A GEOGRAPHIC INFORMATION SYSTEM FOR ESTIMATING THE PREVALENCE OF HELMINTHIASIS INFECTIONS IN THE NORTH-EASTERN THAILAND

Dr. Vladimir Buntilov*[1], Dr. Suwannachai Wattanayingcharoenchai[2],
Dr. Choosak Nithikathkul[3] and Dr. Thitima Wongsaroj[4]

[1] Visiting Professor, Lecturer. Department of Computer Engineering, Mahidol University
25/25 Puttamonton 4 Rd., Salaya, Nakhon Pathom, 73170, Thailand
Tel.: +66 2 889-2138 ext. 6255; E-mail: egbvd@mahidol.ac.th

[2,4] Department of Disease Control
Ministry of Public Health, Nonthaburi, 11000, Thailand
Tel.: +66 2 5918437; E-mail: tmthelma1@yahoo.com

[3] Assistant Professor, Faculty of Medicine, Pre-clinic Department, Mahasarakham University
Mahasarakham, 44150, Thailand
Tel.: +66 43 712992; E-mail: nithikathkul@yahoo.com

KEY WORDS: GIS, public health, Opisthorchis viverrini

ABSTRACT: Human helminthiasis parasites are known to be endemic in several countries in South-eastern Asia. In particular, *Opisthorchis viverrini* (OV) Liver fluke possesses a considerable health threat to human population in North-eastern and Northern Thailand. The goal of the present investigation is to estimate the distribution of OV Liver fluke in humans for 20 provinces in North-eastern part of Thailand (known as Isan region). As a result, an intuitive graphical representation of the spatial distribution of OV parasite across North-eastern part of Thailand has been built. In addition, the prevalence of the parasite has been estimated for various administrative divisions of the country.

The current project utilises the prevalence distribution data of *Opisthorchis viverrini* helminth, received from a country-wide epidemiological study conducted by The Ministry of Public Health of Thailand (MOPH). The associated geographical data from MOPH as well as other publicly available spatial data were imported in a specially designed Geographical Information System (GIS). Major procedures of the project were implemented using open-source software programs.

The developed GIS was used to estimate the number of infected people on a region, province (changwat), district (amphoe) and sub district (tambon) basis. Initial estimates show that the overall prevalence of helminthiasis infection under study in North-eastern Thailand may exceed 7 million people, while the average infection rate constitutes 26%.

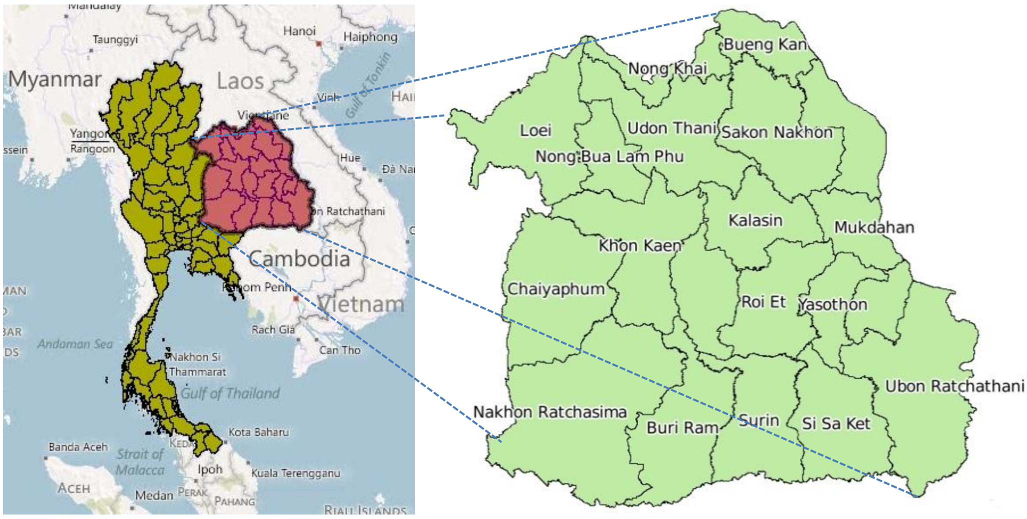
1. INTRODUCTION

Thailand is the fourth most populated country in South-eastern Asia totalling more than 65 million of inhabitants. *Opisthorchis viverrini* (OV) is a Liver fluke parasite which is endemic in several developing countries in SE Asia and constitutes a considerable threat to public health (Shin, et al., 2010; Sithithaworn and Haswell-Elkins, 2003). In The Kingdom of Thailand, this helminth is widely distributed among rural population in North-eastern and Northern parts of the country (Sithithaworn and Haswell-Elkins, 2003; Jongsuksuntigul and Imsomboon, 2003).

From January to May 2009 The Ministry of Public Health of Thailand conducted an extensive study of the distribution of helminths in the the country (Wongsaroj, 2010; Nithikathkul, et. al., 2010). The investigators performed the field studies across the country to collect the information about the infection rate. Besides the medical data recorded for every examined person, the geographic locations of the public health offices were identified as well, for which GPS-enabled devices were used. Upon collecting a sufficient number of specimen samples, the data were aggregated and the cumulative prevalence of the parasites was calculated for certain chosen areas.

The availability of the associated geographical coordinates allowed the researchers to utilise the power of Geographic Information System (GIS) methods for studying the distribution of the parasites. The parasites' prevalence information was imported into a specially designed GIS. The developed system includes several layers of various spatial data: Thailand's administrative division at different levels, land use, population density distribution and other geographical and environmental indicators. The system was built in SAGA 2.0 software (SAGA, 2011), an open-source GIS program. Additional calculations for the project were performed using a Matlab-like language in the open-source GNU Octave environment (GNU Octave, 2011). Visualisation was partly imlpemented with the help of Quantum GIS software (Quantum GIS, 2011).

The geographically limited point-like data from laboratory tests were extrapolated to build the parasite's prevalence maps for the entire territory of the country. Afterwards, the available GIS layers were used to estimate the number of infected people on a region, province (changwat), district (amphoe) and sub district (tambon) basis. Current project limits the investigation to the North-eastern part of the kingdom, shown in Figure 1a,b. This area, known as Isan region, consists of 20 provinces and remains Thailand's poorest region.



(a) (b)

Figure 1: (a) The Kingdom of Thailand, (b) Isan region – the North-eastern part of Thailand consisting of 20 provinces.

The manuscript is organised as follows: The description of the utilised spatial data, the procedures of building the GIS are provided in Section 2. The results of the estimation of the infected population as well as the infection rate for each province of North-eastern Thailand are shown and discussed in Section 3. Finally, Section 4 concludes the paper.

2. DATA AND METHODS

2.1 Epidemiological data

From January to May 2009 The Ministry of Public Health of Thailand (MOPH) conducted an epidemiological study of the national helminthiasis situation throughout the country in the 75 provinces of Thailand (Wongsaroj, 2010; Nithikathkul, et. al., 2010). The investigators from MOPH performed the field studies across the country to collect the information about the prevalence of parasitic infections. The laboratory tests were carried out in the field at provincial levels by MOPH expert teams. Besides stool examination results from individual people, the geographic locations of the district public health offices were identified as well, for which GPS-enabled devices were used.

The medical records from MOPH were received as Microsoft Excel spreadsheet files. After converting them to CSV (Comma Separated Values) files, the data were imported into a GIS. Although the records contain various geographical, environmental and medical data, the current project utilises only geographical coordinates of health offices and the value of parasites' infection rate. The result of mapping the infection rate of OV parasite in North-Eastern Thailand is shown in Figure 2. The mapping was performed with the help of Quantum GIS software (Quantum GIS, 2011).

2.2 Spatial data

The following publicly available spatial data were used in the project:

1. **Administrative boundaries of The Kingdom of Thailand**. The vector layers containing the boundaries at various degrees of details, i.e. country, province (changwat), district (amphoe) and sub-district (tambon), were obtained from the database published by DIVA-GIS project (DIVA-GIS, 2011), while the data originated from Global Administrative Areas project (GADM) (GADM, 2011).

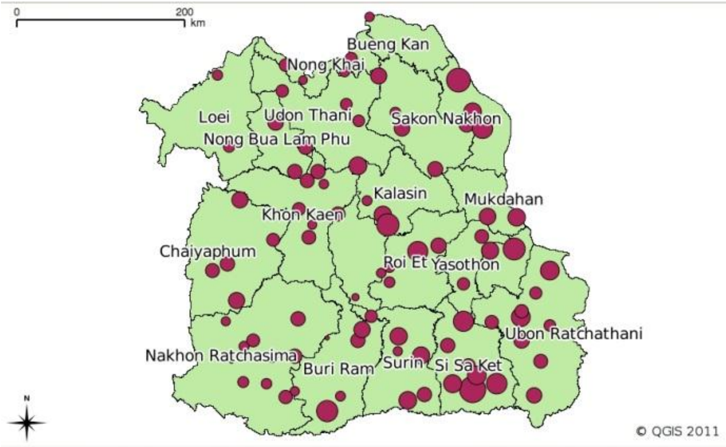


Figure 2: The prevalence of OV parasite for Isan region. The data are provided by The Ministry of Public Health of Thailand (MOPH). The circles show geographical locations of the public health offices in which laboratory tests were carried out. The areas of the circles are proportional to the detected positive infection rate.

2. **Distribution of human population**. Each pixel of the raster map represents the number of people living in the area corresponding to a pixel. The map was obtained from DIVA-GIS project (DIVA-GIS, 2011), which in turn relies on the Gridded Population of the World (GPW) project of the Center for International Earth Science Information Network (CIESIN) (CIESIN, 2005). Although the origin of the population data for Thailand is not specified explicitly in the downloaded files, it is likely that the data originated from the national census of Thailand conducted in 2000.

3. **Land cover**. The pixels in the raster map are categorised into 22 classes representing various land cover. The map was obtained from DIVA-GIS project (DIVA-GIS, 2011), while the data originated from Global Land Cover 2000 project (GLC 2000, 2011).

All raster data were resampled to a common spatial resolution (30 arc-seconds) and cropped to match the extent of the country.

2.3 Procedures

In order to estimate the number of infected people for each administrative division, a map of the parasites' infection rate was obtained. For this, the values of the infection rate from the imported epidemiological records (described in Section 2.1) were extrapolated to the entire territory of Thailand using inverse distance weighted method. Because of the sparsity of the data points, a search radius of 100 km was used. The maximum number of points to be analysed was limited to 10. The result of the infection rate extrapolation with the overlaid original points for North-eastern Thailand is shown in Figure 3.

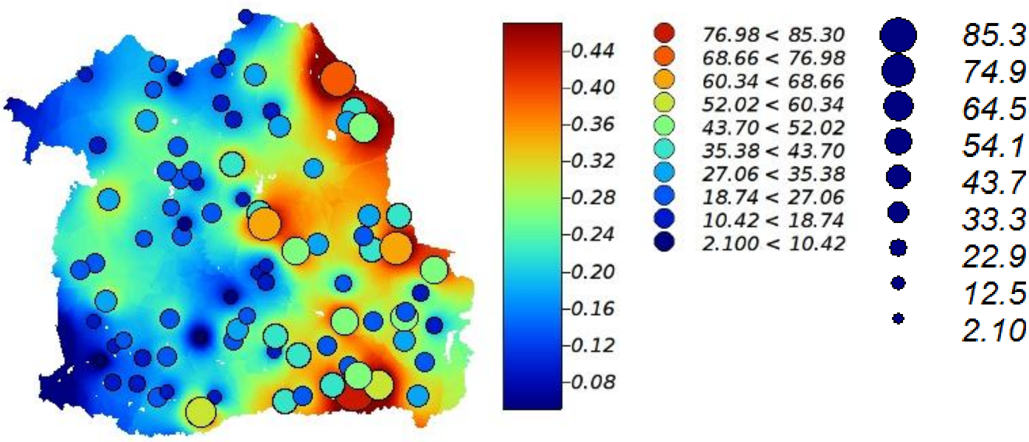


Figure 3: The estimated infection rate map of OV parasite with the overlaid original points of prevalence for the North-eastern part of Thailand. Each pixel in the map represents the percentage of infected people living in an area corresponding to 30×30 arc-second^2 (≈1km^2).

The type of helminths under analysis (*Opisthorchis viverrini*) is known to be widely distributed among people in the rural areas of Thailand because of their habit to certain traditional dishes made from raw fish. The study of MOPH, from which the medical data are used in the project, was focused on the distribution of parasites in rural areas of the kingdom. Therefore, the accuracy of the extrapolated infection rate for urban areas may be low. In addition, the urban areas differ from the rural regions by much higher population densities. Thus, the result of estimating the infection prevalence for urban areas may vary significantly because of inaccuracies of the infection rate map.

To limit the investigation solely to rural areas of the country, a mask of pixels was constructed as following. After examining the available data, it was found that the areas under consideration can be well characterised by relatively low population density and an appropriate type of land cover. A value of 1000 persons per area covered by one pixel, i.e. 30×30 arc-second^2, was chosen as a threshold between the urban and rural areas, while the pixels which are classified as “Water bodies”, “Snow and Ice” and “Artificial surfaces and associated areas” in the land cover map were excluded from the constructed mask.

Afterwards, the values from the map of the infection rate were multiplied by the corresponding values from the population distribution map for the identified rural areas. The multiplication was performed on a pixel by pixel basis, producing a map which shows for each pixel an estimated number of infected persons living in an area corresponding to a pixel. The constructed map for the North-eastern part of Thailand is shown in Figure 4.

Majority of the procedures described in Section 2.2 and Section 2.3 were performed using SAGA GIS software (SAGA, 2011).

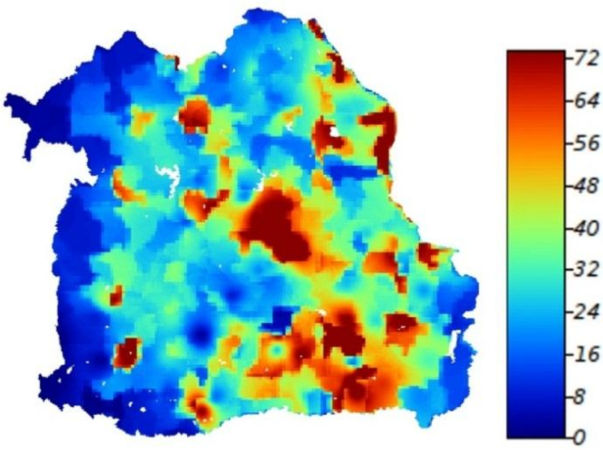


Figure 4: The estimated map of infected population for the North-eastern part of Thailand. Each pixel in the map represents a number of infected people living in an area corresponding to 30×30 arc-second^2 (≈1km^2).

3. RESULTS

The available layers of administrative boundaries together with the distribution map shown in Figure 4 were used to calculate the prevalence of infections for each province, district and sub-district of the country. The results of the estimation for 20 provinces of Isan are shown in Table 1, while the results for the districts and sub-districts are omitted for the sake of brevity.

As can be seen from Table 1, the estimated number of infected people living in Isan region of Thailand is close to 6.5 million (from approximately 25 million total estimated population of Isan). This gives 26% as the positive infection rate for North-eastern part of Thailand.

The numbers of infected people per province amount to hundreds of thousands of people for some provinces. For instance, the estimated infected population of Si Sa Ket and Ubon Ratchathani provinces are 663,198 and 588,022 persons, which makes them two most infected provinces in Isan (in terms of the number of infected people).

The fourth column in Table 1 contains the values of positive infection rate, which are the numbers of infected people normalised with respect to the total estimated population of a province. Nakhon Phanom and Si Sa Ket provinces show the highest percentage of infected people, i.e. approximately 46% and 40%, respectively.

Table 1. Estimation of OV parasite prevalence for 20 provinces in North-eastern Thailand. The highlighted are two highest values of the number of infections and infection rate. The estimation is performed using the human population grid from CIESIN (likely to originate from the national census of Thailand in 2000). The last two columns show the estimations, taking the population growth into consideration.

Province name	***population,* [people]**	***total number of infections,* [people]**	***infection rate,* [%]**	***population* [people]** *(adjusted to 2011)*	***total number of infections* [people]** *(adjusted to 2011)*
Amnat Charoen	468,432	175,078	37	510,819	190,920
Bueng Kan	450,472	130,272	29	491,234	142,060
Buri Ram	1,731,518	385,780	22	1,888,197	420,688
Chaiyaphum	1,321,997	310,508	23	1,441,620	338,605
Kalasin	1,151,492	389,090	34	1,255,686	424,297
Khon Kaen	1,980,162	413,911	21	2,159,340	451,364
Loei	753,522	125,492	17	821,705	136,847
Maha Sarakham	1,105,458	243,965	22	1,205,487	266,041
Mukdahan	403,421	143,780	36	439,925	156,791
Nakhon Phanom	786,376	363,969	46	857,532	396,903
Nakhon Ratchasima	2,980,121	457,177	15	3,249,781	498,545
Nong Bua Lam Phu	598,596	146,171	24	652,761	159,398
Nong Khai	602,437	95,426	16	656,949	104,061
Roi Et	1,543,466	464,083	30	1,683,129	506,076
Sakon Nakhon	1,274,743	374,355	29	1,390,090	408,229
Si Sa Ket	1,656,480	663,198	40	1,806,369	723,208
Surin	1,586,340	448,524	28	1,729,882	489,109
Ubon Ratchathani	1,957,526	588,022	30	2,134,655	641,230
Udon Thani	1,773,798	371,304	21	1,934,302	404,901
Yasothon	636,175	192,186	30	693,740	209,576
total/overall:	**24,762,532**	**6,482,290**	**26**	**27,003,202**	**7,068,849**

It must be noted that the estimated values must be used with caution since the accuracy of the results is difficult to assess. Firstly, the point-like data about the prevalence of the parasite provided by MOPH is relatively sparse, taking into account the large area of the geographical region under consideration. Therefore, straightforward extrapolation of the numerical values of the prevalence to the entire territory of North-eastern Thailand may give inexact results. A denser grid of sample points could be used to increase the accuracy of the approximation. However, at the time of writing additional data points were not available.

The second significant source of inaccuracies may be the human population map described in Section 2.2. Apparently, this map relies on the national census conducted in 2000 (NSO, 2011). Simple GIS procedures to calculate the total population in the kingdom using this map result in 60,274,049 people, which roughly agrees with the official data (60,606,947 people in 2000). Thus, it is reasonable to assume that the estimations of the total population of Isan and the populations of each province (shown in the second column of Table 1) provide realistic values. However, taking into account 0.87% annual population growth rate for North-eastern Thailand (NSO, 2011), the population of Isan region at the time of manuscript writing (2011) may constitute $(1+0.0087)^{10}*24{,}762{,}532=27{,}003{,}202$ people, and, as a result, the estimated total number of infected people may equal 7,068,849. The last two columns in Table 1 provide the recalculated estimated values of the total number of inhabitants and the infected population for each province taking the population growth into consideration.

4. CONCLUSIONS

This study aims at estimating the prevalence of *Opisthorchis viverrini* parasite among human population in rural areas of North-eastern part of the Kingdom of Thailand (Isan). The original data containing the geographical coordinates of the provincial public health offices together with the detected infection rates were provided by Ministry of Public Health of Thailand (MOPH).

The numerical values of the prevalence of the parasites were extrapolated to the entire territory of Thailand using GIS software. Afterwards, the extrapolated values were combined with several publicly available layers of spatial information to estimate the number of infected people in North-eastern part of Thailand (Isan) for each province, district and sub-district. The estimated population of infected people living in Isan is close to 6.5 million. However, the use of the adjusted demographic data suggests that this number may exceed 7 million people in 2011. The overall infection rate is estimated to be 26%. Si Sa Ket and Ubon Ratchathani are the two provinces with the highest amount of infected people, i.e. 663,198 and 588,022 (723,208 and 641,230 for the adjusted for year 2011 demographic data), respectively. The highest infection rate is detected in Nakhon Phanom and Si Sa Ket provinces: 46% and 40%, respectively.

ACKNOWLEDGEMENTS

The corresponding author would like to thank the clergy of Christian Orthodox Church in Thailand for spiritual and moral support.

The corresponding author expresses his gratitude to Mahidol University for financial aids and opportunities for professional development.

The data for the experiments were provided by The Ministry of Public Health of Thailand (MOPH).

REFERENCES

CIESIN, 2005. Center for International Earth Science Information Network (CIESIN), Columbia University; and Centro Internacional de Agricultura Tropical (CIAT). Gridded Population of the World Version 3 (GPWv3): Population Grids. Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University. Online: http://sedac.ciesin.columbia.edu/gpw

DIVA-GIS, 2011. Online: http://www.diva-gis.org

GADM, 2011. Global Administrative Areas project. Online: http://www.gadm.org

GLC 2000, 2011. European Commission, Joint Research Centre. Global Land Cover 2000 Project (GLC 2000) database. Online: http://ies.jrc.ec.europa.eu/global-land-cover-2000

GNU Octave, 2011. Online: http://www.octave.org

Jongsuksuntigul, P. and Imsomboon, T., 2003. Opisthorchiasis control in Thailand. Acta Tropica, 88(3), pp. 229 - 232.

Nithikathkul, C., Royal, L., Wongsaroj, T., Sangthong, S., Bootjinda, T., Reungsang, P., 2010. Opisthorchis Viverrini: Mapping And Surveillance Study In Northeast Thailand. In proceedings of 10th International Conference on Molecular Epidemiology and Evolutionary Genetics of Infectious. 3-5 November 2010. Amsterdam, The Netherlands.

NSO, 2011. Economic and Social Statistics Bureau, National Statistical Office of Thailand. Population and Housing Census 2000. Online: http://web.nso.go.th/eng/en/pop2000/prelim_e.htm

Quantum GIS, 2011. Online: http://qgis.org

SAGA, 2011. SAGA - System for Automated Geoscientific Analyses. Online: http://www.saga-gis.org

Shin, H.-R., Oh, J.-K., Masuyer, E., Curado, M.-P., Bouvard, V., Fang, Y.-Y., et al., 2010. Epidemiology of cholangiocarcinoma: An update focusing on risk factors. Cancer Science, 101(3), pp. 579-585.

Sithithaworn, P. and Haswell-Elkins, M., 2003. Epidemiology of Opisthorchis viverrini. Acta Tropica, 88(3), pp. 187-194.

Wongsaroj, T., 2010. National survey of helminthiasis and health behavior associated with geographic information in Thailand. Presentation, Joint International Tropical Medicine Meeting 2010 and International Malaria Colloquium 2010.