuantan Hospital found that Malays were commonly diagnosed at later stage (54.3%) while Chinese were diagnosed with earlier stage (58.1%) (p = 0.011). However, a more recent study by Laiyemo *et al.* (2010) and Magaji *et al.* (2014) found that there was no significant association between race and occurrence of the late stage at diagnosis of colorectal cancer.

Geographical area whether it is urban area or rural area also has been shown to have association with late stage at diagnosis of colorectal cancer. Study by Fazio *et al.* (2005) found that greatest odd of late stage colorectal cancer with those living in rural area (Adj.OR=1.48, 95%CI 1.01, 2.17). However study in US by Paquette and Finlayson (2007) found that urban patient are more likely to present with later stage of colorectal cancer (p <0.001). A more recent study done in Georgia looking at the differences in late stage diagnosis between rural and urban patient found that there was no association between geographical area and presentation of late stage at diagnosis of colorectal cancer

(Hines and Markossian, 2012). Study by Beckmann *et al.* (2015) looking at the sociodemographic difference among colorectal cancer patient in Australia also found no significance between rural and urban patient to develop late stage colorectal cancer.

As far as cancer subsite is concern, a study conducted in North America using data from 1992 to 1997 found that, cancer originating from proximal colon are more prone to present with later stage at diagnosis of colorectal cancer. Study in Japan also found that an advanced stage was found more frequently in the proximal colon rather than in the distal colorectal (p <0.010) (Ikeda *et al.*, 1998). In more recently study Mogoantă *et al.* (2013) found that more colorectal cancer occur in rectum as compared to colon and more than half of them presented in late stage. This finding might be explained with a study by Siegel *et al.* (2012) that reported the occurrence of late stage cancer in proximal colon has reduced due to coverage of colonoscopy that allow examination of the entire colorectal in which previously done by sigmoidoscopy that have limited cancer detection to rectum and left colon only.

In term of morphology of the cancer cell, most of the colorectal cancer are adenocarcinoma. Mogoantă *et al.* (2013) found that 98.0% of the colorectal cancer cases in Romania from 2005 to 2009 presented with adenocarcinoma. Other morphology that usually found in colorectal cancer include mucinous adenocarcinoma, signet ring carcinoma and squamous cell carcinoma and these type of morphology usually found in the more advanced stage of colorectal cancer (Nitsche *et al.*, 2013). Nabi *et al.* (2010) also found that mucin secreting adenocarcinoma and signet ring carcinoma of colon and rectum are high grade tumours and frequently presented at and advanced stage of colorectal cancer.

Low socioeconomic status (SES) also reported to be associated with late stage colorectal cancer. A study by Fitzgerald *et al.* (2014) in US found that patient in lowest socioeconomic status (Index 1) will have 1.12 odds (95% CI 1.09, 1.12) to present with late stage colorectal cancer as compared to other socioeconomic status group. Another study done in US using the Surveillance, Epidemiology, and End Results (SEER) Registries reported the same finding ($p<0.001$) (Fitzgerald *et al.*, 2014). This might have a relation with insurance status of patient. Study in US found that the uninsured patients have OR 1.25 (95% CI 1.22,1.27) to present with late stage colorectal cancer as compared to privately insured (Ward *et al.*, 2010). Gong *et al.* (2012) also found that uninsured residents in Texas associated with higher proportion of late stage colorectal cancer ($p<0.050$).

Other than sociodemographic factors, environmental factors also play an important role in development of colorectal cancer. This include diet, smoking, physical activity, diabetes mellitus and obesity. Diet that high in trans-fatty acid are associated with higher risk of colorectal cancer (Vinikoor *et al.*, 2008). It is believed that presence of trans-fatty acid in the colon may result in irritation of the colonic mucosa, increasing oxidative stress and inflammation that can lead to colorectal cancer. Diet that contain fibre also has been shown to have protective effect toward colorectal cancer (Huxley *et al.*, 2012). It is explained by anticarcinogenic properties of fibre, decreased bowel transit time, dilution of faecal carcinogens and decreased ammonia, phenols, and indoles. Smoking has been shown to have higher risk (OR 1.18; 95% CI 1.11, 1.25) to get colorectal cancer (Botteri *et al.*, 2008). The risk is higher with increase the number of pack-years and cigarette per day. A study by Halle and Schoenberg (2009) showed that physically active persons are

less likely to develop colorectal cancer than physically inactive persons. Physical activity also improved the outcome of patient who already developed the cancer.

A systemic review and meta-analysis of cohort study done by Jiang *et al.* (2011) showed that diabetes mellitus was associated with increased risk of colorectal cancer (Summary RR 1.27; 95% CI 1.21, 1.34). Another meta-analysis by Yuhara *et al.* (2011) also showed that diabetes mellitus was associated with increased risk of colorectal cancer (Summary RR 1.38; 95% CI 1.26, 1.51) and the association remained after adjustment for smoking, obesity and physical activity. Obesity is another factor that commonly reported to be associated with colorectal cancer. It was shown that for a $2kg/m^2$ increment in BMI, the risk of colorectal cancer increased by 7.0% (4.0% to 10.0%) and for a 2cm increased in waist circumference the risk increased by 4.0% (2.0% to 5.0%) (Moghaddam *et al.*, 2007)

As many studies have shown various factors associated with late stage at diagnosis of colorectal cancer, and prevalence of late stage at diagnosis of colorectal cancer in Malaysia also is still high, this study is very important for us to have a better understanding and to strategies our effort in planning a more holistic control measures. Colorectal cancer is highly treatable if patient present in early stage that will lead to good prognosis as mentioned above.

2.5 Conceptual framework

Based on the literature review, several factors such age, sex, race, geographical area, cancer subsite, cancer morphology, socioeconomic status, insurance status, life style and certain comorbid are associated with late stage at diagnosis of colorectal cancer. However, due to limitation of secondary data that were used in this study, socioeconomic

status, insurance status, life style and certain comorbid will be excluded as those data are not available (Figure 2.1).

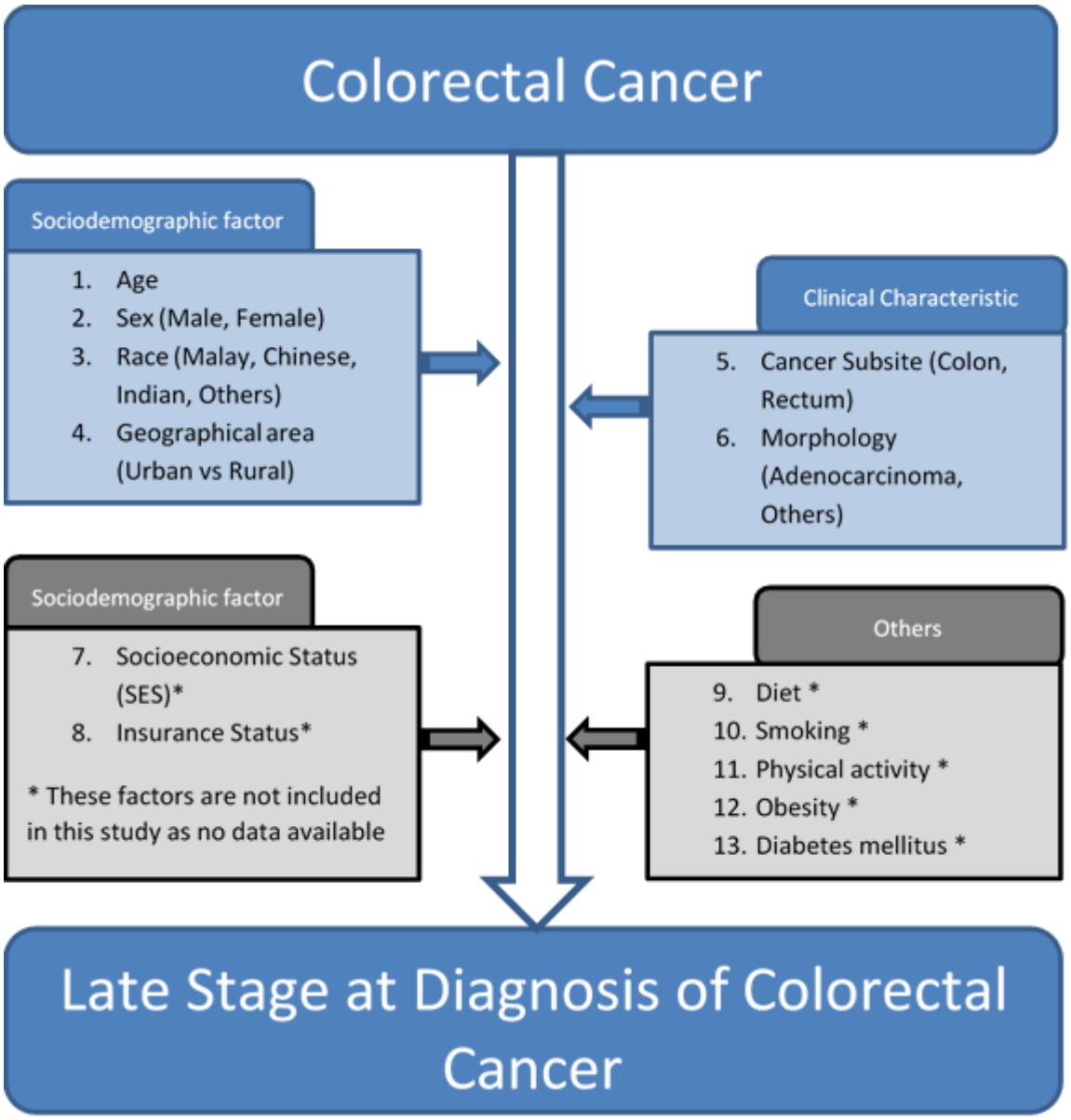

Figure 2.1: Conceptual framework

CHAPTER THREE

METHODOLOGY

3.1 Research design

This study was a cross sectional study design using secondary data from Kedah Cancer Registry from 2007 to 2011.

3.2 Study period

This study was conducted from January 2016 to April 2016.

3.3 Study area

All the data that is necessary for this study was collected from Kedah Cancer Registry run by NCD unit Kedah State Health Department. It registered all cancer cases in which the patient's address is within Kedah state.

Kedah is one of the states located in the northern part of Peninsular Malaysia and has 11 districts comprising of Kubang Pasu, Kota Setar, Padang Terap, Pendang, Baling, Sik, Kuala Muda, Yan, Kulim, Bandar Baharu, and Langkawi. The state covers a total area of over 9,000 km² and consists of the mainland and Pulau Langkawi. The mainland has a relative flat terrain which is used as plantation mainly paddy. Population parameter of 2009 was used as the denominator as the year range from 2007 to 2011 (mid-year population). The estimated Kedah resident in 2009 was 1.93 million (Figure 1). Majority are less than 50 years old (81.7%) and more male as compared to female (50.5% vs 49.5%). Malay contribute the highest proportion of race in Kedah which 75.1% (Table 3.1).

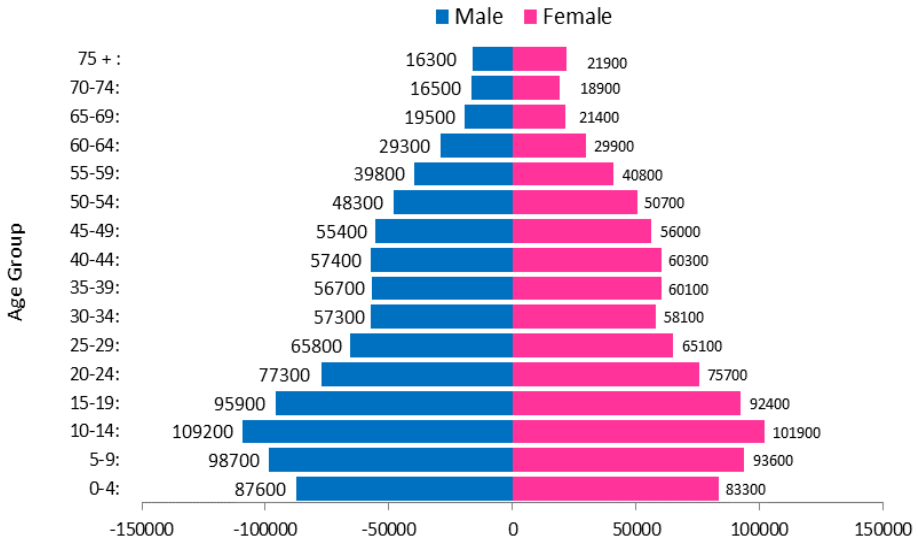

Figure 3.1: Population pyramid of Kedah Residents 2009

(Department of Statistics Malaysia, 2011)

Table 3.1: Distribution according to race and sex for Kedah resident 2009

	Number	%
Race		
Malays	1,446,200	75.1
Chinese	256,200	13.3
Indians	134,800	7.0
Others	23,400	1.2
Total	1,925,300	100.0
Sex		
Male	972,000	50.5
Female	953,300	49.5

(Department of Statistics Malaysia, 2011)

3.4 Reference population

Colorectal cancer patients in Kedah.

3.5 Source population

Colorectal cancer patients in Kedah and registered in Kedah Cancer Registry.

3.6 Sampling frame

Kedah Cancer Registry for colorectal cancer in year 2007 to 2011

3.7 Subject criteria

Data of patient who was diagnosed with colorectal cancer in Kedah in year 2007 to 2011 and registered in Kedah Cancer Registry until 31st December 2015 whom fulfilled the following criteria.

3.7.1 Inclusion criteria : Data of all colorectal cancer cases (ICD-O classification C18.0 - C21.8) with incidence date from 1st January 2007 until 31st December 2011. Incidence date is a date of diagnosis. Since there was no time limit for cancer notification, NCD Unit still receives notification for cancer case diagnosed few years ago.

3.7.2 Exclusion criterion: Data of colorectal cancer with unknown stage at diagnosis.

3.8 Sample size estimation

3.8.1 Objective 2: To determine the proportion of late stage at diagnosis of colorectal cancer in Kedah

Sample size estimation was calculated based on the single proportion formula

$n = (z_\alpha/\Delta)^2 \, p(1-p)$

Absolute precision, $\Delta = 5\%$ (0.05)

p = proportion of patient with late stage at diagnosis of colorectal cancer

= 64% (0.64) (*National Cancer Registry Report 2008*, 2012)

$z = 1.96$

$n = (1.96/0.05)^2 \, (0.64 \times 0.36)$

= 354

To include 20% dropped out,

= 354 x 120/100

= 425

The estimated sample size calculated for objective 2 was 425.

3.8.2 Objective 3: To determine the associated factors of late stage of cancer at diagnosis for colorectal cancer.

Sample size estimation for objective 3 was calculated for each associated factors for late stage cancer at diagnosis of colorectal cancer by using the two proportion formula to compare two independent proportion. PS Software Vr3.1.2 was used to calculate the sample size for diachotomous independent prospective studies where,

$\alpha = 0.05$, value of standard normal distribution cutting off probability α.

power = 0.8, value of the standard normal distribution cutting off probability β.

P0 = proportion of low risk group in late stage colorectal cancer patient (from literature review) (Table 3.2).

P1 = expected proportion high risk group in late stage colorectal cancer patient (Table 3.2).

m = ratio of low risk group to the high risk group in late stage colorectal cancer.

Table 3.2: Summary of sample size calculation for each associated factor

Associated factor	P0	P1	M	Sample Size (+20%)	Literature review
Sex					
Female	0.49	0.65	1	358	(Fazio *et al.*, 2005)
Race					
Malay	0.47	0.62	1	413	(Azmi *et al.*, 2007)
Cancer subsite					
Colon	0.43	0.63	1	233	(Magaji *et al.*, 2014)

As a conclusion, the estimated sample size was 425 that we calculated for objective 2 was chosen for this study because it was the largest sample size calculated.

3.9 Sampling method

There were 797 cases of colorectal cancer that have been registered until December 2015. Simple random sampling was used to select the data sample using Simple Random Sampling Generator using Microsoft Excel V4.0 (Najib, 2015)..

3.10 Research tool

Data was collected from Kedah Cancer Registry database through CanReg V4.0 software. Study variables that were captured from this database include age, sex, race, address, cancer subsite, and cancer staging. Data was collected using data collection form (Appendix B) and entered into Microsoft Excel and IBM SPSS vr 22 software.

3.11 Data collection method

This study involved secondary data collection. Data that is required for this study was retrieved from Kedah Cancer Registry. Data was collected in January 2016. Application to assess the database together was obtained from the Director of Kedah State Health Department. The Principle Assistant Director of Non-communicable Disease (NCD) Control Unit of Kedah State Health Department has been appointed as the co-researcher for this study. Instead of patient's name and identification number, all data were identified by registration number in the registry to ensure the confidentiality. Process of collecting the data was done at NCD Unit itself. All the data was entered into Microsoft Excel before transferred to IBM SPSS V22 for process of data analysis.

3.12 Operational definition

3.12.1 Colorectal cancer

Kedah Cancer Registry is using ICD-O for coding all the notified cancer cases (Table 3.3).

Table 3.3: ICD-O coding for colorectal cancer according to topography

Topography	ICD-O code
Caecum	C18.0
Appendix	C18.1
Ascending colon; right colon	C18.2
Hepatic flexure of colon	C18.3
Transverse colon	C18.4
Splenic flexure of colon	C18.5
Descending colon; left colon	C18.6
Sigmoid colon	C18.7
Overlapping lesion of colon	C18.8
Colon, NOS	C18.9
Rectosigmoid colon	C19.9
Rectum, NOS	C20.9
Anus, NOS (excludes Skin of anus, and Perianal skin C44.5)	C21.0
Anal canal	C21.1
Cloacagenic zone	C21.2
Overlapping lesion of rectum, anus, and anal canal	C21.8

Data for all cancer in Kedah Cancer Registry with ICD-0 code from C18.0 until C21.8 with the incidence date recorded as 20070101 until 20111231 was used in this study.