Chapter 3

Using a logistic regression to model the prevalence of imposex

This chapter presents a preliminary data analysis and the results from fitting the logistic regression model for prevalence of imposex in gastropods in the Gulf of Thailand in 2006. Section 3.1 is the description of variables. Section 3.2 is data characteristics. Section 3.3 is the logistic regression model to imposex outcome. The results presents in this chapter also appeared in Swennen et al (2009) and Khongchouy and Sampantarak (2010).

3.1 Description of variables

The roles of variables are classified as the determinants and outcome. The outcome of interest is imposex (normal female/or female showing imposex), which is binary data type. The determinant (sometimes called the "study factor") is thirteen areas in the Gulf of Thailand where imposex in gastropods are measured. There is also a covariate of determinant: sixteen species groups of gastropods assemblages. Both of them (determinant and covariate) are the nominal data type. The target population is all gastropod assemblages that were collected in the Gulf of Thailand in 2006. Only female gastropods were used for the statistical analysis. These variables and their roles, data types, and categories are listed in Table 3.1.

Table 3.1: Variables and their role, data type, and categories

Role	Variable	Type	Categories
Determinant	Area	Nominal	A: Rayong, B: Namrin, C: Pattaya, D: Si Racha, E: Phet Buri, F: Ban Khau, G: Songkhla, H: Bang Tawa, I: Rusamilae, J: Laem Nok, K: Pattani Bay Mouth, L: Panare, and M: Tak Bai
Covariate	Species group	Nominal	1: *Murex trapa*, 2: *Murex altispira*, 3: *Murex occa*, 4: *Lataxiena blosvillei*, 5: *Semiricinula muricoides*, 6: *Thais bitubercularis*, 7: *Thais lacera*, 8: *Morula musiva*, 9: *Babylonia areolata*, 10: *Nassarius jacksonianus*, 11: *Nassarius siquijorensis*, 12: *Nassarius stolatus*, 13: *Pugilina cochlidium*, 14: *Hemifusus ternatanus*, 15: *Turricula javana*, and 16: other
Outcome	Imposex	Binary	Normal female, Imposex female

3.2 Characteristics of the data

A total of 8,757 specimens were collected in various species and sampling sites in the Gulf of Thailand; 5,044 specimens were female, and 25.2% (n = 3,713) of total females were imposex. Specimens were classified into twenty two species belonging to five families (Muricidae, Buccinidae, Nassariidae, Melongenidae, and Turridae). The female gastropods heights varied between 11.0-131.0 mm. The minimum height of 11.0 mm was found in *Nassaria pusilla* while the maximum height of 131.0 mm was found in *Hemifusus ternatanus*. The names of the species and families are presented in Table 3.2, which also summaries the number and percentage of females, and also percentage of females showing imposex.

Table 3.2: Lists of families and species of the gastropods, with a summary of total number of sample (nTotal), number of females (nFem), percentage of females (%Fem), and percentage of females showing imposex (%Imp) in 2006

Families and species	nTotal	nFem	%Fem	%Imp
Family Buccinidae				
Babylonia areolata (Link, 1807)	314	158	50.3	1.3
Nassaria pusilla (Röding, 1798)	67	53	79.1	1.9
Phos senticosus (Linnaeus, 1758)	2	1	50.0	0.0
Family Muricidae				
Chicoreus banksii (Sowerby II, 1841)	58	39	67.2	12.8
Lataxiena blosvillei (Deshayes, 1832)	721	366	50.8	88.8
Morula musiva (Kiener, 1835)	324	161	49.7	0.6
Murex altispira Ponder & Vokes, 1988	359	286	79.7	0.7
Murex occa Sowerby II, 1824	888	417	47.0	40.8
Murex trapa Röding, 1798	1,156	698	60.4	22.9
Rapana rapiformis (Von Born, 1778)	85	32	37.6	6.3
Semiricinula muricoides (De Blainville, 1832)	753	521	69.2	50.9
Thais bitubercularis (Lamarck, 1822)	427	214	50.1	14.0
Thais clavigera (Kuester, 1860)	75	36	48.0	25.0
Thais lacera (Von Born, 1778)	587	306	52.1	21.9
Thais rufotincta Tan & Sigurdsson, 1996	22	10	45.5	10.0
Family Melongenidae				
Hemifusus ternatanus (Gmelin, 1791)	368	225	61.1	15.6
Pugilina cochlidium (Linnaeus, 1758)	923	413	44.7	9.4
Family Nassariidae				
Nassarius jacksonianus (Quoy & Gaimard, 1833)	233	166	71.2	10.8
Nassarius livescens (Philippi, 1849)	62	29	46.8	6.9
Nassarius siquijorensis (A. Adams, 1852)	260	202	77.7	28.7
Nassarius stolatus (Gmelin, 1791)	533	306	57.4	4.9
Family Turridae				
Turricula javana (Linnaeus, 1767)	540	405	75.0	15.6
Total	8,757	5,044	57.6	25.2

The preliminary results showed that the highest proportion of imposex was found in *Lataxiena blosvillei* (88.8%) followed by *Semiricinula muricoides* (50.9%), Murex

occa (40.8%). Other species were as follows: *Nassarius siquijorensis* (28.7%), *Thais clavigera* (25.0%), *Murex trapa* (22.9%), *Thais lacera* (21.9%), *Hemifusus ternatanus* (15.6%), *Turricula javana* (15.6%), *Thais bitubercularis* (14.0%), *Chicoreus banksii* (12.8%), *Nassarius jacksonianus* (10.8%), *Thais rufotincta* (10.0%), *Pugilina cochlidium* (9.4%), *Nassarius livescens* (6.9%), *Rapana rapiformis* (6.3%), *Nassarius stolatus* (4.9%), *Nassaria pusilla* (1.9%), *Babylonia areolata* (1.3%), *Murex altispira* (0.7%), *Morula musiva* (0.6%), and *Phos senticosus* (which had no imposex females).

3.3 Logistic regression model to imposex outcome

Logistic regression was used to fit a model where the outcome event was the allocation to the treatment group and determinant (area) including covariate (species group) prior to treatment. A total of 56 sampling sites were grouped into 13 geographical areas. They are… A: Rayong, B: Namrin, C: Pattaya, D: Si Racha, E: Phet Buri, F: Ban Khau, G: Songkhla, H: Bang Tawa, I: Rusamilae, J: Laem Nok, K: Pattani Bay Mouth, L: Panare, and M: Tak Bai. (see more details in Appendix 4).

Gastropod species with fewer than 100 individuals (*C. banksii* (n = 58), *T. clavigera* (n = 75), *T. rufotincta* (n = 22), *R. rapiformis* (n = 85), *Nassaria pusilla* (n = 67), *P. senticosus* (n = 2), and *N. livescens* (n = 62)) were combined into a single group refer to as “other” because logistic regression cannot handle grouped data where any cell has zero count. *M. Trapa* (n = 1,156) is the species that had the largest sample size, so it was chosen as referent categories.

Table 3.3 gives the results from fitting a logistic regression model. The results show the evidence of the difference between the two treatment groups with respect to area

or species group. The model gave a residual deviance of 171.45 with 55 df based on the grouped data (the 84 combinations of area and species where samples were obtained). Chi-squared statistics for testing a common imposex frequency between areas and species groups were found to be 438.90 (12 df) and 453.04 (15 df) respectively, with the correspondence being highly statistically significant (Table 3.3).

The coefficients from the model can be related to the differential severity risks for each area and each species group. Imposex varies by area: area D (Si Racha) shows the highest risk with the maximum of coefficient equals to 1.13; second is area C (Pattaya) with the coefficient equals to 0.138. The minimum risk (coef = − 3.253) is found in area E (Phet Buri). As for variation by species, many species of Muricidae are very sensitive prone to imposex e.g. *L. blosvillei* appears to have the most imposex (coef = 1.674), the second is *M. occa* (coef = 0.697). *Morula musiva* seems to have the lowest sensitivity (coef = − 3.646).

From the results as presents in Table 3.3, the adjusted imposex frequency was calculated using formula 2.5 (Chapter 2) with the corresponding 95% confidence intervals (formula 2.2, Chapter 2). Table 3.4 presents imposex frequency before and after adjusting for species.

Before adjusting for the different sensitivities of the species, areas C (Pattaya) and D (Si Racha) had a high imposex frequency (62.7% and 73.6%, respectively), areas F (Ban Khau), H (Bang Tawa), I (Rusamilae), J (Laem Nok), and K (Pattani Bay Mouth) had medium frequency (18.5%, 17.4%, 15.6%, 31.0%, and 18.8%, respectively), whereas areas A (Rayong), B (Namrin), E (Phet Buri), G (Songkhla),

L (Panare), and M (Tak Bai) had low frequency (4.7%, 5.2%, 3.1%, 0.6%, 6.4%, and 1.3%, respectively).

Table 3.3: Results of the logistic model with outcome "Imposex"

Determinant	Estimate	Std. Error	*p*-value
Area			
A: Rayong	− 2.119	0.389	< 0.001
B: Namrin	− 1.429	0.369	< 0.001
C: Pattaya	0.138	0.231	0.551
D: Si Racha	1.138	0.186	< 0.001
E: Phet Buri	− 3.253	0.289	< 0.001
F: Ban Khau	− 1.139	0.164	< 0.001
G: Songkhla	− 2.815	1.110	0.011
H: Bang Tawa	− 1.322	0.214	< 0.001
I: Rusamilae	− 1.491	0.154	< 0.001
J: Laem Nok	− 0.721	0.190	< 0.001
K: Pattani Bay Mouth	− 1.140	0.230	< 0.001
L: Panare	− 1.253	0.462	0.007
M: Tak Bai	− 1.987	1.251	0.112
Species			
01: *Murex trapa* (reference)	0		
02: *Murex altispira*	− 2.961	0.834	< 0.001
03: *Murex occa*	0.697	0.186	< 0.001
04: *Lataxiena blosvillei*	1.674	0.271	< 0.001
05: *Semiricinula muricoides*	− 0.215	0.211	0.308
06: *Thais bitubercularis*	− 0.494	0.289	0.087
07: *Thais lacera*	0.097	0.203	0.633
08: *Morula musiva*	− 3.646	1.069	0.001
09: *Babylonia areolata*	− 2.398	1.243	0.054
10: *Nassarius jacksonianus*	− 0.700	0.291	0.016
11: *Nassarius siquijorensis*	0.373	0.253	0.141
12: *Nassarius stolatus*	− 1.350	0.308	< 0.001
13: *Pugilina cochlidium*	− 1.602	0.231	< 0.001
14: *Hemifusus ternatanus*	− 1.504	0.247	< 0.001
15: *Turricula javana*	− 0.323	0.208	0.120
16: other	− 0.775	0.327	0.018

Table 3.4: Prevalence of imposex (%) by area before and after adjusting for species

Area	Total females	Number of imposex	Unadjusted imposex (%)	Adjusted imposex (%)	
				Prevalence	95% CI
A: Rayong	190	9	4.7	9.0	4.9 - 13.0
B: Namrin	249	13	5.2	16.4	11.8 - 21.0
C: Pattaya	474	297	62.7	48.4	43.9 - 52.9
D: Si Racha	537	395	73.6	71.9	68.1 - 75.7
E: Phet Buri	417	13	3.1	3.1	1.4 - 4.7
F: Ban Khau	286	53	18.5	20.8	16.1 - 25.5
G: Songkhla	171	1	0.6	4.7	1.5 - 7.8
H: Bang Tawa	735	128	17.4	17.9	15.1 - 20.7
I: Rusamilae	916	143	15.6	15.6	13.2 - 17.9
J: Laem Nok	507	157	31.0	28.5	24.5 - 32.4
K: Pattani Bay Mouth	266	50	18.8	20.8	15.9 - 25.6
L: Panare	140	9	6.4	19.0	12.5 - 25.4
M: Tak Bai	156	2	1.3	10.1	5.4 - 14.8

CI: Confidence interval

After adjusting, ignoring species differences, areas C (Pattaya) and D (Si Racha) showed a lower frequency than before, but still had the highest values (48.4%, CI: 43.9%-52.9% and 71.9%, CI: 68.1%-75.7%, respectively), while areas B (Namrin) and L (Laem Nok) moved up to the medium severity level of imposex frequency (16.4%, CI: 11.8%-21.0% and 19.0%, CI: 12.5%-25.4%, respectively). Figure 3.1 shows the areas and the sampling sites with the adjusted imposex prevalence.

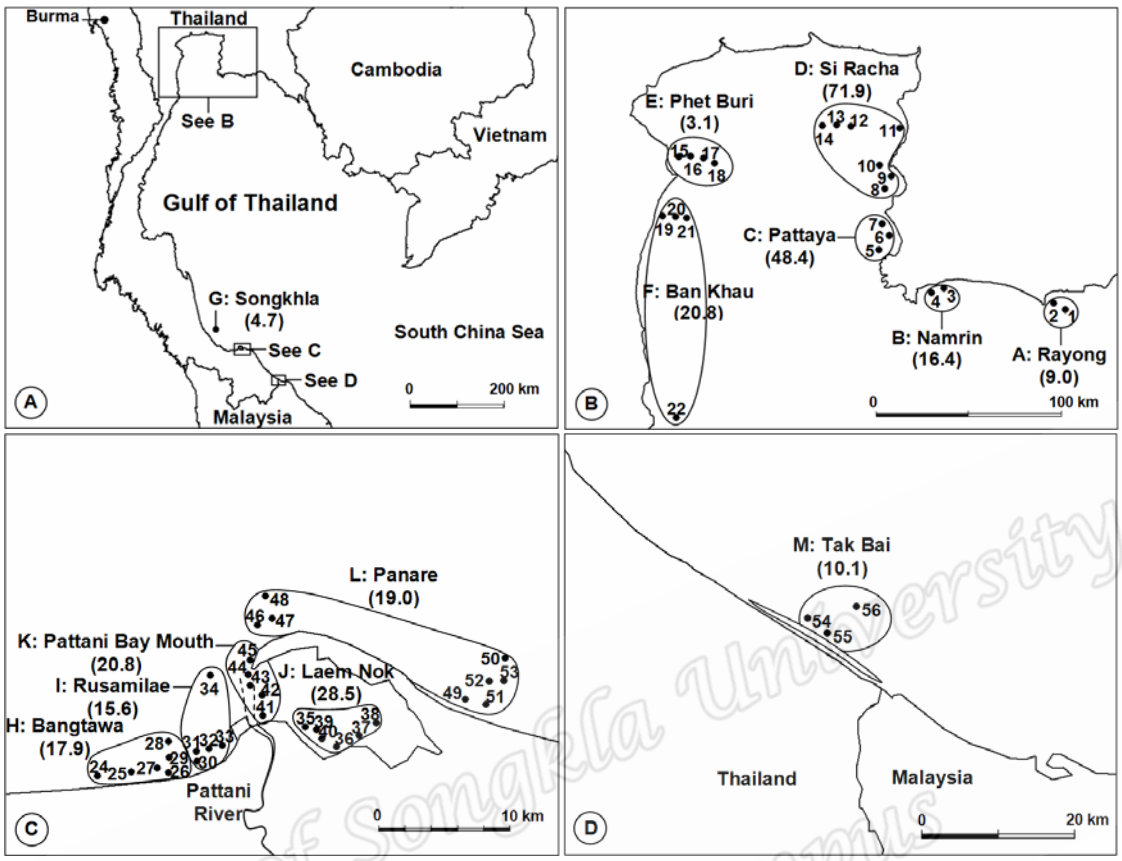

Figure 3.1: The study areas and sampling sites: Ⓐ is Overview; Ⓑ is northern part of the Gulf with study areas A-F; Ⓒ is Pattani Bay and surroundings with study areas H-L; Ⓓ is Narathiwat province with study area M. Dots with their number corresponds with the sampling sites in the detailed of Appendix 4. Value between brackets gives the mean of imposex frequency in the area

Table 3.5 presents the prevalence of imposex by species before and after adjusting for area. The sensitivity to imposex varied with species: *L. blosvillei* was found to be most prone to imposex (88.8%), followed by *S. muricoides* (50.9%), *M. occa* (40.8%) whereas *M. musiva* is least sensitive (0.6%). After adjusting for areas, many species showed a higher sensitivity than before adjusting except for *L. blosvillei* (69.7%, CI: 65.0%-74.4%), *S. muricoides* (25.8%, CI: 22.0%-29.6%), *P. cochlidium* (8.0%, CI:

5.4%-10.6%), and *H. ternatanus* (8.7%, CI: 5.1%-12.4%). Although *L. blosvillei* showed a lower value than before adjusting, it still was the highest sensitivity. In addition, *M. musiva* still was the lowest sensitivity prone to imposex.

Table 3.5: Prevalence of imposex (%) by species before and after adjusting for area

Species groups	Total females	Number of imposex	Unadjusted imposex (%)	Adjusted imposex (%)	
				Prevalence	95% CI
01: *Murex trapa*	698	160	22.9	30.1	26.7 - 33.5
02: *Murex altispira*	286	2	0.7	2.2	0.5 - 3.9
03: *Murex occa*	417	170	40.8	46.4	41.6 - 51.2
04: *Lataxiena blosvillei*	366	325	88.8	69.7	65.0 - 74.4
05: *Semiricinula muricoides*	521	265	50.9	25.8	22.0 - 29.6
06: *Thais bitubercularis*	214	30	14.0	20.8	15.4 - 26.3
07: *Thais lacera*	306	67	21.9	32.2	27.0 - 37.4
08: *Morula musiva*	161	1	0.6	1.1	0.0 - 2.7
09: *Babylonia areolata*	158	2	1.3	3.8	0.8 - 6.7
10: *Nassarius jacksonianus*	166	18	10.8	17.6	11.8 - 23.4
11: *Nassarius siquijorensis*	202	58	28.7	38.5	31.8 - 45.2
12: *Nassarius stolatus*	306	15	4.9	10.1	6.7 - 13.4
13: *Pugilina cochlidium*	413	39	9.4	8.0	5.4 - 10.6
14: *Hemifusus ternatanus*	225	35	15.6	8.7	5.1 - 12.4
15: *Turricula javana*	405	63	15.6	23.8	19.6 - 27.9
16: other	200	20	10.0	16.6	11.4 - 21.7

CI: Confidence interval