Incidence of Cholangiocarcinoma and Prevalence of *Opisthorchis viverrini* Infestation in Songkhla Province, Southern Thailand

Seesai Yeesoonsang

A Thesis Submitted in Partial Fulfillment of the Requirements for the degree of Doctor of Philosophy in Epidemiology (International Program)

Prince of Songkla University

2016

Incidence of Cholangiocarcinoma and Prevalence of *Opisthorchis viverrini* Infestation in Songkhla Province, Southern Thailand

Seesai Yeesoonsang

A Thesis Submitted in Partial Fulfillment of the Requirements for the degree of Doctor of Philosophy in Epidemiology (International Program)

Prince of Songkla University

2016

Thesis Title	Incidence of Cholangiocarcinoma and Prevalence of *Opisthorchis Viverrini* Infestation in Songkhla Province, Southern Thailand
Author	Mr. Seesai Yeesoonsang
Major Program	Epidemiology (International Program)

Major Advisor

……...…...

Assoc.Prof.Hutcha Sriplung, M.D.)

Examining Committee:

…….……....…..Chairperson

(Petcharin Srivatanakul, Ph.D.)

……..…...…Committee

(Assoc.Prof.Hutcha Sriplung, M.D.)

……..…...…Committee

(Prof.Virasakdi Chongsuvivatwong,M.D., Ph.D.)

…….…...…..Committee

(Alan Frederick Geater, Ph.D.)

…….…...…..Committee

(Assoc.Prof.Tippawan Liabsutrakul, M.D., Ph.D)

The Graduate School, Prince of Songkla University, has approved this thesis as partial fulfillment of the requirements for the Doctor of Philosophy Degree in Epidemiology.

……

(Assoc.Prof.Teerapol Srichana, Ph.D.)

Dean of Graduate School

This is to certify that the work here submitted is the result of the candidate's own investigations. Due acknowledgement has been made of any assistance received.

……..Signature

(Assoc.Prof.Hutcha Sriplung, M.D.)
Major Advisor

……..…....Signature

(Mr. Seesai Yeesoonsang)

Candidate

I hereby certify that this work has not already been accepted in substance for any degree, and is not being concurrently submitted in candidature for any degree.

……...……....Signature

(Mr. Seesai Yeesoonsang)

Candidate

Thesis Title Incidence of Cholangiocarcinoma and Prevalence of *Opisthorchis viverrini* Infestation in Songkhla Province, Southern Thailand

Author Mr. Seesai Yeesoosang

Major Program Epidemiology (International Program)

Academic Year 2015

ABSTRACT

Background:

Liver and bile duct (LBD) cancers are classified by the two major histology types: hepatocellular carcinoma (HCC) and cholangiocarcinoma (CCA). Histological specimens are not necessary for diagnosing LBD cancer, resulting in a high percentage of LBD cancers with unknown histology. When the percentage of morphological verification is low, the true incidence rates for each histological type are often underestimated. The first part of this study aims to estimate the incidences of HCC and CCA in Songkhla province between 1989 and 2013 among LBD cancer cases after imputing cases with unknown histology type. The estimated numbers of LBD cancer cases from the first part were used to determine the trend in incidence of LBD cancers from 1989 to 2013 and projected to 2030. *Opisthorchis viverrini* (*O. viverrini*), a fish-

borne trematode parasite, is believed to be the major cause of CCA among people living in northeastern Thailand. To determine if this assumption also exists in southern Thailand, 768 participants from Songkhla province, where cholangiocarcinoma incidence has been increasing in the past decade, were examined in the second part of this study.

Method:

For the first part, a retrospective study was conducted using cases diagnosed with LBD cancer between 1989 and 2003 from the Songkhla cancer registry in southern Thailand. Multivariate imputation by chained equations (mice) was used to re-classify the unknown histologies. Age-standardized rates of HCC and CCA by gender were calculated and the trends were compared. A retrospective descriptive study was also conducted to project trends of these incidence rates. Three different projection models were used to project incidence rates: Joinpoint regression, age-period-cohort, and Nordpred age-period-cohort models. Incidence rates per 10^5 population were calculated and projected to 2030 for HCC and CCA by gender. For the second part, to determine the possible association between CCA and *O. viverrini* infestation in Songkhla population, a cross-sectional parasitological and questionnaire survey was conducted in Songkhla province.

Results:

Of 2,387 LBD cases, 61% had unknown histology. After imputation, the age-standardised incidence rates of HCC in males during 1989 to 2007 increased from

4 to 10 per 100,000 person-years and then decreased after 2007. The rates of CCA increased from 2 to 5.5 per 100,000 person-years. The rates of HCC in females decreased from 1.5 in 2009 to 1.3 in 2013 and that of CCA increased from less than 1 to 1.9 per 100,000 person-years by 2013. Complete case analysis gave somewhat similar, although less dramatic trends. The ratio of males to females was 3:1. Slightly more than half (50.9%) were aged between 50 and 69 years. HCC was the predominant subtype (44.4%), followed by CCA (38.1%). Between 1989 and 2013, the three projection models demonstrated an increase in incidence rate of CCA with a continuing trend until 2030 in both genders whereas the incidence of HCC is predicted to remain stable. The incidence rate of HCC and CCA for males in 2030 is expected range from4.1 – 6.0 and 7.6 – 9.4 cases per 100,000 person-years, respecively whereas the incidence rate of HCC and CCA for females in 2030 is expected to range from 1.3 – 1.5 and 2.7 – 3.9 cases per 100,000 person-years, respectively.

In part 2, the prevalence of *O. viverrini* infestation among 768 participants was 1.94% (0.05% from stool samples and 1.89% from serum samples). The average intensity of *O. viverrini* was 23 eggs per gram among those who tested positive. The intensity of *Trichuris trichiura* and hookworm was 83 and 234 eggs per gram, respectively. Religion, education, district life style, and birthplace were significantly associated with *O. viverrini* infestation ($p<0.05$).

Conclusion:

In Songkhla, the incidence of CCA has been increasing for 20 years but is now appears to be stable whereas the incidence of HCC is now declining. This study suggests that in Songkhla province, the number of cases of LBD cancer will increase. CCA has been increasing and has more affect on the overall rate and frequency of LBD cancer cases than HCC in both genders. The effectiveness of comprehensive cancer prevention and control measures might play an important role in reduction of HCC incidence. There was low prevalence of *O. viverrini* infestation in this province and no association was found between CCA and *O. viverrini* infection. *O. viverrini* may probably not be a risk factor of CCA in Songkhla province. Further studies are needed to elucidate other causes.

ACKNOWLEDGEMENT

Firstly, I would like to express my sincere gratitude to my advisor Assoc. Prof. Hutcha Sriplung for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study.

My sincere thanks also go to FETP-Thailand by Dr. Sophon Iamsirithaworn, Dr. Chuleeporn Jirawongsa, and Dr. Chakrarat Pittayawonganon for allowing me to join this program and supporting my tuition. This study was supported by the Research Chair Grant, the National Science and Technology Development Agency (NSTDA), Thailand by Prof. Virasakdi Chongsuvivatwong. My sincere thanks also included Dr. Alan Geater and Mr. Edward McNeil who guided me a lot on data analysis and narrowed down my ideas on the thesis, Assoc. Prof. Nongyao Sawangjaroen who gave access to the laboratory for sample preparation and research facilities. Without their precious support it would not be possible to conduct this research.

I thank my friends, Epidemiology students. In particular, I am grateful to Mr.Surichai Beenheem for helping me on writing and using R command on my data analysis. Special thanks Miss Saowanee Deemoon, who supported me all the stages of this Ph.D.

Last but not the least, I would like to thank my family: my parents and to my brothers and sister for supporting me spiritually throughout writing this thesis and my life in general.

Seesai Yeesoonsang

..... C

ONTENTS

CHAPTER 1 1

INTRODUCTION 1

1. Background 1

1.1 Study background 1

1.2 Study setting background 2

2. Literature Review 6

2.1 Missing data and methods of handling missing data 6

2.2 Trend and trend projection 14

2.3 Risk factors of cholangiocarcinoma 28

2.4 Liver fluke and prevalence in Thailand 29

3. Rationale 44

4. Research questions 45

5. Objectives 46

CHAPTER 2 47

METHODS 47

6. Conceptual frame work 47

7. Methodology 48

7.1 Methodology of objective-1 48

7.2 Methodology of objective-1 52

7.3 Methodology of objective II 53

8. Ethical consideration 59

CHAPTER 3 60

RESULTS 60

9. Result of objective-I 60

9.1 Part I: Histological types among cases having unknown histology of LBD cancer in Songkhla province from 1989 to 2013 using a multiple imputation technique 60

9.2 Part II: Trend of the HCC and CCA incidence from 1989 to 2013, and projected to 2030 63

10. Result of objective-II : the prevalence of *O. viverrini* infestation and associated factors of parasite infestation among adult population in Songkhla 70

CHAPTER 4 80

DISCUSSION 80

11. Incidence of cholangiocarcinoma in Songkhla 80

11.1Multiple imputation for unknown histology type of LBD cancers 80

11.2Trend and trend projection of cholangiocarcinoma in Songkhla 83

12. Prevalence and associated factors for *Opisthorchis viverrini* infestation in Songkhla 86

13. Strengths and limitations 88

13.1Strengths of the study 88

13.2Limitations of the study 89

14. Conclusions 89

REFERENCES 91

ANNEXES 100

Annex 1: Approved ethical consideration 100

Annex 2: Data abstraction form for LBD cancer cases from Songkhla cancer registry 102

Annex 3: Information sheet and consent form 104

Information Sheet 104

Inform consent 107

Annex 4: Questionnaire for Participant (English and Thai version) 111

Annex 5: Manuscript cxxiv

Manuscript 1 cxxiv

VITAE cxxiv

Tables

Table 1 Summary of studies related to missing data and prediction in multiple imputation of missing data. 10

Table 2 Summary of studies related to trend and trend projection models of CCA incidence. 20

Table 3 Summary of studies related to the prevalence and risk factors of liver fluke infestation. 37

Table 4 Distribution of LBD cancer cases by histological type and sex, Songkhla province, Thailand, 1989-2013. 60

Table 5 Distribution of histologic type of LBD cancer cases by sex after multiple imputation, Songkhla province, Thailand, 1989-2013. 61

Table 6 Distribution of LBD cancer cases by histological types and sex, Songkhla province, Thailand, 1989-2013. 64

Table 7 Number and percentage of two major histology types of LBD cancer cases; HCC and CCA by gender, Songkhla, Thailand, between 1989 and 2013 after imputation (n=2,676). 64

Table 8 LBD cancer incidence rates and trend in Songkhla, Thailand in 1989-2013 using Joinpoint regression. 65

Table 9 Crude and adjusted number and percentage of *O. viverrini* and others parasite infestation in Songkhla province, southern Thailand (n = 768). 71

Table 10 Crude number, adjusted number and proportion, and univariate analysis among potential risk factors and *O. viverrini* infestation in Songkhla province (n = 768). 72

Table 11 Point estimate (SE) of multivariate logistic regression of associated risk factors for *O. viverrini* infestation in Songkhla, Thailand 75

Table 12 Frequency and percentage of risk food items contacts in Songkhla population, Thailand 76

Table 13 Ordinal logistic regression analysis for risk of eating behavior for *O. viverrini* infestation 78

Figures

Figure 1 Percentage of histologic tpyes of liver and bile duct cancer in both sexes by calendar year. 3

Figure 2 The percentage of cases with morphologic verification by calendar year. 5

Figure 3 Opisthorchis life cycle 29

Figure 4 Prevalence of parasite infestation in Thailand 1957-2009 31

Figure 5 Flow of cholangiocarcinoma screening program in Thailand, 2013 [3] 33

Figure 6 Conceptual framework 48

Figure 7 Concept of data imputation for unknown histological types 51

Figure 8 Map of Songkhla and study sites 56

Figure 9 Age standardized incidence rates of LBD cancers in Songkhla from 1989 to 2013 stratified by sex.. 63

Figure 10 Age-Period-Cohort (APC) trend analysis for HCC and CCA by gender 67

Figure 11 Age standardized incidence rates of LBD cancer in Songkhla using 3 projection models by gender in 1989-2013 and projection until 2030. 69

Figure 12 Estimated number of LBD cancer cases in Songkhla in 1989-2013 and projection until 2030. 70

LIST OF ABBREVATIONS

AAPC	Average annual percentage change
AC-P	Age-cohort model
ADP	Age-drift-period model
APC	Age–period–cohort model
AP-C	Age-period model
APCSC	Asia Pacific Cohort Studies Collaboration
ASR	Age standardized rate
C. sinensis	*Clonorchis sinensis*
CCA	Cholangiocarcinoma
cHCC-CC	Combined hepatocellular carcinoma and cholangiocarcinoma
CI	Confidence interval
ELISA	Enzyme-linked immunosorbent assay
EPG	Eggs per gram
FECT	Formalin Ether Concentration Technique
HBV	Hepatitis B virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HPB-c	Liver, gall bladder, bile duct, and pancreas cancer
ICD-10	The International Classification of Diseases, Tenth Edition
ICD-O3	International Classification of Diseases for Oncology , Third Edition
KKCR	Khon Kaen Cancer Registry

LBD	Liver and bile duct
MAR	Missing at random
MCAR	Missing completely at random
MICE	Multivariate Imputation by Chained Equations
MNAR	Missing not at random
MV	Morphological verification
NPV	Negative predictive value
O. felineus	*Opisthorchis felineus*
O. *viverrini*	*Opisthorchis viverrini*
PCR	Polymerase chain reaction
PI	Probability interval.
PLC	Primary liver cancer
PPV	Positive predictive value
ROC	Receiver Operation Characteristics

CHAPTER 1

INTRODUCTION

Chapter 1:Background

1 Study background

Liver and bile duct cancer

Liver and bile duct (LBD) cancer is the second most leading cause of cancer death worldwide.[1] With distinguishing histological characteristics and origins, LBD cancer is classified into two major types, hepatocellular carcinoma (HCC) and cholangiocarcinoma (CCA).

HCC is known to be associated with hepatitis B virus (HBV) and hepatitis C virus (HCV) infections.[2] The majority of HCC cases (84%) and deaths (83%) occur in developing countries. Regions with high rates of HCC tend to be those that are endemic for HBV and HCV infections[3].

CCA, a form of cancer comprised of mutated epithelial cells that originate in the bile ducts, is the second most common primary hepatic malignancy, representing 30% of all primary hepatic malignancies worldwide.[4,5] It is classified as either intrahepatic or extrahepatic, both of which have different epidemiological features, and different biological, pathological and clinical characteristics.[6]

According to the series: Cancer in Thailand[7, 8], HCC and CCA are grouped together as cancers of the LBD. This grouping is in accordance with the Bithmus and Corlette classification. This system provides the description of tumor location and longitudinal extension in the biliary tree.[4] It is limited due to its failure to characterize the radial extension of the tumor lesion. However, it provides a practical

manner for surgical oncologists to describe the lesion, and in turn, the anticipated extent of liver that may need to be resected for complete extirpation of the malignancy.

The incidence of CCA was reported to be increasing in the United States[9] and Australia[10], with age standardized rates (ASR) of around 1.0 per 100,000 in both countries. There have been attempts to classify subtypes of CCA according to the radiographic appearance as described by the Bithmus and Corlette classification of perihilar CCA[11], and based on cells of origin.[12] Both classifications agree that both intrahepatic and extrahepatic cancers are CCA.

In Thailand, during 2004-2006, the ASRs of LBD cancer were 42.8 per 100,000 in males, and 18.2 in females. The rates decreased to 38.6 per 100,000 in males and 14.6 in females during 2007-2009 with variations in the proportions of HCC and CCA being demonstrated. The main risk factor for CCA is liver fluke infestation, specifically the *Opisthorchis viverrini* (OV) species, which is common in Southeast Asia.[13] Infestation from other species of liver fluke, such as *Clonorchis sinensis*, is known to be main risk factor for hepatobiliary cancer in Korea.[14] Reported from Khon Kaen cancer registry in the northeast of Thailand, where the prevalence of *O. viverrini O. viverrini* is very high, the ASRs of HCC were 30.3 in males and 13.1 in females[15] whereas the CCA incidence rates were 62.0 in males and 25.6 in females.[16]

2 Study setting background

Songkhla cancer registry

In 1989, the Songkhla cancer registry was established, the third population-based cancer registry in Thailand. Information of cancer cases is gathered from all hospitals in and around Songkhla province, as well as from laboratories and death certificates. The collected data from cancer patients comprise registry number, name, residential address, date of birth, age, sex, date and method of diagnosis, topographic site, histology and extent of cancer, and vital status of patient.[7, 17]

The primary site and histology of the cancer(s) are coded according to the ICD-O 3rd edition (Fritz et al., 2000), Second and subsequent primary cancers are also entered into the registry.

In the past, the diagnosis of LBD cancer in Songkhla, where liver cancer was not common, required histopathological and/or cytological confirmation, and would only be performed among those who had a good performance status because of the painful side effects of fine needle aspiration (FNA). Due to advances in radiographic techniques and improved image quality, and the use of Bithmus and Corlette classification, histological confirmation of LBD cancer has declined. while the incidence of LBD cancer with unknown histology in the Songkhla cancer registry has increased. The percentage of LBD cancer cases with unknown histological type increased from 40 in 1999 to 70 in 2005 and then plateaued. The percentage of HCC rapidly decreased in 2000 and thereafter has steadily declined while the percentage of CCA increased slightly until 1998 and has remained steady since (Figure 1).

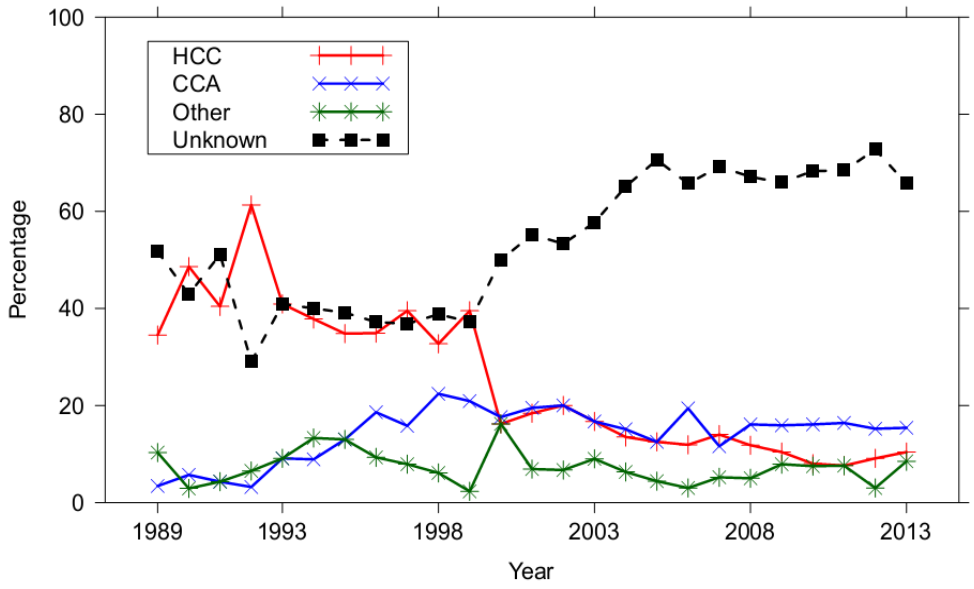

Figure 1 Percentage of histological types of LBD cancer by calendar year in the Songkhla cancer registry.

Cholangiocarcinoma in Thailand and Songkhla province

In Thailand, the association between CCA incidence and the number of cases of liver fluke infestation has been observed for many decades.[18] CCA was observed in a high percentage of liver cancers from the northeastern region where the prevalence of *O. viverrini* *O. viverrini* infestation is higher than in other regions in the country.[19] Other studies showed that the incidence of CCA in the two major regions of Thailand, northeastern and northern, varied at least 12-fold and correlated strongly with the prevalence of *O. viverrini* infestation[20] In contrast, no relationship has been observed between HCC and liver fluke infestation. The geographical pattern of liver fluke infestation in Thailand varies by region, with greatest prevalence in the North (19.3%), and Northeast (15.7%) compared with the central (3.8%) and southern (<1%) regions.[20, 21] There are also regional differences in the ratio of different types of liver cancers. CCA predominates in the Northeast (Khon Kaen) where it comprises 87% of liver cancers (of known histology) whereas in the northern region (Chiang Mai) the percentage is 38.2%. In the South, the proportion of CCA cases is 4%. Moreover, variation also exists within sub-regions. A ten-year cohort study in the Northeastern region[22] showed that the age-standardized incidence of CCA among those aged more than 35 years varied three-fold between districts of Khon Kaen province, from 94 to 318 per 100,000 person-years, with an average of 188. The positive association between prevalence of *O. viverrini* infestation and incidence of CCA was statistically significant at the population level after adjusting for age group, sex and period of sampling.

In southern Thailand, however, there has been an increase in the incidence of LBD over the past decade. In Songkhla province, the annual ASRs of LBD cancer reported in 1998, 2002, 2005, and 2008 were 7.8, 10.9, 16.0, 18.8 per 100,000 population in males, respectively and 2.1, 2.9, 4.6, 5.3 per 100,000 population in females, respectively.[7, 8, 17, 23]

In Songkhla, the percentage of cases with morphological verification (%MV) declined from 60% in 1997 to 20% in 2005 (Figure 2). This occurred after the

adoption of the Bithmus and Corlette classification, which does not require laboratory and histological investigation in diagnostic procedures for LBD cancer cases. Such a classification has higher sensitivity than the pathological diagnosis, thus the number of cases with LBD cancer have increased. Another side effect was that many clinicians did not specify the type of cancer in the medical records.

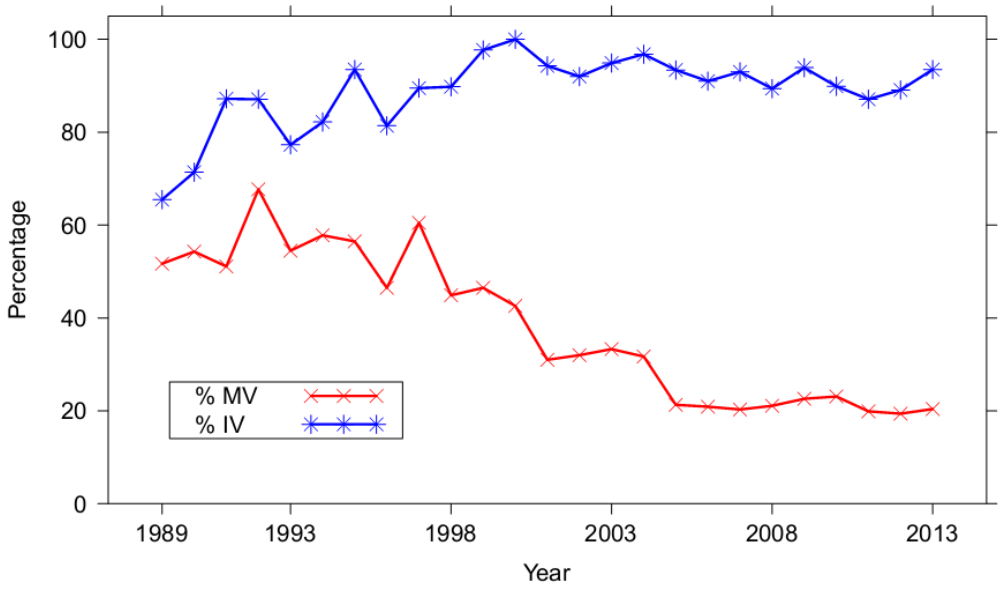

Figure 2 Percentage of morphologic verification (%MV) and imaging verification (%IV) among LBD cases by calendar year, Songkhla province.

When the percentage of morphological verification is low, the true incidence rates for each histological type are often underestimated. Multiple imputation (*multiple imputation*) is a statistical method which can be used for datasets with missing entries.[24-26] *multiple imputation* produces a distribution of plausible values for a missing variable in a record given the values of that record's non-missing covariates.

Chapter 2:Literature Review

This literature review aims to present 1) existing knowledge about imputation models, trends and projections for CCA, and 2) the main risk factors of *O. viverrini* infestation in Thailand from previous studies.

1 Missing data and methods of handling missing data

Based on the Songkhla population-based cancer registry, with a long period of collecting data, the cancer surveillance system provided useful information of the cancer cases. The information of cases, classified according to their site of origin and histology, is mainly used for planning and evaluating cancer control activities. According to limitations of both provider side and payer side, it is impossible to complete all data without unknown information of histology.[27] When classification systems differ in coding histology, temporal inferences by histological categories can lead to incorrect conclusions.

To make use of cancer registry database with incomplete information and improve the cancer prevention and control programs, LBD cancer cases with unknown histology should be treated as missing data.

2.1.1 Missing data

There are three types of missing data[28]:

1) Missing completely at random (MCAR): There is no pattern in the missing data on any variables.
2) Missing at random (MAR): There is a pattern in the missing data but not on the primary dependent variables.
3) Missing not at random (MNAR): There is a pattern in the missing data that affects the primary dependent variables. For example, lower-income participants

are less likely to respond and thus affect the conclusions about

income and likelihood to recommend. Missing not at random is the worst-case

scenario.

2.1.2 Methods of handling missing data

There are many methods for handling missing data during analysis: listwise deletion, recovering the values, educated guessing, average imputation, common-point imputation, regression substitution, and multiple imputation.

The multiple imputation method creates several complete datasets by replacing the missing values with values which are drawn from a distribution specifically modelled for each missing entry. The imputed datasets are then used to estimate the parameter of interest by applying analytic methods. Finally, all parameter estimates are pooled into one estimate, and its variance estimated. Under proper conditions, the pooled estimates are less biased and have the correct statistical properties.[28-30]

A review of studies related to the missing data and prediction in multiple imputation of incomplete variables is summarised in Table 1. The first study explored the implications of perfect prediction for multiple imputation using monotone or the MICE (multiple imputation using chained equations) method.[31] The researchers compared active drug-A data at 5 different doses with a control drug or placebo and the occurrence of pain relief after tooth extraction was an outcome. The result showed that the point estimates and their standard errors varied between the two multiple imputation methods. However, results are not expected to agree precisely, and it is a concern that the standard errors for some of the parameters are substantially larger for the multiple imputation analyses than for the complete-case analysis. The multiple imputation method gives precise estimates when compared with the complete case analysis. However, small sample sizes can lead to perfect predictions.

In the second study, the researchers aimed to impute the missing values for cholesterol and compared eight imputation procedures of 28 cohorts within the Asia Pacific Cohort Studies Collaboration (APCSC).[32] Unconditional and conditional mean, multiple hot deck, expectation maximization, and four different approaches of multiple imputation were used to impute the missing values of cholesterol. For 22 studies for which less than 10% of the values for cholesterol were missing, and for the pooled Asia Pacific Cohort Studies Collaboration, all methods gave similar results. With roughly 10–60% missing values, in such cases multiple imputation would probably be the best choice. Multiple imputation allows for any type and number of variables and gave the most reliable variance estimates when compared with the other imputation methods.

Another study, the researcher try to illustrate bias problems and the role of multiple imputation by generated datasets with speed data MAR for 50% and seat belt use data MAR for 20% of surviving drivers. [28] Multiple imputation was used to create 25 sets of imputed values of each dataset and estimated 2,000 pooled risk ratios using known and imputed values. Risk ratio and variance estimation of complete-case datasets and datasets after using multiple imputation was compared. The multiple imputation method often resulted in a lower bias and a higher power compared with complete-case analysis, however, the amount of missing data in the study and other details of variables was a determining factor.

A study in the USA investigating the patterns of receiving and reporting adjuvant therapies for stage II/III colorectal cancer patients using age, gender, cancer stage at diagnosis, race, marital status, hospital transfer, comorbidity, and income to impute chemotherapy and radiotherapy data.[26] The authors stated that the multiple imputation method could increase the precision of estimates compared to analyzing the survey data alone. According to the limitations of cancer registry system, multiple imputation is one method that can increase precision and power of data in a cancer registry.

Lee (2010) [33] demonstrated the effectiveness of multiple imputation in data with 10% to 80% missing observations using absolute bias and root mean squared error of multiple imputation measured under different types of missing data. The bias and root mean squared error using multiple imputation were found to be much lower than using a complete case analysis. The Markov chain Monte Carlo method as an imputation mechanism can be used for continuous and univariate missing variables.

In conclusion, for improving cancer registry data with a high percentage of unknown histology type in LBD cancer cases, there are many studies explained about missing data and model selection. According to the literature review, incomplete data are common and methods to handle them are important. If we use only complete case analysis, the result can be misleading, with usually an underestimation of the true incidence. The multiple imputation method demonstrated that more precise results can be obtained and provides reasonable values for imputing incomplete data.

Table 1 Summary of studies related to missing data and prediction in multiple imputation of missing data.

Authors, year	White IR, Daniel R, Royston P., 2010 [31]
Country	United Kingdom
Objectives	To explores the implications of perfect prediction for multiple imputation using monotone or MICE imputation method.
Exposure variables	An active drug-A at 5 different doses, a control drug or placebo.
Outcome	The occurrence of pain relief after tooth extraction, 1indicating some or complete relief from pain and 0 indicating no relief
Setting	366 patients, with moderate or severe pain from their third molar extracted, were assigned to one of seven treatment groups randomly.
Methods/Tool/ Analysis	Monotone multiple imputation was used for impute the missing outcomes. The prevalence of pain relief across the 7 treatments at the final time-point.
Results	– The point estimates and their standard errors were varied between the two multiple imputation methods. – These are not expected to agree precisely. However, it is a cause for concern that the standard errors for some of the parameters that substantially larger for the multiple imputation analyses than for the complete-case analysis.
Comment	– Imputation method showed a good precision when compare with complete case analysis. – Multiple imputation method should concern about small sample size which can lead to perfect prediction.

Table 1 (continued)

Authors/year	Barzi F, Woodward M., 2004 [32]
Country	Australia
Objectives	To fill in the missing values for cholesterol and comparison of eight imputation procedures.
Exposure variables	Age, sex, systolic blood pressure, diastolic blood pressure, body mass index, smoking, diabetes, survival time, CHD death, and death from any cause.
Outcome	The missing values for cholesterol.
Setting	The 28 studies in the Asia Pacific Cohort Studies Collaboration (APCSC).
Methods/Tool/ Analysis	Unconditional and conditional mean, multiple hot deck, expectation maximization, and four different approaches of multiple imputation were used for impute the missing values of cholesterol.
Results	– For 22 studies for which less than 10% of the values for cholesterol were missing, and for the pooled APCSC studies, all methods gave similar results. – With roughly 10–60% missing values, multiple imputation would probably give the best result.
Comment	– Compared with the other imputation methods, multiple imputation allows for any type and number of variables and gave the most reliable variance estimates.

Table 1 (continued)

Authors/year	Cummings P., 2013 [28]
Country	United states of America
Objectives	To illustrate bias problems and the role of multiple imputation
Exposure variables	Using seat belt in difference speed of driving
Outcome	Death of driver
Setting	Driver who using seat belt in difference speed of driving
Sample	2,000 hypothetical data sets with speed data missing at random (MAR) for 50% and seat belt use data MAR for 20% of surviving drivers.
Methods/Tool/ Analysis	The adjusted risk ratio for seat belt use was estimated in each data set using complete-case analysis. Multiple imputation was used for create 25 sets of imputed values of each data set and estimated 2,000 pooled risk ratios using known and imputed values. Risk ratio and variance estimation of complete-case data sets and data sets after using multiple imputation was compared.
Results	– Multiple imputation reduced bias and increased power of study.
Comment	– Multiple imputation was not always reduce bias and increased precision compared with complete-case analysis. – Amount of missing data in the study and other details of variables should consider.

Table 1 (continued)

Authors/year	He Y, Yucel R, Zaslavsky AM., 2008. [26]
Country	United states of America
Objectives	To study the patterns of receiving and reporting adjuvant therapies for stage II/III colorectal cancer patients.
Exposure variables	Age, gender, cancer stage at diagnosis, race, marital status, hospital transfer, comorbidity scores, and the median income of the patient's census block group.
Outcome	The adjuvant chemotherapy and radiotherapy data
Setting	Patients age 18 or older who were newly diagnosed with stage III colon cancer or stage II or III rectal cancer and underwent surgery.
Sample	About 12,594 patients from 10 regional cancer registries in California during 1994 to 1997.
Methods/Tool/ Analysis	Cancer registry variables was used for impute the chemotherapy and radiotherapy data. Multiple imputation method using Bayes' theorem underlies the calculation of the probability.
Results	– The multiple imputation analysis can increase the precision compared to analyzing the survey data alone.
Comment	– According to the limitations of cancer registry system, multiple imputation is the one of method that can increase precision and power of data in cancer registry

Table 1 (continued)

Authors/year	J H Lee, 2010 [33]
Country	United states of America
Objectives	To examines the effectiveness of multiple imputation in data with 10% to 80% missing observations using multiple imputation measured under difference type of missing data.
Exposure variables	DBP, height, weight, age, BMI, and cholesterol
Outcome	Systolic blood pressure (the predictive study of coronary heart disease)
Setting	A predictive study of coronary heart disease
Sample	3,154 of coronary heart disease patients
Methods/Tool/ Analysis	An effectiveness of multiple imputation was compared with complete cases analysis.
Results	– The bias and root mean squared error using multiple imputation are much smaller than complete case analysis. – The Markov chain Monte Carlo method as an imputation mechanism can use for continuous and univariate missing variables.
Comment	– When large proportions of data are missing, the number of imputations, imputation mechanisms, and missing data mechanisms need to be considered.

2 Trend and trend projection

2.2.3 Disease trend

Public health officers monitor trends in rates of disease and trends in many other public health aspects, such as in medical, social, and behavioural risk factors for various diseases. Trends in observed rates provide valuable information for program

planning, program evaluation, and policy measures. Examining data over time also facilitates predictions about future frequencies and incidence rates.

For public health, trend data are presented for disease rates over a period of time. These rates are considered as the true underlying population parameters and therefore statistical assessment is rarely engaged. If the rates are assumed to have no error, future trends can be presented and predicted naturally.

2.2.4 Trend projection models

Generally, statistical models are frequently used to estimate future trends in cancer incidence rates by extending past trends. The relationships between the risk factors and the cancer rates are analysed, and projected by applying the future times in the equation. Many different statistical methods in terms of model types and data used have been used to project cancer burden. The results vary depend on the model, for example, simple linear regression, log-linear regression, or age–period–cohort (APC) model.[34, 35] For APC models, the age, period and cohort effects are obtained from generalized linear models. Nordpred APC models are based on 5-year intervals and generalized additive models with spline smoothing.[36] A model is fitted to all available data or their subset for a goodness-of-fit test.[35, 37] The trend is extrapolated based on the assumption of the observed trends which include keeping current rates unchanged with continuing trend and adjusting the extent to the future.

In this study, the trend in incidence of LBD cancer and projection model was analysed by using the 3 models as follows:

1) Joinpoint regression analysis

Joinpoint regression involves fitting a series of joined straight lines or segments on a logarithmic scale to the age-standardized incidence rates.[38] The trends in incidence are reflected by the annual percent change of each segment. The models incorporate estimated standard errors of the rates. The tests of significance use a Monte Carlo method with permutation tests. The estimated slopes from this model are then

transformed back to represent an annual percentage increase or decrease in the rate. If a change point is detected within any period, then the annual percent change is estimated from the trend in the last segment.

2) Age-Period-Cohort model

Age-Period-Cohort models are a class of models for demographic rates usually observed over a broad age range over a reasonably long time period, and classified by current age, follow-up date (period) and birth date (cohort).[35, 39] This can be shown in a Lexis-diagram with data of follow-up time along the *x*-axis, and age along the *y*-axis. A single person's life-trajectory is therefore a diagonal line. Tabulated data enumerates the number of events and the risk time in some subsets of the Lexis diagram, usually subsets classified by age and period in equal intervals.

The Age-Period-Cohort model describes the rates as a sum of age-period- and cohort-effects. The three variables age (at follow-up), a, period (i.e. date of follow-up), p, and cohort (date of birth), c, are related by the equation $a = p - c$. A person's current age can be calculated by subtracting their date of birth from the current date. Hence the three variables used to describe rates are linearly related, and the model can therefore be parametrized in different ways, and still produce the same estimated rates.

In popular terms it is possible to move a linear trend around the three terms, because the age-terms contain the linear effect of age, the period-terms contains the linear effect of period and the cohort effect contains the linear effect of cohort.

3) Nordpred power-5 models

The Nordpred approach was developed as part of a comprehensive analysis of cancer trends in the Nordic countries.[35, 40] This approach is based on a standard APC Poisson regression model and gives more realistic long-term projections. It is now one of the most frequently used methods for worldwide cancer projections.[41-

[43] The log-linear relationship between the rate and the covariates in the standard model produces predictions in which the rates grow exponentially with time. The Nordpred model uses a power- link function instead of the log-link function to lower this growth. The power-link function is an approximation of the log-link function based on the Box-Cox power transformation theory, in which $\lim_{\lambda \to 0} x^{\lambda} = \log(x)$. The Nordpred model is defined as follows:

$$Case_{ap} \sim Poisson(\mu_{ap}),$$

$$R_{ap} = (A_a + D.p + P_p + C_c)^5$$

where R_{ap} is the incidence rate for age group *a* in calendar period *p*, which is the mean count μ_{ap} of $case_{ap}$ divided by the corresponding population $size_{ap}$; A_a, P_p and C_c are the non-linear components of age group *a*, period *p* and cohort *c*, respectively; and D is the common linear drift parameter of period and cohort.[39] A cohort is calculated by subtracting age from period: c = A + p – a, with A = number of age groups.

To extrapolate the model for future periods, two approaches are considered instead of simple continuation of the overall historical trend. The software determines whether the average trend across all observed values, or the slope for the last 10 years of observed values, is used as the drift component D to be projected. The software does this by testing for departure from a linear trend. If the trend across the entire observation period departs significantly from linearity, only the trend in the current decade is used for projection. To attenuate the impact of current trends in future periods, a ‘‘cut trend’’ option is used, which is a vector of proportions indicating how much to cut the trend estimate for each 5-year projection period. A gradual reduction in the drift parameter of 25%, 50%, 75% and 75% in the second, third, fourth and fifth 5- year period, respectively, is used as a default.[35, 40]

A review of studies related to trend and trend projection models is summarised in Table 2. The first study was conducted in Denmark.[44] The researcher examined the changes in incidence, mortality, and prevalence relative survival of liver,

gall bladder, bile duct, and pancreas cancer (HPB-c) patients. The incidence and mortality rates of cancer of the liver and pancreas increased over time while the rates of cancer of the gall bladder and bile duct decreased. About 20% of Denmark cancer registry cases did not have morphological verification. But the rare histological subtypes of specific cancers were associated with either a much worse or better prognosis which reduced generalizability.

The second study aimed to assess the incidence trend of HCC in Khon Kaen, Thailand, from 1990 to 2009.[15] The result showed that gender and year of diagnosis was associated with the incidence of HCC. The third study aimed to estimate the incidence and trends of liver cancer, particularly focusing on CCA in Khon Kaen, Thailand.[16] The authors suggested that the increase in incidence rates in the first 5 to 6 years may have been due to improved completeness of the registry. The rate stabilized in the subsequent 10 to 12 years. Gender and year of diagnosis were associated with trend of HCC. The issues of prime concern included data collection, data processing, statistical analysis/synthesis and reporting reliable and complete data.

The fourth study aimed to forecast mortality of HCC in Taiwan using an APC model.[45] The age and period effects were quite notable of increasing HCC mortality of Taiwanese in both sexes while the cohort effect decreased the HCC mortality. Gender, age, year of diagnosis, and year of birth were associated with the trend of HCC mortality. Although the APC forecasting seems to be a good fit for trend, researchers should not project trends for more than 15 years in the future due to forecasting method assumptions.

Trends of PLC should be presented by histologic subtypes because different histologies require different measures and policy for disease control. A study from Minnesota reported a trend of diagnosing earlier stages of the cancer.[46] The age and period effects were quite notable of increasing HCC mortality of Taiwanese in both sexes while the cohort effect decreased the HCC mortality. Gender, age, year of diagnosis, and year of birth were associated with HCC mortality.

In summary, findings from the literature review suggest that handling of incomplete data is very important. If only complete-cases are used in the analysis, the result can be misleading with usually under estimation of the incidence. Among the projection models, APC model, Joinpoint regression, and the Nordpred model were used in many countries for model projection, especially on cancer trend. These multiple imputation methods and three trend projection models will be used in this study.

Table 2 Summary of studies related to trend and trend projection models of CCA incidence.

Authors/year	Bjerregaard JK, Mortensen MB, Pfeiffer P., 2016 [44]
Country	Denmark
Objectives	To examine the changes in incidence, mortality, prevalence relative survival of HPB-c patients with particular on the elderly
Explanatory factors on LBD cancer incidence	Age and sex of patient
Outcome	incidence, mortality, prevalence of HPB-c patients
Setting	HPB-c patients in Danish Cancer Registry
Sample	All cases of the Danish Cancer Registry and the Danish Cause of Death Registry with follow-up for death or emigration from 1978 to 2013.
Methods/Tool/ Analysis	Sex- and age-specific relative survival proportion ratios were calculated for each of the diagnostic groups for the age groups 0–69, 70–79, 80–89 and 90 + years and for the five-year periods of diagnosis 1968–1972, 1973–1977, ..., 2003–2007 and 2008–2012. Incidence and mortality, age group specific numbers and rates per 100 000 person years of diagnosis.
Results	The incidence and mortality rates of cancer of the liver and pancreas increased over time while the rates of cancer of the gall bladder and bile duct decreased. All

	HBP-c were more frequent in person age over 70 than in younger persons.
Comment	About 20% of Denmark cancer registry was missing of microscopic verification. But the rare histological subtypes of specific cancers associated with either a much worse or better prognosis which reduced generalizability.

Table 2 (Trend, continued)

Authors/year	Wiangnon S, Kamsa-ard S, Suwanrungruang K, Promthet S, Mahaweerawat S, Khuntikeo N., 2012 [15]
Country	Thailand
Objectives	To assess the incidence trend of hepatocellular carcinoma in Khon Kaen, Thailand, between 1990 and 2009.
Explanatory factors on LBD cancer incidence	Gender and year of diagnosis
Outcome	The age standardized rate (ASR) and annual percent change.
Setting	The data of all cancers in Khon Kaen can be retrieved from database of the Khon Kaen Cancer Registry (KKCR).
Sample	All 7,859 liver cancer cases (ICD10 code C22.0) according to International Classification of

	Disease for Oncology were selected from population-based data set of KKCR.
Methods/Tool/ Analysis	Jointpoint analysis was used for trend incidence analysis and identified a significant changing point of the trend. The age standardized rates (ASR) were used to describe the incidence rates.
Results	The incidence trends decreased in both sexes. Males were affected two times more likely than females. The most common age group of cases was 50 and 69 years. The ASR was 13.1 to 49.8 per 100,000 among males and 4.8 to 38.4 per 100,000 among females since 1985.
Comment	Having a long period of data are useful for trend analysis

Table 2 (Trend, continued)

Authors/year	Kamsa-ard S, Wiangnon S, Suwanrungruang K, Promthet S, Khuntikeo N, Kamsa-ard S, Mahaweerawat S., 2011 [16]
Country	Thailand
Objectives	To perform a statistical estimation of the incidence trends of liver cancer, particularly focusing on cholangiocarcinoma (CCA)
Explanatory factors on LBD cancer incidence	Gender, year of diagnosis, and ASR by calendar year.
Outcome	ASR and annual percent change of liver cancer.
Setting	Cases of CCA were retrieved from the KKCR and all those with a specific ICD-O-3rd diagnosis with a coding of C22.1, C24.0, C24.8 and C24.9 were selected.
Sample	18,589 cases of liver cancer registered between 1985 and 2009.
Methods/Tool/ Analysis	The ASR and 95% CI were used for describe the incidence rates of CCA. Incidence trends were calculated using the generalized linear model method (GLM). Jointpoint analysis was used to identify the best fitting model.
Results	The rate increase in the first 5 to 6 years may be due to improved completeness of the registry. The rather stable rate in the subsequent 10 to 12 years later.
Comment	The issues of prime concern include: data collection, data processing, statistical analysis/synthesis and reporting reliable and complete data.

Table 2 (Trend, continued)

Authors/year	Tzeng IS, Lee WC., 2015 [45]
Country	Taiwan
Objectives	To forecast hepatocellular carcinoma mortality in Taiwan using an age-period-cohort (APC) model.
Explanatory factors on LBD cancer incidence	Gender, age, year of diagnosis, and year of birth
Outcome	Mortality of hepatocellular carcinoma
Setting	HCC mortality of Taiwanese age more than 40 year old.
Sample	Individual health records derived from the Department of Health, Taiwan between 1976 and 2005.
Methods/Tool/ Analysis	The age model, the age-period model, and the APC model.
Results	The age effects and period effect were quite notable of increasing HCC mortality of Taiwanese in both sexes while cohort effect decreased the HCC mortality.
Comment	Although the APC forecasting seems to be a good fit for trend, researcher should not project trend more than 15 years ahead due to forecasting method assumption.

Table 2 (Trend, continued)

Authors/year	Zhang Y, Ren JS, Shi JF, Li N, Wang YT, Qu C et al., 2015 [46]
Country	China
Objectives	To give a longer-term and more recent comprehensive picture on the current status of PLC worldwide.
Explanatory factors on LBD cancer incidence	Age, gender, year of diagnosis
Outcome	Incidence and changing trend in incidence
Setting	PLC cases including HCC, CCA, and combined of HCC and CCA (cHCC-CC).
Sample	The data from Cancer Incidence in Five Continents of 24 countries in Americas, Asia, Europe, and Oceania between 1973 and 2007.
Methods/Tool/ Analysis	Joinpoint regression analysis was used. The significance tests of changing trend using permutation method. The significant of annual percentage change (APC) and average annual percentage change (AAPC) was examined.
Results	The ASRs for PLC were declining in several Asian populations between 1973 and 2007. In contrast, ASRs for PLC were increasing in some European, American and Oceanian populations.
Comment	Trend of the PLC should present by histologic subtypes because difference histology are different measures and policy for disease control.

3 Risk factors of cholangiocarcinoma

The risk of developing bile duct cancer is mainly from chronic inflammations of the bile duct.[4] In western countries, primary sclerosing cholangitis may cause by many conditions of the liver or bile duct.[47] In addition, hepatolithiasis, chronic non-alcoholic liver disease, cirrhosis, and obesity are other known risk factors for CCA. Other rarer conditions associated with CCA development are bile-duct adenoma, multiple biliary papillomatosis, congenital fibropolycystic liver, choledochal cysts, Caroli's disease, and exposure to the radiopaque medium thorium dioxide.[48, 49] However, approximately 90% of CCA cases cannot be assigned to any of these recognized risk factors.[47, 50] Comparison of these risk factors in the West, hepatolithiasis is a very uncommon disease. In contrast, intra- and extrahepatic bile duct stones are much more common in Eastern Asia.[51] In Japan, hepatolithiasis is strongly associated with cholangiocarcinoma with proportions varying from 10% to more than 60%.[52] In Taiwan, similarly other risk factors of cholangiocarcinogenesis, as bacterial infection and bile stasis, which virtually demonstrate in all patients, and underlie cholangiocarcinoma development. This carcinogenic process is not limited to the intrahepatic lesions.[51, 53]

The situation of liver fluke infestation in East Asian countries indicates that it is endemic in Korea, China, Vietnam, Thailand, Laos, and Cambodia. In addition, people with *C. sinensis* infestation were infrequently reported from Singapore and Malaysia due to many being infested from other countries with travelling or through eating imported fish[54]. Moreover, Asian immigrants from endemic areas contribute liver fluke infestation in all parts of the world. In Thailand however, CCA is strongly associated with infestations from liver fluke parasites which occur in the bile ducts, gallbladder, and liver as a result of eating raw fish.[6, 50, 55, 56]

In 1994 the International Agency for Research on Cancer has recognized *O. viverrini* as "definitely carcinogenic to humans", based on evidence from experimental studies on animals and humans.[57] In 2009, the same classification was

given to *C. sinensis*.[58] Several studies in Thailand indicated that estimations of the relative risk associated with *O. viverrini* infestation although none of them evaluated the specific risk of CCA.

4 Liver fluke and prevalence in Thailand

2.4.5 Species and life cycle

Liver fluke infestation in Thailand is due to the *O. viverrini Opisthorchis viverrini* parasite.[59] The characters of *O. viverrini* adult worms are flat, leaf-shaped, and transparent and the adult worm lives in the hepatic bile ducts of a definitive host, such as humans, dogs, and cats. Their first intermediate host is snails in Bithynia genus and cyprinoids fishes are the second intermediate host. A schematic overview of the *Opisthorchis* life cycle is given in Figure 3.

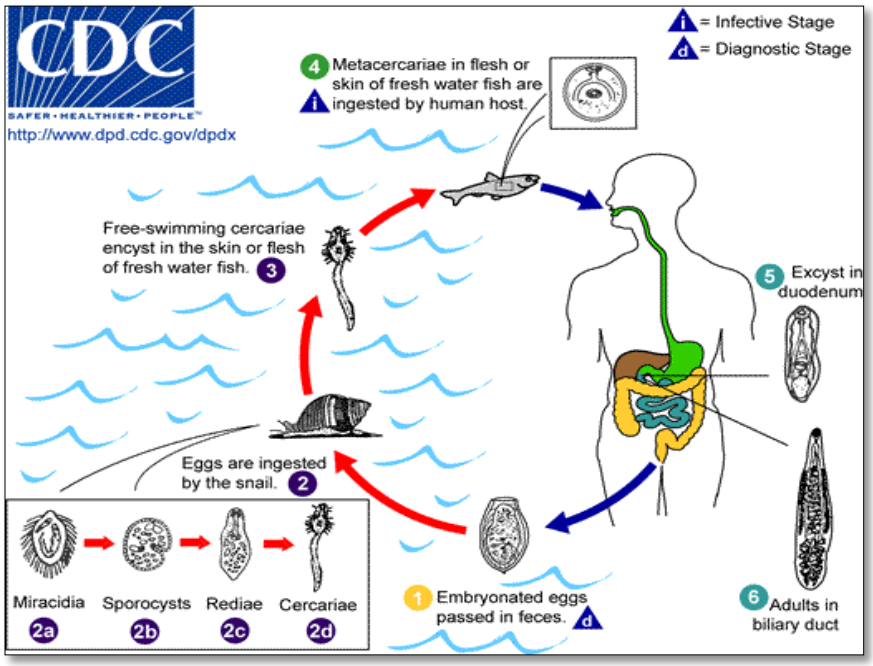

Figure 3 Opisthorchis life cycle

(Modified from http://www.cdc.gov/parasites/opisthorchis)

In the life cycle of opisthorchiasis from adult flukes inhibit the bile duct where they lay many eggs. Fully developed eggs are then deposited into the environment via the feces. After ingestion by a suitable host (first intermediate host) such as a snail, the eggs develop inside the snail in several stages and consequently become cercariae, a tadpole-shaped larval trematode worm. Then, the cercariae are released from the snail and penetrate freshwater fish (second intermediate host) and encyst as metacercariae in the muscles or under the scales. Mammals, such as cats, dogs, and humans as a definitive host become infested by ingesting undercooked fish containing these metacercariae. The metacercariae undergo excystation in the duodenum after ingestion. They then ascend through the intestinal wall and liver parenchyma into the biliary ducts where they attach and develop into adults. They can lay eggs after 3 to 4 weeks in the bile duct. The adult flukes reside in the biliary and pancreatic ducts of the mammalian host, where they attach to the mucosa.[5]

2.4.6 Geographic distribution of liver fluke

Approximately 25 million people in Asia and Europe are infested with three human liver flukes: O. *viverrini*, *C. sinensis* and *O. felineus*. Among these, humans are mostly infested with *C. sinensis* (15 million) followed by *O. viverrini* O. *viverrini* (10 million) and *O. felineus* (1.2 million).More than 700 million people worldwide are at risk of being infested by these three liver flukes.[60, 61] The regions mostly affected by *O. viverrini* O. *viverrini* are in the areas of northeast Thailand, Lao PDR, Cambodia and Vietnam while endemic areas for *C. sinensis* are Korea, China, Taiwan, Vietnam, and Japan. Poland, Germany, Russia, Federation of Kazakhstan and Western Siberia are endemic for *O. felineus*.[62] An outbreak of *O. felineus* reoccurred in central Italy between 2003 and 2011 after consuming raw fillets of tench (*Tinca tinca*).[63] Because of extensive traveling and also importation of fresh water aquaculture products from Asia, it is now increasingly common to find infested people in non-endemic areas.[54]

2.4.7 Prevalence of liver fluke in Thailand

The prevalence of parasite infestation in Thailand in 1957, 1981, 1991, 1996, 2001 and 2009 was reported by the Bureau of General Communicable Diseases, Department of Disease Control, Thai Ministry of Public Health.[64, 65] Previous surveys used the Kato thick smear and Formalin-ether concentration technique (FECT) for stool examination. The parasite infestations prevalence has been decreasing along the survey years (Figure 4). The last nationwide survey in Thailand during 2009 found an overall prevalence of *O. viverrini* infestation of 8.7%; the Northeast (16.6%), the North (10%), the Central (1.3%), and the South (0.01%) regions.[65]

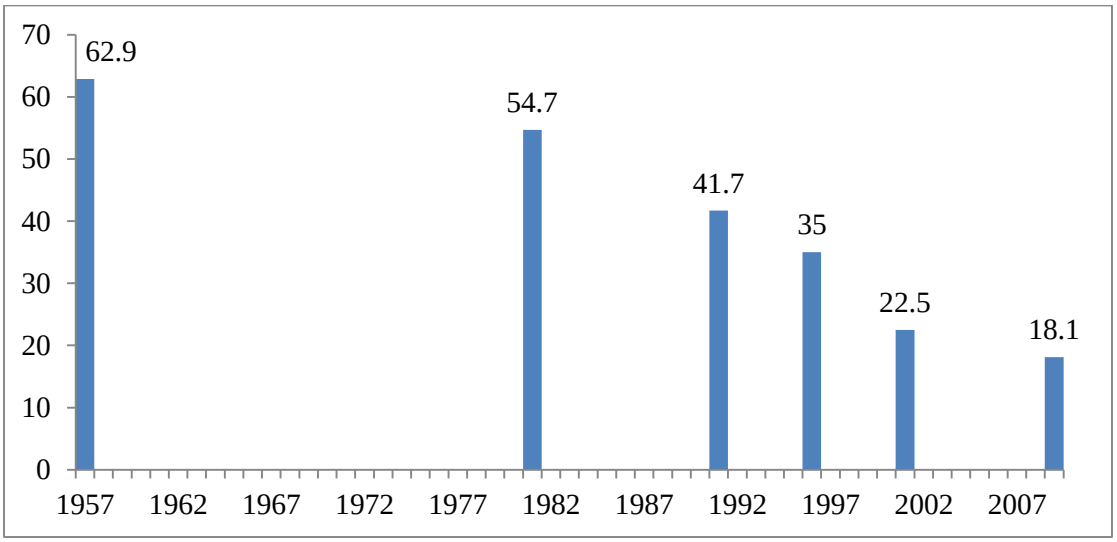

Figure 4 Prevalence of parasite infestation in Thailand 1957-2009 [64]
(Source: http://thaigcd.ddc.moph.go.th)

2.4.8 Liver fluke control program in Thailand

O. viverrini is a carcinogenic parasite and the primary prevention of CCA can be achieved by parasite control, especially mass drug administration, which, in northeast Thailand, was initiated in the early 1980s.[21, 66] In the past, the historical prevalence of *O. viverrini* in Thailand was as high as 80%, and then declined to 25% in 1990. In 1997, it reduced even further to 15–20%. After that it was more or less stable

until 2013. The highest prevalence of *O. viverrini* infestation was found in Northeast Thailand, while it also has an extensive mass drug administration program. A trend analysis of CCA during 1985–2000 in Khon Kaen province showed that the significant reduction of CCA incidence in the few past years may be due to a real decrease in risk factors such as liver flukes.[16] Not only the liver fluke control but other factors may also lead to a reduction of the CCA incidence and it is not easy to evaluate the direct impact of mass drug administration and health education on the occurrence of liver cancer. However, likely parasite control and health education could be the essential factors. Further analysis is required to ascertain this reduction of CCA rate.

2.4.9 Screening program under the campaign "Strategy for liver flukes and cholangiocarcinoma elimination, the E-sarn agenda"

The Thai Ministry of Public Health has been putting more effort on the control of liver and bile duct cancer by initiating the campaign "Liver flukes and cholangiocarcinoma elimination, The E-sarn agenda".[67] This project was launched in 2012 in the northeastern region and has expanded to all other regions countrywide. In this campaign, every provincial health office screened at risk groups. Public health officers would examine stool for presence of *O. viverrini* infestation by microscopy and ultra-sound. Figure 5 displays a flow diagram of the procedures that should be followed by the public health officers.[64, 67]

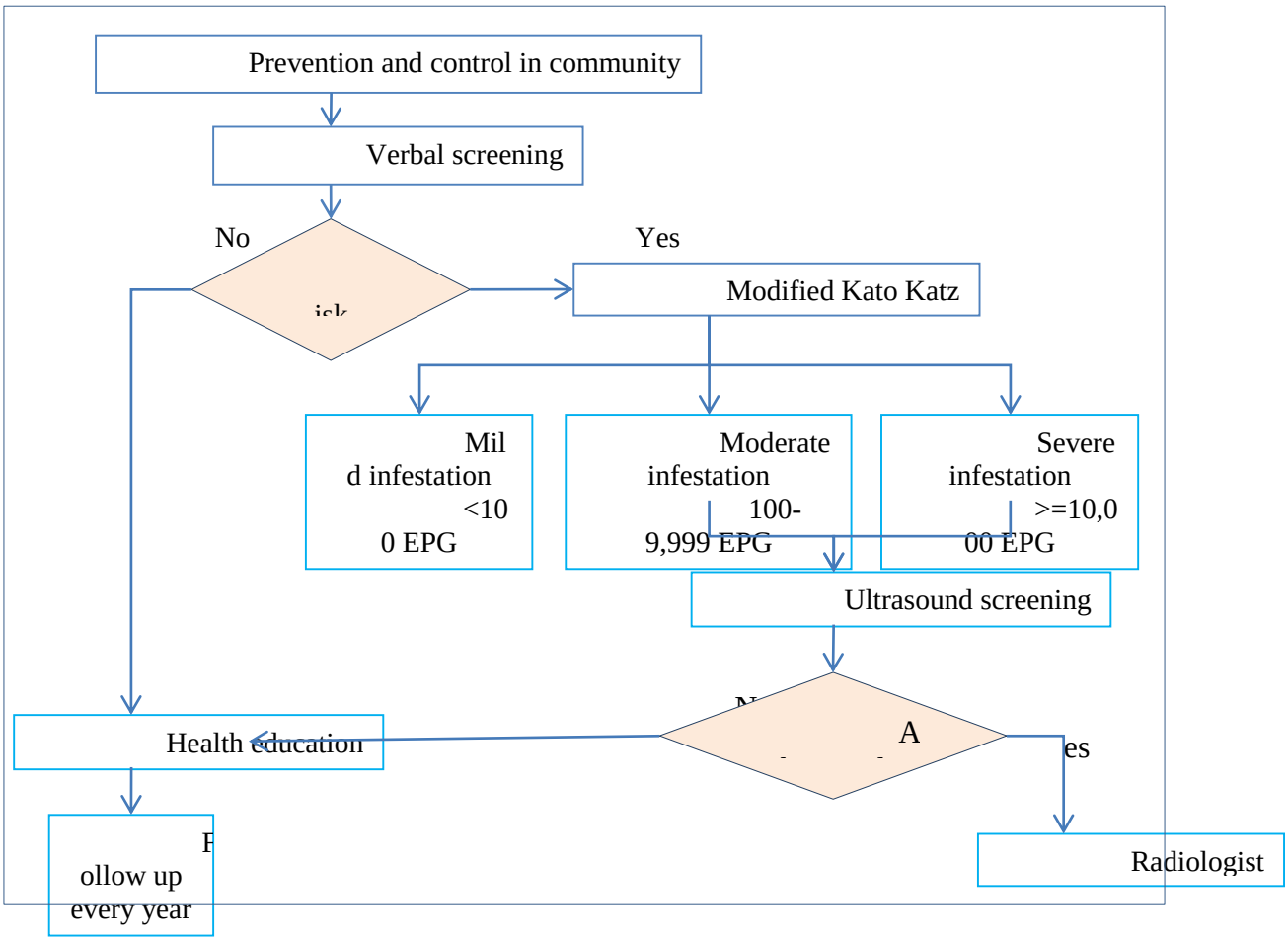

Figure 5 Flow of cholangiocarcinoma screening program in Thailand, 2013 [64]

2.4.10 Diagnosis of human opisthorchiasis

The human opisthorchiasis can be diagnosed by clinical manifestations, parasitological, molecular biological, or immunological methods.[68] Clinical manifestations of the patients are practically indistinguishable from those of other liver diseases. The features of the *O. viverrini* eggs are, by light microscopy, difficult to differentiate from those of other minute intestinal fluke eggs. Polymerase chain reaction is very complicated, needs special and expensive materials, and is time-consuming; however, it is highly sensitive and specific. Immunological testing is the method of choice: the techniques are applicable to both routine laboratory work and field or

epidemiological studies. Of these tests, enzyme-linked immunosorbent assay (ELISA) and immune electro transfer blot assay are often used for the detection of *O. viverrini* -specific antigens and antibodies (IgM, IgG, IgA, or IgE). Monoclonal antibodies are prepared to detect copro antigens, while the crude somatic and excretory-secretory antigens from the adult worms, metacercariae, eggs, and snail intermediate hosts are prepared in order to detect antibodies in sera [68].

The diagnostic method of liver fluke infestation, i.e. opisthorchiasis caused by *O. viverrini*, is by detecting the parasite ova in faecal samples. However, in light infestations or when ova output is low or nil as a result of biliary obstruction, faecal examination for *O. viverrini* eggs may be falsely negative.[69]

Alternative methods for diagnosis of opisthorchiasis include the use of deoxyribonucleic acid probe for detecting complementary deoxyribonucleic acid of the fluke in stool, serum by using antibody-based ELISA for antigen detection and various antibody detection assays.[70] However, an indirect ELISA may be as sensitive as the detection of eggs in stool but cross-reaction with other parasitic infestations is commonly and frequently found when crude parasite extract was used. The value of the antibody detection assay in opisthorchiasis is limited by the cross-reactive nature of the crude antigen used.[71]

A review of studies related to prevalence and associated factors of *O. viverrini* infestation is summarised in Table 3. The first study aimed to determine the association between infestations with *O. viverrini* with incidence of CCA in Khon Kaen province, Thailand. [22] There was no association among CCA and *O. viverrini* infestation. CCA incidence was measured in the entire district population, while prevalence of *O. viverrini* was estimated from a sample. Correlation of *O. viverrini* infested person and may have been underestimated due to previous treatment and light infestations or when ova in feces is low.

The second study aimed to identify risk factors and predict the prevalence of *O. viverrini* at non-surveyed locations in Lao PDR. [55] The overall

prevalence of *O. viverrini* infestation was 61.1% (95% CI: 59.5–62.8%).
Rice farmers were at highest risk of infestation (81.3%, 95% CI: 79.2–83.4%), followed by workers of the tertiary sector (73.8%, 95% CI: 64.6–83.1%). Age, low land area, and farmer and workers of the tertiary sector had a higher risk of *O. viverrini* infestation.

The third study was conducted in Nakhon Ratchasima province, Northeast Thailand an endemic area of human *O. viverrini*. [72] About 2.48% of participants were infested with *O. viverrini*. Males were more likely to be infested than females. *O. viverrini* infestation was more common in those aged 51-60 years and was positively associated with education (primary school) and occupation (agriculture). However, the study did not discuss about limitations of using a modified Kato thick smear technique because it is difficult to differentiate by egg morphology between minute intestinal flukes and *O. viverrini*.

The fourth study focused on using indirect ELISA for measuring *O. viverrini* specific immunoglobulins in serum, urine and saliva in endemic areas of opisthorchiasis, Khon Kaen province, northeastern Thailand.[70] Seropositive cases were found in both parasite egg-negative as well as *O. viverrini* egg-positive groups. The levels of serum IgG correlated with intensity of *O. viverrini* infestation. Diagnostic sensitivities based on serum and salivary IgG, IgA were also positively associated with the intensity of infestation.

The last reviewed study was also conducted in Thailand.[71] The study aimed to evaluate indirect ELISA for identifying *O. viverrini* infestation. Sensitivities of the indirect ELISA were 92% and 91% using the first and second batches of crude extracts prepared from adult *O. viverrini* derived from two distinct sources, respectively. Four out of 53 patients of groups 1 and 2 gave false negative ELISA when the first batch of antigen was used. The ELISA test could be used as a screening method for opisthorchiasis. However, the test could not be used alone for diagnosis as it could

not differentiate between the present and past cure infestations and the false positive rate was high.

In summary, there have been many surveys conducted to determine the prevalence of *O. viverrini* in Thailand and endemic areas and its association with CCA. The finding from the literature review suggests that due to the difficulty of differentiating egg morphology between minute intestinal flukes and *O. viverrini* more accurate diagnostic techniques such as ELISA should be introduced. However, cross-reaction of antibody of the infestation is a concern.

Table 3 Summary of studies related to the prevalence and risk factors of liver fluke infestation.

Authors/year	Sriamporn S. et al., 2004 [22]
Country	Thailand
Objectives	To determine the association between infection with *O. viverrini* with incidence of CHCA in the same and different district of Khon Kaen province.
Exposure variables	Age, sex, living district, year of infected
Outcome	*O. viverrini* infection, number of *O. viverrini* eggs per gram (EPG)
Setting	Khon Kaen population, based on Khon Kaen cancer registry.
Sample	The cohort population in selected villages in Khon Kaen province from 1990 to 2001.
Methods/Tool/ Analysis	FECT method was used for examine *O. viverrini* infestation. The association between the incidence of CCA and the prevalence of OV-infected population was estimated. A Poisson model was fitted between the number of incidence cases, age group, and sex.
Results	There was no association among CCA and *O. viverrini* infection (r=0.009) after adjusted age, sex and period.

Risk factors for *O. viverrini* infection	Age, sex, living district, year of infected
Comment	Correlation of *O. viverrini* infected person and might been under estimated due to previous treatment and light infections or when ova in feces is low.

Table 3 (continued)

Authors/year	A Forrer et al., 2012 [55]
Country	Lao PDR
Objectives	To map the distribution of *O. viverrini* in southern Lao PDR, identify underlying risk factors, and predict the prevalence of *O. viverrini* at non-surveyed locations.
Related variables	Age, Lao Loum ethnic group, educational attainment, occupation, and unsafe drinking water.
Outcome	OV infection
Setting	Population in Champasack, Lao PDR, the plains and flatland dominated by the Mekong River.
Sample	About 4,380 samples in 51 villages were selected. Participant selection was achieved using a two-stage sampling method.
Methods/Tool/ Analysis	A two-stage sampling method was used. The association between infection risk and covariates was

	assessed using non spatial bivariate logistic regressions and logistic geostatistical models.
Results	The prevalence of *O. viverrini* infection was 61.1% (95% CI: 59.5–62.8%). Rice farmers were at highest risk (81.3%, 95% CI: 79.2–83.4%), followed by workers of the tertiary sector (73.8%, 95% CI: 64.6–83.1%).
Risk factors for *O. viverrini* infection	Age, low land area, and farmer and workers of the tertiary sector
Comment	Geographical variable such as distance from river should be in consideration. The study did not present sampling frame and sampling technics which can affect to prevalence of the *O. viverrini* infection.

Table 3 (continued)

Authors/year	Kaewpitoon SJ, Rujirakul R, Kaewpitoon N., 2012 [72]
Country	Thailand
Objectives	To determine actual levels of *O. viverrini* infection in Nakhon Ratchasima province, Northeast Thailand.
Outcome	Prevalence of *O. viverrini* infection.
Setting	A cross-sectional parasitological and questionnaire survey was conducted in 32 districts of Nakhon Ratchasima province.

Authors/year	Kaewpitoon SJ, Rujirakul R, Kaewpitoon N., 2012 [72]
Country	Thailand
Sample	A total of 1,168 villagers (516 male and 652 female).
Methods/Tool/ Analysis	A socio-demographic data were collected using questionnaires. *O. viverrini* infection test using modified Kato's Katz technique. Chi-square test was performed to determine association between socio-demographic and *O. viverrini* infection.
Results	The prevalence of *O. viverrini* infestation was 2.48%. Males were more likely to be infected than females, but the different was not statistically significant. *O. viverrini* infection was most found in age 51-60 year and was positively associated with education (primary school) and occupation (agriculture).
Risk factors for *O. viverrini* infection	Age, education, occupation
Comment	This study did not present the process of sampling technic. Equal sample size (40 respondents) in every district may induce some bias. Researcher did not discussed about limitation of a modified Kato's thick smear technique.

Table 3 (continued)

Authors/year	Sawangsoda P, Sithithaworn J, Tesana S, Pinlaor S, Boonmars T, Mairiang E, et al., 2012 [70]
Country	Thailand
Objectives	To evaluate indirect Enzyme Immunosorbent Assay (ELISA) for measure *O. viverrini* specific immunoglobulins in serum, urine and saliva and assessed efficacies in diagnosis of opisthorchiasis in a sample population in Khon Kaen.
Outcome	Diagnostic efficacy of clinical specimens and relationship between antibody levels and intensity of *O. viverrini* infection.
Setting	Participants from endemic area of opisthorchiasis in Khon Kaen Province, northeast of Thailand.
Sample	A total of 207 individuals were sampled. Those were having an aged between 28 and 92 years old.
Methods/Tool/ Analysis	Using ELISA procedures to detect *O. viverrini* specific antibodies in urine, saliva and serum while FECT technique was used as a gold standard diagnosis. Sensitivity, specificity, and the likelihood ratio for a given result were calculated.
Results	Seropositive cases were found in both parasite egg-negative as well as *O. viverrini* egg-positive groups. The levels of serum IgG correlated with intensity of *O. viverrini* infection ($P < 0.05$). Diagnostic sensitivities

	based on serum and salivary IgG, IgA positively associated with the intensity of infection.
Comment	Although there were no ova or light infections the ELISA can use for identified *O. viverrini* infection by using serum sample. However, cross-reaction should be concern.

Table 3 (continued)

Authors/year	Y. Sakolvaree et al., 1997 [71]
Country	Thailand
Objectives	To evaluate indirect Enzyme Immunosorbent Assay (ELISA) for identified *O. viverrini* infection.
Outcome	Antibodies test positive for *O. viverrini* infection
Setting	Adult *O. viverrini* worms were from two different sources. First batch were collected from infected people whereas the second batch was from infected hamsters.
Sample	Antibodies sample collected from 4 group of samples; group1=22: only *O. viverrini* infected, group2=31: *O. viverrini* with other parasites infected, group3=141: only other parasites infected (no OV), and group4=24: healthy sample (parasite free)
Methods/Tool/ Analysis	Two batches of crude antigens extracted from adult *O. viverrini* worms were compared. One was derived from adult worms harvested from the livers of laboratory infected hamsters and another was obtained from worms sedimented from the feces of *opisthorchiasis* patients following treatment with Praziquantel.
Results	Sensitivities of the indirect ELISA were 92% and 91 % using the first and second batches of crude extracts prepared from adult *O. viverrini* derived from

	two distinct sources, respectively. Four out of 53 patients of groups 1 and 2 gave false negative ELISA when the first batch of antigen was used.
Comment	The ELISA test could be used as a screening method for *opisthorchiasis*. However, the test could not be used alone for diagnosis as it could not differentiate between the present and past cure infections and could give also some false positive results.

Chapter 3:Rationale

Cholangiocarcinoma caused by *O. viverrini* infestation is a major problem in Thailand. The epidemic regions of the disease are mainly in the northeast and north of the country. Reports and investigations of cholangiocarcinoma and *O. viverrini* infestation in southern Thailand are limited.

The incidence of LBD cancer by histological type has been changing. Since 2000 CCA in Songkhla seems to be increasing while the percentage of cases with morphological verification (%MV) of LBD cancer declined from 60% in 1997 to 20% in 2005. This occurred after the adoption of the Bithmus and Corlette classification system, which does not require laboratory and histological investigation in diagnostic procedures for LBD cancer cases. Such a classification has higher sensitivity than the pathological diagnosis, thus the number of cases with LBD cancer increased. Another side effect was that many clinicians did not specify the type of cancer in the medical records. When the percentage of morphological verification is low, the true incidence rates for each histological type are often underestimated. Thus, this study aims to estimate the incidences of HCC and CCA in Songkhla province from 1989 to 2013 using a multiple imputation technique to determine histological type among cases having unknown histology. After imputation, this study also aims to fill the knowledge

gap on disease trends which could enhance our understanding of the incidence of LBD cancer subtypes in the future. Furthermore, the changing demographic profiles and the control measures might affect the HCC and CCA incidence.

The reason for the rising trend of CCA in Songkhla is unknown. According to a report from the Department of Disease Control, Ministry of Public Health, Thailand, the prevalence of *O. viverrini* infestation in southern Thailand was low.[65] Thus, there have been no reports or investigations of CCA and *O. viverrini* infestation in southern Thailand.

A cross-sectional survey was conducted to determine the possible association between CCA and *O. viverrini* infestation in Songkhla province. The results of this study may shed some light for the increasing CCA trend in Songkhla province facilitating appropriate strategies for CCA control programs.

Chapter 4:Research questions

The number of unknown histological type of LBD cancers in Songkhla province have been increasing. The prevalence of CCA, the second most common LBD cancer, seems to be increasing. *O. viverrini* infestation is believed to be the major cause of CCA in Thailand. Thus, the following research questions are proposed:

1) What is the incidence of CCA in Songkhla province?
2) What are the reasonable predicting factors for unknown histology type for LBD cancer in Songkhla cancer registry?
3) What is the future trend of HCC and CCA incidence in Songkhla province?
4) What is the prevalence of *O. viverrini* infestations in Songkhla province?
5) What are the principal factors associated with *O. viverrini* infestation in Songkhla province?

Chapter 5:Objectives

The objectives of this study are:

1) To determine the histological type among cases having unknown histology of LBD cancer in Songkhla province from 1989 to 2013 using a multiple imputation technique.
2) To estimate the incidence and trends of HCC and CCA incidence in the Songkhla population from 1989 to 2013 and projected to 2030.
3) To determined the prevalence of *O. viverrini* infestation and identified associated factors of *O. viverrini* infestation among the adult population in Songkhla.

CHAPTER 2

METHODS

Chapter 6:Conceptual frame work

The methodology of this study was separated to two parts according to the objectives as follows:

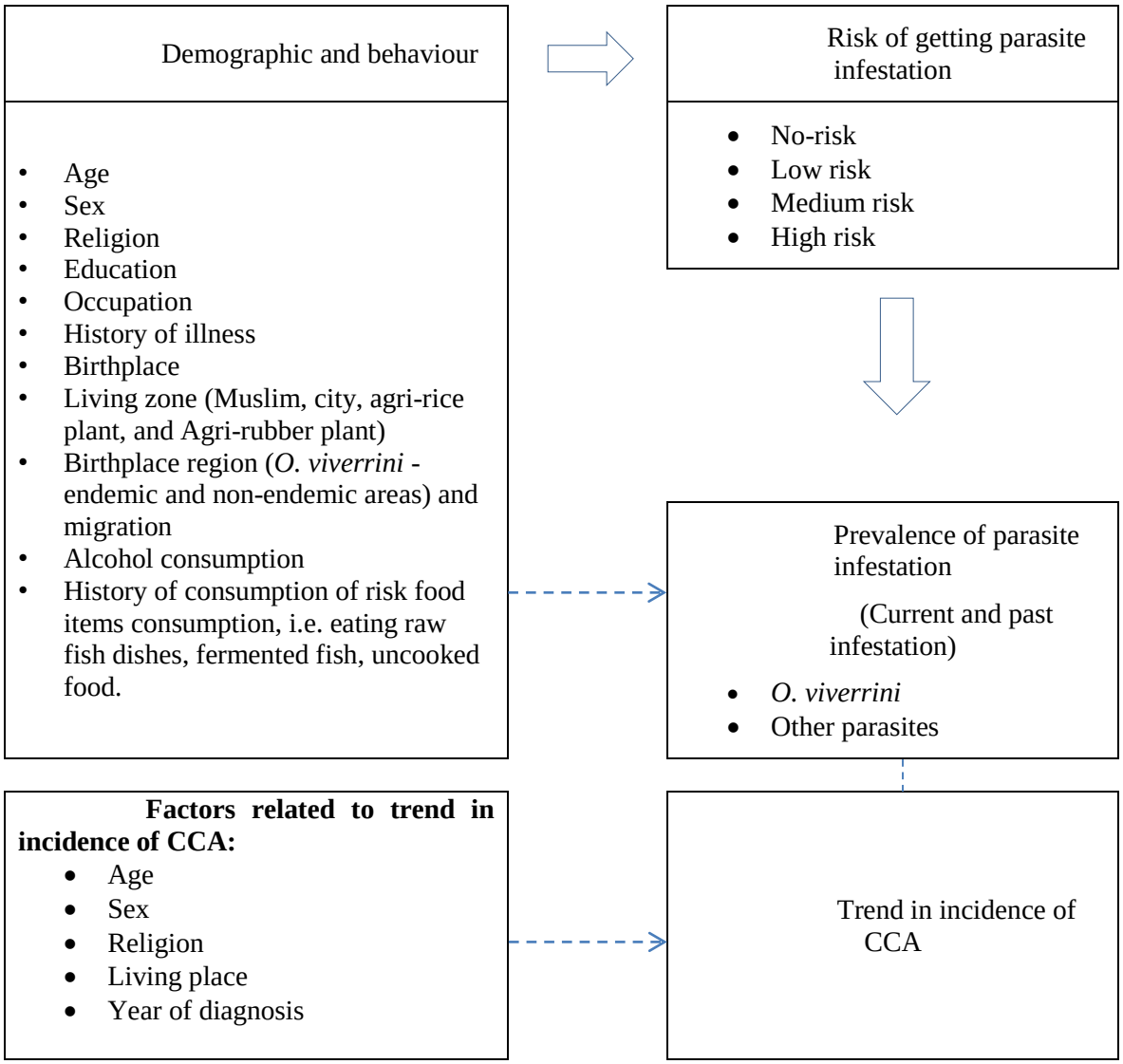

Figure 6 Conceptual framework

Chapter 7:Methodology

1 Methodology of objective 1

7.1.1 Objective 1 part I

To estimate the incidences of HCC and CCA in Songkhla province from 1989 to 2013 by imputing unknown cell type of LBD cancer in the Songkhla population-based cancer registry using variables that predict the known cell types.

7.1.2 Study design

Retrospective descriptive study

7.1.3 Study setting

All records from Songkhla population-based cancer registry.

7.1.4 Study setting

Liver cancer patients in Songkhla population-based cancer registry between 1989 and 2013 who had one of three histological types: Hepatocellular (HCC), Cholangiocarcinoma (CCA), and other specified histology.

7.1.5 Study subjects

This study included 2,387 liver cancer cases with available information for each case, including registration area, age or birth date, gender, place of residence, date of diagnosis, basis of diagnosis, International Classification of Diseases (ICD-10) code, topography code, and morphology code.

Cases were classified into four initial groups based on the third edition of the International Classification of Diseases for Oncology (ICD-O3) [73] as follows: Group 1: HCC (topography (T) code C22.0 and morphology (M) codes 8170-8176); Group 2: CCA (T C22.1, and C24.x, excluding C24.1 and M 8050, 8140-8141, 8160-8161, 8260, 8440, 8480-8500, 8570-8572); Group 3: Other specified LBD cancers: (T C22.0 with any M and T C24.1); Group 4: LBD cancers with unknown histology: (T C22.0 and M 8000-8005).

There have been attempts to classify subtypes of CCA according to the radiographic appearance as the above-mentioned Bithmus and Corlette classification of perihilar CCA, and based on cells of origin.

7.1.6 Inclusion criteria

All newly diagnosed patients with primary liver cancer during 1989-2013 were included in this study.

7.1.7 Exclusion criteria

Any individual with missing values of explanatory variables included in the model were excluded from the analysis.

7.1.8 Data collection

Data were obtained from the Songkhla population-based cancer registry using a data collecting form (Annex 2). All patients diagnosed between 1989 and 2013 were included.

7.1.9 Population

The population denominator used for the calculation of age-adjusted rates were estimated from 1990, 2000, and 2010 censuses published by the National Statistical Office, Thailand, which provided annual estimates by age group and sex for each province. Inter-census estimates of the population by age group and sex were

calculated with the assumption of an exponential change in the population between the three censuses.

7.1.10 Multiple imputation method

The imputations were performed using the 'mice' (Multivariate Imputation by Chained Equations) package [25] in R[74] . Cases with unknown histology were imputed with one of the other known histological categories according to the probability distribution of the groups among those who had known histology obtained by the chained equation method plus a degree of random error. Since the outcome in this case was a multiple categorical variable, a multinomial logistic regression model was used to generate the distribution according to the predictive ability of variables present in the registry database. These variables included sex, age, year of diagnosis, religion, and residential district. The model is given by

$$f(k, i) = \beta_k + x_i, \qquad (1)$$

where β_k is the set of regression coefficients associated with histological type k (HCC, CCA, or other), and x_i is the set of predictor variables associated with observation i. The method described by White et al (2011) was applied to avoid bias due to perfect prediction.[75] We repeated 1,000 iterations of multiple imputation to obtain the 95% Bayesian probability intervals (PI) obtained from the quantiles of the posterior distribution for the three histological types.

As the data in the cancer registry also shows that the number of LBD cancer cases diagnosed during 1989-2006 was lower than in the period after 2006, we therefore included a random sample of cancer cases with unknown primary in the abdomen (C76.2) or unknown primary (C80.9) into Group 4, stratified by age group, sex and year of diagnosis. The optimal number of cases randomly selected from the unknown primary group to include in group 4 differed each year ranging from 30% for the period 1989 to 1997 down to 0% from 2007 onwards or the total number of

unknown primary cancers in each age/sex/year strata, whichever was the lowest. The reason for including these unknown primary cancer cases into group 4 is that we believe some of them were misclassified due to incomplete investigations (for example, the patient died or refused to have any surgical procedure (Figure 7).

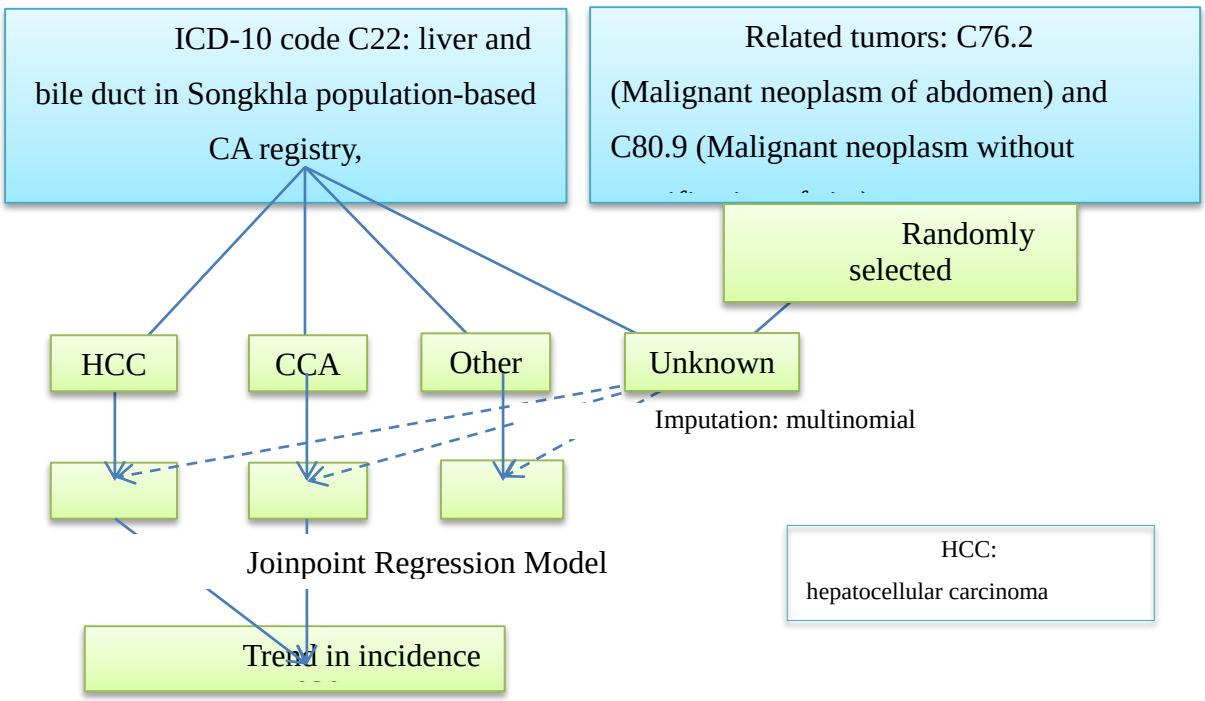

Figure 7 Concept of data imputation for unknown histological types of LBD cancer

7.1.11 Data analysis

Since comparison of the proportion of HCC and CCA over a long period can be biased by the change in the age structure of the population, we used age-standardized incidence rates for the two groups to illustrate the effect of time on the probability of imputation. The rates were standardized to the world population as proposed by Doll (1966) and calculated for each of the 24 calendar years between 1989 and 2013.

After imputation, descriptive statistics such as frequencies and percentages were presented. We compared temporal trends of HCC and CCA based on

three models: Model 1: LBD cancers with known histology only; Model 2: All LBD cancers with imputation of unknown histology, and Model 3: All LBD cancers plus cases with unknown primary, both in the abdomen and not otherwise specified, with imputation of unknown histology.

2 Methodology of objective-1

7.2.12 Objective 1 part II

To determine trends of HCC and CCA incidence from 1989 to 2013 and projected to 2030.

7.2.13 Study design

Retrospective descriptive study

7.2.14 Study setting

All records used from Songkhla population-based cancer registry.

7.2.15 Study setting

Liver cancer patients in Songkhla population-based cancer registry diagnosed between 1989 and 2013. Cases were classified by histological type: HCC, CCA or other.

7.2.16 Study subjects

This study included 2,676 cases derived from the previous multiple imputation models in the first part. The results showed reasonable number of the three subtypes: HCC, CCA and other, which was appropriate for the trend analysis.

7.2.17 Population

The population denominator used for the calculation of age-standardized rates was estimated from 1990, 2000 and 2010 censuses published by the National Statistical Office, Thailand.

7.2.18 Data analysis

The age-standardized incidence rates (ASRs) per 100,000 population of HCC and CCA from 1989 to 2013 were analysed by the direct method using the Segi world standard[76] in each 5-year age group (age 20 to 85+ years). The trends in incidence of LBD cancer in Songkhla by histological subtype and sex were estimated using a Poisson regression model and the ASR projections were calculated until the year 2030 in 5-year groups. The ASR and number of LBD cancer cases was estimated using three models; Joinpoint regression analysis[77], age-period-cohort (APC) model [78], and Nordpred age-period-cohort model[37]. The number of LBD cancers and projections until 2030 were compared between the three models. To avoid overestimation of cases from the multiplicative model, a power function in the Nordpred model was used to attenuate the linear drift by 21.6%, 48.3%, 65.9%, and 77.6% respectively, for the first period of projection trend (2014–2018), second (2019–2023), third (2024–2028), and fourth (2029-2030). The natural spline method[79] was used to smooth the non-linear curves of the age-period-cohort model and extend the period and cohort effects into the future via non-linear interpolation for yearly ASR projection.

3 Methodology of objective II

This method aimed to determined the prevalence of *O. viverrini* infestation and identify associated factors of parasite infestation among the adult population in Songkhla province.

7.3.19 Study design

A cross-sectional parasitological and questionnaire survey

7.3.20 Study setting

This study was conducted in Songkhla province, southern Thailand between January 2015 and May 2015.

7.3.21 Study subjects

Thai residents registered in Songkhla province.

7.3.22 Population and sample size

The population denominators in each stratum; district, sub-district, and villages used were taken from the population survey of Songkhla provincial Statistics Office in 2012. The age- and district-specific populations were used to adjust sample weighting.

7.3.23 Sample size calculation for prevalence estimation

The required sample size was calculated using the following formula:

$$n = \frac{Z^2_{\alpha/2}\, p \times (1-p)}{d^2}$$

where p = estimated prevalence, obtained from a previous study[65] of *O. viverrini* and assumed to be 10 per 100 population,.

d = error in estimating prevalence (margin of error), assumed to be 0,03

α = probability of type I error, assumed to be 0.05

Z = standard normal deviate corresponding to α, equal to 1.96

With these parameters, the required sample size was deemed to be 691. We assumed a non-response rate of 10%, resulting in a sample size of 768 individuals.

Since individuals from the same village have a similar risk of infestation, a design effect of 2 was applied to account for the clustering of individuals within villages. The total sample size was thus increased to 768 individuals.

7.3.24 Sampling

A multi-stage sampling technique was applied. First, province was stratified into four regions based on characteristics and lifestyles of the residents. The first region has a high proportion of Muslims and comprises three districts, namely; Chana, Thepha, and Saba Yoi. The risk of liver fluke among Muslim populations is low because of their diet. The second region comprises two districts, namely; Songkhla city and Hat Yai. These two regions have a high proportion of Buddhists and high immigration. The third region comprises five districts, namely; Ranot, Singhanakhon, Sathing Phra, Khuan Niang, Krasae Sin, and contain mostly rice farms, thus the residents have the highest risk of *O. viverrini* infestation. The last group comprises 6 districts, namely; Sadao, Rattaphum, Na Thawi, Na Mom, Bang Klam, Khlong Hoi Khong, and consists mainly of rubber plantations (Figure 8).

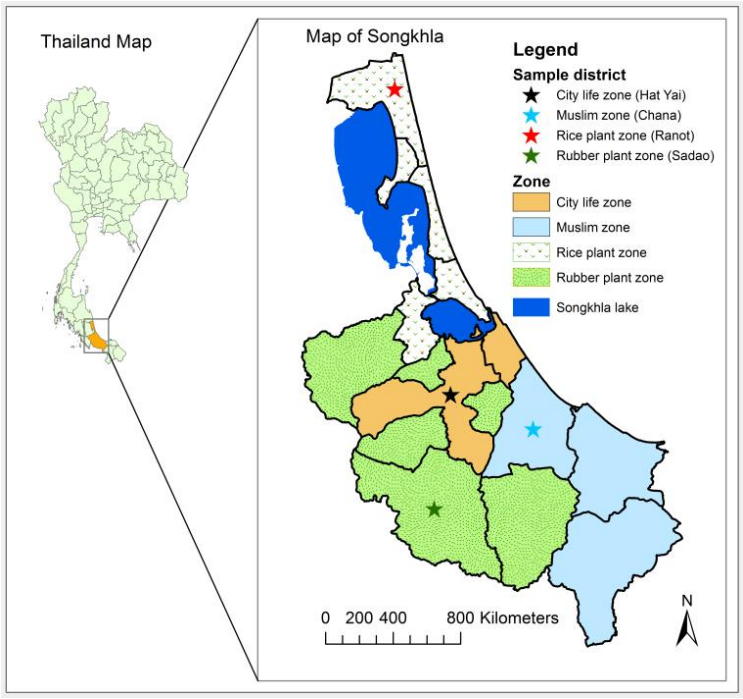

Figure 8 Map of Songkhla and study sites

Within each region, one district was selected using simple random sampling. Within each district, the sub-districts were selected using probability proportional to size. Within each selected sub-district, villages were selected using simple random sampling. Thai residents aged 20 years old or higher, registered and lived in Songkhla province for more than 3 months before the survey were eligible. Finally, 768 subjects (64 clusters, 12 persons each) from 52 villages, 33 sub-districts, and 4 districts were recruited.

7.3.25 Inclusion criteria

Thai residents aged 20 years old or higher, registered and having lived in Songkhla province for more than 3 months before the survey were eligible.

7.3.26 Exclusion criteria

Residents who were unable provide either serum or stool for parasitic infestation testing were excluded.

7.3.27 Data collection

The close-ended and rating scale questions were used in the questionnaire in Thai language (Annex 4). The questionnaire comprised of socio-demographic characteristics and related data such as history of illness, history of exposure to *O. viverrini*, cooking and dietary habits, alcohol consumption and history of travel to *O. viverrini* endemic regions. Participants were asked to report the frequency of eating specific foods such as raw fish, fermented fish and other uncooked food. Respondents could respond with never, sometimes, and often.

Data collection was done by well-trained interviewers. Stool (10 grams) and blood (5 mL) samples were collected from all participants.

7.3.28 Laboratory study

Sera specimens from subjects were examined for antibody to *O. viverrini* infestation using Enzyme-Linked Immunosorbent Assay (ELISA) test. The test was done by the Department of Parasitology, Faculty of Medicine, Khon Kaen University, Thailand. For feces examination, the Formalin Ether Concentration Technique (FECT), and Kato Katz's modified thick smear technique were done in parallel at the Office of Disease Prevention and Control 9, Phitsanulok province, Thailand. Those protocol as follows;

1) Enzyme-Linked Immunosorbent Assay

Indirect ELISA was used to measure *O.viverrini*-specific immunoglobulins in serum at the Department of Parasitology, Faculty of Medicine, Khon Kaen University, Thailand. The protocol has been described elsewhere.[70]

2) Formalin ether concentration technique (FECT)

In brief, FECT methods use centrifugation to concentrate the protozoa, helminth ova and larva in the bottom of the tube.[80] Ether was used as an extractor of debris and fat from the feces. Formalin acts both a fixative and preservative of protozoan eggs, larvae and cysts. The specific gravity of protozoan cysts and helminthes eggs was greater than that of water. Fecal debris was extracted into the ether phase so that the parasitic forms can be separated and then sediment by centrifugation. After the process of sediment preparation is finished, the sediment is examined by the following steps: prepare a saline mount by adding 1 drop sediment to 1 drop saline. Add a coverslip. Scan the entire area under the coverslip under low power (10×) objective for eggs and larvae. Then add 1 drop iodine to the edge of the coverslip to assist in the observation of cysts. Examine under the high-dry objective (40×-45×).

3) Kato Katz's modified thick smear technique

The modified Kato thick smear[81] was prepared and processed according to the method of Kato and Miura (1954). In brief, the glycerin-malachite green solution was mixed with 1 ml of 3% malachite green, 100 ml of 6% phenol and 100 ml of pure glycerin. The cellophane strips, each 22×40 mm, were soaked in this solution for at least 24 hours before use. Additionally, in order to eliminate fibers or seed, the technique was modified by pressing a 105-mesh stainless steel grid onto the sample which was then filtered, transferred to slides covered by the cellophane soaked cover slips and allowed to stand for 30 minutes. All preparations were initially screen with a low-power (10×) objective lens. Suspected parasitic objects were subsequently examined under a high-power (40×) objective. The stool samples were preserved in 10% formalin for later confirmation.

7.3.29 Data management

Questionnaires were double entry using Epidata software.

7.3.30 Data analysis

R language and environment[74] with survey package[82] was used for data analysis. Descriptive statistics were used to describe the overall prevalence of *O. viverrini* infestation and the prevalence by gender, age groups, occupation, knowledge on *O. viverrini*, habit of eating raw fish dishes, place where the consumption of raw fish took place, history of *O. viverrini* infestation and treatment, and intensity (EPG) from feces egg count. Each participant who was positive for Opisthorchis-like eggs was assumed to be infested with *O. viverrini*. Univariate analysis was performed to determine risk factors of *O. viverrini* infestation. Variables with a p-value ≤ 0.2 were included in the initial multivariate logistic regression modelling.

Chapter 8:Ethical consideration

The protocol was approved by the Ethical Committee, Faculty of Medicine, Prince of Songkhla University (REC number: 57-336-18-5, Annex 1).

CHAPTER 3

RESULTS

Chapter 9:Result of objective I

1 Part I: Histological types among cases having unknown histology of LBD cancer in Songkhla province from 1989 to 2013 using a multiple imputation technique

From 1989 to 2013, there were 2,387 LBD cancers in the Songkhla cancer registry. There was a high proportion of males (74.6%) and about half of the cases were aged between 50 and 69 years. Cases with unknown histology accounted for about 61% in both sexes; 64.9% in males, and 49.6% in females. As shown in Table 4, the proportion of HCC was higher in males (18.5%) than in females (13.1%) whereas CCA had a higher proportion in females (24.9%) compared to males (11.6%). LBD cancers with other known histology comprised about 7% in both sexes.

Table 4 Distribution of LBD cancer cases by histological type and sex, Songkhla province, Thailand, 1989-2013.

Histological type	All n=2,387	Male n=1,776	Female n=611	value
	n (%)	n (%)	n (%)	
HCC	409 (17.1)	329 (18.5)	80 (13.1)	0.
CCA	358 (15.0)	206 (11.6)	152 (24.9)	0.001

Other	165 (6.9)	89 (5.0)	76 (12.4)	0.001
Unknown	1,455 (61.0)	1,152 (64.9)	303 (49.6)	0.001

HCC: hepatocellular carcinoma, CCA: cholangiocarcinoma,

The strongest predictors for histological type among the cases with known histology from the multinomial logistic regression was sex, followed by year of diagnosis, age group, and district of residence. The estimated incidence and percentage of HCC, CCA and LBD cancers with other known histology after imputation are shown in Table 5. Compared to model 2, the percentage of HCC cases was a slightly higher in model 3 among both males (51.0% vs 50.9%) and females (24.7% vs 24.4%). However, the percentage of CCA cases was not different (34.1% vs 34.2% for males and 50.1% vs 50.2% for females). Model 2 and model 3 also gave similar results for LBD cancers with other histology.

Table 5 Distribution of histologic type of LBD cancer cases by sex after multiple imputation, Songkhla province, Thailand, 1989-2013.

	Model 2: LBD cancers with unknown histology only		Model 3: LBD cancers unknown histology + C76.2 and C80.9	
Histologic type	Male (n=1,152)	Female (n=303)	Male (n=1,323)	(
HCC				
Number (95% PI)	586 (552-620)	74 (64-84)	675 (639-711)	(8
Percent (95% PI)	50.9 (47.9-53.8)	24.4 (21-27.9)	51 (48.3-53.8)	(21.
CCA				

Number (95% PI)	394 (362-426)	152 (140-164)	451 (363-423)	(14
Percent (95% PI)	34.2 (31.4-37)	50.2 (46.2-54.2)	34.1 (31.5-36.7)	(46.
Other				
Number (95% PI)	172 (148-196)	77 (66-88)	197 (149-194)	
Percent (95% PI)	14.9 (12.8-17)	25.4 (21.9-28.9)	14.9 (12.9-16.8)	(21.

HCC: hepatocellular carcinoma, CCA: cholangiocarcinoma, C76.2: unknown primary in the abdomen, C80.9: unknown primary site, NOS. PI: Probability interval.

The average ASR of HCC in males throughout the observed years increased after imputation from 2.3/100,000 to 5.9 (252%) in model 2 and to 6.6 (281%) in model 3. The average ASR of HCC in females increased from 0.5/100,000 to 0.9 (180%) in model 2 and 1.1 (226%) in model 3. The average ASR of CCA among males with known histology increased from 1.2/100,000 to 3.6 (293%) in model 2 and 4.0 (330%) in model 3, respectively. The average ASRs per 100,000 of CCA in females increased from 0.8 to 1.5 and 1.9 in model 2 and model 3, respectively. The annual ASRs of HCC and CCA in both sexes and the three models (model 1 = complete case analysis) are shown in Figure 9.

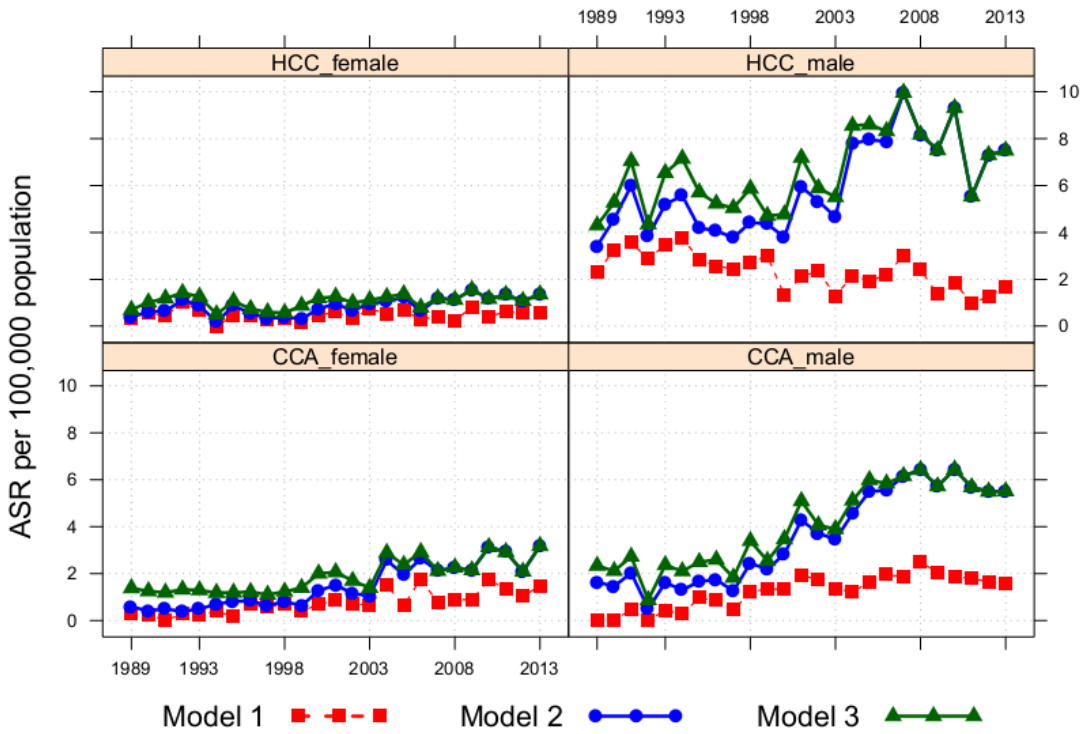

Figure 9 Age standardized incidence rates of LBD cancers in Songkhla from 1989 to 2013 for the two major histologic categories; HCC and CCA, stratified by sex. Model 1 = complete-case analysis, Model 2 = multiple imputation with LBD cancers with unknown histology only, model 3 = multiple imputation of LBD cancers with unknown histology combined with cancer of unknown primary in the abdomen (C76.2) and unknown primary site, NOS (C80.9).

2 Part II: Trend of HCC and CCA incidence from 1989 to 2013, and projected to 2030

From 1989 to 2013, there were 2,676 LBD cancer cases in the Songkhla cancer registry. The ratio of males to females was 3:1. Table 6 shows the distribution of LBD cancer cases by histological type and sex. HCC was the most common histological type overall (44.4%) followed by CCA (38.1%). Among males, HCC

shared a greater proportion (51.6%) than CCA (33.7%). Among females, CCA had a higher proportion (49.8%) than HCC (25.2%).

Table 6 Distribution of LBD cancer cases by histological types and sex, Songkhla province, Thailand, 1989-2013.

Histological type	All n=2,676	Male n=1,947	Female n=729	value
	n (%)	n (%)	n (%)	
HCC	1,188 (44.4)	1,004 (51.6)	184 (25.2)	<(
CCA	1,020 (38.1)	657 (33.7)	363 (49.8)	
Others	468 (17.5)	286 (14.7)	182 (25.0)	

HCC: hepatocellular carcinoma, CCA: cholangiocarcinoma

Among males, the highest proportion of cases with HCC was found in the 50-69 year age group (57%) while the highest proportion of CCA was found in those aged 50-79 years (67%). Among females, the highest proportion of HCC was found in the 60-79 year age group (51%) while the highest proportion of CCA was found among those aged 60-79 years (49%). With 5-year periods observed, the number of HCC cases fluctuated in both sexes while the number of CCA cases increased. The proportion of CCA in both sexes showed a rapid increase from 1999 to 2013. This period of time shared more than 80% of CCA cases in both sexes (Table 7).

Table 7 Number and percentage of two major histology types of LBD cancer cases; HCC and CCA by gender, Songkhla, Thailand, between 1989 and 2013 after imputation (n=2,676).

Variables	Male			
	HCC (n=1,004)		CCA (n=657)	
	Number	%	number	%
Age group (year)				
< 40	91	9.1	43	6.5
40-49	176	17.5	105	16.0
50-59	298	29.7	158	24.0
60-69	280	27.9	156	23.7
70-79	141	14.0	125	19.0
≥ 80	18	1.8	70	10.7
Period of diagnosis				
1989-1993	122	12.2	45	6.8
1994-1998	143	14.2	64	9.7
1999-2003	158	15.7	106	16.1
2004-2008	292	29.1	204	31.1
2009-2013	289	28.8	238	36.2

Table 8 shows the trends in incidence grouped by histological subtype and sex using Joinpoint regression. Among males, there were two clear trends for the incidence of HCC. From 1989-2007, the HCC incidence increased by 3.2% (1.3% − 5.1%) per year, then, a 2.8% (-9.8% − 4.8%) declining trend occurred after 2007. The incidence of CCA showed a significantly increasing trend of about 5.2% (3.8% − 6.6%) per year. Among females, the trend in incidence of HCC between 1989 and 2013 increased by 1.4% (0% − 2.7%) per year while the incidence of CCA significantly increased by 4.4% (3.1% − 5.7%) per year.

Table 8 LBD cancer incidence rates and trend in Songkhla, Thailand in 1989-2013 using Joinpoint regression

Gender	Histological subtypes	Trend	Period	Annual percent change

					(95% CI)
			Trend	1989-	3.2
Male	HCC	1		2007	(1.3,5.1)*
			Trend	2007-	-2.8 (-
		2		2013	9.8, 4.8)
			Trend	1998-	5.2
	CCA	1		2013	(3.8, 6.6)*
			Trend	1998-	1.4
Female	HCC	1		2013	(0.0, 2.7)
			Trend	1998-	4.4
	CCA	1		2013	(3.1, 5.7)*

Abbreviations: HCC = hepatocellular carcinoma, CCA = cholangiocarcinoma

* Annual percent change is significantly different from zero ($p < 0.05$)

Figure 9 shows the APC trend analysis for HCC and CCA by gender. According to the projected APC model, among males, age-distribution of next years of HCC cases was at increasing risk until age 75, and thereafter it decreased. A high rate ratio was evident in birth cohort 1945-1960 and declined in later cohorts, whereas the cohort effect was present during 2000-2007 and declined in later periods (Fig 10A). In contrast, the risk of CCA increased with age peaking at age 80-85. The AP-C model presented the risk of CCA in males increased after 1995. A high rate ratio presented in birth cohort 1945 and later (Fig 10B). Among females, the risk of HCC increased with age, peaking at age 75-80. A high rate ratio was evident in birth cohort 1950-1955 and declined in later cohorts. The effect of age-period model was low in this group (Fig 10C). The risk of CCA increased with age, peaking at age 80-85. A high rate ratio was evident in birth cohort 1950 and later (Fig 10D).

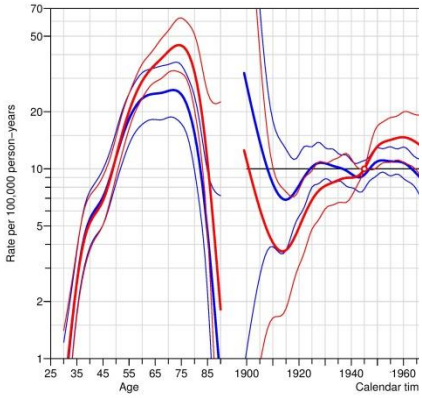

Fig 10A, male-HCC

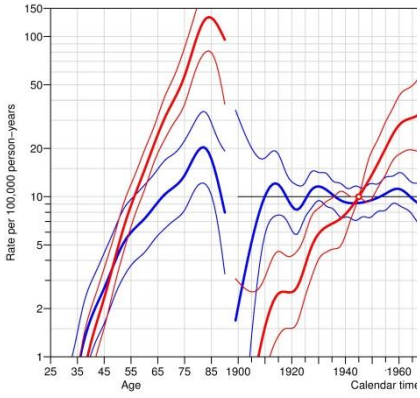

Fig 10B, male-CCA

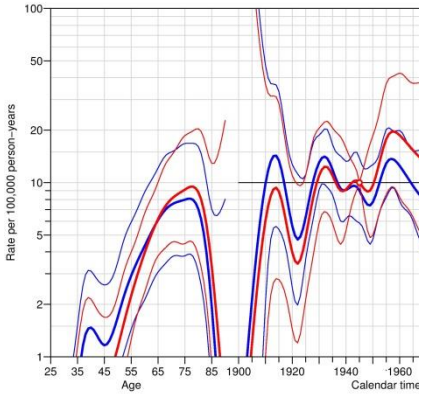

Fig 10C, female-HCC

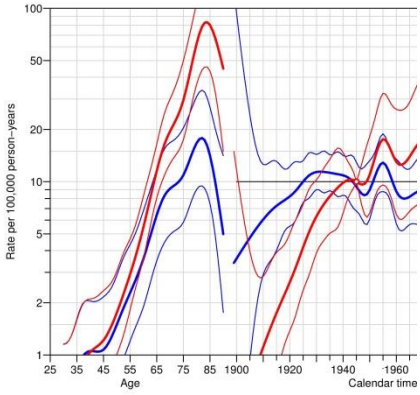

Fig 10D, female-CCA

Figure 10 Age-Period-Cohort trend analysis for HCC and CCA by gender

(Fig 10A= male-HCC, Fig 10B= male-CCA, Fig 10C= female-HCC, and Fig 10D= female-CCA which included Age-period: (AP-C) model (blue) and Age-cohort: AC-P (red) models. Incidence rates are plotted in the log scale (left y-axis))

Figures 11 and 12 show the age-standardized incidence rates of HCC and CCA using 3 projection models. Among males, using the Joinpoint model, HCC decreased and is expected to continue to decrease to 5.4 cases per 100,000 person-year or about 50 cases per year in 2030. Using the APC model the rate is expected to decrease to 6.0 cases per 100,000 or 48 cases per year, in 2030 while using the Nordpred model,

the rate is expected to decrease to 4.2 cases per 100,000 or 50 cases per year in 2030. The CCA incidence rates among males were increasing in all three models. For the Joinpoint model, the rate increased from 5.5 in 2013 to 9.4 cases per 100,000 or 113 cases per year, in 2030. Using the APC model, the rate is expected to increase to 7.6 cases per 100,000 or 90 cases per year, while using the Nordpred model, the rate is expected to increase to 8.2 cases per 100,000 or 101 cases per year in 2030.

Among females, using the Joinpoint models, the rate of HCC is quite stabled and is expected to increase slightly to 1.4 cases per 100,000 or 16 cases per year in 2030. Using the APC model, the rate is expected to decrease to 1.5 cases per 100,000 or 17 cases per year, in 2030 while using the Nordpred model, the rate is expected to decrease to 1.3 cases per 100,000 or 18 cases per year in 2030. Similarly to CCA in males, the CCA incidence rates among females are expected to increase in all three models. Using the Joinpoint model, the rate is expected to increase from 3.2 in 2013 to 3.9 cases per 100,000 or 55 cases per year, in 2030. Using the APC model, the rate is expected to increase to 3.3 cases per 100,000 or 43 cases per year, in 2030 while using the Nordpred model, the rate is expected to increase to 2.7 cases per 100,000 or 43 cases per year in 2030.

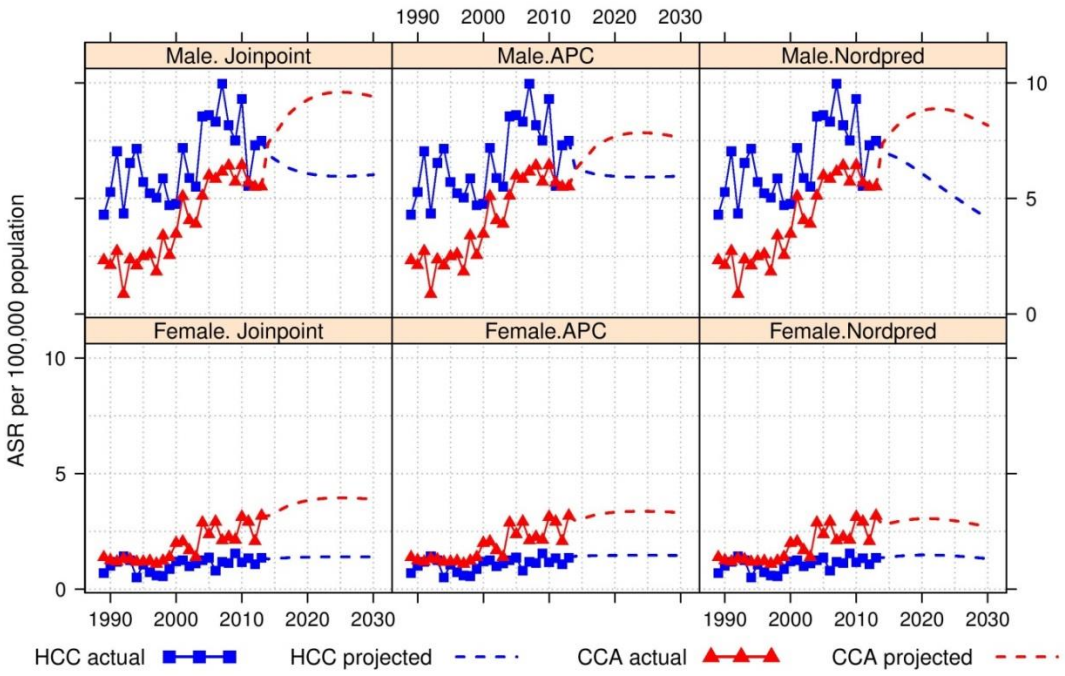

Figure 11 Age standardized incidence rates of LBD cancer in Songkhla by the major histologic categories; HCC and CCA using 3 projection models; Joinpoint analysis, Age-period-cohort (APC), and Nordpred by gender in 1989-2013 and projection until 2030.

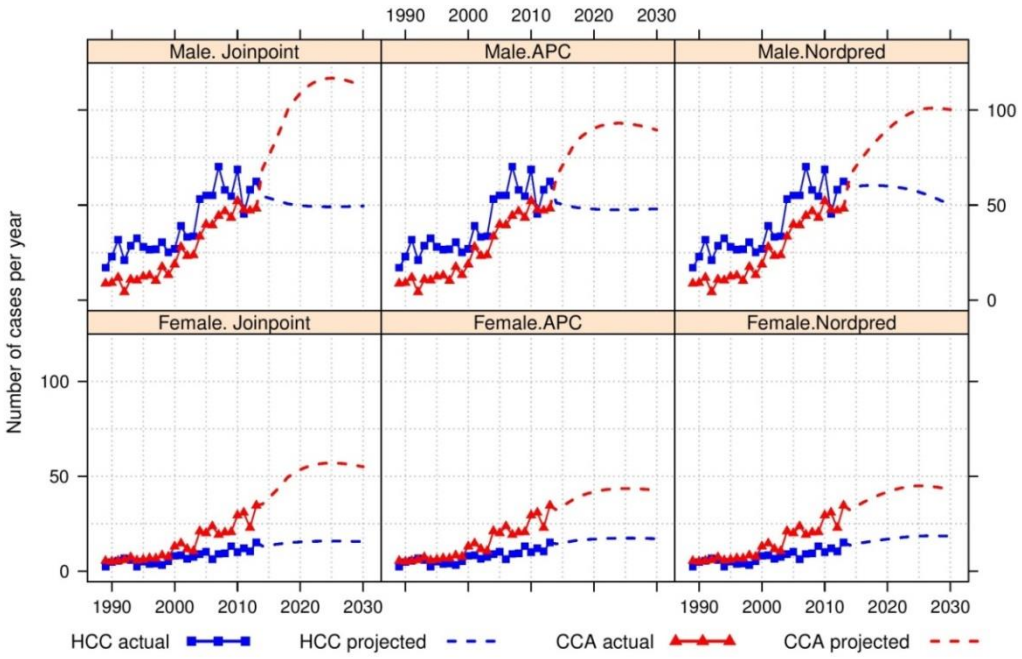

Figure 12 Estimated number of LBD cancer cases in Songkhla by the major histologic categories; HCC and CCA using 3 projection models; Joinpoint analysis, Age-period-cohort (APC), and Nordpred by sex in 1989-2013 and projection until 2030.

Chapter 10:Result of objective II : the prevalence of *O. viverrini* infestation and associated factors of parasite infestation among adult population in Songkhla

The crude prevalence of *O. viverrini* infestation was 3% (23/768) with 0.13% (1/768) stool samples containing *O. viverrini* eggs and 2.9% (22/768) was antibody positive to *O. viverrini* infestation, while the prevalence of *Trichuris trichiura* and *Hookworm* were 0.65% and 2.2%, respectively. After adjusting for age group and lifestyle zone (district) population, the prevalence of *O. viverrini* infestation decreased to 2.14% (0.05% *O. viverrini* egg from stool samples and 2.09% antibody positive to

O. viverrini infestation). The intensity of *O. viverrini* egg counts was 23 per gram feces while *Trichuris trichiura* and Hookworm were 83 and 234 egg per gram feces (EPG), respectively (Table 9).

Table 9 Crude and adjusted number and percentage of *O. viverrini* and others parasite infestation in Songkhla province, southern Thailand (n = 768, N=971,070).

Parasite	Crude number (%)	Estimated prevalence (95% CI)	Average (EPG)*
OV (egg)	1 (0.13)	0.05 (0.00-0.11)	23
OV (antibody)	22 (2.86)	2.09 (1.09-3.09)	-
OV (Overall)	23 (2.99)	2.14 (1.14-3.13)	23
Trichuris trichiura	5 (0.65)	0.79 (0.03-1.56)	83
Hookworm	17 (2.21)	2.40 (0.19-4.62)	234

* Egg per gram feces: Number of egg per gram feces using modified Kato's Katz method

The overall crude prevalence of *O. viverrini* infestation was rather higher in females (2.9%) than in males (1.2%). The highest proportion of *O. viverrini* positive cases was found in the 40-59 year age group (3.3%) followed by age 60 years or higher (2.8%).

The univariate analysis showed that religion, education, lifestyle, and birthplace region (Table 10) were significantly associated with *O. viverrini* infestation ($p<0.05$). Gender, age group, area type, marital status, income, Songkhla birthplace, history of travelling to another region, drinking alcohol, history of diseases with cross-reaction of antibody for *O. viverrini* infestation, history of eating raw fish dishes,

fermented fish dishes, and uncooked foods were not significantly associated with *O. viverrini* infestation.

Table 10 Crude number, adjusted number and proportion, and univariate analysis among potential risk factors and *O. viverrini* infestation in Songkhla province (n = 768).

Variable	Category	Crude number OV +ve (%)	Estimated %OV +ve (SE)	(95% C
Gender	Male	8 (1.04)	0.55 (0.21)	
	Female	15 (1.95)	1.59 (0.47)	(0.90-
Age group (years)	20-39	2 (0.26)	0.50 (0.40)	
	40-59	16 (2.08)	1.17 (0.29)	(0.60-
	60+	5 (0.65)	0.46 (0.19)	(0.40-1
Area type	Urban	13 (1.69)	1.15 (0.48)	
	Rural	10 (1.30)	0.99 (0.33)	(0.40-
Religion	Muslim	1 (0.13)	0.06 (0.07)	
	Non-Muslim	22 (2.86)	2.08 (0.50)	.0 (1.20-8
Marital status	Single	2 (0.26)	0.15 (0.10)	
	Married	18 (2.34)	1.68 (0.51)	(0.80-1
	Divorced/widowed	3 (0.39)	0.31 (0.18)	(0.80-2
Education	Primary school	14 (1.82)	1.47 (0.46)	
	High school	7 (0.91)	0.45 (0.16)	(0.10-
	Vocational college	1 (0.13)	0.06 (0.07)	(0-

			Bachelor or higher	1 (0.13)	0.16 (0.16)	(0.10-
	Occupation	Officer	Gov./business	1 (0.13)	0.16 (0.16)	
			Business Owner	6 (0.78)	0.4 (0.13)	(0.10-1
			Freelance/Other	6 (0.78)	0.62 (0.38)	(0.10-1
			Rice-Farmer	4 (0.52)	0.55 (0.31)	(0.30-3
			Rubber tapper	6 (0.78)	0.41 (0.19)	(0.10-
	Income (Thai-Baht)		<=4000	2 (0.26)	0.25 (0.18)	
			4,001- 6,000	5 (0.65)	0.44 (0.27)	(0.30-1
			6,001- 9,000	5 (0.65)	0.31 (0.12)	(0.30-
			9,001-12,000	6 (0.78)	0.78 (0.40)	(0.70-2
			>12,000	5 (0.65)	0.35 (0.12)	(0.40-1
	District zone		City zone	14 (1.82)	0.91 (0.33)	
			Muslim zone	0 (0)	0 (0)	
			Rice farm	6 (0.78)	0.97 (0.43)	(0.7(
			Rubber plantation	3 (0.39)	0.25 (0.18)	(0.10-
	Songkhla birthplace		Yes	19 (2.47)	1.58 (0.41)	
			No	4 (0.52)	0.55 (0.35)	(0.50-
	Birthplace region	Non-endemic for *O. viverrini*		19 (2.47)	1.58 (0.41)	
		O. viverrini - endemic area		4 (0.52)	0.55 (0.35)	(2.50-4
	History of regional migration		No	19 (2.47)	1.52 (0.33)	

	Yes	4 (0.52)	0.61 (0.42)	(0.10-
Alcohol consumption	No	14 (1.82)	1.55 (0.50)	
	Yes	9 (1.17)	0.59 (0.18)	(0.20-
Use anti-anthelmintic drugs	No	18 (2.34)	1.82 (0.50)	
	Yes	5 (0.65)	0.32 (0.14)	(0.90-
History of past infection from other disease cross-reaction with OV	No	18 (2.34)	1.82 (0.50)	
	Yes	5 (0.65)	0.32 (0.14)	(0.50-
Eats raw fish dishes (1)	No	18 (2.34)	0.31 (0.18)	
	Yes	5 (0.65)	1.82 (0.53)	(2.80-
Eats fermented fish dishes (2)	No	11 (1.43)	0.32 (0.12)	
	Yes	12 (1.56)	1 (0.34)	(6.(
Eats uncooked foods (3)	No	14 (1.82)	1.13 (0.46)	
	Yes	9 (1.17)	1.46 (0.50)	(4.50-
Eats (1) and (2) and (3) (CCA-risk)	No	18 (2.34)	1.82 (0.53)	
	Yes	5 (0.65)	0.68 (0.23)	(0.60-

Non-endemic region = south and central region, Thailand, Endemic region = Northern and north region, Thailand

Cross-reaction effect: the disease that can present cross-reaction effect for *O. viverrini* antibody positive, including malaria, parasite infestation (i.e. infested trematodes, nematodes, cestodes, and protozoa)

Table 11 shows point estimate of multivariate logistic regression associated risk factors for *O. viverrini* infestation in Songkhla, Thailand. Having birthplace in an *O. viverrini* endemic region (northeast and north region) in Thailand, living zone, and education were the potential factors associated with *O. viverrini infestation*.

Those born in *O. viverrini* endemic regions (northeast and north) in Thailand had odds of being infested with *O. viverrini* that was 5 (1.75-14.25) times higher than who were born in the south and central region. Those who graduated from high school had odds of being infested with *O. viverrini* that was 0.42 (0.2-0.89) in comparison to those educated in primary school only. Those who lived in a Muslim zone had an odds of *O. viverrini* infestation of 0 ($p < 0.01$) in comparison to those who lived in another zone.

Table 11 Point estimate (SE) of multivariate logistic regression of associated risk factors for *O. viverrini* infestation in Songkhla, Thailand

Variable	categories	OR (95% CI)	P-value
Gender (ref: male)	Female	1.31 (0.43-3.98)	0.63
Religion (ref: Muslim)	non-Muslim	2.83 (0.4 -19.87)	0.30
Education (ref: primary school)	High school	0.42 (0.2-0.89)	0.03*
	Under graduate	0.19 (0.02-1.94)	0.17
	Graduate or higher	0.52 (0.07-3.73)	0.52
Lifestyle zone (ref: City-HatYai)	Muslim zone	0 (0-0)	<0.01*
	Rice plant zone	1.41 (0.46-4.3)	0.55

	Rubber plant zone	0.52 (0.11-2.42)	0.41
Birthplace region (ref: non *O. viverrini*-endemic region)	*O. viverrini* -endemic region	5.00 (1.75-14.25)	0.01*
Moved to another region (ref: no)	Yes	0.52 (0.15-1.81)	0.31
Drinking alcohol (ref: no)	Yes	0.72 (0.24-2.2)	0.56

* Statistic significant with p-value < 0.05

Table 12 Frequency and percentage of risk food items contacts in Songkhla population, Thailand

ategories	Local name of food items	I on't know food items (%)	N ever (%)	S ometimes (%)	ften	ean	.D.
	Raw fish salad	47 (32.2)	67 (60.8)	1 (6.6)	(0.4)	.75	.59
(1)	Minced fish salad	58 (33.6)	40 (57.3)	6 (8.6)	(0.5)	.76	.62
(2)	Minced fish salad	53 (32.9)	85 (63.2)	7 (3.5)	(0.4)	.71	.55
(3)	Minced fish salad	47 (19.1)	84 (76.0)	5 (4.6)	(0.3)	.86	.48
	Raw fish uncooked	4 (12.2)	58 (85.7)	4 (1.8)	(0.3)	.90	.38
intestine spicy dip	Fermented fish	33 (43.4)	17 (54.3)	6 (2.1)	(0.3)	.59	.55
spicy dip	Fermented fish	31 (30.1)	74 (61.7)	3 (6.9)	0 (1.3)	.79	.62

	Minced spicy pickled fish	152 (19.8)	550 (71.6)	54 (7)	2 (1.6)	.90	.57
	Papaya salad with fermented fish	66 (8.6)	459 (59.8)	213 (27.7)	0 (3.9)	.27	.67
	Fermented fish	255 (33.2)	475 (61.8)	34 (4.4)	(0.5)	.72	.57
	Fermented fish with rice	214 (27.9)	524 (68.2)	28 (3.6)	(0.3)	.76	.52
	Fermented fish with rice	90 (11.7)	655 (85.3)	22 (2.9)	(0.1)	.91	.38
	Spicy shrimp salad	109 (14.2)	550 (71.6)	100 (13)	(1.2)	.01	.57
	Raw pork salad	70 (9.1)	656 (85.4)	37 (4.8)	(0.7)	.97	.41
	Raw beef salad	63 (8.2)	664 (86.5)	36 (4.7)	(0.7)	.98	.39
	Other uncooked menu	80 (10.4)	640 (83.3)	47 (6.1)	(0.1)	.96	.41
	Average	166.4 (21.7)	543.6 (70.8)	52.1 (6.8)	.9 (0.8)	.87	.52

Categories: 1= raw-fish dishes, 2= fermented-fish dishes, 3= uncooked dishes

Unknown= participant does not know about that kind of food, Never= participant does know the food but never taste, Sometimes= less than 6 times/year, Often= 6 times/year or higher

Most (70.8%) respondents knew food items which were popular in the northeast and northern regions which are the dishes known to be contaminated with *O. viverrini* metacercariae but had never ate these dishes. One-fifth of participants never knew about these dishes while less than 10% had ever eaten these dishes (Table 12).

Table 13 shows factors associated with *O. viverrini* infestation. Religion, income, birthplace region, history of moving from Songkhla province to live in other regions, and drinking alcohol were the potential risk factors associated with *O. viverrini* infestation. Having income 4,001-6,000 baht/month and never moved to

another province, the odds of "no-risk" applying versus "low risk", "medium risk", or "high risk" applying combined are about 50% lower. Buddhist/other religion, the odds of "no-risk" applying versus "low risk", "medium risk", or "high risk" applying is 2.64 times greater than Muslim. Having birthplace in north or northeast region Thailand, the odds of "no-risk" applying versus "low risk", "medium risk", or "high risk" applying is 23.6 (7.1-77.9) times higher than south or central region. Drinking alcohol, the odds of "no-risk" applying versus "low risk", "medium risk", or "high risk" applying is 1.4 times greater than who do not drink.

Table 13 Ordinal logistic regression analysis for risk of eating behaviour for *O. viverrini* infestation

Variables	categories	OR (95% CI)
Religion (ref: Muslim)	Buddhist/Other	2.6 (1.6-4.4)*
Income (ref: <=4000 Baht)	4,001- 6,000	0.5 (0.30-0.82)*
	6,001- 9,000	1.2 (0.6-2.2)
	9,001-12,000	0.7 (0.4-1.4)
	>12,000	0.7 (0.4-1.4)
Songkhla birthplace (ref: Yes)	No (migrant)	1.3 (0.7-2.4)
Birthplace region (ref: south/central)	North/northeast	23.6 (7.1-77.9)*
Moved to another region (ref: Yes)	No	0.5 (0.3-0.9)*
Drinking alcohol (ref: no)	Yes	1.4 (1.0-1.9)*

No-risk| low risk	0.03 (0.01-0.1)*
Low risk| medium risk	1.2 (0.9-1.4)
Medium risk| high risk	227.5 (200.5-258.2)*

Estimated prevalence and number of O. viverrini infestation

Based on sensitivity (97.1%) and specificity (14.6%) of ELISA technique[70] we estimated the possible prevalence of *O. viverrini* infestation from positive predictive value (PPV) of the confirmed laboratory examination, sensitivity, and specificity of *O. viverrini* infestation in Songkhla used the following formula:

$$\text{PPV} = \frac{\text{sensitivity} \times \text{prevalence}}{(\text{sensitivity} \times \text{prevalence} + ((1 - \text{specificity}) \times (1 - \text{prevalence})))}$$

Where PPV = 0.0214 (0.016-0.026)

Sensitivity = 0.971

Specificity = 0.146, thus

$$\mathbf{0.0214} = \frac{\mathbf{0.971x}}{(\mathbf{0.971x} + ((\mathbf{1} - \mathbf{0.146}) * (\mathbf{1} - \mathbf{x})))}$$

The estimated prevalence of *O. viverrini* infestation using microscopic test was 0.05% (0.00-0.11%) while using ELISA test was 1.89% (1.44%-2.34%). Thus, the final prevalence of *O. viverrini* infestation in Songkhla province was 1.94% (1.44%-2.45%).

CHAPTER 4

DISCUSSION

Chapter 11:Incidence of cholangiocarcinoma in Songkhla province

1 Multiple imputation for unknown histology type of LBD cancers

This study used a population-based cancer registry database which has a good quality of data and provided 24 years of continuous records. However, a serious limitation was the high rate of incomplete or missing histological type. To reduce this limitation, the missing histological types were imputed using a multiple imputation method and multinomial logistic regression. Results of the multiple imputation depend on the distribution of the predictive factors and their relative risk ratios from the multinomial regression.

The findings demonstrated a change in ASR of the two major histological sub-types of LBD cancer throughout the study period. Based on Model 1 - the complete cases analysis - a decline in HCC incidence among males after 2007 was not evident. We would expect to see a decline in the incidence since the nationwide program of HBV immunization to all newborns in Songkhla was initiated in 1991[83] and a large proportion of children and adults were immunized both before and after 1991. We would also expect to see a slight decline in the incidence of HCC well before 2007 as tests for HBV and HCV infections have routinely been performed in blood donors since 1985[84]. Such a decline after 2007 was not observed among females in any of the

three models. As model 3 gave a rather stable trend in incidence of HCC in both sexes before 2007, it seems to give a better prediction capability than model 2. In other words, including a random number of unknown primary cases prior to the imputation process may have been justified. The U-shaped decline in HCC incidence during 1995-2005 as seen in model 2 reflects the changes in diagnostic methods of LBD cancer during this time period.

Multiple imputation performs well when the responses are missing at random (MAR). However, the assumption of MAR cannot usually be verified.[28] Multiple imputation methods can avoid bias only if enough predictor variables are included in the imputation process.[85] One study demonstrated that the multiple imputation method works well when the percentage of missing values is between 10 and 60%.[32] In this study, the percentage of cases with missing histology ranged from 60% to 67%. When the number of cases in the dataset is reasonably high, the MICE method for a binomial outcome gives low variation of coefficients.[86] However, there has been no study investigating how well multiple imputation works for multinomial outcomes.

There was no solid evidence of misclassification bias evident among the cases with unknown histology. The chance of death among all histologic types is theoretically non-differential.[87] Most of the cases with unknown histological type were diagnosed by death certificate or clinical investigation and all of the major sub-types of LBD cancers do not differ in their prognosis and survival chances.[44] In addition, clinicians may not perform cytology and/or biopsy for reasons mainly due to the performance status of patients and their compliance, especially in controlling for intraperitonial bleeding after a biopsy.[88] It is possible that the patients with jaundice who are likely to be CCA rather than HCC have a higher chance of being investigated by imaging. However, there is no reason that this phenomenon would affect the multiple imputation process.

In Songkhla, the hepatitis B vaccine has been included in the natinoal expanded program of immunization since 1991.[83] The prevalence of *O. viverrini* has been very low in the southern region of Thailand.[16, 21] Results based on models 2 and 3 showed that the incidence of HCC among males started to decrease in 2007, which is 16 years after the incorporation of hepatitis B vaccine into the national immunization program. Such a phenomenon is consistent with trends from Taiwan.[89] However, the decrease was not observed among females in which the rates were much lower. In addition, *O. viverrini* infestations have not increased in the southern Thai population[13], thus the continuous increase in incidence of CCA in Songkhla province during the past two decades cannot be explained by *O. viverrini* infestations. The increase is more likely to be due to the increased facilities for diagnosing LBD cancers and also the real increase in incidence which has also been observed in the US and Australia.[9, 10] However, the estimated ASRs among males and females in Songkhla, as well as the rate of increase in incidence rates, was much higher than the rates in the United States and Australia (around 1 per 100,000).

The method used in surveys of *O. viverrini* infestation is the stool egg count with formalin-ethyl acetate concentration technique, which can be negative in mild parasite infestations and in those who had an infestation in the past without reinfestation.[90] Eating raw fish is not the dietary habit of people in the southern region of Thailand. Immigrants to Songkhla from those who reside in regions with a high prevalence of *O. viverrini* infestation and residents of Songkhla who visited regions with a high prevalence of *O. viverrini* infestation can be tested negative with this technique even if they are infested. Thus, it is possible that these people were exposed to *O. viverrini* in the past but were negative for stool *O. viverrini* egg count. Surveys utilising newly developed techniques that have high sensitivity and specificity and can detect past infestations are needed to confirm the true proportion of people who were ever exposed to *O. viverrini* and therefore explain the increasing trend in incidence of CCA in Songkhla province.

2 Trend and trend projection of cholangiocarcinoma in Songkhla

This study presents trends in LBD cancer incidence in Songkhla province by histological subtypes. The rates have increased significantly since 1989, seemingly due to a combination of changes in demographics and the risk profile of the population. The model projections were presented agreeable evident that the burden of CCA in Songkhla has been rising and is expected to continue to 2030 while the incidence of HCC has decreased. The potential expected number of HCC and CCA cases for males in 2030 were about 49 (48–51) and 101 (90-113) cases per year, respectively. The potential expected rate of HCC and CCA for females in 2030 were about 17 (16-19) and 47 (43-55) cases per year, respectively. However, the magnitude of the increase could be affected by future healthcare planning and other cancer control programs in southern Thailand.[37, 91, 92]

The study showed that the older age group had a higher incidence rate of both HCC and CCA, contributing significantly to the overall trend for both genders. However, only HCC in males seems to have the increasing of age-specific incidence in younger age (50-59) than others. Thailand is facing a demographic change which has led to an epidemiological transition. The predominance of infectious diseases is shifting to non-communicable diseases. For example, the average life expectancy is expected to increase for males from 77.5 years in 2010 to 80.8 years in 2030 and for females from 70.4 years in 2010 to 74.4 years in 2030.[35] According to the changing age demographics, the percentage of southern Thais aged 50 and over increased from 15.3% in 1989 to 24.5% and then to 35.9% in 2013 and 2030. [35, 93] There is considerable evidence that the risk of LBD cancer for both HCC and CCA is highest for those aged 50 and over.

The trend of LBD cancer in Songkhla is similar to the international trends in Europe, America, Oceania, and Asia which in the last decade (2004-2013) is

likely to increase more than before.[94] However, Asian countries, for example, China, Taiwan and Singapore, provide evidence of effective HCC prevention after implementation of universal HBV immunization.[95] This study demonstrated that the incidence of HCC in males in southern Thailand is decreasing while HCC in females is stable. The ASR in the since 2007 in both genders seems relatively low when compared to the ASR of HCC in the whole country which were 30.3 and 13.1 per 100,000 person-year in males and females, respectively.[16] This is likely a consequence of the effectiveness of comprehensive cancer prevention and control measures implemented in Thailand. Firstly, the effective control measures in two major modes of transmission of HBV, namely blood transfusion and sexual contact. HBV screening of donated blood in Thailand has been conducted for many years, and the seroprevalence of hepatitis B surface antigen was recently reported at around 5%.[96-98] The HBV immunization programme in Thailand was implemented in the early 1990s, thus a gradual decline in the incidence of HCC is anticipated in this decade. The evaluation of the HBV vaccination program in Bangkok showed a high prevalence of HBV immunity after 20-years vaccination[91] in Thai children and 1.2% of natural HBV infection was found in the area that some children had never received vaccination.[99] HCC incidence was significantly lower in Thai children who received the hepatitis B vaccine at birth.[100] Secondly, effective government policies and campaigns for prohibiting alcohol consumption are very important measures to prevent HCC. This can slow the increasing rate of alcohol consumption. Thus, the percentage of alcohol drinkers in 2013 decreased and the proportion of drinkers by gender did not differ from the previous decade. Males still play a role as a high risk group with a 5-fold higher risk than females, even though the percentage of new drinkers among females is higher.[101] Similar stable incidence rates have been reported in Germany[102] while many countries have reported that alcohol and HBV infections might have contributed to the increasing trend in LBD cancer incidence rates in males more than females.[46, 103, 104]

The trends in incidence of CCA in Thailand are largely influenced by the liver fluke, *O. viverrini*.[7, 17] From our study, the trend of CCA in Songkhla has been

increasing whereas a declining trend of LBD cancer has been observed in endemic areas, such as the northeastern region Thailand. Those reflected shading of successful in controlling *O. viverrini* infestation in the past decade.[105] The increasing trend of liver cancer may indicate an increase in CCA, since people from northeastern regions migrate to work in, and carry *O. viverrini* to, other parts of the country where *O. viverrini* control is not intensive. Although many people from endemic areas also migrate to Songkhla, in the southern region, the prevalence of *O. viverrini* infestation in Songkhla in the last survey from Ministry of Public Health, Thailand was still low.[65] The increasing rate either relates to the *O. viverrini* infestation in the past or other risk factors.

To estimate the future trend of LBD cancer in Southern Thailand, the projection was done using three different methods. Each method has their own strengths and weaknesses; thus, it was decided to include all of them with making the conclusions based on trends from various methods. All three methods demonstrated an increase in the projected incidence rate of CCA through to 2030 in both genders whereas the incidence of HCC was predicted to either remain stable (methods 1 and 2) or decrease (method 3). With the similarly of the trends, the Joinpoint predictions made with 1-year periods, the age-period-cohort made with 1-year periods, and the aggregated 5-year periods using Nordpred age-period-cohort model with spline method for 1-years periods were aggregated. The potential reaching rate of HCC and CCA for males in 2030 about 4.1–6.0 and 7.6-9.4 cases per 100,000 person-year, respectively whereas the potential reaching rate of HCC and CCA for females in 2030 about 1.3-1.4 and 2.7-3.9 cases per 100,000 person-year, respectively.

The application of the APC model was illustrated for the LBD cancer data in Figure 9. The ASRs over time calculated using the observed rates are nearly identical to the ASRs calculated using the APC-fitted rates (not shown). Apart from age effects, “period” significantly affected the trend of HCC and CCA incidence in both genders during 2004-2008 and dipped slightly before rising again at a lower rate.The

LBD cancer trends in Songkhla are influenced by period effects, cohort effects or both since the APC effects are combined with the highly correlated predictor variables. However, the pointwise confidence intervals for the fitted rates were substantially narrower over the study period and the fitted values of the age-period model indicated a better fit than the age-cohort model. The results showed that both period and cohort effects are both relevant in shaping the trends and so both models were presented and projections made with each one (Figures 10A-10D). However, consistent with reports from other countries[106], there are indications of a stronger cohort effect. Small sample sizes in the youngest and oldest cohorts and before 2004 may affect the level of significance.

This study provides the first report on the trend of HCC and CCA in southern Thailand. Because there are differences in etiologic agents and risk factors between HCC and CCA and changing risk profiles of the population in this region, the study needed to explore some these risk factors, i.e. prevalence of liver fluke infestation, factors associated with cholangiocarcinoma. Future studies should focus on risk factors of CCA in southern Thailand to direct appropriate strategies to control the burden of LBD cancer in this region.

Chapter 12:Prevalence and associated factors for *Opisthorchis viverrini* infestation in Songkhla province

An overall prevalence of 2.14% (1.6-2.6%) of *O. viverrini* infestations was found in the investigated population in Songkhla province, southern Thailand, which is considered one of the lowest prevalence among all Thai provinces.[62] This study found a higher prevalence of *O. viverrini* infestation in Songklha when compared to the last survey conducted by the Thai CDC in 2009.[65] However, the detection of cases using the ELISA technique was highly sensitive for detecting *O. viverrini*

infestation than the other previous study which only investigated the presence of *O. viverrini* eggs in stool samples.[70]

No association was found between *O. viverrini* infestation and history of eating raw fish dishes. Raw fish dishes are not commonly eaten in Southern Thailand. People who migrated from the north and northeast regions of Thailand, which are endemic for *O. viverrini*, and then settled in Songkhla, are prone to keep their old cooking styles and eating behaviours during visits to their hometowns, where they have opportunity to consume the raw or insufficiently cooked fish.

Although this study found a higher prevalence of *O. viverrini* infestation in females compared to males, no association was found between *O. viverrini* infestation and gender, a result in contrast with other studies in *O. viverrini* endemic areas which found that males had a similar or slightly higher prevalence of infestation than females.[13, 45, 72]

About 90% of respondents never ate food items that are likely to be contaminated with *O. viverrini* metacercariae while about 10% of respondents were at risk of CCA because they had eaten both raw fish and fermented fish dishes.

Many studies have found an association between *O. viverrini* infestation and CCA using microscopy techniques such as Kato's Katz and FECT; however, those techniques can only detect current infestations. Our study used both the ELISA technique, which is sensitive for detecting past infestations, and microscopy (FECT and Kato's Katz) method for current infestation, which gives more accuracy and can better explain the prevalence of CCA, a cancer that is well-established within 7 to 15 years after infestation of *O. viverrini*[13], which the severity needed both intensity and duration of infestation.[108] The ELISA technique is an updated method with high sensitivity to detect *O. viverrini* infestation.[70] It was possible that the antibody positive to *O. viverrini* was disturbed from other diseases, particularly from other parasitic infestations[69] and possibly leading to an overestimation of the infestation. However, the analysis by

subgroup of who had/no experience to contact the diseases which can affect to *O. viverrini* antibody, there was no cross-reaction effect found in this study.

The prevalence of *O. viverrini* infestation from stool samples in this study was low. Due to many step are needed in the *O. viverrini* life cycle, included both primary and secondary intermediate hosts, it difficult to transmission in the non-endemic area.[109] However, an outbreak of *O. felineus* in central Italy [63] occurred because of extensive traveling and importation of freshwater fish from Asia; thus it is now increasingly common to find infested people in non-endemic areas.[54]

Similar to other studies,[21] there was low evidence to support liver fluke transmission in the southern Thai region since there was a low prevalence of *O. viverrini* infestation from stool samples.

This study may potentially have bias for leading to an overestimation of *O. viverrini* infestation since we used the ELISA method for detecting the *O. viverrini* infested cases which is high sensitive measurement and *O. viverrini* infestation antibody have the cross-reaction with other intestinal parasites. However, the study shows the potential reaching prevalence and number of *O. viverrini* infestation after adjusting for the sensitivity and specificity of the test with the *O. viverrini* prevalence, thus, the expected numbers of *O. viverrini* infested are believed to be close to the real situation.

Chapter 13:Strengths and limitations

1 Strengths of the study

This study used data from a cancer registry which has been collecting cancer cases over a 25-year period. The quality of the data in the registry is also high, making it suitable for model prediction. Cancer registry database is good for unique

characteristic which can explain trend in incidence and generalized the finding to people in local population.

2 Limitations of the study

13.2.1 The type of missing data of LBD cancer histology in the Songkhla cancer registry could not be ascertained.

13.2.2 High sensitivity and cross-reaction of the ELISA test to other parasitic infestations might affect the expected number and prevalence of *O. viverrini* infestations.

Chapter 14:Conclusions

In Songkhla province, the incidence of HCC has been decreasing among males since 2007 while the incidence of CCA has been continuously increasing. The effect of hepatitis B vaccination in newborns in reduction of HCC incidence was demonstrated. Although multiple imputation can provide more accurate estimates of incidence rates and the trends in incidence of LBD cancer, verification of missing at random is needed by reviewing the radiographic images. With the review, details on subtypes of CCA can also be summarized. There is also a need to conduct case control studies to elucidate the role of *O. viverrini* infestations and other risk factors of CCA.

Our study suggests that in Songkhla province, the number of cases of LBD cancer will increase. The incidence of CCA has been increasing and takes more affect to rate and number of LBD cancer cases than HCC in both genders. Effective and comprehensive cancer prevention and control measures might play important roles in the reduction of HCC incidence. Since liver fluke infestation is a major cause of CCA in Thailand, prevalence of liver fluke infestation and other associated factors in this area should be further investigated.

The study found a low prevalence of *O. viverrini* infestation in Songkhla province. Lack of correlation between CCA incidence and the *O. viverrini* infestation means we may not be able to explain the increasing trend of CCA in this region. Regional migration may play an important role in *O. viverrini* infestation since people who are born in endemic regions (north and northeast) and moved to settle in Songkhla are more likely to be infested by *O. viverrini* than who are the localized. However, there is a lack of knowledge related to the prevalence of *O. viverrini* in the environment such as in fish, snails, and mammals. The major cause of CCA in southern Thailand might be due to others factors which need further study.

REFERENCES

1. GLOBOCAN 2012 (IARC). Liver Cancer Estimated Incidence, Mortality and Prevalence Worldwide in 2012 [Internet]. Section of Cancer Surveillance; 2004 [cited 2015 April, 10]; Available from: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx?cancer=liver.

2. Srivatanakul P, Sriplung H, Deerasamee S. Epidemiology of liver cancer: an overview. Asian Pac J Cancer Prev. 2004 Apr-Jun;5(2):118-25.

3. World Health Organization. Hepatitis B fact sheet [Internet]. July, 2015 [cited 2016 May, 10]; Available from: http://www.who.int/mediacentre/factsheets/fs204/en.

4. Chiveerawattana A. Guidelines for screening, diagnosis and treatment of liver cancer and bile duct. In: National Cancer Institute. Deapartment of Medicine MoPH, Thailand, editor.: National Cancer Institute; 2011.

5. CDC. Parasites - Opisthorchis Infection. 2012 [updated 2012, January 10; cited 2014, September 1]; Available from: http://www.cdc.gov/parasites/opisthorchis/.

6. Bragazzi MC, Cardinale V, Carpino G, Venere R, Semeraro R, Gentile R, et al. Cholangiocarcinoma: Epidemiology and risk factors. Translational Gastrointestinal Cancer. 2011;1(1):21-32.

7. Khuhaprema T. SP, Sriplung H.,Wiangnon S., Sumitsawan Y., Attasara P. Cancer in Thailand, 2004-2006.2011.

8. Khuhaprema T. SP, Sriplung H.,Wiangnon S., Sumitsawan Y., Attasara P. Cancer in Thailand, 2007-2009.2014.

9. Altekruse SF, Petrick JL, Rolin AI, Cuccinelli JE, Zou Z, Tatalovich Z, et al. Geographic variation of intrahepatic cholangiocarcinoma, extrahepatic cholangiocarcinoma, and hepatocellular carcinoma in the United States. PLoS One. 2015;10(3):e0120574.

10. Luke C, Price T, Roder D. Epidemiology of cancer of the liver and intrahepatic bile ducts in an Australian population. Asian Pac J Cancer Prev. 2010;11(6):1479-85.

11. Ghouri YA, Mian I, Blechacz B. Cancer review: Cholangiocarcinoma. J Carcinog. 2015;14:1.

12. Cardinale V, Semeraro R, Torrice A, Gatto M, Napoli C, Bragazzi MC, et al. Intra-hepatic and extra-hepatic cholangiocarcinoma: New insight into epidemiology and risk factors. World J Gastrointest Oncol. 2010 Nov 15;2(11):407-16.

13. Sithithaworn P, Yongvanit P, Duenngai K, Kiatsopit N, Pairojkul C. Roles of liver fluke infection as risk factor for cholangiocarcinoma. J Hepatobiliary Pancreat Sci. 2014 May;21(5):301-8.

14. Song HN, Go SI, Lee WS, Kim Y, Choi HJ, Lee US, et al. Population-Based Regional Cancer Incidence in Korea: Comparison between Urban and Rural Areas. Cancer Res Treat. 2016 Apr;48(2):789-97.

15. Wiangnon S, Kamsa-ard S, Suwanrungruang K, Promthet S, Mahaweerawat S, Khuntikeo N. Trends in incidence of hepatocellular carcinoma, 1990-2009, Khon Kaen, Thailand. Asian Pac J Cancer Prev. 2012;13(3):1065-8.

16. Kamsa-ard S, Wiangnon S, Suwanrungruang K, Promthet S, Khuntikeo N, Mahaweerawat S. Trends in liver cancer incidence between 1985 and 2009, Khon Kaen, Thailand: cholangiocarcinoma. Asian Pac J Cancer Prev. 2011;12(9):2209-13.

17. Khuhaprema T. SP, Sriplung H.,Wiangnon S., Sumitsawan Y., Attasara P. Cancer in Thailand, 2001-2003.2008.

18. Sripa B, Kaewkes S, Sithithaworn P, Mairiang E, Laha T, Smout M, et al. Liver fluke induces cholangiocarcinoma. PLoS Med. 2007 Jul;4(7):e201.

19. Srivatanakul P. Epidemiology of Liver Cancer in Thailand. Asian Pac J Cancer Prev. 2001;2(2):117-21.

20. Sripa B, Pairojkul C. Cholangiocarcinoma: lessons from Thailand. Curr Opin Gastroenterol. 2008 May;24(3):349-56.

21. Jongsuksuntigul P, Imsomboon T. Opisthorchiasis control in Thailand. Acta Trop. 2003 Nov;88(3):229-32.

22. Sriamporn S, Pisani P, Pipitgool V, Suwanrungruang K, Kamsa-ard S, Parkin DM. Prevalence of Opisthorchis viverrini infection and incidence of cholangiocarcinoma in Khon Kaen, Northeast Thailand. Trop Med Int Health. 2004 May;9(5):588-94.

23. T. Khuhaprema PS, H. Sriplung, S. Wiangnon, Y. Sumitsawan, P. Attasara. Cancer in Thailand, 1998-2000.2005.

24. White IR, Daniel R, Royston P. Avoiding bias due to perfect prediction in multiple imputation of incomplete categorical variables. Comput Stat Data Anal. 2010 Oct 1;54(10):2267-75.

25. Van Buuren SaG-O, C. G. M. mice: Multivariate imputation by chained equations in R. J of Statistical Software. 2011;45(3):1-67.

26. He Y, Yucel R, Zaslavsky AM. Misreporting, Missing Data, and Multiple Imputation: Improving Accuracy of Cancer Registry Databases. Chance (N Y). 2008 Sep;21(3):55-8.

27. Parkin DM. The evolution of the population-based cancer registry. Nat Rev Cancer. 2006 Aug;6(8):603-12.

28. Cummings P. Missing data and multiple imputation. JAMA Pediatr. 2013 Jul;167(7):656-61.

29. Lee KJ, Simpson JA. Introduction to multiple imputation for dealing with missing data. Respirology. 2014 Feb;19(2):162-7.

30. Chevret S, Seaman S, Resche-Rigon M. Multiple imputation: a mature approach to dealing with missing data. Intensive Care Med. 2015 Feb;41(2):348-50.

31. White IR, Daniel R, Royston P. Avoiding bias due to perfect prediction in multiple imputation of incomplete categorical variables. Computational Statistics & Data Analysis. 2010;54(10):2267-75.

32. Barzi F, Woodward M. Imputations of missing values in practice: results from imputations of serum cholesterol in 28 cohort studies. Am J Epidemiol. 2004 Jul 1;160(1):34-45.

33. Jin Hyuk Lee and John Huber Jr. Multiple imputation with large proportions of missing data: How much is too much? [Internet]. Texas A&M Health Science Center; 2011 [cited 2015 Feb, 1]; Available from: https://ideas.repec.org/p/boc/usug11/23.html.

34. Holford TR. The estimation of age, period and cohort effects for vital rates. Biometrics. 1983 Jun;39(2):311-24.

35. Moller B, Fekjaer H, Hakulinen T, Sigvaldason H, Storm HH, Talback M, et al. Prediction of cancer incidence in the Nordic countries: empirical comparison of different approaches. Stat Med. 2003 Sep 15;22(17):2751-66.

36. Zhang Z. Monte Carlo based statistical power analysis for mediation models: methods and software. Behav Res Methods. 2014 Dec;46(4):1184-98.

37. Moller B, Fekjaer H, Hakulinen T, Tryggvadottir L, Storm HH, Talback M, et al. Prediction of cancer incidence in the Nordic countries up to the year 2020. Eur J Cancer Prev. 2002 Jun;11 Suppl 1:S1-96.

38. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. Stat Med. 2000 Feb 15;19(3):335-51.

39. Clayton D, Schifflers E. Models for temporal variation in cancer rates. II: Age-period-cohort models. Stat Med. 1987 Jun;6(4):469-81.

40. Xie L, Semenciw R, Mery L. Cancer incidence in Canada: trends and projections (1983-2032). Health Promot Chronic Dis Prev Can. 2015 Spring;35 Suppl 1:2-186.

41. Olsen AH, Parkin DM, Sasieni P. Cancer mortality in the United Kingdom: projections to the year 2025. Br J Cancer. 2008 Nov 4;99(9):1549-54.

42. Moller H, Fairley L, Coupland V, Okello C, Green M, Forman D, et al. The future burden of cancer in England: incidence and numbers of new patients in 2020. Br J Cancer. 2007 May 7;96(9):1484-8.

43. Coupland VH, Okello C, Davies EA, Bray F, Moller H. The future burden of cancer in London compared with England. J Public Health (Oxf). 2010 Mar;32(1):83-9.

44. Bjerregaard JK, Mortensen MB, Pfeiffer P. Trends in cancer of the liver, gall bladder, bile duct, and pancreas in elderly in Denmark, 1980-2012. Acta Oncol. 2016;55 Suppl 1:40-5.

45. Tzeng IS, Lee WC. Forecasting hepatocellular carcinoma mortality in Taiwan using an age-period-cohort model. Asia Pac J Public Health. 2015 Mar;27(2):NP65-73.

46. Zhang Y, Ren JS, Shi JF, Li N, Wang YT, Qu C, et al. International trends in primary liver cancer incidence from 1973 to 2007. BMC Cancer. 2015;15:94.

47. de Groen PC, Gores GJ, LaRusso NF, Gunderson LL, Nagorney DM. Biliary tract cancers. N Engl J Med. 1999 Oct 28;341(18):1368-78.

48. Tyson GL, El-Serag HB. Risk factors for cholangiocarcinoma. Hepatology. 2011 Jul;54(1):173-84.

49. Rerknimitr R, Angsuwatcharakon P, Ratanachu-ek T, Khor CJ, Ponnudurai R, Moon JH, et al. Asia-Pacific consensus recommendations for endoscopic and interventional management of hilar cholangiocarcinoma. J Gastroenterol Hepatol. 2013 Apr;28(4):593-607.

50. Ben-Menachem T. Risk factors for cholangiocarcinoma. Eur J Gastroenterol Hepatol. 2007 Aug;19(8):615-7.

51. Okuda K, Nakanuma Y, Miyazaki M. Cholangiocarcinoma: recent progress. Part 1: epidemiology and etiology. J Gastroenterol Hepatol. 2002 Oct;17(10):1049-55.

52. Kubo S, Kinoshita H, Hirohashi K, Hamba H. Hepatolithiasis associated with cholangiocarcinoma. World J Surg. 1995 Jul-Aug;19(4):637-41.

53. Chen MF, Jan YY, Jeng LB, Hwang TL, Wang CS, Chen SC, et al. Intrahepatic cholangiocarcinoma in Taiwan. J Hepatobiliary Pancreat Surg. 1999;6(2):136-41.

54. Yossepowitch O, Gotesman T, Assous M, Marva E, Zimlichman R, Dan M. Opisthorchiasis from imported raw fish. Emerg Infect Dis. 2004 Dec;10(12):2122-6.

55. Forrer A, Sayasone S, Vounatsou P, Vonghachack Y, Bouakhasith D, Vogt S, et al. Spatial distribution of, and risk factors for, Opisthorchis viverrini infection in southern Lao PDR. PLoS Negl Trop Dis. 2012;6(2):e1481.

56. Yoon HJ, Ki M, Eom K, Yong TS, Chai JY, Min DY, et al. Risk Factors for Opisthorchis viverrini and Minute Intestinal Fluke Infections in Lao PDR, 2009-2011. Am J Trop Med Hyg. 2014 Aug 6;91(2):384-8.

57. IARC. Infection with liver flukes (Opisthorchis viverrini, Opisthorchis felineus and Clonorchis sinensis). IARC Monogr Eval Carcinog Risks Hum. 1994;61:121-75.

58. IARC. A review of human carcinogens--Part B: biological agents. Lyon: World Health Organization, International Agency for Research on Cancer. 2012;Volume 100B 341-65.

59. Wykoff DE, Harinasuta C, Juttijudata P, Winn MM. Opisthorchis Viverrini in Thailand--the Life Cycle and Comparison with O. Felineus. J Parasitol. 1965 Apr;51:207-14.

60. Keiser J, Utzinger J. Food-borne trematodiases. Clin Microbiol Rev. 2009 Jul;22(3):466-83.

61. Sripa B, Kaewkes S, Intapan PM, Maleewong W, Brindley PJ. Food-borne trematodiases in Southeast Asia epidemiology, pathology, clinical manifestation and control. Adv Parasitol. 2010;72:305-50.

62. Sithithaworn P, Andrews RH, Nguyen VD, Wongsaroj T, Sinuon M, Odermatt P, et al. The current status of opisthorchiasis and clonorchiasis in the Mekong Basin. Parasitol Int. 2012 Mar;61(1):10-6.

63. Pozio E, Armignacco O, Ferri F, Gomez Morales MA. Opisthorchis felineus, an emerging infection in Italy and its implication for the European Union. Acta Trop. 2013 Apr;126(1):54-62.

64. Yankawin O. Strategies of liver fluke prevention and screening in Thailand. In: Control THCPBoGCDDoD, editor.: Bereau of General Communicable Disease.; 2013.

65. Wongsaroj T. A study of Helminthiasis and Protozoan Diseases in Thailand, 2009: The Helminth Control Program, Bereau of General Communicable Disease., Deapartment of Disease Control MoPH, Thailand.;2009.

66. Jongsuksuntigul P, Imsomboon T. Epidemiology of opisthorchiasis and national control program in Thailand. Southeast Asian J Trop Med Public Health. 1998 Jun;29(2):327-32.

67. MoPH T. E-san agenda: "Liver fluke reduction and CCA elimination" In: Khon Kaen University NSO, and MoPH, Thailand, Ministry of Public Health; 2012.

68. Jamornthanyawat N. The diagnosis of human opisthorchiasis. Southeast Asian J Trop Med Public Health. 2002;33 Suppl 3:86-91.

69. Wongsaroj T, Sakolvaree Y, Chaicumpa W, Maleewong W, Kitikoon V, Tapchaisri P, et al. Affinity purified oval antigen for diagnosis of Opisthorchiasis viverrini. Asian Pac J Allergy Immunol. 2001 Dec;19(4):245-58.

70. Sawangsoda P, Sithithaworn J, Tesana S, Pinlaor S, Boonmars T, Mairiang E, et al. Diagnostic values of parasite-specific antibody detections in saliva and urine in comparison with serum in opisthorchiasis. Parasitol Int. 2012 Mar;61(1):196-202.

71. Sakolvaree Y, Ybanez L, Chaicumpa W. Parasites elicited cross-reacting antibodies to Opisthorchis viverrini. Asian Pac J Allergy Immunol. 1997 Jun;15(2):115-22.

72. Kaewpitoon SJ, Rujirakul R, Kaewpitoon N. Prevalence of Opisthorchis viverrini infection in Nakhon Ratchasima province, Northeast Thailand. Asian Pac J Cancer Prev. 2012;13(10):5245-9.

73. Bray F, Ferlay J, Laversanne M, Brewster DH, Gombe Mbalawa C, Kohler B, et al. Cancer Incidence in Five Continents: Inclusion criteria, highlights from Volume X and the global status of cancer registration. Int J Cancer. 2015 Nov 1;137(9):2060-71.

74. R Development Core Team. R: A language and environment for statistical computing. In: Computing RFfS, editor. Vienna, Austria: ISBN 3-900051-07-0; 2011.

75. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011 Feb 20;30(4):377-99.

76. Gorlitz L, Gao Z, Schmitt W. Statistical analysis of chemical transformation kinetics using Markov-Chain Monte Carlo methods. Environ Sci Technol. 2011 May 15;45(10):4429-37.

77. National Cancer Institute. Surveillance Research: Joinpoint Regression Program. In: Institute NC, editor. United States2014.

78. Crawford JR, Garthwaite PH. Methods of testing for a deficit in single-case studies: Evaluation of statistical power by Monte Carlo simulation. Cogn Neuropsychol. 2006 Sep;23(6):877-904.

79. Carstensen B. Age-period-cohort models for the Lexis diagram. Stat Med. 2007 Jul 10;26(15):3018-45.

80. Anamnart W, Intapan PM, Maleewong W. Modified formalin-ether concentration technique for diagnosis of human strongyloidiasis. Korean J Parasitol. 2013 Dec;51(6):743-5.

81. WHO. Field tools - Measuring the severity of disease. 2008 [updated 2008, Febuary; cited 2014, September 1]; Available from: http://www.who.int/neglected_diseases/preventive_chemotherapy/pctnewsletter11.pdf?ua=1.

82. Carriere GM, Sanmartin C, Bryant H, Lockwood G. Rates of cancer incidence across terciles of the foreign-born population in Canada from 2001-2006. Can J Public Health. 2013 Nov-Dec;104(7):e443-9.

83. Chub-uppakarn S, Panichart P, Theamboonlers A, Poovorawan Y. Impact of the hepatitis B mass vaccination program in the southern part of Thailand. Southeast Asian J Trop Med Public Health. 1998 Sep;29(3):464-8.

84. Chimparlee N, Oota S, Phikulsod S, Tangkijvanich P, Poovorawan Y. Hepatitis B and hepatitis C virus in Thai blood donors. Southeast Asian J Trop Med Public Health. 2011 May;42(3):609-15.

85. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393.

86. J. Hardt et al. Multiple Imputation of Missing Data: A Simulation Study on a Binary Response. Open Journal of Statistics. 2013(3):370-8.

87. The International Agency for Research on Cancer. WHO Classification of Tumours of the Digestive System (IARC WHO Classification of Tumours) 4th ed. 2010:1-418.

88. Chhieng DC. Fine needle aspiration biopsy of liver - an update. World J Surg Oncol. 2004;2:5.

89. Hung GY, Horng JL, Yen HJ, Lee CY, Lin LY. Changing incidence patterns of hepatocellular carcinoma among age groups in Taiwan. J Hepatol. 2015 Dec;63(6):1390-6.

90. Worasith C, Kamamia C, Yakovleva A, Duenngai K, Wangboon C, Sithithaworn J, et al. Advances in the Diagnosis of Human Opisthorchiasis: Development of Opisthorchis viverrini Antigen Detection in Urine. PLoS Negl Trop Dis. 2015;9(10):e0004157.

91. Sanchez-Ramirez DC, Colquhoun A, Parker S, Randall J, Svenson LW, Voaklander D. Cancer incidence and mortality among the Metis population of Alberta, Canada. Int J Circumpolar Health. 2016;75:30059.

92. Poovorawan Y, Chongsrisawat V, Theamboonlers A, Leroux-Roels G, Kuriyakose S, Leyssen M, et al. Evidence of protection against clinical and chronic

hepatitis B infection 20 years after infant vaccination in a high endemicity region. J Viral Hepat. 2011 May;18(5):369-75.

93. Gunsoy NB, Garcia-Closas M, Moss SM. Estimating breast cancer mortality reduction and overdiagnosis due to screening for different strategies in the United Kingdom. Br J Cancer. 2014 May 13;110(10):2412-9.

94. Nishioka M, Kan D, Nawata J, Takemoto T, Fukuda T, Miyaji T. [Clinicopathological studies in patients with hepatocellular carcinoma. III. Clinical application of prognosis forecasting by computer (author's transl)]. Rinsho Byori. 1980 Mar;28(3):285-7.

95. Chang ET, Nguyen BH, So SK. Attitudes toward hepatitis B and liver cancer prevention among Chinese Americans in the San Francisco Bay Area, California. Asian Pac J Cancer Prev. 2008 Oct-Dec;9(4):605-13.

96. Sont WN, Zielinski JM, Ashmore JP, Jiang H, Krewski D, Fair ME, et al. First analysis of cancer incidence and occupational radiation exposure based on the National Dose Registry of Canada. Am J Epidemiol. 2001 Feb 15;153(4):309-18.

97. Luksamijarulkul P, Thammata N, Tiloklurs M. Seroprevalence of hepatitis B, hepatitis C and human immunodeficiency virus among blood donors, Phitsanulok Regional Blood Center, Thailand. Southeast Asian J Trop Med Public Health. 2002 Jun;33(2):272-9.

98. Corsten MJ, Hearn M, McDonald JT, Johnson-Obaseki S. Incidence of differentiated thyroid cancer in Canada by City of residence. J Otolaryngol Head Neck Surg. 2015;44:36.

99. Anderson DL. Oral cancer incidence and mortality in Canada and abroad. J Can Dent Assoc (Tor). 1969 Apr;35(4):192-7.

100. Wichajarn K, Kosalaraksa P, Wiangnon S. Incidence of hepatocellular carcinoma in children in Khon Kaen before and after national hepatitis B vaccine program. Asian Pac J Cancer Prev. 2008 Jul-Sep;9(3):507-9.

101. Gaudette LA, Holmes TM, Laing LM, Morgan K, Grace MG. Cancer incidence in a religious isolate of Alberta, Canada, 1953-74. J Natl Cancer Inst. 1978 Jun;60(6):1233-8.

102. Penner DW. Cancer incidence studies in Canada, with emphasis on a study in the Province of Manitoba. Pathol Microbiol (Basel). 1961;24:707-10.

103. Ladep NG, Khan SA, Crossey MM, Thillainayagam AV, Taylor-Robinson SD, Toledano MB. Incidence and mortality of primary liver cancer in England and Wales: changing patterns and ethnic variations. World J Gastroenterol. 2014 Feb 14;20(6):1544-53.

104. Palmer EP. The incidence of cancer among the Indians of the United States and Canada with specific reference to Arizona. Acta Unio Int Contra Cancrum. 1953;9(2):373-91.

105. Sripa B, Tangkawattana S, Laha T, Kaewkes S, Mallory FF, Smith JF, et al. Toward integrated opisthorchiasis control in Northeast Thailand: The Lawa project. Acta Trop. 2014 Aug 4.

106. Westlake S. Cancer incidence and mortality in the United Kingdom and constituent countries, 2003-05. Health Stat Q. 2008 Winter(40):91-7.

107. Dao TT, Bui TV, Abatih EN, Gabriel S, Nguyen TT, Huynh QH, et al. Opisthorchis viverrini infections and associated risk factors in a lowland area of Binh Dinh Province, Central Vietnam. Acta Trop. 2016 May;157:151-7.

108. Kaewpitoon N, Kaewpitoon SJ, Pengsaa P, Sripa B. Opisthorchis viverrini: the carcinogenic human liver fluke. World J Gastroenterol. 2008 Feb 7;14(5):666-74.

109. Sithithaworn P, Haswell-Elkins M. Epidemiology of Opisthorchis viverrini. Acta Trop. 2003 Nov;88(3):187-94.

ANNEXES

Annex 1: Approved ethical consideration

AF/18-03/01.1

Faculty of Medicine, Prince of Songkla University

This document is to certify that

REC Number:	57-336-18-5		
Project entitled:	Prevalence of *Opisthorchis viverrini* infestation and trend in incidence of cholangiocarcinoma in Songkhla population		
Principal Investigator:	Mr. Seesai Yeesoonsang	**Affiliation**	Epidemiology Unit, Faculty of Medicine, Prince of Songkla University
Sub-investigator:	Assoc.Prof. Hutcha Sriplung	**Affiliation**	Epidemiology Unit, Faculty of Medicine, Prince of Songkla University

Document acceptance:

1. Submission form version 1.0 date January 9, 2015
2. Study protocol version 1.0 date January 9, 2015
3. Participant Information Sheet version 1.0 date January 9, 2015
4. Informed Consent Form version 1.0 date January 9, 2015
3. Clinical record form version date 1.0 date January 9, 2015
4. Curriculum Vitae

have been reviewed by the Research Ethics Committee is in full compliance with the Declaration of Helsinki and the International Conference on Harmonization in Good Clinical Practice (ICH-GCP) Guidelines. Please submit the progress report every 12 months.

.....

(Assoc.Prof. Boonsin Tangtrakulwanich, M.D.)
Chairman of Research Ethics Committee

Date of approval: January 13, 2015
Date of expiration: January 12, 2016

Office of Human Research Ethics Committee
Faculty of Medicine Prince of Songkla University
15 Karnjanavanit Rd. Hat YaiSongkhla 90110
Tel. +66 7445-1149, +667445-1157
Fax 66 7421-2900

Annex 2: Data abstraction form for LBD cancer cases from Songkhla cancer registry

Number	Variable	Meaning
Var01	RegNo	Recode ID
Var02	Sex	Gender
Var03	Age	Age
Var04	Rel	Religion
Var05	Top	Topography
Var06	Mor	Morphology
Var07	Beh	Behavior
Var08	Grad	Grade of patient
Var09	Stag	Stage
Var10	Ext	Extend
Var11	I10	Code of ICD 10

Var12	Basis	Basis
Var13	AddCode	Address code
Var14	BirthD	Birthdate
Var15	DiagD	Date of diagnosis
Var16	DLS	Date of last seen
Var17	Stat	Patient status

Annex 3: Information sheet and consent form

Information Sheet

1. Title of Protocol: Incidence of cholangiocarcinoma and prevalence of *Opisthorchis Viverrini* infestation in Songkhla, Southern Thailand
2. Name and title of principle investigator: Mr.Seesai Yeesoonsang
3. Office address: Epidemiology unit, Faculty of medicine, Prince of Songkla University

    - Telephone Number Office:
    - Mobile: 66 89 439 1677
    - Fax: 66 7442 9754

4. Essential Information for volunteers include:

    4.1. Rationale for research

Cholangiocarcinoma is an important public health concern in East Asian countries especially Thailand which was the highest incidence occurred. However, highest incidence of CCA and *O. viverrini* infestation are in the northeast and northern region of Thailand. The increasing of cholangiocarcinoma patients in Songkhla, province, southern region of Thailand is unusual.

Our study aims to examine prevalence of *O. viverrini* infestation, and identify factors associated to infestation among Songkhla population, and estimated incidence and trend of cholangiocarcinoma in Songkhla population-based cancer registry.

    4.2. Purpose of research

1) To estimated trend incidence of cholangiocarcinoma in Songkhla population-based cancer registry.

2) To survey prevalence *O. viverrini* infestation and identify factors associated to cholangiocarcinoma in Songkhla province.

4.3. Methods of the research

Since *O. viverrini* has been proof for the risk of choloangiocarcinoma in Thailand, the prevalence of *O. viverrini* infested will be examine by antibody testing to present the situation of *O. viverrini* infestation in Songkhla population. The association of prevalence of *O. viverrini* infestation and the risk behaviour will be determined. The study design will be cross-sectional study including serology test for *O. viverrini* infestation. All participants will be used for calculation of prevalence whereas the comparison of case and non-case will be used to identify risk of *O. viverrini* infestation. The risk factor of main interest in this study is living history and eating behavior because most of patients occurred in non-endemic are usually are imported cases. Therefor we expect to find the prevalence of *O. viverrini* positive from antibody testing among person who has risk behaviour such as eating raw fish higher than who have no experience. In addition, the clinical profiles of patient can support the information of CCA patient in study area that differ or similar in endemic area. Whereas situation of CCA patient in the past, present and future in Songkhla province will be estimate and model for support cancer control program officer.

4.4. The expected duration of the subject's participation

Each subject will have not more than 4 month participation in the study.

4.5. Total number of subject's participation: 768 participants.

4.6. The benefits that might reasonably be expected to result to subject or to others as an outcome of the research.

Participants will have benefit to know their parasite infested status if they have any parasite infested he will have opportunity to be treated by get anthelmintic drug during the participation in the research.

4.7. Cancer registry record review

Cancer registry record will be extracted. The information of patient including demographic data, visited date, diagnosed date, and status will be collected.

4.8. Prevention of anticipated risks and measures prepared to cope with problem.

The information sheet and the informed consent will be explained by researcher. The subjects will have time for decision to participation. The subjects are free to participate and free to stop if they feel uncomfortable. Hospital medical record and cancer registry record of the consent data-authorized will be extracted. The information of patient, underlying disease, treatment, and patient status will be collected.

4.9. The extent to which confidentiality of records in which the subject is identified will be maintained.

The researcher will maintain the confidentiality of data, especially with respect to the information about the participant which would otherwise be known only to the researcher but would now be available to the entire research team.

4.10. The anticipate prorated payment, if any, to the subject.

No payment to subjects.

4.11. Withdrawal from the research

The patient is free to refuse to participate and will be free to withdraw from the research at any time without penalty or loss of benefits to which he or she would otherwise be entitled.

4.12. Name, address, telephone numbers of physicians or other research team to be contacted conveniently 24 hrs both during on or out of duty in case of need or emergency.

Name ……

Address……

……

Telephone No.……

4.13. Request for of subjects' rights and benefits

The subjects can request for rights and benefits at the Ethical committee, Ministry of Public Health, Thailand

Telephone number: …….

Fax Number: ……...

Inform consent

Title of Protocol: Cholangiocarcinoma, trend in incidence, and prevalence of *Opisthorchis viverrini* infestation in Songkhla population, Southern Thailand

Date of consent……

(Day/month/year)

Before I consent to participate of the study, I have accurately read out the information sheet and understand the whole methodology of the study.

I was given an opportunity to ask questions about the study, and all the questions have been answered correctly. I confirm that I have not been forced into giving consent, and the consent has been given freely and voluntarily.

I am voluntary and include the right to withdraw from the study without prior notice. The withdrawal from the study will not be affected to any other service or therapy which regardless of participation.

The research team will maintain the confidentiality of data, especially with respect to the information about the participant which would otherwise be known only to the physician but would now be available to the entire research team.

Researcher contact name with 24 hrs telephone number ……..……

Researcher contact name with 24 hrs telephone number ……..……

Signature of Researcher /person taking the consent

Date..……/……/……

(Day/month/year)

I have read the foregoing information, or it has been read to me. I have had the opportunity to ask questions about it and any questions that I have asked have been

answered to my satisfaction. I consent voluntarily to participate as a participant in this research.

Signature of Participant……

Date..……/……/……
(Day/month/year)

Signature of Researcher...…….

Signature of Witness……

Signature of Witness……..……

If illiterate, I have witnessed the accurate reading of the consent form to the potential participant, and the individual has had the opportunity to ask questions. I confirm that the individual has given consent freely.

Signature of participant……

Date..……/……/……
(Day/month/year)

Signature of Researcher...…….

Signature of Witness……

Signature of Witness……..……

Signature of parent, family, relative, guardian, or a person appointed to act as an aforesaid status by Thai court.

Date..……/……/……

(Day/month/year)

Signature of Researcher...…….

Signature of Witness……

Signature of Witness……

If non conscious, the participant must be consented from parent, family, relative, guardian, or a person appointed to act as an aforesaid status by Thai court.

Signature of parent, family, relative, guardian, or a person appointed to act as an aforesaid status by Thai court……

Date..……/……/……

(Day/month/year)

Signature of Researcher...…….

Signature of Witness……

Signature of Witness……

Annex 4: Questionnaire for Participant (English and Thai version)

Title: Prevalence of *Opisthorchis viverrini* infestation and trend in incidence of cholangiocarcinoma in Songkhla population, 2015

The study aimed to determine the prevalence of *O. viverrini* infestation, risk behaviour of *O. viverrini* infestation and trend of cholangiocarcinoma in the Songkhla population-based cancer registry in year 2015. The researchers would like to ask for your participation to answer this questionnaire. If there are any question that you not willing or feel uncomfortable to answer, you can ignore that question. There will be no impact to you at all. Your information will be kept confidential and will be presented in the research aggregate without identification of the respondents. The purposes of this study are for academic research, planning, and prevention and control of cholangiocarcinoma in Songkhla.

The questionnaire composts of 5 sections.

Section 1: General information of respondent	12	questions
Section 2: Birthplace and migration history	2	questions
Section 3: History of illness and risk behavior	10	questions
Section 4: Eating behavior	19	questions
Section 5: Information of laboratory specimen	2	questions

If you have any question, comments, or suggestion related to this research, please feel free to contact the researcher at Epidemiology Unit, Faculty of Medicine, Prince of Songkla University, Telephone: 66 7445 1165, Fax: 66 7442 9754

1. Mr Seesai Yeesoonsang Project Leader Mobile phone no. 089-439-1677
2. Mrs. Nussara Sirichote co-researcher Mobile phone no. 080-539-0270

3. Mr. Preecha Noofong co-researcher Mobile phone no. 081-957-5009
4. Assoc. Prof. Hucha Sriprung, M.D. Project consultant

1. ID : [ ][ ][ ][ ]	ID [ ][ ][ ][ ]
2. Address: Village…..……. Moo ..... ❒❒	Moo [ ] [ ]
Sub-District..... District.....	Sub_D [ ] [ ]
	Distr [ ] [ ]
3. Area type: ❒ (1) Municipal area ❒ (2) Local authority administration area	Area_typ [ ]
4. Date of interview: ❒❒/ ❒❒/ 2015	Indate [ ][ ][ ] [ ]
Section 1: General information of respondent	
1. Gender ❒ (1) Male ❒(2) Female	Gender [ ]
Tel no.(if available): _____	
2. Age [ ] [ ] year	Age [ ] [ ]
3. Weight (Kg.) [ ] [ ] [ ]	Weigh [ ] [ ] [ ]
4. Height (cm.) [ ] [ ] [ ]	Heigh [ ] [ ] [ ]
5. Religion	Religion [ ]
❒ (0) none ❒ (1) Buddhist ❒ (2) Muslim	
❒ (3) Christianity ❒ (4) Other	
6. Marital Status	Marital [ ]
❒ (1) Single ❒ (2) Married	
❒ (3) Widowed ❒ (4) Divorced/Separated	
7. Education	Educa [ ]
❒ (1) Primary or lower ❒ (2) Secondary school	
❒ (3) Diploma or equal ❒ (4) Bachelor	
❒ (5) Higher Bachelor	
8. Main occupation [no: …] Alt-occupation (if available) [no: …]	Occ1 [ ]
❒ (1) rice farm ❒ (2) rubber plant	Occ2 [ ]
❒ (3) Oth agriculture ❒ (4) Business owner	
❒ (5) Employee ❒ (6) Government officer	
❒ (7) Labor ❒ (8) Others (Pls. indicated.....)	
9. Monthly income (Baht) [ ][ ][ ],[ ][ ][ ]	Income
	[ ][ ][ ][ ][ ][ ]

ID □□□□

10. 1) Number of HH member [ ] [ ] person, — HH_memb [ ][ ]
    2) Number of HH member age >= 20 yrs. [ ] [ ] peron — HH_20 [ ][ ]
11. Born in Songkhla province — SK_pop [ ]
    ❒ (1) Yes ❒ (2) No
12. How long have you lived in Songkhla — SK_live_yy [ ][ ]
    [ ][ ] year [ ][ ] month — SK_live_mm [ ][ ]

Section 2: Birthplace and migration history

13. Have you ever moved to stay other region more than 3 month () — His1mov0 [ ]
    ❒ (0) Never — His1mov1 [ ]
    ❒ (1) Southern (Outside Songkhla) ❒ (2) Central — His1mov2 [ ]
    ❒ (3) Northeastern ❒ (4) Northern — His1mov3 [ ]
    His1mov4 [ ]
14. Your household member move from northeastern/northern region — His2mov2 [ ][ ]
    ❒ (0) No
    ❒ (1) Yes, please identify number of [ ] person

Section 3: History of illness and risk behavior

15. Have experienced for examined parasite infestation? — Stool_ex [ ][ ]
    ❒ (0) Never
    ❒ (1) Yes (Please identify)
    ❒ (1) 1 ❒ (2) 2 times ❒ (3) 3 times
    ❒ (4) > 3 times ❒ (5) N/A
16. Have experienced for examined liver fluke infestation? — OV_infec [ ]
    ❒ (1) Never ❒ (2) Yes, negative result
    ❒ (3) Yes, positive result ❒ (4) N/A
17. Have experienced of took drug for liver fluke treatment? — Treat [ ][ ]
    ❒ (0) Never
    ❒ (1) Yes (Pls. identify)
    ❒ (1) 1 ❒ (2) 2 times ❒ (3) 3 times
    ❒ (4) > 3 times ❒ (5) N/A

ID □□□□

18. Your relatives have history of illness with cholangiocarcinoma. his_ill [ ][ ]

❒ (0) No

❒ (1) Yes (pls. identify the relationship of relative)

❒ (1) Grandparents (Paternal) ❒ (2) Grandparents (Maternal)

❒ (3) uncle, aunt (Paternal) ❒ (4) uncle, aunt (Maternal)

❒ (5) Parents ❒ (6) Sibling

❒ (7) Son/daughter ❒ (8) Grandchild

19. Do you have any history of following diseases:

1) Hepatitis B	❒ (1) Yes	❒ (2) No	❒ (3) Unsure	Hep_B [ ]
2) Hepatitis C	❒ (1) Yes	❒ (2) No	❒ (3) Unsure	Hep_C [ ]
3) DM	❒ (1) Yes	❒ (2) No		DM [ ]
4) Malaria	❒ (1) Yes	❒ (2) No		Malaria [ ]
5) Parasite	❒ (1) Yes	❒ (2) No		Para_inf [ ]

20. Smoking Smoke [ ]

❒ (0) Never ❒ (1) Yes ❒ (2) Ever

21. Alcohol Alc [ ]

❒ (0) Never ❒ (1) Yes ❒ (2) Ever

22. Have latrine ❒ (0) No ❒ (1) Yes Latrin [ ]

23. Have experience of pass stool outside latrine Latr_par [ ]

❒ (0) Never ❒ (1) sometimes ❒ (2) Always

24. Your family member have experience of pass stool outside latrine Latr_hhm [ ]

❒ (0) Never ❒ (1) sometimes ❒ (2) Always

ID □□□□

Section 4: Eating behavior, sources of food, ingredients, and cooking information

Direction: Please fill ✓ in the box that match your behavior

Food items	History of eating				
	Unknown (0)	Never (1)	Sometimes (<6 time/year) (2)	Often (≥6 time/year) (3)	
1) Koy-pla-dib					F_item01 [ ]
2) Som-pla-noi					F_item02 [ ]
3) Som-pla-dib					F_item03 [ ]
4) Lhab-pla-dib					F_item04 [ ]
5) Mum-pla					F_item05 [ ]
6) Jaw-bong-pla-ra-dib					F_item06 [ ]
7) Pla-ra-sub-kreung-kang-dib					F_item07 [ ]
8) Som-tam with raw fermented fish					F_item08 [ ]
9) Pla-jao					F_item09 [ ]
10) Raw fish (Pla-dib, no cooked)					F_item10 [ ]
11) Koy-kung					F_item11 [ ]
12) Raw pork salad					F_item12 [ ]
13) Raw meat salad					F_item13 [ ]
14) Pla-jom					F_item14 [ ]
15) Pla-som-dib					F_item15 [ ]
16) Unknown food, raw fish cooked					F_item16 [ ]

Remark: In the list, if any answer of sometime or often. Please answer the following 3 questions. If you answer never know or never eat, you no need to answer the following 3 questions.

ID □□□□

25. Where the foods in the list that you have ate come from? (multiple choice)

- ❒ (0) Fishing and cooked by myself — F_src1 [ ]
- ❒ (1) Buy ingredients (market) and cook by myself — F_src2 [ ]
- ❒ (2) Relative/neighbors shared the food — F_src3 [ ]
- ❒ (3) Buy/order from restaurant — F_src4 [ ]

26. Where is the place that you eat the food in the list? (multiple choice) — Eat_loc1 [ ]

- ❒ (0) Home/Village
- ❒ (1) Other place in Songkhla — Eat_loc2 [ ]
- ❒ (2) Southern, outside Songkhla — Eat_loc3 [ ]
- ❒ (3) Other region with ever visit — Eat_loc4 [ ]

27. Who <u>mostly</u> ate with you when you ate food in the list? — Eat_mate1 [ ]

- ❒ (0) Ate alone — Eat_mate2 [ ]
- ❒ (1) Ate with family member
- ❒ (2) Ate with friends or neighbors — Eat_mate3 [ ]

Section 5: <u>Staff only</u> Information of laboratory specimen

28. Taking blood for examine parasite infection. — Blood_ex [ ]

- ❒ (0) No
- ❒ (1) Yes
  - ❒ (0) Negative
  - ❒ (1) Positive

29. Collected stool sample for parasite eggs seeking. — Stool_ex [ ]

- ❒ (0) No collect
- ❒ (1) Collect to detect parasite eggs
  - ❒ (0) not found
  - ❒ (1) Found

=== Thank you very much for your kind cooperation in completing the interview ===

ID: .....

แบบสัมภาษณ์

ความชุกของการติดเชื้อพยาธิใบไม้ตับและแนวโน้มของอุบัติการณ์โรคมะเร็งท่อน้ำดีในประชากรจังหวัดสงขลา (Prevalence of *Opisthorchis viverrini* infestation and trend in incidence of cholangiocarcinoma in Songkhla population, 2015)

การศึกษานี้ มีจุดมุ่งหมาย เพื่อสำรวจความชุกของการติดเชื้อพยาธิใบไม้ตับและระบุปัจจัยที่มีความสัมพันธ์ต่อการติดเชื้อพยาธิใบไม้ตับในประชากรจังหวัดสงขลา รวมถึงเพื่อคาดประมาณแนวโน้มของอุบัติการณ์โรคมะเร็งท่อน้ำดีในจังหวัดสงขลา ผู้วิจัยขอความร่วมมือจากท่าน กรุณาตอบแบบสอบถามโดยให้ข้อมูลที่ตรงกับความเป็นจริงมากที่สุด หากมีคำถามข้อใดที่ท่านไม่ทราบคำตอบหรือไม่สะดวกที่จะตอบ ท่านสามารถละเว้นการตอบคำถามข้อนั้นได้โดยจะไม่มีผลกระทบใด ๆ ต่อท่านทั้งสิ้น โดยคำตอบของท่านจะถูกปกปิดเป็นความลับ ผลการศึกษาจะถูกนำเสนอในภาพรวมของชุมชน โดยไม่ระบุตัวบุคคลของผู้ตอบ งานวิจัยนี้ใช้เป็นส่วนหนึ่งของการศึกษาระดับปริญญาเอก ซึ่งสามารถนำไปใช้ในการวางแผนป้องกันและควบคุมโรคมะเร็งตับและท่อน้ำดี ในจังหวัดสงขลา และภาคใต้ต่อไป

แบบสอบถามมีทั้งหมด โดยแบ่งเป็น 5 ส่วนดังนี้

ส่วนที่ 1: ข้อมูลทั่วไปของผู้ตอบแบบสัมภาษณ์	12	คำถาม
ส่วนที่ 2: ข้อมูลเกี่ยวกับการเกิดและประวัติการย้ายถิ่น	2	คำถาม
ส่วนที่ 3: ประวัติการเจ็บป่วยพฤติกรรมและ ประวัติเสี่ยงจากโรคมะเร็ง	10	คำถาม
ส่วนที่ 4: พฤติกรรมการบริโภค แหล่งอาหาร และผู้สัมผัสร่วม	19	คำถาม
ส่วนที่ 5: ข้อมูลเกี่ยวกับการส่งสิ่งตัวอย่าง	2	คำถาม

หากท่านมีข้อเสนอแนะเพิ่มเติมหรือข้อสงสัยในการศึกษาวิจัยครั้งนี้ ท่านสามารถติดต่อคณะผู้วิจัยได้ที่หน่วยระบาดวิทยา คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ หมายเลขโทรศัพท์: 074451165 โทรสาร: 07442 9754

1. นายสีใส ยี่สุ่นแสง หัวหน้าโครงการ โทร. 089-439-1677
2. นางนุชศรา ศิริโชติ ผู้ร่วมวิจัย สำนักงานสาธารณสุขจังหวัดสงขลา โทร. 080-539-0270
3. นายปรีชา หนูฟอง ผู้ร่วมวิจัย สำนักงานป้องกันควบคุมโรคที่ 12 สงขลา โทร. 081-957-5009
4. รองศาสตราจารย์ นพ.หัชชา ศรีปลั่ง ที่ปรึกษาโครงการ

ID □□□□

---

1. รหัส (ID) : [ ][ ][ ][ ] | ID [ ][ ][ ][ ]
2. ที่อยู่ : ชื่อหมู่บ้าน/ชุมชน..... | Moo [ ][ ]
   หมู่ที่ ..... [ ][ ]..... ตำบล ..... อำเภอ ..... | Sub_D [ ][ ]
   | Distr [ ][ ]
3. ประเภทพื้นที่ ที่ตั้งครัวเรือน:
   ❒ (1) ในเขตเทศบาล ❒ (2) นอกเขตเทศบาล | Area_typ [ ]
4. วันที่สัมภาษณ์: [ ][ ]/ [ ][ ]/ 2558 | Indate [ ][ ][ ][ ]

ส่วนที่ 1: ข้อมูลทั่วไปของผู้ตอบแบบสัมภาษณ์

1. เพศ ❒ (1) ชาย ❒ (2) หญิง | Gender [ ]
   หมายเลขโทรศัพท์ (ถ้ามี): _____
2. อายุ [ ][ ] ปี | Age [ ][ ]
3. น้ำหนัก [ ][ ][ ] กิโลกรัม | Weigh [ ][ ][ ]
4. ส่วนสูง [ ][ ][ ] เซนติเมตร | Heigh [ ][ ][ ]
5. ศาสนา | Religion [ ]
   ❒ (0) ไม่มี ❒ (1) พุทธ ❒ (2) อิสลาม
   ❒ (3) คริสต์ ❒ (4) อื่นๆ
6. สถานภาพสมรส | Marital [ ]
   ❒ (1) โสด ❒ (2) แต่งงาน
   ❒ (3) หม้าย ❒ (4) แยก/หย่า
7. การศึกษาสูงสุด | Educa [ ]
   ❒ (1) ประถมศึกษาหรือต่ำกว่า ❒ (2) มัธยมศึกษา หรือเทียบเท่า
   ❒ (3) อนุปริญญาหรือเทียบเท่า ❒ (4) ปริญญาตรี
   ❒ (5) สูงกว่าปริญญาตรี
8. อาชีพหลัก (ระบุหมายเลข) [.....] อาชีพรอง (ระบุหมายเลข ถ้ามี) [.....] | Occ1 [ ]
   ❒ (1) ทำนา ❒ (2) ทำสวนยาง | Occ2 [ ]
   ❒ (3) เกษตรอื่นๆ ❒ (4) ค้าขาย/ธุรกิจส่วนตัว
   ❒ (5) ลูกจ้าง/พนักงานบริษัท ❒ (6) รับราชการ/รัฐวิสาหกิจ
   ❒ (7) รับจ้างทั่วไป ❒ (8) อื่นๆ ระบุ.....

ID □□□□ หน้า 3/6

9. รายได้ต่อเดือนเฉลี่ย [ ][ ][ ],[ ][ ][ ] บาท — Income [ ][ ][ ][ ][ ][ ]

10. 1) สมาชิกที่อยู่ในครัวเรือนเดียวกันทั้งหมด [ ] [ ] คน, — HH_memb [ ][ ]
    2) สมาชิกที่อยู่ในครัวเรือนที่อายุ 20 ปีหรือมากกว่า [ ] [ ] คน — HH_20 [ ][ ]

11. ท่านเป็นคนจังหวัดสงขลาโดยกำเนิด — SK_pop [ ]
    ❒ (1) ใช่ ❒ (2) ไม่ใช่

12. ระยะเวลาที่ท่านอาศัยอยู่ในจังหวัดสงขลา — SK_live_yy [ ][ ]
    ❒ จำนวน [ ][ ] ปี [ ][ ] เดือน — SK_live_mm [ ][ ]

ส่วนที่ II: Information of the Birth and Immigration

13. ท่านเคยย้ายไปพักอาศัยที่ภาคอื่นมากกว่า 3 เดือน (ตอบได้มากกว่า 1 ข้อ) — His1mov0 [ ]
    ❒ (0) ไม่เคยย้าย — His1mov1 [ ]
    ❒ (1) ภาคใต้ (นอกสงขลา) ❒ (2) ภาคกลาง — His1mov2 [ ]
    ❒ (3) ภาคตะวันออกเฉียงเหนือ ❒ (4) ภาคเหนือ — His1mov3 [ ]
    His1mov4 [ ]

14. มีคนในครอบครัวท่านย้ายมาหรือเคยย้ายไปพักอาศัยที่ภาคเหนือ/ภาคตะวันออกเฉียงเหนือ — His2mov2 [ ][ ]
    ❒ (0) ไม่มี
    ❒ (1) มี ระบุจำนวน [ ] คน

ส่วนที่ 3: ประวัติการเจ็บป่วย การสัมผัส

15. ท่านเคยส่งอุจจาระตรวจหาหนอนพยาธิหรือไม่ — Stool_ex [ ][ ]
    ❒ (0) ไม่เคย
    ❒ (1) เคย (กรุณาระบุ)
    ❒ (1) 1 ครั้ง ❒ (2) 2 ครั้ง ❒ (3) 3 ครั้ง
    ❒ (4) > 3 ครั้ง ❒ (5) จำไม่ได้

16. ท่านเคยมีประวัติตรวจหาพยาธิใบไม้ตับหรือไม่ — OV_infec [ ]
    ❒ (1) ไม่เคย ❒ (2) เคยตรวจ, ไม่พบเชื้อ
    ❒ (3) เคยตรวจ, พบเชื้อ ❒ (4) จำไม่ได้

ID □□□□ หน้า 4/6

17. ท่านเคยทานยารักษาโรคพยาธิใบไม้ตับหรือไม่ Treat [ ][ ]

❒ (0) ไม่เคย

❒ (1) เคย (กรุณาระบุ)

❒ (1) 1 ครั้ง ❒ (2) 2 ครั้ง ❒ (3) 3 ครั้ง

❒ (4) > 3 ครั้ง ❒ (5) จำไม่ได้

18. ท่านมีญาติที่ป่วยเป็นมะเร็งตับและท่อน้ำดีหรือไม่ his_ill [ ][ ]

❒ (0) ไม่มี

❒ (1) มี กรุณาระบุความเกี่ยวข้อง...

❒ (1) ปู่-ย่า ❒ (2) ตา-ยาย

❒ (3) ลุง ป้า น้า อา (ญาติฝ่ายพ่อ) ❒ (4) ลุง ป้า น้า อา (ญาติฝ่ายแม่)

❒ (5) พ่อ-แม่ ❒ (6) พี่-น้อง

❒ (7) ลูก ❒ (8) หลาน

19. ท่านมีประวัติเจ็บป่วยด้วยโรคเหล่านี้หรือไม่

1) ตับอักเสบ ชนิด บี ❒ (1) ใช่ ❒ (2) ไม่ใช่ ❒ (3) ไม่แน่ใจ Hep_B [ ]

2) ตับอักเสบ ชนิด ซี ❒ (1) ใช่ ❒ (2) ไม่ใช่ ❒ (3) ไม่แน่ใจ Hep_C [ ]

3) เบาหวาน ❒ (1) ใช่ ❒ (2) ไม่ใช่ DM [ ]

4) มาลาเรีย ❒ (1) ใช่ ❒ (2) ไม่ใช่ Malaria [ ]

5) โรคหนอนพยาธิ อื่นๆ ❒ (1) ใช่ ❒ (2) ไม่ใช่ Para_inf [ ]

20. ท่านสูบบุหรี่หรือไม่ Smoke [ ]

❒ (0) ไม่สูบ ❒ (1) สูบ ❒ (2) เคยสูบ

21. ท่านดื่มเครื่องดื่มที่มีแอลกอฮอล์ผสมหรือไม่ Alc [ ]

❒ (0) ไม่เคย ❒ (1) ดื่ม ❒ (2) เคยดื่ม

22. บ้านท่านมีส้วมหรือไม่ ❒ (0) ไม่มี ❒ (1) มี Latrin [ ]

23. ท่านเคยถ่ายอุจจาระนอกส้วมหรือไม่ Latr_par [ ]

❒ (0) ไม่เคย ❒ (1) บางครั้ง ❒ (2) ประจำ

24. สมาชิกในครอบครัวท่านเคยถ่ายอุจจาระนอกส้วมหรือไม่ Latr_hhm [ ]

❒ (0) ไม่เคย ❒ (1) บางครั้ง ❒ (2) ประจำ

ID □□□□ หน้า 5/6

ส่วนที่ 4: พฤติกรรมการบริโภคปลาดิบ แหล่งอาหาร และผู้สัมผัสร่วม

คำแนะนำ: กรุณาใส่เครื่องหมาย ✓ ในช่องให้ตรงกับความเป็นจริง

รายการอาหาร	ความถี่ของการทาน				
	ไม่รู้จัก (0)	ไม่เคย (1)	บางครั้ง (<6ครั้ง/ปี) (2)	บ่อย (≥6ครั้ง/ปี) (3)	
1. ก้อยปลาดิบ					F_item01 [ ]
2. ส้มปลาน้อย					F_item02 [ ]
3. ส้มปลาดิบ					F_item03 [ ]
4. ลาบปลาดิบ					F_item04 [ ]
5. หม่ำปลา					F_item05 [ ]
6. แจ่วบองปลาร้าดิบ					F_item06 [ ]
7. ปลาร้าสับเครื่องแกงดิบๆ					F_item07 [ ]
8. ส้มตำใส่ปลาร้าที่ไม่ได้ต้ม					F_item08 [ ]
9. ปลาเจ่า					F_item09 [ ]
10. ปลาทั้งตัวดิบๆ					F_item10 [ ]
11. ก้อยกุ้ง					F_item11 [ ]
12. ลาบหมูดิบ					F_item12 [ ]
13. ลาบเนื้อดิบ					F_item13 [ ]
14. ปลาจ่อม					F_item14 [ ]
15. ปลาส้มดิบ					F_item15 [ ]
16. อาหารที่ทำจากปลาดิบ แต่ไม่ทราบชื่อ					F_item16 [ ]

หมายเหตุ: หากท่านตอบ <u>บางครั้งหรือบ่อย ในข้อ 1-16</u> กรุณาตอบคำถามข้อ 24-26 ถ้าไม่ใช่ ให้ข้ามไป ข้อ 28

สำหรับเจ้าหน้าที่

ID □□□□ หน้า 6/6

25. รายการอาหารที่ท่านทาน ได้มาจากที่ใด (ตอบได้มากกว่า 1 ข้อ)

- ❒ (0) จับปลามาจากแหล่งน้ำเอง และปรุงอาหารเอง — F_src1 [ ]
- ❒ (1) ซื้อปลามา แต่ปรุงอาหารเอง — F_src2 [ ]
- ❒ (2) ญาติ/เพื่อนบ้าน ปรุงมาให้ทาน — F_src3 [ ]
- ❒ (3) สั่งซื้อจากร้านอาหาร — F_src4 [ ]

26. ท่านทานอาหารเหล่านั้นที่ไหน (ตอบได้มากกว่า 1 ข้อ) — Eat_loc1 [ ]

- ❒ (0) บ้าน/ในหมู่บ้าน
- ❒ (1) ในจังหวัดสงขลา — Eat_loc2 [ ]
- ❒ (2) ภาคใต้ นอกจังหวัดสงขลา — Eat_loc3 [ ]
- ❒ (3) ภาคอื่นๆ — Eat_loc4 [ ]

27. ผู้ที่ร่วมรับประทานอาหารเหล่านั้นกับท่าน (ตอบได้มากกว่า 1 ข้อ) — Eat_mate1 [ ]

- ❒ (0) ทานคนเดียว
- ❒ (1) ทานกับคนอื่นๆ ในครอบครัว — Eat_mate2 [ ]
- ❒ (2) ทานกับเพื่อน ญาติ หรือเพื่อนบ้าน — Eat_mate3 [ ]

<u>สำหรับเจ้าหน้าที่</u>: ข้อมูลการจัดเก็บสิ่งส่งตรวจ

28. เจาะเลือดสำหรับตรวจหาประวัติพยาธิ — Blood_ex [ ]

- ❒ (0) ไม่เจาะ
- ❒ (1) เจาะ

29. เก็บอุจจาระสำหรับตรวจหาไข่พยาธิ — Stool_ex [ ]

- ❒ (0) ไม่ได้เก็บ
- ❒ (1) เก็บ

== ขอขอบพระคุณท่านที่ให้ความร่วมมือในการตอบคำถาม ==

Annex 5: Manuscript

Manuscript 1

pISSN 1598-2998, eISSN 2005-9256

Cancer Res Treat. 2016 May 18 [Epub ahead of print] http://dx.doi.org/10.4143/crt.2016.045

Original Article Open Access

Estimation of the Incidence of Hepatocellular Carcinoma and Cholangiocarcinoma in Songkhla, Thailand, 1989-2013, Using Multiple Imputation Method

Seesai Yeesoonsang, BPH, MSc[1]
Surichai Bilheem, BSc[1]
Edward McNeil, MSc[1]
Sophon Iamsirithaworn, MD, PhD[2]
Chuleeporn Jiraphongsa, MD, PhD[3]
Hutcha Sriplung, MD[1]

[1]*Epidemiology Unit, Faculty of Medicine, Prince of Songkla University, Hat Yai,* [2]*Department of Disease Control, Ministry of Public Health, Bangkok,* [3]*Thailand MOPH - U.S. CDC Collaboration (TUC), Ministry of Public Health, Nonthaburi, Thailand*

Correspondence: Seesai Yeesoonsang, BPH, MSc
Epidemiology Unit, Faculty of Medicine,
Prince of Songkla University, Thailand
Tel: 66-7445-1165
Fax: 66-7442- 9754
E-mail: hutcha.s@psu.ac.th

Received January 29, 2016
Accepted May 2, 2016

Purpose
Histological specimens are not required for diagnosis of liver and bile duct (LBD) cancer, resulting in a high percentage of unknown histologies. We compared estimates of hepatocellular carcinoma (HCC) and cholangiocarcinoma (CCA) incidences by imputing these unknown histologies.

Materials and Methods
A retrospective study was conducted using data from the Songkhla Cancer Registry, southern Thailand, from 1989 to 2013. Multivariate imputation by chained equations (mice) was used in re-classification of the unknown histologies. Age-standardized rates (ASR) of HCC and CCA by sex were calculated and the trends were compared.

Results
Of 2,387 LBD cases, 61% had unknown histology. After imputation, the ASR of HCC in males during 1989 to 2007 increased from 4 to 10 per 100,000 and then decreased after 2007. The ASR of CCA increased from 2 to 5.5 per 100,000, and the ASR of HCC in females decreased from 1.5 in 2009 to 1.3 in 2013 and that of CCA increased from less than 1 to 1.9 per 100,000 by 2013. Results of complete case analysis showed somewhat similar, although less dramatic, trends.

Conclusion
In Songkhla, the incidence of CCA appears to be stable after increasing for 20 years whereas the incidence of HCC is now declining. The decline in incidence of HCC among males since 2007 is probably due to implementation of the hepatitis B virus vaccine in the 1990s. The rise in incidence of CCA is a concern and highlights the need for case control studies to elucidate the risk factors.

Key words
Hepatocellular carcinoma, Cholangiocarcinoma, Thailand, Incidence, Estimation techniques

Introduction

Hepatocellular carcinoma (HCC) is the most common type of liver cancer worldwide. It is known to be associated with hepatitis B and C infections [1]. The incidence of HCC has declined since the introduction of hepatitis B vaccination, particularly in Taiwanese adolescents and young adults [2]. Increasing incidence of cholangiocarcinoma (CCA), a rare cancer of bile duct epithelial lining, has been reported in the United States [3] and Australia [4], with age standardized rates (ASR) of around 1.0 per 100,000 in both countries. Attempts have been made to classify subtypes of CCA according to the radiographic appearance as described by the Bithmus and Corlette classification of perihilar CCA [5], and based on cells of origin [6]. Both classifications agree that both intra- and extrahepatic cancer are CCA.

According to the series: "Cancer in Thailand" [7,8], HCC and CCA are grouped together as cancers of the liver and bile duct (LBD). This grouping is in accordance with the

Cancer Res Treat. 2016 May 18 [Epub ahead of print]

above-mentioned Bithmus and Corlette classification. During 2004-2006, the ASRs of LBD cancer were 42.8 per 100,000 in men, and 18.2 in women. The rates decreased to 38.6 per 100,000 in men and 14.6 during 2007-2009 with variations in the proportions of HCC and CCA. The main risk factor for CCA is liver fluke infection, specifically the *Opisthorchis viverrini* (OV) species, which is common in Southeast Asia [9]. Infection from other species of liver fluke, including *Clonorchis sinensis*, is the main risk factor for hepatobiliary cancer in Korea [10]. In a study reported from Khon Kaen cancer registry in the northeast of Thailand, where the prevalence of OV is very high, the ASRs of HCC were 30.3 in males and 13.1 in females [11] whereas the CCA incidence rates were 62.0 in men and 25.6 in women [12].

Songkhla, a province in the south of Thailand, occupies an area of 7,392 square km and is situated on the eastern side of the Malay Peninsula adjoining Malaysia to the south. There are 16 districts in Songkhla and approximately 25% of the population is Muslim. The incidence of LBD cancers in Songkhla has increased in the past decade. In contrast with the national average, the ASRs in Songkhla have increased from 16.0 per 100,000 in men and 4.4 in women during 2004-2006 to 18.4 in men and 5.3 in women in 2007-2009.

In the past, the diagnosis of LBD cancer in Songkhla, where liver cancer was not common, required histopathological and/or cytological confirmation, and would only be made in those with a good performance status. Due to advances in radiographic techniques and improved image quality, as well as the use of Bithmus and Corlette classification, histological confirmation of LBD cancer has declined while the incidence of LBD cancer has increased.

When the percentage of morphological verification is low, the true incidence rates for each histological type are often underestimated. Multiple imputation (MI) is a statistical method which can be used for datasets with missing entries [13-15]. MI produces a distribution of plausible values for a missing variable in a record given the values of that record's non-missing covariates.

The purpose of this study is to estimate the incidences of HCC and CCA in Songkhla province from 1989 to 2013 using a MI technique to determine histological type among cases having unknown histology.

Materials and Methods

1. Cancer cases

This study was conducted using data on LBD cancer cases registered in the population-based cancer registry of Songkhla province. The study protocol was approved by the Ethics Committee of the Faculty of Medicine, Prince of Songkla University. All cases diagnosed with LBD cancer between 1989 and 2013 with a basis of diagnosis were included in the analysis. Prior to imputation, four initial groups of patients were defined based on the third edition of the International Classification of Diseases for Oncology (ICD-O3) [16] as follows: group 1, HCC (topography [T] code C22.0 and morphology [M] codes 8170-8176); group 2, CCA (T C22.1, and C24.x, excluding C24.1 and M 8050, 8140-8141, 8160-8161, 8260, 8440, 8480-8500, 8570-8572); group 3, other specified LBD cancers (T C22.0 with any M and T C24.1); group 4, LBD cancers with unknown histology (T C22.0 and M 8000-8005).

As shown in Fig. 1, the percentage of LBD cancer cases with unknown histological type increased from 16 cases (40%) in 1999 to 96 cases (70%) in 2005 and then plateaued. The percentage of HCC decreased rapidly in 2000 and declined steadily thereafter. As shown in Fig. 2, the percentage of cases with morphological verification declined from 60% in 1997 to 20% in 2005, which occurred after the adoption of the Bithmus and Corlette classification, a radiological classification system which does not require laboratory and histological investigation in diagnostic procedures for LBD cancer cases. Such a classification has higher sensitivity than the pathological diagnosis, thus the number of cases with LBD cancer increased. Another side effect was that many clinicians did not specify the type of cancer in the medical records.

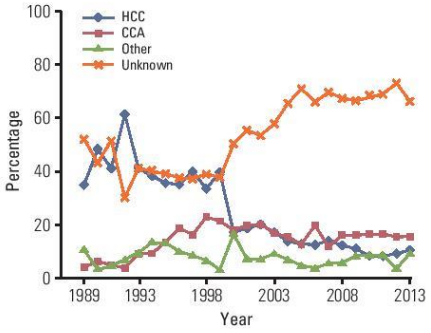

Fig. 1. Percentage of histologic tpyes of liver and bile duct cancer in both sexes by calendar year. The percentage of cases with unknown histologic type (orange line) increased from around 40% in 1999 to 70% in 2005 and then plateaued. The percentage of hepatocellular carcinoma (blue line) decreased rapidly in 2000 and then showed a steady decline.

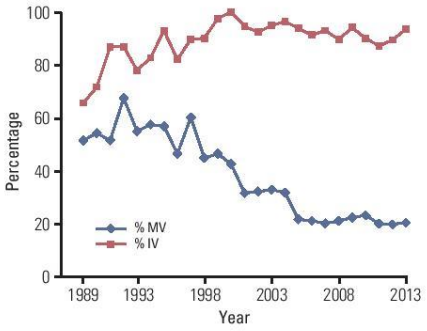

Fig. 2. The percentage of cases with morphologic verification (% MV; basis of diagnosis code, 5-8) and the percentage of cases with imaging verification (% IV; basis of diagnosis code, 2-8) by calendar year. The pathological diagnosis declined from 1998 to 2005 and then plateaued while the use of imaging in diagnosis of liver and bile duct cancer increased from 1989 through around 2000 and then plateaued at around 90%.

As the data in the cancer registry also show that the number of LBD cancer cases diagnosed during 1989-2006 was lower compared with the period after 2006, we therefore included a random sample of cancer cases with unknown primary in the abdomen (C76.2) or unknown primary (C80.9) in group 4, stratified by age group, sex, and year of diagnosis. The optimal number of cases randomly selected from the unknown primary group to include in group 4 differed each year ranging from 30% for the period 1989 to 1997 down to 0% from 2007 onwards or the total number of unknown primary cancers in each age/sex/year strata, whichever was the lowest.

The reason for including these unknown primary cancer cases in group 4 is that we believe some of them were misclassified due to incomplete investigations (for example, the patient died or refused to undergo any surgical procedure).

2. Population denominators

The population denominators used for the calculation of 5-year age-specific and age-standardized rates were estimated from the 1990, 2000, and 2010 censuses published by the National Statistical Office (NSO), Thailand, which provides annual estimates by age group and sex. Intercensus populations for the years in between were estimated using a log-linear function between two consecutive censuses. The populations beyond 2010 were estimated and reported by the Office of the National Economic and Social Development Board [17].

3. MI method

Multivariate Imputation by Chained Equations (MICE) package [13] in R [18] was used in performance of the imputations. Cases with unknown histology were imputed with one of the other known histological categories according to the probability distribution of the groups among those with known histology obtained by the chained equation method plus a degree of random error. Since the outcome in this case was a multiple categorical variable, a multinomial logistic regression model was used to generate the distribution according to the predictive ability of existing variables in the registry database. These variables included sex, age, year of diagnosis, religion (Budhist, Muslim, and other), and residential district. The model is given by

$$f(k, i) = \beta_k \times x_i, \qquad (1)$$

, where β_k is the set of regression coefficients associated with histological type k (HCC, CCA, or other), and x_i is the set of predictor variables associated with observation i. The method described by White et al. [14] was applied to avoid bias due to perfect prediction. We repeated 1,000 iterations of MI to obtain the 95% Bayesian probability intervals (PI) obtained from the quantiles of the posterior distribution for the three histological types.

Because the imputation method cannot produce 95% confidence intervals, the imputations were iterated 1,000 times to obtain 95% PI for the estimates.

4. Computation of age-standardized incidence rates

Comparison of the proportion of HCC and CCA over a long period can be biased by the change in the age structure of the population, therefore age-standardized incidence rates were used for both groups to illustrate the effect of time on the probability of imputation. The rates were standardized to the world population as proposed by Doll et al. [19] and calculated for each of the 24 calendar years between 1989 and 2013.

After imputation, descriptive statistics including frequencies and percentages were presented. Temporal trends of HCC and CCA were compared based on three models: model 1, LBD cancers with known histology only; model 2, All LBD cancers with imputation of unknown histology, and model 3, All LBD cancers plus cases with unknown primary, both in the abdomen and not otherwise specified, with imputation of unknown histology.

Results

From 1989 to 2013, there were 2,387 LBD cancers in the Songkhla registry. A high proportion of males (74.6%) and approximately half of the cases were aged between 50 and 69 years. Cases with unknown histology accounted for approximately 61% in both sexes; 64.9% in males, and 49.6% in females. As shown in Table 1, a higher proportion of HCC was observed in males (18.5%) than in females (13.1%) whereas a higher proportion of CCA was observed in females (24.9%) compared to males (11.6%). LBD cancers with other known histology comprised approximately 7% in both sexes.

Among the cases with known histology from the multinomial logistic regression, the strongest predictor for histological type was sex, followed by year of diagnosis, age group, and district of residence. The estimated incidence and percentage of HCC, CCA, and LBD cancers with other known histology after imputation are shown in Table 2. Compared to model 2, the percentage of HCC cases was slightly higher in model 3 among both males (51.0% vs. 50.9%) and females (24.7% vs. 24.4%), however, the percentage of CCA cases was not different (34.1% vs. 34.2% for males and 50.1% vs. 50.2% for females). Model 2 and model 3 also showed similar results for LBD cancers with other histology.

The average ASR of HCC in males throughout the observed years increased after imputation from 2.3/100,000 to 5.9 (252%) in model 2 and to 6.6 (281%) in model 3. The average ASRs of HCC in females increased from 0.5/100,000 to 0.9 (180%) in model 2 and 1.1 (226%) in model 3. The average ASR of CCA among males with known histology increased from 1.2/100,000 to 3.6 (293%) in model 2 and 4.0 (330%) in model 3, respectively. The average ASRs per 100,000 CCA in females increased from 0.8 to 1.5 and 1.9 in model 2 and model 3, respectively. The annual ASRs of HCC and CCA in both sexes and the three models (model 1, complete case analysis) are shown in Fig. 3.

Table 1. Distribution of liver and bile duct cancer cases by histological type and sex, Songkhla, Thailand, 1989-2013

Histological type	Total (n=2,387)	Male (n=1,776)	Female (n=611)	p-value
HCC	409 (17.1)	329 (18.5)	80 (13.1)	0.003
CCA	358 (15.0)	206 (11.6)	152 (24.9)	< 0.001
Other	165 (6.9)	89 (5.0)	76 (12.4)	< 0.001
Unknown	1,455 (61.0)	1,152 (64.9)	303 (49.6)	< 0.001

Values are presented as number (%). HCC, hepatocellular carcinoma; CCA, cholangiocarcinoma.

Table 2. Distribution of histologic type of LBD cancer cases by sex after multiple imputation, Songkhla, Thailand, 1989-2013

Histologic type	Model 2: LBD cancers with unknown histology only		Model 3: LBD cancers with unknown histology+C76.2 and C80.9	
	Male (n=1,152)	Female (n=303)	Male (n=1,323)	Female (n=421)
HCC				
95% PI (n)	586 (552-620)	74 (64-84)	675 (639-711)	104 (89-119)
95% PI (%)	50.9 (47.9-53.8)	24.4 (21-27.9)	51 (48.3-53.8)	24.7 (21.3-28.2)
CCA				
95% PI (n)	394 (362-426)	152 (140-164)	451 (363-423)	211 (140-164)
95% PI (%)	34.2 (31.4-37)	50.2 (46.2-54.2)	34.1 (31.5-36.7)	50.1 (46.1-54.1)
OSCA				
95% PI (n)	172 (148-196)	77 (66-88)	197 (149-194)	106 (66-87)
95% PI (%)	14.9 (12.8-17)	25.4 (21.9-28.9)	14.9 (12.9-16.8)	25.2 (21.7-28.7)

C76.2, unknown primary in the abdomen; C80.9, unknown primary site, not otherwise specified; LBD, liver and bile duct; HCC, hepatocellular carcinoma; PI, probability interval; CCA, cholangiocarcinoma; OSCA, other specified histology cancers.

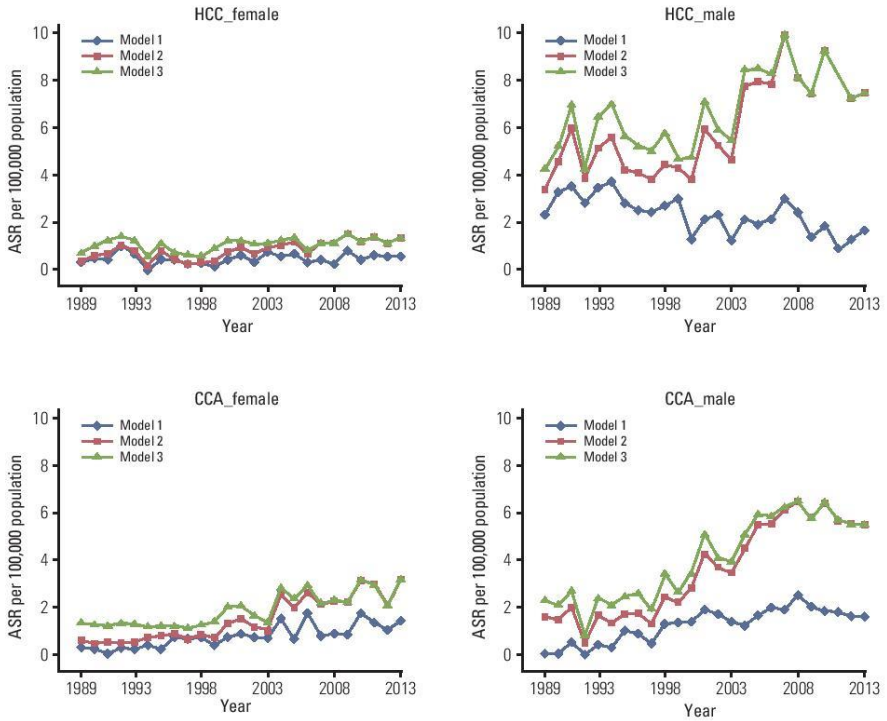

Fig. 3. Age standardized incidence rates (ASR) of liver and bile duct (LBD) cancers in Songkhla from 1989 to 2013 for the two major histologic categories; hepatocellular carcinoma (HCC) and cholangiocarcinoma (CCA), stratified by sex. Model 1, complete-case analysis; model 2, multiple imputation (MI) with LBD cancers with unknown histology only; model 3, MI of LBD cancers with unknown histology combined with cancer of unknown primary in the abdomen (C76.2) and unknown primary site, not otherwise specified (C80.9).

Discussion

In this study, missing histological types of LBD cancers were imputed using multinomial logistic regression. Results of the MI depend on the distribution of the predictive factors and their relative risk ratios from the multinomial regression.

Changes in ASR of the two major histological sub-types of LBD cancer throughout the study period are shown in Fig. 3. Based on model 1—the complete cases analysis—a decline in HCC incidence among males after 2007 was not evident. We would expect to see a decline in the incidence since the nationwide program of hepatitis B virus (HBV) immunization in all newborns in Songkhla was initiated in 1991 [20] and a large proportion of children and adults were immunized both before and after 1991. Because testing for HBV and hepatitis C virus infection has been routinely performed in blood donors since 1985, we would also expect to see a slight decline in the incidence of HCC well before 2007 [21]. Such a decline after 2007 was not observed among females in any of the three models. Model 3 showed a rather stable trend in incidence of HCC in both sexes before 2007, and therefore appears to have a better prediction capability than model 2. In other words, inclusion of a random number of

unknown primary cases prior to the imputation process may have been justified. The U-shaped decline in HCC incidence during 1995-2005, as seen in model 2, reflects the changes in diagnostic methods of LBD cancer during this time period.

MI performs well when the responses are missing at random (MAR). However, the assumption of MAR cannot usually be verified [22]. Bias can only be avoided in MI analyses if enough predictor variables are included in the imputation process [23]. One study demonstrated that the MI method works well when the percentage of missing values is between 10% and 60% [24]. In this study, the percentage of cases with missing histology ranged from 60% to 67%. When the number of cases in the dataset is reasonably high, the MICE method for a binomial outcome gives low variation of coefficients [25]. However, no study investigating MI for multinomial outcomes has been reported.

There was no solid evidence of misclassification bias among the cases with unknown histology. The chance of death among all histologic types is theoretically non-differential [26]. Most cases with unknown histological type were diagnosed by death certificate or clinical investigation and there was no difference in prognosis and survival in all of the major sub-types of LBD cancers [27]. In addition, clinicians may not perform cytology and/or biopsy for reasons mainly due to the performance status of patients and their compliance, particularly in controlling for intraperitonial bleeding after a biopsy [28]. It is possible that examination by imaging is performed more often in patients with jaundice who are likely to be CCA rather than HCC. However, this phenomenon would not affect the MI process.

In Songkhla, the hepatitis B vaccine has been included in the Expanded Program of Immunization (EPI) since 1991 [20], and the prevalence of OV has been very low in the southern region of Thailand [12,29]. Results based on models 2 and 3 showed that the ASR of HCC among males started to decrease in 2007, 16 years after the incorporation of hepatitis B vaccine into the national EPI program. Such a phenomenon is consistent with trends from Taiwan [2]; however, the decrease was not observed among females in which the rates were much lower. OV infections have not increased in the southern Thai population [9], thus the continuous increase in incidence of CCA in Songkhla province during the past two decades cannot be explained by OV infections. The increase is more likely due to the increased facilities for diagnosing LBD cancers as well as the real increase in incidence, which was also observed in the United States and Australia [3,4]. However, the estimated ASRs among males and females in Songkhla, as well as the rate of increase in incidence rates, was much higher than the rates reported in the United States and Australia (around 1 per 100,000).

The stool egg count with formalin-ethyl acetate concentration technique, which can be negative in mild parasite infections and in those who had an infection in the past without reinfection, is the method used in surveys of OV infection [30]. Eating raw fish is not the habit of people in the southern region of Thailand. Immigrants to Songkhla from those who reside in regions with a high prevalence of OV infection and residents of Songkhla who visited regions with a high prevalence of OV infection can test negative with this technique even if they are infected. Thus, it is possible that these people were exposed to OV in the past but were negative for stool OV egg count. Surveys utilising newly developed techniques with high sensitivity and specificity which can detect past infections are needed to confirm the true proportion of people who were ever exposed to OV and therefore explain the increasing trend in incidence of CCA in Songkhla province.

Conclusion

In Songkhla province, the incidence of HCC has decreased among males since 2007 while the incidence of CCA has shown a continuous increase. The effect of hepatitis B vaccination in newborns in reduction of HCC incidence was demonstrated. MI can provide more accurate estimates of ASR and the trends in incidence of LBD cancer; however, MAR should be verified by review of the radiographic images, and details on subtypes of CCA can also be summarized. Case-control studies are also needed to elucidate the role of OV infections and other risk factors of CCA.

Conflicts of Interest

Conflict of interest relevant to this article was not reported.

Acknowledgments

This study was supported by the Research Chair Grant from the National Science and Technology Development Agency (NSTDA: P-10-10307), the National Research University Grant, Prince of Songkla University (MED580635S), and the Bureau of Epidemiology, Department of Disease Control, Ministry of Public Health, Thailand. We would like to thank the Songkhla cancer registry for providing information pertaining to liver and bile duct cancer.

References

1. Srivatanakul P, Sriplung H, Deerasamee S. Epidemiology of liver cancer: an overview. Asian Pac J Cancer Prev. 2004;5:118-25.
2. Hung GY, Horng JL, Yen HJ, Lee CY, Lin LY. Changing incidence patterns of hepatocellular carcinoma among age groups in Taiwan. J Hepatol. 2015;63:1390-6.
3. Altekruse SF, Petrick JL, Rolin AI, Cuccinelli JE, Zou Z, Tatalovich Z, et al. Geographic variation of intrahepatic cholangiocarcinoma, extrahepatic cholangiocarcinoma, and hepatocellular carcinoma in the United States. PLoS One. 2015;10:e0120574.
4. Luke C, Price T, Roder D. Epidemiology of cancer of the liver and intrahepatic bile ducts in an Australian population. Asian Pac J Cancer Prev. 2010;11:1479-85.
5. Ghouri YA, Mian I, Blechacz B. Cancer review: Cholangiocarcinoma. J Carcinog. 2015;14:1.
6. Cardinale V, Semeraro R, Torrice A, Gatto M, Napoli C, Bragazzi MC, et al. Intra-hepatic and extra-hepatic cholangiocarcinoma: New insight into epidemiology and risk factors. World J Gastrointest Oncol. 2010;2:407-16.
7. Khuhaprema T, Attasara P, Sriplung H, Wiangnon S, Sumitsawan Y, Sangrajrang S. Cancer in Thailand Vol. VI, 2004-2006. Bangkok: National Cancer Institute; 2012.
8. Khuhaprema T. Attasara P, Sriplung H, Wiangnon S, Sangrajrang S. Cancer in Thailand Vol VII, 2007-2009. Bangkok: National Cancer Institute, 2013.
9. Sithithaworn P, Yongvanit P, Duenngai K, Kiatsopit N, Pairojkul C. Roles of liver fluke infection as risk factor for cholangiocarcinoma. J Hepatobiliary Pancreat Sci. 2014;21:301-8.
10. Song HN, Go SI, Lee WS, Kim Y, Choi HJ, Lee US, et al. Population-Based Regional Cancer Incidence in Korea: Comparison between Urban and Rural Areas. Cancer Res Treat. 2016;48:789-97.
11. Wiangnon S, Kamsa-ard S, Suwanrungruang K, Promthet S, Mahaweerawat S, Khuntikeo N. Trends in incidence of hepatocellular carcinoma, 1990-2009, Khon Kaen, Thailand. Asian Pac J Cancer Prev. 2012;13:1065-8.
12. Kamsa-ard S, Wiangnon S, Suwanrungruang K, Promthet S, Khuntikeo N, Mahaweerawat S. Trends in liver cancer incidence between 1985 and 2009, Khon Kaen, Thailand: cholangiocarcinoma. Asian Pac J Cancer Prev. 2011;12:2209-13.
13. Van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations in R. J Stat Softw. 2011;45:1-67.
14. White IR, Daniel R, Royston P. Avoiding bias due to perfect prediction in multiple imputation of incomplete categorical variables. Comput Stat Data Anal. 2010;54:2267-75.
15. He Y, Yucel R, Zaslavsky AM. Misreporting, missing data, and multiple imputation: improving accuracy of cancer registry databases. Chance (N Y). 2008;21:55-8.
16. Forman D, Bray F, Brewster DH, Gombe Mbalawa C, Kohler B, Pineros M, et al. Cancer incidence in five continents. Vol. X. IARC Scientific Publications No. 164. Lyon: International Agency for Research on Cancer; 2014.
17. Population Projection Working Group, Office of the National Economic and Social Development Board. Population Projections for Thailand 2010-2040. Bangkok: Office of the National Economic and Social Development Board; 2013.
18. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna: R Foundation for Statistical Computing; 2014 [cited 2014 Dec 1]. Available from: http://www.r-project.org/.
19. Doll R, Payne P, Waterhouse JA. Cancer incidence in five continents, Vol. I. Geneva: Union Internationale Contre le Cancer; 1966.
20. Chub-uppakarn S, Panichart P, Theamboonlers A, Poovorawan Y. Impact of the hepatitis B mass vaccination program in the southern part of Thailand. Southeast Asian J Trop Med Public Health. 1998;29:464-8.
21. Chimparlee N, Oota S, Phikulsod S, Tangkijvanich P, Poovorawan Y. Hepatitis B and hepatitis C virus in Thai blood donors. Southeast Asian J Trop Med Public Health. 2011;42:609-15.
22. Cummings P. Missing data and multiple imputation. JAMA Pediatr. 2013;167:656-61.
23. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393.
24. Barzi F, Woodward M. Imputations of missing values in practice: results from imputations of serum cholesterol in 28 cohort studies. Am J Epidemiol. 2004;160:34-45.
25. Hardt J, Herke M, Brian T, Laubach W. Multiple imputation of missing data: a simulation study on a binary response. Open J Stat. 2013;3:370-8.
26. Bosman FT, Carneiro F, Hruban RH, Theise ND. WHO classification of tumours of the digestive system. 4th ed. Geneva: World Health Organization; 2010.
27. Bjerregaard JK, Mortensen MB, Pfeiffer P. Trends in cancer of the liver, gall bladder, bile duct, and pancreas in elderly in Denmark, 1980-2012. Acta Oncol. 2016;55 Suppl 1:40-5.
28. Chhieng DC. Fine needle aspiration biopsy of liver: an update. World J Surg Oncol. 2004;2:5.
29. Jongsuksuntigul P, Imsomboon T. Opisthorchiasis control in Thailand. Acta Trop. 2003;88:229-32.
30. Worasith C, Kamamia C, Yakovleva A, Duenngai K, Wangboon C, Sithithaworn J, et al. Advances in the Diagnosis of Human Opisthorchiasis: Development of Opisthorchis viverrini Antigen Detection in Urine. PLoS Negl Trop Dis. 2015;9:e0004157.

VITAE

Name Seesai Yeesoonsang

Student ID 5410330006

Educational Attainment

Degree	Name of Institution	Year of Graduation
Master of Science (Natural Resources and Environmental Management)	**Naresuan University**	**2004**
Bachelor of Public Health (Public Health Administration)	**Sukhothai Thammathirat Open University**	**1994**

Scholarship Awards during Enrolment

1. Field Epidemiology Training Program (FETP) Thailand, under Bureau of Epidemiology, Department of Disease Control, Ministry of Public Health, Thailand

2. The National Science and Technology Development Agency through Research Chair Grant for Prof. Virasakdi Chongsuvivatwong

Work – Position and Address

Public Health Technical Officer, Senior Professional Level

The office of Disease Prevention and Control 2th, Phitsanulok, Department of Disease Control, Ministry of Public Health, Thailand

List of Publication and Proceeding

Conference presentation

Yeesoonsang S, Beenheem S, Jiraphongsa C, Sriplung H. (2014). Estimations and trend projection of lung cancer incidence by histological type in Songkhla province, Thailand. Poster presentation at the ASEAN Cancer Registry Forum 2014 conference. Bangkok.

Yeesoonsang S, Beenheem S, Sriplung H. (2016). Liver and bile duct cancer incidence in southern Thailand, 1989-2030. Oral presentation at the 10th Graduate Forum on Health Systems and Policy, Yogyakarta, Indonesia.

Manuscript publication

Yeesoonsang S, Bilheem S, McNeil E, Iamsirithaworn S, Jiraphongsa C, Sriplung H. Estimation of the incidence of hepatocellular carcinoma and cholangiocarcinoma in Songkhla, Thailand, 1989-2013, using multiple imputation method. *Cancer Res Treat*. 2016. DOI: http://dx.doi.org/10.4143/crt.2016.045.