



OPEN

Spatial analysis of cholangiocarcinoma in relation to diabetes mellitus and *Opisthorchis viverrini* infection in Northeast Thailand

Kavin Thinkhamrop^{1,2,3✉}, Kulwadee Suwannatrai⁴, Matthew Kelly⁵ & Apiporn T. Suwannatrai^{1,2,6}

Cholangiocarcinoma (CCA) exhibits a heightened incidence in regions with a high prevalence of *Opisthorchis viverrini* infection, with previous studies suggesting an association with diabetes mellitus (DM). Our study aimed to investigate the spatial distribution of CCA in relation to *O. viverrini* infection and DM within high-risk populations in Northeast Thailand. Participants from 20 provinces underwent CCA screening through the Cholangiocarcinoma Screening and Care Program between 2013 and 2019. Health questionnaires collected data on *O. viverrini* infection and DM, while ultrasonography confirmed CCA diagnoses through histopathology. Multiple zero-inflated Poisson regression, accounting for covariates like age and gender, assessed associations of *O. viverrini* infection and DM with CCA. Bayesian spatial analysis methods explored spatial relationships. Among 263,588 participants, *O. viverrini* infection, DM, and CCA prevalence were 32.37%, 8.22%, and 0.36%, respectively. The raw standardized morbidity ratios for CCA was notably elevated in the Northeast's lower and upper regions. Coexistence of *O. viverrini* infection and DM correlated with CCA, particularly in males and those aged over 60 years, with a distribution along the Chi, Mun, and Songkhram Rivers. Our findings emphasize the association of the spatial distribution of *O. viverrini* infection and DM with high-risk CCA areas in Northeast Thailand. Thus, prioritizing CCA screening in regions with elevated *O. viverrini* infection and DM prevalence is recommended.

Cholangiocarcinoma (CCA), a form of cancer impacting the bile ducts, exhibits a noteworthy prevalence in Thailand. Recent investigations indicate that Thailand has the highest recorded incidence of CCA globally. Specifically, the age-standardized rate (ASR) is reported at 33.4 per 100,000 in men and 12.3 per 100,000 in women^{1,2}. The disease burden associated with intrahepatic CCA in Thailand is also substantial, with an incidence rate of 14.6 per 100,000 population per year during the period 2009–2013³. This elevated incidence in Thailand is closely tied to the widespread occurrence of *Opisthorchis viverrini* infection^{4–7}. *Opisthorchis viverrini* poses a significant public health challenge, with a documented prevalence in Thailand, Laos, Cambodia, Myanmar, and Vietnam^{8,9}. The primary mode of infection stems from the consumption of raw or undercooked freshwater cyprinid fish^{1,9–11}. This dietary practice is deeply embedded in the cultural norms of the people in northeastern Thailand and the broader Lower Mekong Region¹². Previous studies conducted in northeastern Thailand revealed that the prevalence of *O. viverrini* infection was estimated at 17% in 2009⁹. A subsequent study in 2014 reported an increase to about 23%, with a higher prevalence observed among males and individuals aged 40–49 years¹³. Moreover, following the treatment of *O. viverrini* infections, reinfection rates were approximately 10%, attributed to unchanged

¹Cholangiocarcinoma Screening and Care Program (CASCAP), Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand. ²Cholangiocarcinoma Research Institute (CARI), Khon Kaen, Thailand. ³Health and Epidemiology Geoinformatics Research (HEGER), Faculty of Public Health, Khon Kaen University, Khon Kaen, Thailand. ⁴Faculty of Science and Technology, Sakon Nakhon Rajabhat University, Sakon Nakhon, Thailand. ⁵National Centre for Epidemiology and Population Health, College of Health and Medicine, Australian National University, Canberra, Australia. ⁶Department of Parasitology, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand. ✉email: kavith@kku.ac.th

dietary habits¹⁴. This parasitic infection triggers chronic inflammation, resulting in oxidative DNA damage to the infected biliary epithelium and subsequent malignant transformation¹. Notably, recent investigations have indicated a declining trend in the incidence of CCA in Thailand. Both males and females have experienced a significant reduction in the incidence rate, marking a positive shift in the epidemiological landscape¹⁵.

The incidence of CCA exhibits significant regional variability, with distinct risk factors identified across diverse populations. Previous studies in Thailand have outlined specific risk factors for CCA, including *O. viverrini* infection, the consumption of raw cyprinid fish, a family history of cancer, the use of praziquantel (PZQ), hepatitis C virus, hepatitis B virus, cirrhosis, obesity, alcohol consumption, tobacco smoking, and diabetes mellitus (DM)^{2,16,17}. Furthermore, a recent study conducted in northeastern Thailand has shed light on an intriguing correlation. This research uncovered that individual infected with *O. viverrini* not only faced an increased risk of CCA but that this risk was further amplified in individuals who also had DM¹⁸. This dual association suggests a potential synergistic effect between *O. viverrini* infection and DM in elevating the susceptibility to CCA.

While prior research has explored the association of various factors with the development of CCA, a noteworthy revelation pertains to the connection between the combined presence of *O. viverrini* infection and DM and the development of CCA. Despite these insights, the scarcity of spatial analyses aimed at pinpointing geographic areas with a high prevalence of CCA, while considering associated factors, remains evident. This gap is particularly pronounced in regions characterized by a heightened prevalence of both *O. viverrini* infection and DM.

Extensive research has delved into the spatial analysis of CCA concerning *O. viverrini* infection and DM in Northeast Thailand. The heightened prevalence of liver and biliary tract abnormalities in this region, predominantly attributed to the distribution of *O. viverrini*, has been a focal point of spatial analysis studies^{19,20}. These investigations aim to unravel the geographical distribution and clustering of CCA cases in Northeast Thailand, offering valuable insights into the environmental and regional factors contributing to the CCA prevalence²⁰. These studies have played a pivotal role in emphasizing the pressing need for eradicating the liver fluke and identifying high-risk populations to effectively address the elevated incidence of CCA in the region²⁰. Additionally, the investigation into the correlation between DM and CCA has garnered attention, as DM is considered a potential risk factor for CCA^{18,21–23}. Previous studies have suggested that individuals with DM were more likely to be incident cases of CCA²⁴, and another study demonstrated an association between DM and shorter survival among CCA patients²⁵. Understanding the spatial distribution of CCA cases in relation to DM and *O. viverrini* infection is paramount for crafting targeted prevention and control strategies. Moreover, this knowledge is crucial for enhancing early detection and treatment outcomes, particularly in high-risk areas such as Northeast Thailand^{18,26}. This region is endemic for *O. viverrini*, the primary cause of abnormalities in the biliary system examined in this study and a key factor on the etiological pathway leading to the development of CCA²⁷.

Geographic Information System (GIS) analysis has emerged as a prevalent tool in epidemiological research, facilitating the identification of disease clustering, spatial patterns, and the analysis of interactions between environmental factors and health. Nevertheless, the authors are not aware of any studies employing these methods in the investigation of the distribution of CCA, DM, and the history of *O. viverrini* infection. Therefore, our study aimed to explore the spatial distribution of CCA in relation to *O. viverrini* infection and DM within high-risk populations in Northeast Thailand.

Materials and methods

Study area

The study was conducted in Northeast Thailand, situated between latitudes 14.50°N and 17.50°N, and longitudes 102.12°E and 104.90°E, encompassing an approximate area of 168,854 km² (Fig. 1). This region comprises 20 provinces (Supplementary Fig. 1), and data for this study was systematically gathered from all 2678 sub-districts distributed across these provinces.

Study design and data sources

The cross-sectional study was reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement. The data for this study were obtained from the Cholangiocarcinoma Screening and Care Program (CASCAP) in northeastern Thailand. CASCAP represents the pioneering initiative for CCA screening within a high-risk population, employing a community-based, bottom-up approach²⁷. The CASCAP screening program endeavors to encompass all residents of northeastern Thailand, achieving this through various methods and settings, such as tertiary-care hospitals, district-level hospitals, and mobile screening sessions at the sub-district level. For this study, we included all participants who underwent CCA screening and were subsequently enrolled in the CASCAP database.

Study population

Northeastern Thailand boasts a population of approximately 21 million individuals, accounting for roughly one-third of Thailand's total population. Our study encompasses all individuals who participated in CASCAP across 20 provinces and received information regarding their history of *O. viverrini* infection and diagnosis of DM. Specifically, our study sample comprises individuals who were enrolled in the CASCAP database between 2013 and 2019, totaling 263,588 subjects. Population data for the sub-districts during the study period (2013–2019) were acquired from the Thailand Official Statistics Registration Systems website (<https://stat.bora.dopa.go.th>).

Study outcome and independent variables

The primary outcome in this study was the diagnosis of CCA, which was determined through a combination of ultrasonography (US), computed tomography (CT)/magnetic resonance imaging (MRI), and histopathological examination. Positive US findings were defined as those indicating liver masses and/or bile duct dilatation,

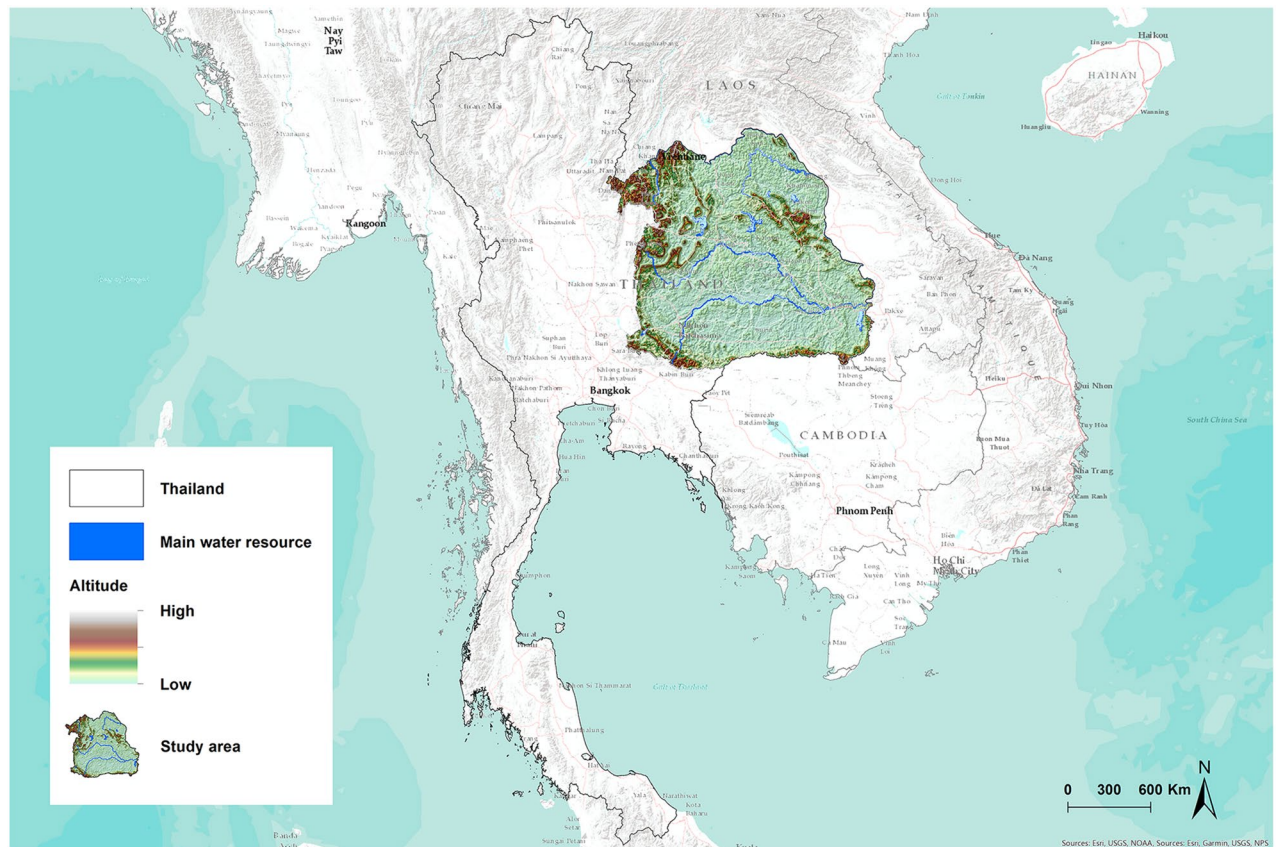


Figure 1. Map of the study area. Map was created using ArcGIS Pro software version 3.2 (ESRI: <https://www.esri.com/en-us/home>).

indicative of suspected CCA. Individuals falling into this category were subsequently referred for either CT or MRI scans. The results of the CT/MRI scans were categorized as either positive or negative for CCA. For confirmation of positive CCA cases, biopsy procedures were performed. Participants who did not exhibit suspicious findings on US, CT/MRI scans, or had histologically negative results were classified as negative for CCA. The diagnosis of CCA was only confirmed through biopsy results. The independent variables in this study encompassed the key factors of interest, namely the history of *O. viverrini* infection and the diagnosis of DM, both of which were categorized into two groups (negative/positive). Additionally, the covariates considered for analysis consisted of gender and age at the time of enrollment. These factors were obtained through interviews with study participants using the CASCAP questionnaires.

Statistical analysis

The prevalence of CCA, *O. viverrini* infection, and DM was calculated as percentages, with the number of positive cases as the numerator and the total number of participants as the denominator. In order to assess differences in the prevalence of CCA across different categorical factors, we utilized the chi-square test statistic, employing a significance level of 0.05. The prevalence is reported along with its corresponding 95% confidence interval (CI). This CI is determined by calculating the lower bound as the proportion value minus the product of the z-score for the 95% CI (1.96) and the standard error (SE). The SE is computed as the square root of the proportion value multiplied by one minus the proportion value, divided by the number of samples. Subsequently, the upper bound is obtained by the proportion value plus the product of the z-score and the SE. These values are then multiplied by 100 and presented as the 95% CI of the prevalence. All statistical analyses were conducted using STATA version 18 (StataCorp, College Station, TX, USA).

We accounted for differences in population sizes by calculating raw standardized morbidity ratios (SMR) for the overall positive of CCA in each sub-district of Northeast Thailand from 2013 to 2019, using the following formula:

$$Y_i = \frac{O_i}{E_i}$$

where Y_i represents the SMR in sub-district i , O_i stands for the observed number of cases for CCA in the sub-district, and E_i represents the expected number of cases in the sub-district over the study period. We calculated

the expected number of cases by multiplying the population of each sub-district by the overall crude rates of these conditions during the study period²⁸.

Spatial analysis

This analysis covered a region that included 2678 sub-districts in Northeast Thailand. The administrative boundaries for these sub-districts were defined using polygon shapefiles obtained from the DIVA-GIS website (www.diva-gis.org). The spatial dataset, which included SMR for CCA were imported into ArcGIS Pro software version 3.2 (ESRI Inc., Redlands, CA, USA) and projected onto the Universal Transverse Mercator (UTM) coordinate system zone 48 N. The polyline shapefiles representing main water resources in Northeast Thailand were obtained from the DIVA-GIS website (www.diva-gis.org). The administrative boundary map encompassed a total of 2678 sub-district-level areas²⁰.

We identified cases where *O. viverrini* infection history was present alongside CCA or DM, cases where DM co-occurred with CCA, and cases where all three conditions coexisted as OVCCA, DMCCA, and OVDMCCA, respectively. These cases, along with the spatial distributions of demographic data, were integrated into ArcGIS Pro software. The datasets were geospatially linked based on the administrative boundary map of Northeast Thailand. This allowed us to aggregate and extract data by sub-district areas and create variables for subsequent statistical analyses. The process of linking and analyzing the data followed established procedures previously employed in spatial analyses within this region²⁰. The cut-off points reported in the figure legends were determined using the Natural Breaks (Jenks) classification method in the ArcGIS Pro software.

We initially conducted bivariate Poisson regression analyses to explore associations between CCA cases and various independent variables, including gender, age group, history of *O. viverrini* infection and DM. This step aimed to identify covariates supported by evidence from prior articles, indicating associations with the development of CCA or demonstrating significant relationships with CCA prevalence (with a threshold of *P* value < 0.25) in the bivariate models. Variables meeting this criterion were subsequently incorporated into the final models. To assess collinearity among all included variables, we performed Pearson correlation analyses. All preliminary statistical analyses were carried out using STATA software version 18.0 (Stata Corporation, College Station, TX, USA).

Spatial poisson regression analysis

We conducted a Spatial Poisson regression analysis in our study. To address the presence of zero cases in some sub-districts for the CCA, we opted for a zero-inflated Poisson (ZIP) regression model rather than a standard Poisson regression. The ZIP regression was implemented using a Bayesian framework in WinBUGS software version 1.4.3, developed by the Medical Research Council in Cambridge, UK. Three distinct models were constructed to account for different sources of variability. The first model, referred to as Model I, incorporated unstructured random effects for gender, age groups, history of *O. viverrini* infection, and DM. In the second model, Model II, we introduced spatially structured random effects alongside the same covariates as Model I. Finally, the third model, Model III, was a convolution model that included both spatially structured and unstructured random effects, while maintaining the same covariate set as the previous two models.

The convolution model, with an outcome of observed cases of CCA (numbers), *Y*, for the *i*th sub-district in Northeast Thailand (*i* = 1 to 2678), for the *j*th gender group and *k*th age group was structured as follows:

$$P(Y_{ijk} = y_{ijk}) = \begin{cases} \omega + 1(1 - \omega)e^{-\mu}, & y_{ijk} = 0 \\ (1 - \omega)e^{-\mu} \mu^{y_{ijk}} / y_{ijk}!, & y_{ijk} > 0; \end{cases}$$

$$Y_{ijk} \sim \text{Poisson}(\mu_{ijk})$$

$$\log(\mu_{ijk}) = \log(E_{ijk}) + \theta_{ijk}$$

$$\theta_{ijk} = \alpha + \beta_1 \times \text{gender}_j + \beta_2 \times \text{age}_k + \beta_3 \times O. \text{Viverrini infection}_i + \beta_4 \times \text{DM}_i + u_i + s_i$$

where E_{ijk} is the expected number of CCA cases (offsetting population size) in sub-district *i*, gender *j*, age group *k*, μ_{ijk} is the parameter of distribution, θ_{ijk} is the mean log relative risk (RR) and $\omega \in [0,1]$ relates to proportion of extra zero; α is the intercept, and β_1 , β_2 , β_3 , and β_4 are the coefficients for age (≤ 60 years as the reference category), gender (female as the reference category), *O. viverrini* infection history, DM; u_i and s_i are the unstructured and structured random effects (assuming a mean of zero and variances of σ_u^2 and σ_s^2 respectively).

In the modelling the spatially structured random effects, we implemented a conditional autoregressive (CAR) prior structure. This structure was employed to account for spatial relationships among sub-districts within the study area. To represent these relationships, we utilized an adjacency weights matrix, where a weight of 1 denoted sub-districts sharing a border, and 0 indicated no shared border. For the model's priors, we employed a flat distribution for the intercept and a normal distribution for the coefficients, with a mean of zero and a precision (the inverse of variance) set at 0.0001²⁰. The precision of both unstructured and spatially structured random effects was modelled using non-informative gamma distributions, characterized by shape and scale parameters of 0.001. These methodologies have previously been applied in spatial analyses of CCA in Northeast Thailand.

We initiated the analysis by discarding the initial 10,000 iterations as a burn-in period. Subsequently, we executed sets of 15,000 iterations and scrutinized them for signs of convergence in the Monte Carlo chains. This process was iterated three times, with the outcomes of the first two iterations being discarded. At this

juncture, we conducted a visual examination of the posterior density plots to ascertain that convergence had indeed been achieved. The final model selection hinged on striking a balance between model fit and parsimony. We employed the deviance information criterion (DIC) to guide our decision-making, ultimately selecting the model that yielded the lowest DIC score. To establish statistical significance of the covariates, we set an α -level of 0.05, determined by analyzing the 95% credible intervals (CrI) for mean posterior relative risks (RRs). Variables were considered statistically significant if the 95% CrI for the RRs did not encompass 1^{20,29}. Finally, to visually represent the spatial distribution of posterior means for both unstructured and structured random effects, we utilized ArcGIS Pro software to generate maps.

Ethical considerations

The research protocol was approved by Khon Kaen University Ethics Committee for Human Research, reference number HE631061. The data were provided from the CASCAP. The CASCAP data collection was conducted according to the principles of Good Clinical Practice, the Declaration of Helsinki, and national laws and regulations about clinical studies. It was approved by the Khon Kaen University Ethics Committee for Human Research under the reference number HE551404. All patients gave written informed consent for the study.

Results

Descriptive summary

A total of 263,588 participants undergoing screening for CCA were included in our study. Approximately two-thirds of the participants were females (60.7%), with the majority being aged 60 years or less, reflecting a mean age of 55.7 (standard deviation = 9.2) years (Table 1). The prevalence of *O. viverrini* infection was 32.37% (95% CI: 32.19–32.54), and the prevalence of those diagnosed with DM was 8.22% (95% CI: 8.12–8.33).

Distribution of cholangiocarcinoma

Table 2 illustrates the prevalence of CCA among a total of 263,588 participants, with 0.36% diagnosed with CCA. Elevated prevalence was observed in individuals aged over 60 years (0.66%; 95% CI: 0.60–0.72), those diagnosed with DM (0.60%; 95% CI: 0.49–0.70), males (0.59%; 95% CI: 0.54–0.63), and those infected with *O. viverrini* (0.47%; 95% CI: 0.42–0.51), respectively.

Over the seven-year period under examination, the overall crude incidence rate of CCA was calculated to be 2.64 cases per 100,000 population. At the sub-district level, the incidence rates exhibited considerable variation, spanning from 0 to 38 cases per 100,000 population. Particularly noteworthy were the substantial spatial disparities observed in the incidence rate of CCA among different sub-districts. Figure 2 visually depicts the spatial distribution of SMR for CCA in Northeast Thailand at the sub-district level.

Figures 3 depict the distribution of prevalence for *O. viverrini* infection history, DM, CCA, OVCCA, OVDM, DMCCA, and OVDMCCA within the Northeast Thailand. The spatial representation highlights Chaiyaphum, Khon Kaen, Kalasin, Yasothon, and Sisaket Provinces as having the highest prevalence of *O. viverrini* (Fig. 3A). Similarly, the prevalence of DM is most pronounced in Nakhon Ratchasima, Khon Kaen, Maha Sarakham, Surin, Ubon Ratchathani, Kalasin, Nakhon Phanom, Bueng Kan, and Nong Khai Provinces (Fig. 3B). Concerning CCA, Nong Bua Lamphu, Maha Sarakham, Kalasin, and Nakhon Phanom Provinces exhibit the highest prevalence (Fig. 3C). Furthermore, the comorbidity OVCCA is prominently present in Nong Bua Lamphu, Maha Sarakham, Surin, Srisaket, and Nakhon Phanom Provinces (Fig. 3D). Additionally, DMCCA is notably prevalent in Nong Bua Lamphu, Maha Sarakham, Kalasin, Surin, and Nakhon Phanom Provinces (Fig. 3E). Lastly, OVDMCCA is most prevalent in Nong Bua Lamphu, Maha Sarakham, Kalasin, and Chaiyaphum Provinces (Fig. 3F).

Characteristics	Frequency	Percentage
Gender		
Female	160,007	60.70
Male	103,581	39.30
Age groups (years)		
≤ 60 years	185,015	70.19
> 60 years	78,573	29.81
Mean (standard deviation)	55.7 (9.2)	
<i>O. viverrini</i> infection		
No	178,275	67.63
Yes	85,313	32.37
Diabetes mellitus		
No	241,917	91.78
Yes	21,671	8.22

Table 1. Baseline characteristics and health history of participants in the CASCAP study 2013–2019. Percentage—The percentage per 100.

Factors	Participants	Cholangiocarcinoma			
		Number of cases	Prevalence	95% CI	P value
Overall	263,588	944	0.36	N/A	N/A
Gender					
Female	160,007	336	0.21	0.19–0.23	< 0.001
Male	103,581	608	0.59	0.54–0.63	
Age groups					
≤ 60 years	185,015	426	0.23	0.21–0.25	< 0.001
> 60 years	78,573	518	0.66	0.60–0.72	
<i>O. viverrini</i> infection					
No	178,275	544	0.31	0.28–0.33	< 0.001
Yes	85,313	400	0.47	0.42–0.51	
Diabetes mellitus					
No	241,917	815	0.34	0.31–0.36	< 0.001
Yes	21,671	129	0.60	0.49–0.70	

Table 2. Prevalence of cholangiocarcinoma for each factor. N/A, Not applicable; 95% CI, 95% Confidence interval of the prevalence; Prevalence, The prevalence per 100; P value, Probability values from chi-square tests.

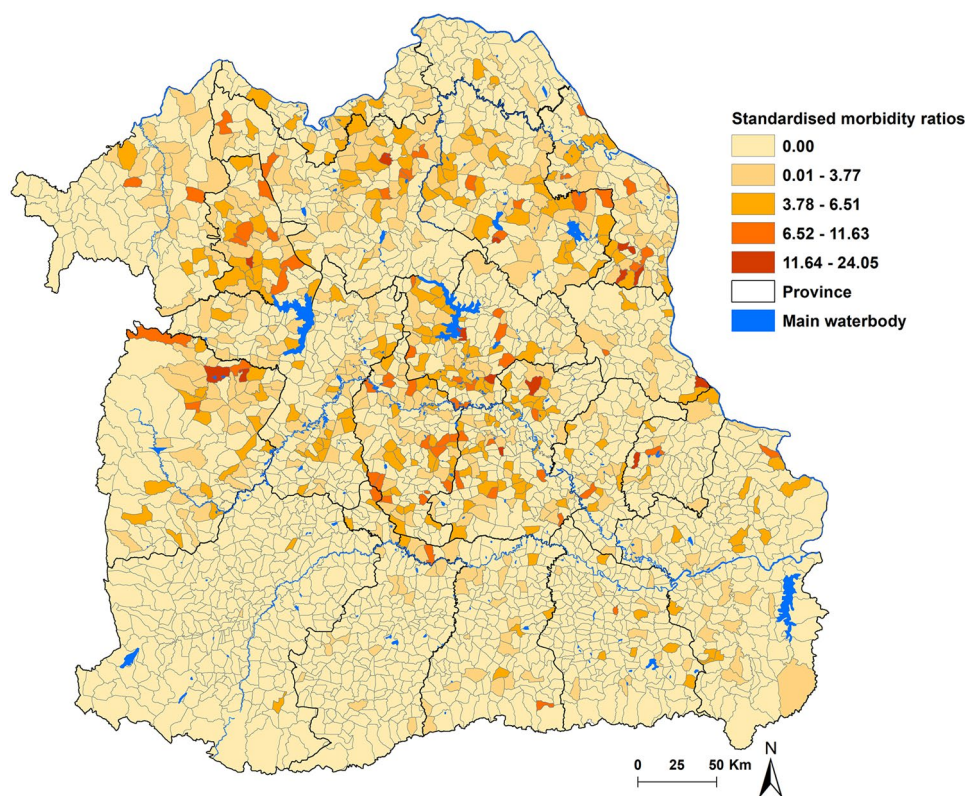


Figure 2. Raw standardized morbidity ratios of cholangiocarcinoma by sub-districts in Northeast Thailand. Map was created using ArcGIS Pro software version 3.2 (ESRI: <https://www.esri.com/en-us/home>).

Spatial poisson regression analysis

Table 3 presents the Bayesian spatial and non-spatial models applied to CCA. Based on the DIC statistics, the most fitting model was identified as Model III, incorporating both unstructured and structured random effects. Within Model III, gender, age, *O. viverrini* infection history, and DM exhibited significant associations with CCA. Males were found to be 2.52 times more likely to have CCA compared to females (95% CrI: 2.20–2.90). Individuals aged over 60 years had a 2.61 times higher likelihood of having CCA than those aged 60 years or younger (95% CrI: 2.28–2.98). The presence of *O. viverrini* infection history was associated with an increased risk of CCA (mean posterior relative risks (RRs) = 1.01; 95% CrI: 1.00–1.02), as was the case for DM patients (RRs = 1.02; 95% CrI: 1.01–1.03). Figure 4A showed spatial clustering was evident when accounting for the model covariates, as illustrated in the map depicting the posterior means of the spatially structured random

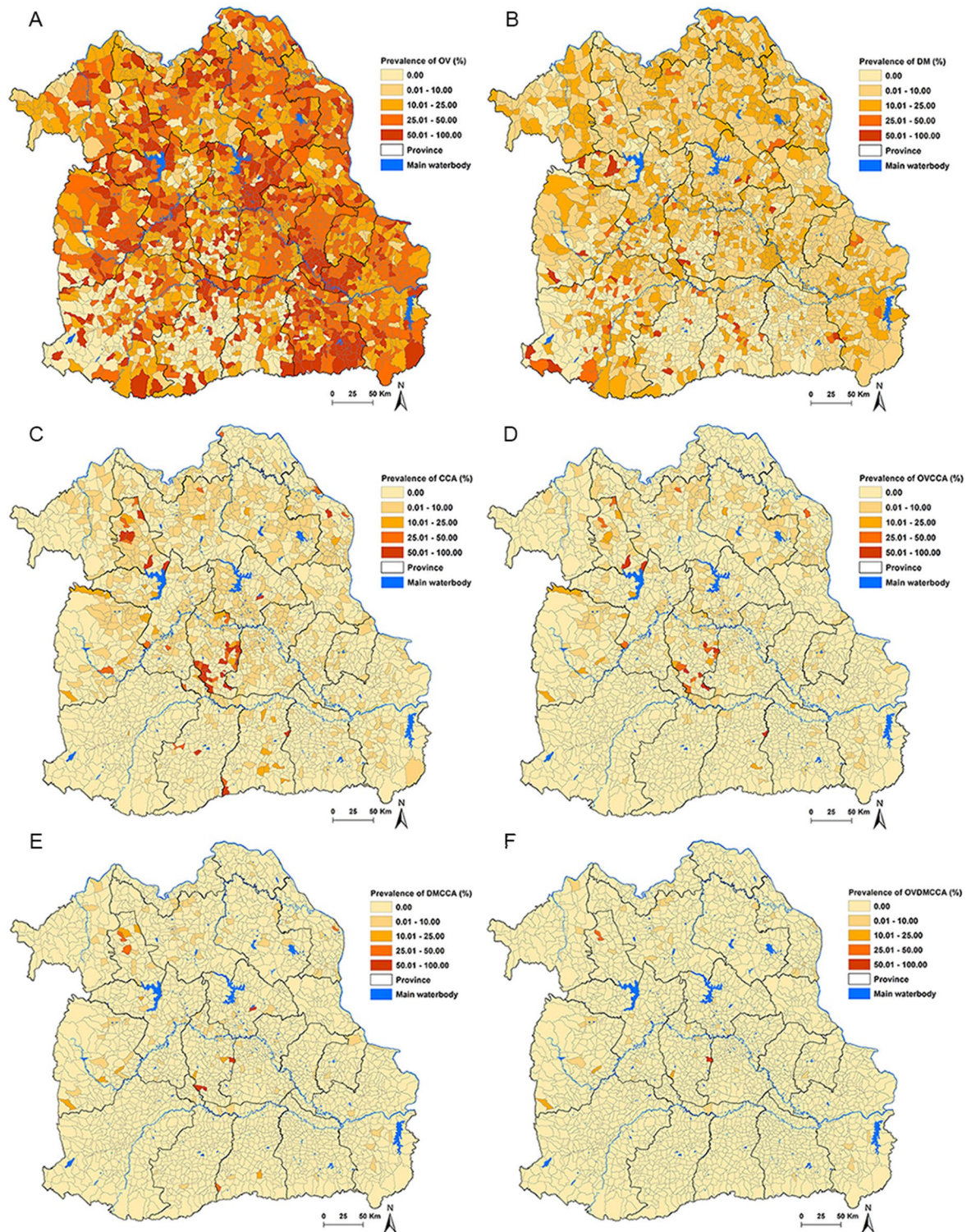


Figure 3. Prevalence of: (A) *Opisthorchis viverrini* infection (B) diabetes mellitus (C) cholangiocarcinoma (D) *O. viverrini* + cholangiocarcinoma (E) diabetes mellitus + cholangiocarcinoma and (F) *O. viverrini* + diabetes mellitus + cholangiocarcinoma. OV *Opisthorchis viverrini*; DM Diabetes mellitus, CCA Cholangiocarcinoma. Maps were created using ArcGIS Pro software version 3.2 (ESRI: <https://www.esri.com/en-us/home>).

effects. Specifically, Maha Sarakham Province, situated between the Chi and Mun rivers, exhibited a significant cluster of CCA. Other provinces, including Nong Bua Lamphu, Khon Kaen, Kalasin, Udon Thani, Surin, Nakhon Phanom, Sakon Nakhon, and Chiyaphum, also showed areas of elevated CCA risk, although these clusters

Model/variables	Coefficient, posterior mean (95% CrI)	RRs, posterior mean (95% CrI)
Model I (Unstructured)		
α (Intercept)	-7.03 (-7.22, -6.84)	
Gender ¹	0.94 (0.80, 1.08)	2.56 (2.23, 2.94)
Age ²	0.99 (0.85, 1.12)	2.68 (2.35, 3.06)
<i>O. viverrini</i> infection history	0.01 (0.00, 0.02)	1.01 (1.00, 1.02)
DM	0.03 (0.02, 0.03)	1.03 (1.02, 1.03)
Heterogeneity		
Structured (variance)		
Unstructured (variance)	0.44 (0.36, 0.53)	
DIC	5856.58	
Model II (Structured)		
α (Intercept)	-7.09 (-7.30, -6.90)	
Sex ¹	0.93 (0.79, 1.06)	2.52 (2.20, 2.90)
Age ²	0.96 (0.83, 1.09)	2.61 (2.28, 2.98)
<i>O. viverrini</i> infection history	0.01 (0.00, 0.02)	1.01 (1.00, 1.02)
DM	0.02 (0.01, 0.03)	1.02 (1.01, 1.03)
Heterogeneity		
Structured (variance)	0.17 (0.14, 0.22)	
Unstructured (variance)		
DIC	5695.79	
Model III (Structured and unstructured)*		
α (Intercept)	-7.13 (-7.34, -6.92)	
Sex ¹	0.92 (0.79, 1.06)	2.52 (2.20, 2.90)
Age ²	0.96 (0.82, 1.09)	2.61 (2.28, 2.98)
<i>O. viverrini</i> infection history	0.01 (0.00, 0.02)	1.01 (1.00, 1.02)
DM	0.02 (0.01, 0.03)	1.02 (1.01, 1.03)
Heterogeneity		
Structured (variance)	0.21 (0.15, 0.29)	
Unstructured (variance)	4.74 (1.30, 23.04)	
DIC	5673.73*	

Table 3. Regression coefficients, relative risks, and 95% credible interval from Bayesian models for cholangiocarcinoma in Northeast Thailand. CrI, Bayesian credible intervals; RRs, Relative risks; DIC, Deviance information criterion. ¹Female was reference; ²Age \leq 60 years was reference. *Best fit model.

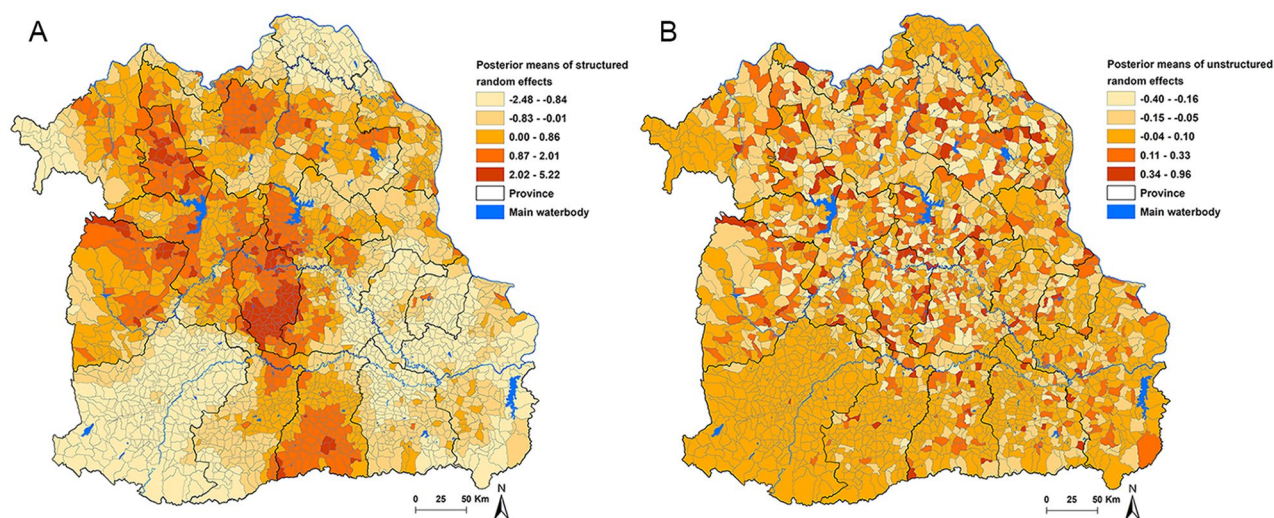


Figure 4. Spatial distributions of the posterior means of cholangiocarcinoma in Northeast Thailand in Model III: (A) Spatially structured random effects and (B) Spatially unstructured random effects. Maps were created using ArcGIS Pro software version 3.2 (ESRI: <https://www.esri.com/en-us/home>).

were more localized and included some high-risk sub-districts. The unstructured random effects displayed the expected random spatial pattern. As expected, the map displaying the posterior means of unstructured random effects, which characterized the spatially random effects of unmeasured risk factors, showed no discernible geographic pattern (Fig. 4B).

Discussion

This study stands as a pioneering endeavor, being the first to undertake a comprehensive spatial analysis of CCA concerning its correlation with DM and *O. viverrini* infection. Notably, these health factors are recognized as significant and prevalent issues in the Northeast Thailand. By exploring the spatial dynamics of CCA in conjunction with DM and *O. viverrini*, this research contributes to a deeper understanding of the interplay between these health concerns, shedding light on potential geographic patterns and correlations that can inform targeted interventions and public health strategies in Northeast Thailand. Our study revealed an association between *O. viverrini* infection and individuals diagnosed with DM and CCA. These observed effects remained significant even after adjusting for the influence of gender and age, both identified as significant covariates in this investigation. This discovery aligns with prior research indicating that gender and age factors are linked to CCA, particularly among males and the elderly^{18,26,30,31}.

The outcomes of our investigation indicate a widespread prevalence of CCA, *O. viverrini* infection, and DM across nearly all regions in northeastern Thailand. This aligns with prior research conducted in 2011, which identified a strong association between *O. viverrini* infection and DM with CCA¹⁸. Furthermore, a prospective follow-up study in 2020, involving *O. viverrini*-positive subjects in Thailand, revealed a significant increase in serum levels of HbA1c and HDL during the follow-up period³². The spatial distribution of CCA is conspicuously apparent in the upper, middle, and specific lower regions of the studied area. This distribution aligns with a previous spatial analysis of CCA in northeast Thailand conducted in 2019, which identified significant clusters of CCA in the upper part (Nong Bua Lamphu, Udon Thani, Sakon Nakhon, and Nakhon Phanom Provinces), middle part (Maha Sarakham and Khon Kaen Provinces), and lower part (Chiyaphum Province)²⁰.

On the contrary, *O. viverrini* infection and DM exhibit a widespread distribution throughout the region, consistent with a study conducted in 2021 that demonstrated a high proportion of *O. viverrini* infection and DM in a high-risk group of CCA in northeastern Thailand¹⁸. Moreover, in our study, the spatial distribution of CCA, *O. viverrini* infection, and DM was notably elevated, particularly in proximity to main water sources such as the Mekong, Songkhram, Chi, and Mun rivers. This finding aligns with previous research associating the prevalence of *O. viverrini* with populations residing near water sources. This observed pattern corresponds with earlier investigations, such as a 2019 study that identified a large high-risk area cluster of CCA in provinces located in the Chi and Mun River Basin, with more isolated high-risk sub-districts found in provinces situated in the Mekong River Basin²⁰. Additionally, a study in 2020 revealed a significant association between geographic factors and *O. viverrini* infection³³. Furthermore, a comprehensive review in 2018, which systematically searched five international and seven Thai research databases to identify studies relevant to risk factors for CCA in the Lower Mekong Region, recommended addressing *O. viverrini* infection as a key strategy to mitigate CCA incidence in affected regions³⁴.

Our findings on Bayesian spatial analysis revealed a pronounced cluster of CCA in Maha Sarakham Province, which is situated between the Chi and Mun rivers in the middle part of the region. Subsequent notable clusters were identified in Khon Kaen and Kalasin, which are also located in the middle part, as well as provinces in the upper part (Nong Bua Lamphu, Udon Thani, Nakhon Phanom, and Sakon Nakhon), and provinces in the lower part (Surin and Chiyaphum). These findings align with prior research that has similarly identified the prevalence and spatial distribution of CCA in the northeastern region of Thailand^{19,20}.

One limitation of our study pertains to the data collection method for factors related to *O. viverrini* infection and DM, which relied on participant interviews without the physical examinations or diagnostic procedures. This approach introduces the potential for data reliability issues. Furthermore, this information bias could result in underestimation or overestimation of study results, potentially impacting the overall reliability of the study's conclusions. However, the substantial size of our study's sample mitigates this concern, as the large sample size contributes to reduced variance in the data. Consequently, this diminishes the likelihood of estimation errors and strengthens the reliability of our findings, promoting more conclusive and robust outcomes.

Conclusions

Our findings elucidate a pervasive distribution of CCA, *O. viverrini* infection, and DM across the entirety of the northeastern region of Thailand, exhibiting a congruent spatial pattern. Specifically, regions characterized by a high distribution of CCA often coincide with elevated occurrences of both *O. viverrini* infection and DM. Consequently, prioritizing CCA screening initiatives in areas exhibiting high rates of *O. viverrini* infection and prevalent DM should be accorded top priority. This targeted approach aligns with the observed spatial relationships, optimizing resource allocation for effective and region-specific health interventions.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Received: 18 December 2023; Accepted: 3 May 2024

Published online: 07 May 2024

References

- Sripa, B. & Pairojkul, C. Cholangiocarcinoma: lessons from Thailand. *Curr. Opin. Gastroenterol.* **24**, 349–356. <https://doi.org/10.1097/MOG.0b013e3282fb9b3> (2008).
- Chaiterakij, R. *et al.* Characteristics and outcomes of cholangiocarcinoma by region in Thailand: a nationwide study. *World J. Gastroenterol.* **23**, 7160–7167. <https://doi.org/10.3748/wjg.v23.i39.7160> (2017).
- Treepasertsuk, S. *et al.* A significant cancer burden and high mortality of intrahepatic cholangiocarcinoma in Thailand: a nationwide database study. *BMC Gastroenterol.* **17**, 3. <https://doi.org/10.1186/s12876-016-0565-6> (2017).
- Watanapa, P. & Watanapa, W. B. Liver fluke-associated cholangiocarcinoma. *Br. J. Surg.* **89**, 962–970. <https://doi.org/10.1046/j.1365-2168.2002.02143.x> (2002).
- Poomphakwaen, K. *et al.* Risk factors for cholangiocarcinoma in Khon Kaen, Thailand: a nested case-control study. *Asian Pac. J. Cancer Prev.* **10**, 251–258 (2009).
- Songserm, N. *et al.* Risk factors for cholangiocarcinoma in high-risk area of Thailand: role of lifestyle, diet and methylenetetrahydrofolate reductase polymorphisms. *Cancer Epidemiol.* **36**, e89–94. <https://doi.org/10.1016/j.canep.2011.11.007> (2012).
- Sithithaworn, P., Yongvanit, P., Duenngai, K., Kiatsopit, N. & Pairojkul, C. Roles of liver fluke infection as risk factor for cholangiocarcinoma. *J. Hepatobiliary Pancreat. Sci.* **21**, 301–308. <https://doi.org/10.1002/jhbp.62> (2014).
- Aung, W. P. P. *et al.* First report and molecular identification of *Opisthorchis viverrini* infection in human communities from Lower Myanmar. *PLoS ONE* **12**, e0177130. <https://doi.org/10.1371/journal.pone.0177130> (2017).
- Sithithaworn, P. *et al.* The current status of opisthorchiasis and clonorchiasis in the Mekong Basin. *Parasitol. Int.* **61**, 10–16. <https://doi.org/10.1016/j.parint.2011.08.014> (2012).
- Sripa, B. *et al.* Opisthorchiasis and opisthorchis-associated cholangiocarcinoma in Thailand and Laos. *Acta Trop.* **120**(Suppl 1), S158–168. <https://doi.org/10.1016/j.actatropica.2010.07.006> (2011).
- Prasongwatana, J., Laummaunwai, P., Boonmars, T. & Pinlaor, S. Viable metacercariae of *Opisthorchis viverrini* in northeastern Thai cyprinid fish dishes—as part of a rational program for control of *O. viverrini*-associated cholangiocarcinoma. *Parasitol. Res.* **112**, 1323–1327. <https://doi.org/10.1007/s00436-012-3154-9> (2013).
- Grundy-Warr, C. *et al.* Raw attitudes, wetland cultures, life-cycles: socio-cultural dynamics relating to *Opisthorchis viverrini* in the Mekong Basin. *Parasitol. Int.* **61**, 65–70. <https://doi.org/10.1016/j.parint.2011.06.015> (2012).
- Thaewngiew, K. *et al.* Prevalence and risk factors for *Opisthorchis viverrini* infections in upper Northeast Thailand. *Asian Pac. J. Cancer Prev.* **15**, 6609–6612. <https://doi.org/10.7314/apjcp.2014.15.16.6609> (2014).
- Saengsawang, P., Promthet, S. & Bradshaw, P. Reinfection by *Opisthorchis viverrini* after treatment with praziquantel. *Asian Pac. J. Cancer Prev.* **17**, 857–862. <https://doi.org/10.7314/apjcp.2016.17.2.857> (2016).
- Kamsa-ard, S. *et al.* Decreasing trends in cholangiocarcinoma incidence and relative survival in Khon Kaen, Thailand: an updated, inclusive, population-based cancer registry analysis for 1989–2018. *PLOS ONE* **16**, e0246490. <https://doi.org/10.1371/journal.pone.0246490> (2021).
- Kamsa-ard, S. *et al.* Risk factors for cholangiocarcinoma in thailand: a systematic review and meta-analysis. *Asian Pac. J. Cancer Prev.* **19**, 605–614. <https://doi.org/10.22034/apjcp.2018.19.3.605> (2018).
- Tyson, G. L. & El-Serag, H. B. Risk factors for cholangiocarcinoma. *Hepatology* **54**, 173–184. <https://doi.org/10.1002/hep.24351> (2011).
- Thinkhamrop, K. *et al.* Association of comorbidity between *Opisthorchis viverrini* infection and diabetes mellitus in the development of cholangiocarcinoma among a high-risk population, northeastern Thailand. *PLoS Negl. Trop. Dis.* **15**, e0009741. <https://doi.org/10.1371/journal.pntd.0009741> (2021).
- Thinkhamrop, K. *et al.* Spatial analysis of hepatobiliary abnormalities in a population at high-risk of cholangiocarcinoma in Thailand. *Sci. Rep.* **10**, 16855. <https://doi.org/10.1038/s41598-020-73771-0> (2020).
- Suwannatrai, A. T. *et al.* Bayesian spatial analysis of cholangiocarcinoma in Northeast Thailand. *Sci. Rep.* **9**, 14263. <https://doi.org/10.1038/s41598-019-50476-7> (2019).
- Li, J., Han, T., Xu, L. & Luan, X. Diabetes mellitus and the risk of cholangiocarcinoma: an updated meta-analysis. *Prz. Gastroenterol.* **10**, 108–117. <https://doi.org/10.5114/pg.2015.49004> (2015).
- Tsai, M. S., Lee, P. H., Lin, C. L., Peng, C. L. & Kao, C. H. Type II diabetes mellitus is associated with a reduced risk of cholangiocarcinoma in patients with biliary tract diseases. *Int. J. Cancer* **136**, 2409–2417. <https://doi.org/10.1002/ijc.29292> (2015).
- Saengboonmee, C., Seubwai, W., Lert-Itthiporn, W., Sanlung, T. & Wongkham, S. Association of Diabetes mellitus and cholangiocarcinoma: update of evidence and the effects of antidiabetic medication. *Can. J. Diabetes* <https://doi.org/10.1016/j.cjcd.2020.09.008> (2020).
- Jing, W. *et al.* Diabetes mellitus and increased risk of cholangiocarcinoma: a meta-analysis. *Eur. J. Cancer Prev.* **21**, 24–31. <https://doi.org/10.1097/CEJ.0b013e3283481d89> (2012).
- Liu, R. Q. *et al.* Prognosis of the intrahepatic cholangiocarcinoma after resection: hepatitis B virus infection and adjuvant chemotherapy are favorable prognosis factors. *Cancer Cell. Int.* **13**, 99. <https://doi.org/10.1186/1475-2867-13-99> (2013).
- Thinkhamrop, K. *et al.* Associations between ultrasound screening findings and cholangiocarcinoma diagnosis in an at-risk population. *Sci. Rep.* **12**, 13513. <https://doi.org/10.1038/s41598-022-17794-9> (2022).
- Khuntikeo, N. *et al.* Cohort profile: cholangiocarcinoma screening and care program (CASCAP). *BMC Cancer* **15**, 459. <https://doi.org/10.1186/s12885-015-1475-7> (2015).
- Alene, K. A., Viney, K., McBryde, E. S. & Clements, A. C. Spatial patterns of multidrug resistant tuberculosis and relationships to socio-economic, demographic and household factors in northwest Ethiopia. *PLoS ONE* **12**, e0171800. <https://doi.org/10.1371/journal.pone.0171800> (2017).
- Restrepo, A. C., Baker, P. & Clements, A. C. National spatial and temporal patterns of notified dengue cases, Colombia 2007–2010. *Trop. Med. Int. Health* **19**, 863–871. <https://doi.org/10.1111/tmi.12325> (2014).
- Moungthard, H. *et al.* Association of cholangiocarcinoma among people with periductal fibrosis. *Asian Pac. J. Cancer. Prev.* **24**, 4147–4154. <https://doi.org/10.31557/APJCP.2023.24.12.4147> (2023).
- Tawarungruang, C. *et al.* Survival after surgery among patients with cholangiocarcinoma in Northeast Thailand according to anatomical and morphological classification. *BMC Cancer* **21**, 497. <https://doi.org/10.1186/s12885-021-08247-z> (2021).
- Muthukumar, R. *et al.* Effects of *Opisthorchis viverrini* infection on glucose and lipid profiles in human hosts: a cross-sectional and prospective follow-up study from Thailand. *Parasitol. Int.* **75**, 102000. <https://doi.org/10.1016/j.parint.2019.102000> (2020).
- Pengput, A. & Schwartz, D. G. Risk factors for *Opisthorchis viverrini* infection: a systematic review. *J. Infect. Public Health* **13**, 1265–1273. <https://doi.org/10.1016/j.jiph.2020.05.028> (2020).
- Steele, J. A. *et al.* Thinking beyond *Opisthorchis viverrini* for risk of cholangiocarcinoma in the lower Mekong region: a systematic review and meta-analysis. *Infect. Dis. Poverty* **7**, 44. <https://doi.org/10.1186/s40249-018-0434-3> (2018).

Acknowledgements

The authors are truly thankful for all members of CASCAP, particularly the cohort members and staff from all participating institutions including the Ministry of Public Health, Ministry of Interior, and Ministry of Education of Thailand. This research was supported by National Science, Research and Innovation Fund (NSRF) under

the Basic Research Fund of Khon Kaen University through Cholangiocarcinoma Research Institute. Additional support was also granted by the Fundamental Fund of Khon Kaen University which has received funding support from the NSRF (Project ID 161735).

Author contributions

K.T., K.S., M.K., and A.T.S. conceived and designed this study. K.T. and A.T.S. performed the analysis. K.T., K.S., M.K., and A.T.S. wrote the manuscript. K.T. and A.T.S. collected the data and generated the clinical database. All authors have been involved in revising the manuscript, and all authors have read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-61282-1>.

Correspondence and requests for materials should be addressed to K.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024