

This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike License](https://creativecommons.org/licenses/by-nc-sa/4.0/). Your use of this material constitutes acceptance of that license and the conditions of use of materials on this site.



Copyright 2006, The Johns Hopkins University and John McGready. All rights reserved. Use of these materials permitted only in accordance with license rights granted. Materials provided "AS IS"; no representations or warranties provided. User assumes all responsibility for use, and all liability related thereto, and must independently review all materials for accuracy and efficacy. May contain materials owned by others. User is responsible for obtaining permissions for use from third parties as needed.



JOHNS HOPKINS
BLOOMBERG
SCHOOL *of* PUBLIC HEALTH

*Tying it All Together: Examples of Logistic
Regression and Some Loose Ends*

John McGready
Johns Hopkins University

Examples for PH literature

Different presentation styles for results from LR

Logistic regression and case control studies

Summary



JOHNS HOPKINS
BLOOMBERG
SCHOOL *of* PUBLIC HEALTH

Section A

Examples from the Literature

Predictors of Injury among Adult Recreational In-line Skaters: A Multi-City Study

★ *Seldes, et al. (1999), AJPH, Vol. 89: 238–241*

See: Predictors of Injury among Adult Recreational In-line Skaters: A Multi-City Study *Seldes, et al. (1999), AJPH, Vol. 89: 238–241*

Table 3N – Risk Factors for Injury from In-line Skating

Example of Logistic Regression

Increased Cesarean Section Rates and Emergency Patterns of Health Insurance in Shanghai, China

★ *Cai, et al., AJPH, 1998, 88: 777–780*

See: Increased Cesarean Section Rates and Emergency Patterns of Health Insurance in Shanghai, China Cai, et al., AJP, 1998, 88: 777–780

Table 2 – Women Who Had a Cesarean Section, 1988 through 1993

See: Increased Cesarean Section Rates and Emergency Patterns of Health Insurance in Shanghai, China Cai, et al., *AJPH*, 1998, 88: 777–780

Table 3 – Proportion of Births Delivered by Cesarean Section, 1988 through 1993

Footnote details in *Increased Cesarean Section Rates and Emergency Patterns of Health Insurance in Shanghai, China* Cai, et al., *AJPH*, 1998, 88: 777–780 Table 2 – Women Who Had a Cesarean Section, 1988 through 1993 reference groups and units for continuous predictors.

Teenage Smoking, Attempts to Quit, and School Performance

★ *Hu, et al. (1998), AJP, 88: 940–943*

This study examined the relationship between school performance, smoking, and quitting attempts among teenagers.

Data were derived from the 1990 California Youth Tobacco Survey.

See: Teenage Smoking, Attempts to Quit, and School Performance *Hu, et al. (1998), AJPH, 88: 940–943*

Table 2 – Logistic Regression of Factors Associate with Teenage Smoking Status (Former Smoker = 1, Current Smoker = 0)

The footnote in *Teenage Smoking, Attempts to Quit, and School Performance* Hu, et al. (1998), *AJPH*, 88: 940–943
Table 2 – Logistic Regression of Factors Associate with Teenage Smoking Status (Former Smoker = 1, Current Smoker = 0) holds the keys to interpreting the results.



JOHNS HOPKINS
BLOOMBERG
SCHOOL *of* PUBLIC HEALTH

Section B

Logistic Regression and Case Control Studies

We can always translate coefficient estimates into (adjusted) odds ratios

Odds ratios can be estimated regardless of what type of study we have

Logistic regression is very useful in case-control studies for controlling confounding, estimating, adjusted relationships

In cohort studies, we can translate our logistic regression results into estimates of prevalence (recall curves from sepsis example) and we can calculate adjusted RR

We can only calculate adjusted OR for case control studies

Example

- ★ *Risk factors for esophageal cancer*
- ★ *Choose 200 cases and 778 controls*
- ★ *Ask about exposures—tobacco and alcohol*

We over-sampled cases

Thus we cannot estimate the probability of getting cancer

So we can't draw graphs like in the sepsis example from the last lecture that display the probability of being a case

We can still do a logistic regression where case/control is the outcome variable and tobacco and alcohol are the x's and obtain odds ratios

We can calculate adjusted odds ratios

Esophageal Cancer Example

Portion of results table related to tobacco

Tobacco (gm/day)	Adjusted odds ratio[*]	P-value
< 10	1.00	—
10–19	1.58	.03
20–29	1.73	.02
30+	4.57	.007

* Also adjusted for age (linear term), which was significant ($p < .001$)

Continued

Esophageal Cancer Example

Portion of results table related to tobacco

Alcohol (gm/day)	Adjusted odds ratio[*]	P-value
0–39	1.00	—
40–79	4.62	< .001
80–119	8.76	< .001
120+	35.60	< .001

* Also adjusted for age (linear term), which was significant ($p < .001$)

Esophageal Cancer Example

After accounting for tobacco and age, the odds ratio associated with high alcohol consumption (> 120+ grams/day) is 36 relative to low consumption (0–39 gm/day)

Esophageal Cancer Example

High alcohol consumption is associated with an almost 36 fold increase in risk of esophageal cancer even after accounting for the effects of tobacco and age

Adjusting for Confounders in Case-Control Studies

Include the confounder as a variable in the logistic regression

- ★ *For example, age was included in the esophageal cancer example*

Adjusting for Confounders in Case-Control Studies

Match cases with controls

- ★ *For example, each case that is identified could be matched to a control who is of the same age*
- ★ *In this way, the cases and the controls will have comparable ages*

Adjusting for Confounders in Case-Control Studies

Match cases with controls

- ★ *Of course, if you do that you cannot determine if age is a risk factor because you forced the cases and controls to have similar ages*

Conditional Logistic Regression

A special kind of logistic regression is used for the analysis of matched case-control studies

This is called conditional logistic regression

Conditional Logistic Regression

Conditional logistic regression does an analysis very similar to the usual logistic regression except it keeps track of which case was matched with which control

The computer output and presentation of conditional logistic regression looks very similar to usual logistic regression

Case-Control Study of Exposure to Medication and the Risk of Injurious Falls Requiring Hospitalization among Nursing Home Residents

★ *Mustard, et al. (1997), AJE, 145: 738–745*

N = 2,972

- ★ *1,486 cases with injurious falls*
- ★ *1,486 matched controls*

Controls drawn from population of nursing home residents, matched on

- ★ *Sex*
- ★ *Year of residence*
- ★ *Age*
- ★ *Level of dependency*
- ★ *Duration of residence in home*
- ★ *Chronic disorder*

	OR	95% CI
Inotropic agents	0.69**	0.54-0.89
Calcium channel blocking agents	1.02	0.68-1.51
Beta blocking agents	1.04	0.64-1.63
Antihypertensive agents	0.91	0.68-1.26
Vasodilating agents	0.86	0.64-1.17
Other cardiovascular agents	0.64	0.29-1.40
Nonsteroidal anti-inflammatory agents	1.11	0.92-1.33
Antidepressants	0.92	0.75-1.12
Antipsychotic agents	1.31**	1.06-1.61
Anxiolytics/sedatives/hypnotics	1.35**	1.09-1.68
Diuretics	0.97	0.82-1.15
Narcotic/opiod analgesic	1.18	0.96-1.45
* p<0.05; ** p<0.01; *** p<0.001.		
OR – odds ratio; CI – confidence interval		



JOHNS HOPKINS
BLOOMBERG
SCHOOL *of* PUBLIC HEALTH

Section C

Quick Example of Interaction in Logistic Regression

Example of Logistic Regression

Prospective Psychosocial, Interpersonal, and Behavioral Predictors of Handgun Carrying among Adolescents

★ *Simon, et al. (1998), AJPH, 88: 960–963*

“This study identified behavioral and psychosocial/interpersonal factors in young adolescence that are associated with handgun carrying in later adolescence.”

See: Prospective Psychosocial, Interpersonal, and Behavioral Predictors of Handgun Carrying among Adolescents Simon, et al. (1998), *AJPH*, 88: 960–963

Table 3 – Logistic Regression Analysis of Behavioral Variables Measured in 9th Grade Predicting Handgun Carrying in 12th Grade among Students in San Diego and Los Angeles Counties