

Does age and latitude effect odds of parasite infestation?

The parasite *Lernaeocera branchialis*
on cod *Gadus morhua* in the Northwest Atlantic

Anna Heiða Ólafsdóttir

October 26th, 2005

B7932

The Data for Age

- Parasite either present or absent (binomial)
- Host cod age 1 – 16 years old
 - Use ages 2,3.....,10+
 - expecting greater odds of parasite with increased age
- Total 11.940 fishes (drop age1)
- Total w/parasite 1047 fishes

Age	Fish w/parasite	Total # fish
1	0	46
2	24	439
3	155	1669
4	233	2523
5	216	2012
6	183	1909
7	112	1159
8	75	1098
9	30	572
10-16	19	559

The Data for Latitude

- Sampling area NW-Atlantic:
N 45.00° – 54.00°
 - $\leq 46^\circ, 47^\circ, 48^\circ, 49^\circ, \geq 50^\circ$
 - expecting greater odds of parasite as move south
- Total 11,940 fishes
- Total w/parasite 1047 fishes
- Sampling period 1996 – 2000

Lat.	Fish w/parasite	Total # fish
45-46	5	596
47	280	3258
48	579	5594
49	42	302
50-54	141	2190

The Model

- **Odds** = $p/1-p$, p =proportion infected with parasite
- **Odds Ratio (OR)** = $\text{odds}_A/\text{odds}_B$ hence $\text{odds}_B = \text{OR} * \text{odds}_A$
 - Logistic regression considers the log odds as a function of explanatory variables: $\ln(\text{Odds}_B) = \text{Odds}_A + \ln(\text{OR})$
- **Formal model:** $\text{Odds} = e^{\beta_0} e^{\beta_{\text{age}}} e^{\beta_{\text{lat}}} e^{\beta_0 + \beta_{\text{age}}} + \text{error}$
- **Dependent variables**
 - odds of parasite
 - odds ratio of parasite
- **Explanatory variables**
 - “interaction latitude and age”
 - latitude
 - age
- **Odds and Odds Ratio in SAS**
 - e^{β_0} = odds of parasite for the intercept
 - $e^{\beta_{\text{age}}}$ = odds ratio relative to intercept
 - $e^{\beta_0 + \beta_{\text{age}}}$ = odds of parasite for age or latitude
- **Residuals:** categorical binomial data assume binomial res. do not check

Analysis Interaction Age & Lat.

The GENMOD Procedure Model Information

Data Set	SAVE.AGEGENMOD
Distribution	Binomial
Link Function	Logit-odds
Dependent Variable	LERNAE
Number of Observations Read	11940
Number of Observations Used	11940
Number of Events	1047
Number of Trials	11940

Class Level Information

Class	Levels	Values
LATDEC	5	46 47 48 49 50
AGE	9	2 3 4 5 6 7 8 9 10

Response Profile

Ordered	Total	
Value	LERNAE	Frequency
1	1	1047
2	0	10893

PROC GENMOD is modeling the probability that LERNAE='1'.

Results interaction lat-age

LR Statistics For Type 1 Analysis

Source	Deviance	DF	Chi-Square	Pr > ChiSq
Intercept	7096.1			
LATDEC	6979.5	4	116.64	<.0001
AGE	6893.4	8	86.05	<.0001
LATDEC*AGE	6853.1	30	40.33	0.0986

Results Latitude SAS output

Parameter	DF	Estimate	Standard Error	Wald 95% Confi. Limits		Chi-Square	Pr > ChiSq
Interc($\geq 50^\circ$)	1	-2.6763	0.0871	-2.8470	-2.5057	944.93	<.0001
LAT 49°	1	0.8533	0.1877	0.4854	1.2212	20.67	<.0001
LAT 48°	1	0.5175	0.0975	0.3264	0.7086	28.17	<.0001
LAT 47°	1	0.3121	0.1072	0.1021	0.5222	8.48	0.0036
LAT $\leq 46^\circ$	1	-2.0960	0.4575	-2.9926	-1.1994	20.99	<.0001

Odds intercept : $e^{-2.6763} = \mathbf{0.0688}$

odds of parasite in cod north of 50° are 1:0.07

Odds at 49° : $e^{-2.6763+0.8533} = \mathbf{0.1615}$

odds of parasite in cod at 49° are 1:0.16

Odds ratio of 49° relative to ≥ 50 : $e^{0.8533} = 2.3474$

or $\mathbf{0.1615/0.0688} = 2.3474$

Check odds calculations using excel

- Model: Odds= $((\#para/total)/(\#nopara/total))$
- Odds $\geq 50^\circ$: 141 para & 2050 nopara, total 2191 fishes
 - Odds = $((141/2190) / (2049/2190)) = \mathbf{0.0688}$
 - $0.0688 = e^{-2.6763}$
- Odds 49° : 42 para and 260 nopara, total 302 fishes
 - Odds = $((42/302) / (260/302)) = \mathbf{0.1615}$
 - $0.1615 = e^{-2.6763+0.8533}$
- Odds ratio 49° relative to $\geq 50^\circ$
 - $\mathbf{0.1615 / 0.0688 = 2.3474}$
 - $\mathbf{2.3474 = e^{-0.8533}}$

Confidence Limits Latitude

Parameter	DF	Estimate	Wald 95% Confi Limits		Chi-Square	Pr > ChiSq
Interc($\geq 50^\circ$)	1	-2.6763	-2.8470	-2.5057	944.93	<.0001
LAT 49	1	0.8533	0.4854	1.2212	20.67	<.0001
LAT 48	1	0.5175	0.3264	0.7086	28.17	<.0001
LAT 47	1	0.3121	0.1021	0.5222	8.48	0.0036
LAT 46	1	-2.0960	-2.9926	-1.1994	20.99	<.0001

CL odds for $\geq 50^\circ$: $e^{-2.8470}=0.0580$, $e^{-2.5057}=0.0816$

CL odds for 49° : $e^{-2.8470+0.4854}=0.0943$, $e^{-2.5057+1.2212}=0.2768$

Calculated Odds with Confidence Limits and Proportion (%)

LAT	Prop. (%)	Odds	Lower 95%CL	Upper 95%CL	P-value SAS
≥50°	6.44	1:0.0688	0.0580	0.0816	<.0001
49°	13.91	1:0.1615	0.0943	0.2768	<.0001
48°	10.35	1:0.1155	0.0804	0.1658	<.0001
47°	8.59	1:0.0941	0.0643	0.1376	0.0036
≤46°	0.84	1:0.0085	0.0029	0.0246	<.0001

P-value: significantly different from the **odds / odds ratio** of 1:1

Results Latitude

LR Statistics For Type 1 Analysis

Source	Deviance	DF	Chi-Square	Pr > ChiSq
Intercept	7104.5317			
LATDEC	6988.0277	4	116.64 ΔG	<.0001

Wald Statistics For Type 3 Analysis

Source	DF	Chi-Square	Pr > ChiSq
LATDEC	4	68.45	<.0001

- Only 1 explanatory variable: type1=type3
- LR=likelihood ratio
 - compares regression lines
 - usually more reliable than WALD
- Wald
 - compares odds ratios
 - needs very large sample size to be reliable

Results Age SAS output

Parameter	DF	Estimate	Standard Error	Wald 95% Confi. Limits		Chi-Square	Pr > ChiSq
Interc. age2	1	-2.8502	0.2099	-3.2617	-2.4387	184.31	<.0001
Age 3	1	0.5711	0.2262	0.1277	1.0146	6.37	0.0116
Age 4	1	0.5650	0.2209	0.1320	0.9979	6.54	0.0105
Age 5	1	0.7322	0.2220	0.2927	1.1672	10.88	0.0010
Age 6	1	0.6061	0.2239	0.1674	1.0449	7.33	0.0068
Age 7	1	0.6150	0.2323	0.1598	1.0703	7.01	0.0081
Age 8	1	0.2372	0.2416	-0.2364	0.7108	0.96	0.3262
Age 9	1	-0.0438	0.2815	-0.5956	0.5079	0.02	0.8762
Age 10	1	-0.4696	0.3139	-1.1122	0.1184	2.51	0.1135

Calculated Odds & Confidence Limits

Age	Odds	Lower 95%CL	Upper 95%CL	P-value SAS
2	1:0.0578	0.0383	0.0873	<.0001
3	1:0.1024	0.0435	0.2407	0.0116
4	1:0.1018	0.0437	0.2367	0.0105
5	1:0.1203	0.0516	0.2804	0.0010
6	1:0.1060	0.0453	0.2481	0.0068
7	1:0.1070	0.0450	0.2545	0.0081
8	1:0.0733	0.0303	0.1777	0.3262
9	1:0.0554	0.0211	0.1450	0.8762
10	1:0.0352	0.0126	0.0982	0.1135

Results Age

LR Statistics For Type 1 Analysis

Source	Deviance	DF	Chi-Square	Pr > ChiSq
Intercept	7096.1055			
Age	7034.6925	8	61.41 ΔG	<.0001

Wald Statistics For Type 3 Analysis

Source	DF	Chi-Square	Pr > ChiSq
Age	4	51.11	<.0001

Conclusion

- Odds of parasite with latitude:
 - **Accept H_{lat} : odds ratio = $e^{\beta_{lat}} \neq 1$ - expected**
 - **Max odds at 49° - surprise**
 - **South of 46° less odds than north of 50° - surprise**
- Odds of parasite with age
 - **Accept $H_{age/lat}$: odds ratio = $e^{\beta_{age}} \neq 1$ - expected**
 - **Max odds 5 years old - surprise**
 - **Decrease from 6 – 10 years old - surprise**
- Happy with my model? Yes and No
 - **Yes: there was a significant increase in goodness of fit ΔG**
 - **No**
 - **age & lat did not have the effect I expected**
 - **New explanatory variables needed**
 - **separated cod stocks**
 - **longitude**
 - **inherent susceptibility and resistant of fish to parasite**
 - **intermediate host (lumpfish)**