Chapter 13

Estimating the Modified Odds Ratio

Modified odds ratio vis-à-vis modified mean difference

To a large extent, this chapter replicates the content of Chapter 10 (Estimating the modified mean difference), with one notable difference: the dependent variable changes from "mean Y" to "log odds (Y=1)". As a result, regression coefficients will be interpreted as log odds ratios, and their exponential form will deliver modified odds ratios, rather than modified mean differences. We switch from effect modification on the additive scale to effect modification on the multiplicative scale.

Consider, for example, the basic interaction model on the additive scale

$$\text{Mean Y} = \beta_0 + \beta_1 M + \beta_2 E + \beta_3 (M \times E)$$

which was re-written as

$$\text{Mean Y} = \beta_0 + \beta_1 M + (\boldsymbol{\beta_2 + \beta_3 M}) E$$

The parallel model on the log-odds scale is

$$\text{Log odds (Y=1)} = \beta_0 + \beta_1 M + \beta_2 E + \beta_3 (M \times E)$$

which may be re-written as

$$\text{Log odds (Y=1)} = \beta_0 + \beta_1 M + (\boldsymbol{\beta_2 + \beta_3 M}) E$$

As was the case in chapter 10, the effect of E on Y is no longer assumed to be constant. It has turned into a function of M. When M=0, the effect of one-unit increment in E is β_2. But when M=1, the effect of one-unit increment in E is $\beta_2 + \beta_3$. Recalling that the dependent variable is log-odds, rather than mean Y, these two estimates (β_2, $\beta_2+\beta_3$) should be interpreted as a difference between two log-odds, which is equivalent to log odds ratio. To obtain the modified odds ratio, one more step is needed: exponentiating β_2 and $\beta_2+\beta_3$.

When M=0, the effect of one-unit increment in E is estimated by OR=exp(β_2)
When M=1, the effect of one-unit increment in E is estimated by OR=exp($\beta_2+\beta_3$).

Because we estimate a ratio, and not a difference, we inevitably commit to examining effect modification on the multiplicative scale. That is true for all models that specify a log function on the left hand-side of a regression equation, including linear regression: $\log(Y) = \beta_0 + \beta_1 M + \beta_2 E + \beta_3 (M \times E)$. In fact, every idea in this chapter perfectly overlaps an idea in Chapter 11, substituting the words "odds ratio" for "geometric mean ratio". Unfortunately, many fail to recognize the similarity between a linear regression model of a log-transformed Y and all other log-based models (logistic, Poisson, Cox). They usually see in that linear model no more than a method "to normalize a skewed distribution of Y".

Mortality after hospitalization for heart failure

The modified odds ratio will be estimated on a sample of 1,011 patients who were hospitalized due to heart failure in 1995. By the end of 2000, many of these patients have died. Several causal variables will be examined: carrying a diagnosis of heart failure before the index hospitalization, diabetes status, and ejection fraction (the percentage of

blood volume in the left ventricle that is ejected during a single contraction.) Table 13–1 shows these variables, as well as a categorical version of ejection fraction (good, low, or very low). The effect of interest is death.

Table 13–1. Variables and their coded values

Variable Name	Variable Values	
CHF (previous congestive heart failure)	1 = yes	0 or 2 = no*
EF (ejection fraction)	10%-85% (continuous)	
EF_CAT	good, low, very low (categorical)	good = 50% or greater low = 25%-50% very low = smaller than 25%
DIABETES	1 = diabetes	0 = no diabetes
DEATH	1 = dead	0 = alive

* For SAS-related technical reasons, I used the value of 2 in tabular analysis and the value of 0 in regression models

As shown in the 2x2 table below, patients known to have heart failure before their index hospitalization were more likely to die in the next five years than patients who developed overt heart failure for the first time. The incidence odds ratio of death, mislabeled "Case-Control (Odds Ratio)" on the output, was 2.5.

```
PROC FREQ;
 TABLES chf*death/NOCOL NOPERCENT RELRISK;
run;
```

```
                    The FREQ Procedure

                  Table of CHF by DEATH

CHF
                     DEATH(Vital status)

Frequency
Row Pct   |dead     |alive    |  Total
----------+---------+---------+
        1 |      378|      132|    510
          |    74.12|    25.88|
----------+---------+---------+
        2 |      267|      234|    501
          |    53.29|    46.71|
----------+---------+---------+
Total           645       366     1011


          Estimates of the Relative Risk (Row1/Row2)

Type of Study                    Value       95% Confidence Limits
-------------------------------------------------------------------

Case-Control (Odds Ratio)       2.5097        1.9260         3.2703
```

Diabetes modifies the effect of previous heart failure

To explore effect modification by diabetes status, we compare the odds ratio of interest in two strata: diabetic patients and non-diabetic patients.

```
PROC FREQ;
 TABLES diabetes*chf*death/NOCOL NOPERCENT RELRISK;
run;
```

The FREQ Procedure

Table 1 of CHF by DEATH
Controlling for DIABETES=0

CHF
DEATH(Vital status)

Frequency Row Pct	dead	alive	Total
1	207 66.99	102 33.01	309
2	181 50.56	177 49.44	358
Total	388	279	667

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	1.9846	1.4485	2.7190

Table 2 of CHF by DEATH
Controlling for DIABETES=1

CHF
DEATH(Vital status)

Frequency Row Pct	dead	alive	Total
1	171 85.07	30 14.93	201
2	86 60.14	57 39.86	143
Total	257	87	344

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	3.7779	2.2631	6.3066

Evidently, having a previous diagnosis of heart failure had a stronger effect on death in patients who were diabetics (OR=3.8) than in their non-diabetic counterparts (OR=2.0). To estimate the heterogeneity of the effect, we may compute the *ratio* of these two odds ratio (not the difference!): 3.8/2.0=1.9. One effect is almost twice as strong as the other.

Identical results may be obtained from a logistic regression model, regressing the log-odds of death on diabetes status, previous heart failure diagnosis, and their product:

```
PROC LOGISTIC DESCENDING;
 MODEL death = diabetes chf diabetes*chf;
run;
```

The LOGISTIC Procedure

Model Information

Response Variable	DEATH	Vital status
Number of Response Levels	2	
Model	binary logit	

Number of Observations Used 1011

Response Profile

Ordered Value	DEATH	Total Frequency
1	dead	645
2	alive	366

Probability modeled is DEATH='dead'.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error
Intercept	1	0.0223	
DIABETES	1	0.3889	0.2009
CHF	1	0.6854	0.1607
DIABETES*CHF	1	0.6436	0.3069

The regression equation is therefore:

Log odds (DEATH=1) =
= 0.0223 + 0.3889 x DIABETES + 0.6854 x CHF + 0.6436 x DIABETES x CHF

Following re-arrangement, we get

Log odds (DEATH=1) = 0.0223 + 0.3889 x DIABETES + (0.6854 + 0.6436 x DIABETES) x CHF

The CHF effect

Recall that the term called "The CHF effect" is estimating a difference between two log odds, or log odds ratio. Therefore the modified odds ratio takes the following form:

Modified OR = exp(0.6854 + 0.6436 x DIABETES)

In non-diabetics DIABETES=0, so the modified OR=exp(0.6854)=1.98
In diabetics DIABETES=1, so the modified OR=exp(0.6854 + 0.6436 x 1)=3.78

How do we interpret the coefficient of the interaction term?
In this particular case of two binary variables, exponentiating that coefficient provides the same measure of heterogeneity we computed by tabular methods: exp(0.6436)=1.9 = 3.78/1.98. When the variables are not binary, the coefficient of the product term might not have such a simple interpretation.

It is worthwhile to recall the reciprocal property of effect modification. If diabetes modifies the effect of having heart failure before the index hospitalization, then having heart failure before the index hospitalization modifies the effect of diabetes. To compute the modified odds ratio for the diabetes effect, we have to re-arrange the model differently, isolating a multiplier of the diabetes variable:

Log odds (DEATH=1) = 0.0223 + 0.6854 x CHF + (0.3889 + 0.6436 x CHF) x DIABETES

The Diabetes effect

Modified OR = exp(0.3889 + 0.6436 x CHF)

When CHF=0, the modified OR for the diabetes effect =exp(0.3889)=1.48
When CHF=1, the modified OR for the diabetes effect=exp(0.3889 + 0.6436 x 1)=2.81

The magnitude of the heterogeneity of these two effects of diabetes is identical to the magnitude of the heterogeneity of the CHF effect: 2.81/1.48=1.9. The diabetes effect on death is almost twice as strong in the presence of a previous diagnosis of heart failure than in its absence. Again, in this special case of two binary variables, we may also get that ratio measure of heterogeneity by exponentiating the coefficient of the interaction term: exp(0.6436)=1.9.

Back to the modified effect of heart failure. Let's write the model again—in notation:

$$\text{Log odds (DEATH=1)} = \beta_0 + \beta_1 \text{ DIABETES} + (\beta_2 + \beta_3 \textbf{ DIABETES}) \text{ CHF}$$

If we wish to compute confidence intervals for two modified odds ratios, we need two standard errors: SE(β_2) for the effect of heart failure in non-diabetics and SE($\beta_2 + \beta_3$) for the same effect in diabetics. SE(β_2) is available on the output: it is the standard error of the coefficient of CHF (0.1607). The other standard error requires variance arithmetic, and cannot be computed from the available output. Nonetheless, the SAS code below generates exactly the output we would like to see, namely, two CHF coefficients (one in non-diabetics and another in diabetics) and their standard errors. Ignore the intercept and the coefficient of diabetes.

```
PROC LOGISTIC DESCENDING;
 CLASS diabetes;
 MODEL death = diabetes chf(diabetes);
run;
```

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error
Intercept		1	0.2168	
DIABETES	0	1	-0.1945	0.1004
CHF(DIABETES)	**0**	**1**	**0.6854**	**0.1607**
CHF(DIABETES)	**1**	**1**	**1.3290**	**0.2614**

In non diabetics, the modified OR=exp(0.6854)=1.98
95% CI: exp(0.6854 $\pm$ 1.96x0.1607) = [1.45, 2.72]

In diabetes, the modified OR=exp(1.329)=3.78
95% CI: exp(1.329 $\pm$ 1.96x0.2614) = [2.26, 6.30]

Flip back to the output from tabular analysis. Both methods have produced the same results: the same points estimates and the same confidence limits.

Diabetes modifies the effect of ejection fraction (categorical)

We turn next to ejection fraction, and specifically to its categorical version: good, low, or very low. Tabular analysis (below) show expected results: the frequency of death increased with worsening of ejection fraction.

```
PROC FREQ;
TABLES ef_cat*death/NOCOL NOPERCENT;
run;
```

The FREQ Procedure

Table of EF_CAT by DEATH

EF_CAT
DEATH(Vital status)

Frequency Row Pct	dead	alive	Total	
good	148	118	266	
	55.64	44.36		REF
low	343	184	527	
	65.09	34.91		OR=1.49
very low	154	64	218	
	70.64	29.36		OR=1.92
Total	645	366	1011	

Does diabetes modify that effect, too?

Again, stratified analysis should help to answer the question.

```
PROC FREQ;
TABLES diabetes*ef_cat*death/NOCOL NOPERCENT RELRISK;
run;
```

The FREQ Procedure

Table 1 of EF_CAT by DEATH
Controlling for DIABETES=0

EF_CAT
DEATH(Vital status)

Frequency Row Pct	dead	alive	Total	
good	94	86	180	
	52.22	47.78		Reference
low	189	144	333	
	56.76	43.24		OR=1.20
very low	105	49	154	
	68.18	31.82		OR=1.96
Total	388	279	667	

Table 2 of EF_CAT by DEATH
Controlling for DIABETES=1

EF_CAT
DEATH(Vital status)

Frequency Row Pct	dead	alive	Total	
good	54	32	86	
	62.79	37.21		Reference
low	154	40	194	
	79.38	20.62		OR=2.28
very low	49	15	64	
	76.56	23.44		OR=1.94
Total	257	87	344	

Focusing first on the effect of low ejection fraction (versus normal), we see a stronger, harmful effect in diabetics (2.28) than in non-diabetics (1.20). In contrast, the estimated odds ratio for very low ejection fraction (vs. normal) is almost identical in the two strata: 1.94 vs. 1.96. How come?

When ejection fraction is very low, it might not matter anymore whether diabetes is present in the background. That may be the case, of course, but the critical mind might question the credibility of the number 1.94. That estimated effect of very low ejection fraction was derived from the smallest group (N=64 patients) and was based on the smallest number of deaths (49 deaths). Moreover, the estimated odds ratio of 1.94 contradicts the expected monotonicity of the dose-response function in diabetics. It is *smaller* than the estimated effect of low ejection fraction, which was computed from a larger sample. As always, every observation is compatible with more than one theory (Chapter 4).

Logistic regression can easily replicate these tabular analyses. To do so, we may create two dummy variables for the three levels of ejection fraction in a data step, or request SAS to create those for us by introducing the variable EF_CAT in the class statement. I used the options **PARAM=REF REF=FIRST** to ensure that "good ejection fraction" will serve as the reference category.

```
PROC LOGISTIC DESCENDING;
 CLASS ef_cat/PARAM=REF REF=FIRST;
 MODEL death = ef_cat;
run;
```

The LOGISTIC Procedure
Model Information

Response Variable	DEATH	Vital status
Number of Response Levels	2	
Model	binary logit	

Number of Observations Used 1011

Response Profile

Ordered Value	DEATH	Total Frequency
1	dead	645
2	alive	366

Probability modeled is DEATH='dead'.

Class Level Information

Class	Value	Design Variables	
EF_CAT	good	0	0
	low	1	0
	very low	0	1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error
Intercept		1	0.2265	
EF_CAT	low	1	0.3963	0.1536
EF_CAT	very low	1	0.6515	0.1933

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
EF_CAT low vs good	1.486	1.100	2.008
EF_CAT very low vs good	1.918	1.314	2.802

Notice the dummy coding in the section "design variables". Each of the two columns corresponds to a dummy variable, but they were not named there. Later on, these dummy variables carry long, identifying names under the heading "Parameter": one is called "EF_CAT low" and the other is called "EF_CAT very low". The output shows the odds ratios, and even states the reference category (vs good). Both point estimates are identical to those computed by tabular analysis.

Turning next to modeling effect modification by diabetes. Again, we could create two dummy variables to represent the three categories of ejection fraction, and add two product terms. Alternatively, SAS will create the interaction terms for us if we introduce the variables in the class statement.

```
PROC LOGISTIC DESCENDING;
 CLASS diabetes ef_cat/PARAM=REF REF=FIRST;
 MODEL death = diabetes ef_cat diabetes*ef_cat;
run;
```

Analysis of Maximum Likelihood Estimates

Parameter			DF	Estimate	Standard Error
Intercept			1	0.0889	
DIABETES	1		1	0.4343	0.2684
EF_CAT	low		1	0.1830	0.1857
EF_CAT	very low		1	0.6732	0.2285
DIABETES*EF_CAT	1	low	1	0.6418	0.3402
DIABETES*EF_CAT	1	very low	1	-0.0127	0.4348

To write the regression equation, let's shorten the lengthy names of the dummy variables and the interaction terms

Log odds (DEATH=1) = 0.0889 + 0.4343 x DIABETES + 0.183 x LOW + 0.6732 x VERY LOW + 0.6418 x DIABETES x LOW – 0.0127 x DIABETES x VERY LOW

After re-arrangement, we can easily identify the effects of interest, shown in parentheses.

Log odds (DEATH=1) = 0.0889 + 0.4343 x DIABETES +
(**0.1830 + 0.6418 x DIABETES**) x LOW +
(**0.6732 – 0.0127 x DIABETES**) x VERY LOW

When diabetes is absent (DIABETES=0),
Modified OR (low vs. normal) = exp(0.183)=1.20
Modified OR (very low vs. normal) = exp(0.6732)=1.96

When diabetes is present (DIABETES=1),
Modified OR (low vs. normal) = exp(0.183 + 0.6418)= exp(0.8248)= 2.28
Modified OR (very low vs. normal) = exp(0.6732 – 0.0127)= exp(0.6605)=1.94

Compare these results to those computed by tabular analysis. They are identical.

Again, alternative SAS code (below) generates exactly the output we would like to see, namely, a pair of coefficients (for LOW and VERY LOW) in non-diabetics and another pair in diabetics, as well as four standard errors. Ignore the intercept and the coefficient of diabetes.

```
PROC LOGISTIC DESCENDING;
 CLASS diabetes ef_cat/PARAM=REF REF=FIRST;
 MODEL death = diabetes ef_cat(diabetes);
run;
```

Class Level Information

Class	Value	Design Variables	
DIABETES	0	0	
	1	1	
EF_CAT	good	0	0
	low	1	0
	very low	0	1

Analysis of Maximum Likelihood Estimates

Parameter			DF	Estimate	Standard Error
Intercept			1	0.0889	
DIABETES	1		1	0.4343	0.2684
EF_CAT(DIABETES)	low	0	1	**0.1830**	0.1857
EF_CAT(DIABETES)	very low	0	1	**0.6732**	0.2285
EF_CAT(DIABETES)	low	1	1	**0.8248**	0.2851
EF_CAT(DIABETES)	very low	1	1	**0.6605**	0.3699

The coefficients in bold print are identical to those we calculated by hand from the previous output that contained interaction terms.

The nature of effect modification may be grasped better, if we display the two dose-response functions for the effect of ejection fraction on death: one in diabetics and another in non-diabetic. Each of these functions is, of course, step-like. To display them, we start with the interaction model

Log odds (DEATH=1) = 0.0889 + 0.4343 x DIABETES +
(**0.1830 + 0.6418 x DIABETES**) x LOW +
(**0.6732 – 0.0127 x DIABETES**) x VERY LOW

And write the model separately by diabetes status. In non-diabetics (DIABETES=0), the regression equation is reduced to

$$\text{Log odds (DEATH=1)} = 0.0889 + 0.1830 \times \text{LOW} + 0.6732 \times \text{VERY LOW}$$

Whereas in diabetics (DIABETES=1), the regression is reduced to

$$\text{Log odds (DEATH=1)} = 0.5232 + 0.8248 \times \text{LOW} + 0.6605 \times \text{VERY LOW}$$

Using these equations, we can compute the log-odds of death for each ejection fraction category, in diabetics and non-diabetics (Table 13–2).

Table 13–2. Log odds of death, by ejection fraction category and diabetes status

Ejection fraction category	Values of dummy variables	Log odds (death)	
		Non-diabetics	Diabetics
Good (≥50%)	LOW = 0; VERY LOW=0	0.0889	0.5232
Low (25%-50%)	LOW = 1; VERY LOW=0	0.2719	1.3480
Very low (<25%)	LOW = 0; VERY LOW=1	0.7621	1.1837

Finally, we display the two step functions in a single graph (Figure 13–1). As expected, the log odds of death usually increase as we move from right to left on the X-axis, that is, when ejection fraction decreases. To interpret the graphs more quantitatively, we should recall that each vertical distance between two horizontal lines is a difference between two log-odds, or log odds ratio. Low ejection fraction (versus good ejection fraction) has a stronger effect on death in diabetics than in non-diabetics, which may be inferred by comparing the respective vertical distances. In contrast, the vertical distance between very low ejection fraction and good ejection fraction is similar in the two groups. The reciprocal property of effect modification is evident, too: the effect of diabetes is strongest in patients with low ejection fraction (middle vertical distance), but seems almost identical in patients with very low ejection fraction and in patients with good ejection fraction.

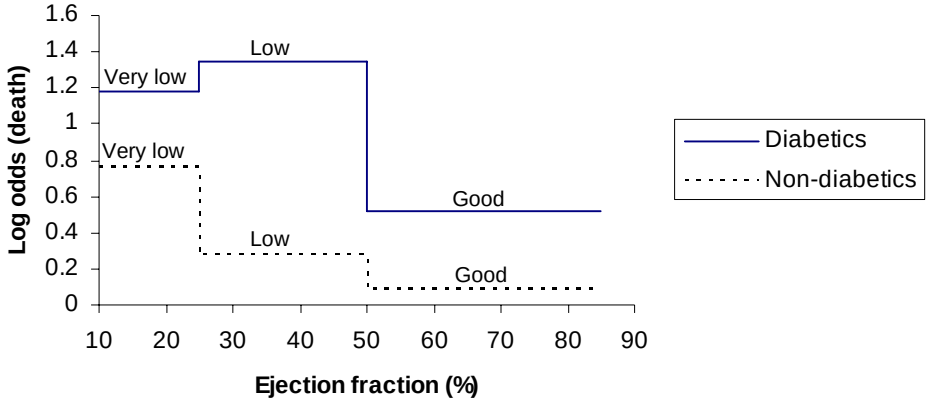

Figure 13–1. Two dose-response functions for the effect of ejection fraction on death: in diabetics (solid line) and in non-diabetics (dashed line)

Diabetes modifies the effect of ejection fraction (continuous)

Although linearity on the log scale is far from apparent, especially in diabetics, we will examine an interaction model with ejection fraction in its original, continuous form.

```
PROC LOGISTIC DESCENDING;
 MODEL death = diabetes ef diabetes*ef;
run;
```

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error
Intercept	1	0.8093	0.1962
DIABETES	1	1.0926	0.3954
EF	1	-0.0128	0.00477
DIABETES*EF	1	-0.00889	0.00952

The regression equation is therefore
Log odds (DEATH=1) =
= 0.8093 + 1.0926 x DIABETES – 0.0128 x EF – 0.00889 x DIABETES x EF

Following re-organization, we get

Log odds (DEATH=1) = 0.8093 + 1.0926 x DIABETES + (–0.0128 – 0.00889 x DIABETES) x EF

The ejection fraction effect

According to this model, the effect of ejection fraction is a function of diabetes status. The log odds of death changes by "–0.0128 – 0.00889 x DIABETES" per 1 percentage point increment in ejection fraction. If we are interested in the effect of declining ejection fraction, we just have to reverse the sign. The log odds of death changes by "0.0128 + 0.00889 x DIABETES" per 1 percentage point *decrement* in ejection fraction. Recall that a change in the log-odds means log odds ratio, so the term "0.0128 + 0.00889 x DIABETES" is also a term for the modified odds ratio per 1 percentage point decrement in ejection fraction. A more meaningful estimate may be computed for a decrement of 10 percentage points. For example:

In non diabetics, the modified OR=exp(0.0128x10)=1.13
In diabetes, the modified OR=exp([0.0128+0.0089]x10)= exp(0.0217x10)= 1.24

Stratum-specific estimates, and their confidence limits may also be generated by alternative SAS code (below).

```
PROC LOGISTIC DESCENDING;
 CLASS diabetes;
 MODEL death = diabetes ef(diabetes);
run;
```

Class Level Information

Class	Value
DIABETES	0
	1

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error
Intercept		1	1.3556	
DIABETES	0	1	-0.5463	0.1977
EF(DIABETES)	0	1	-0.0128	0.00477
EF(DIABETES)	1	1	-0.0217	0.00824

Again, recall that the negative coefficient estimate log odds ratio per 1 percentage point increment in ejection fraction; larger ejection fraction decreases the log odds of death. The effect of worsening ejection fraction may be computed by reversing the negative sign: smaller ejection fraction increases the log odds of death. Starting with the interaction model, we can easily write the two dose-response functions.

Log odds (DEATH=1) = 0.8093 + 1.0926 x DIABETES + (–0.0128 – 0.00889 x DIABETES) x EF

In diabetics: Log odds (DEATH=1) = 1.9019 – 0.0217 x EF
In non-diabetics: Log odds (DEATH=1) = 0.8093 – 0.0128 x EF

The two lines are displayed in Figure 13–2. When ejection fraction decreases (moving from right to left on the X-axis), the log odds of death increases more steeply in diabetics than in non-diabetics.

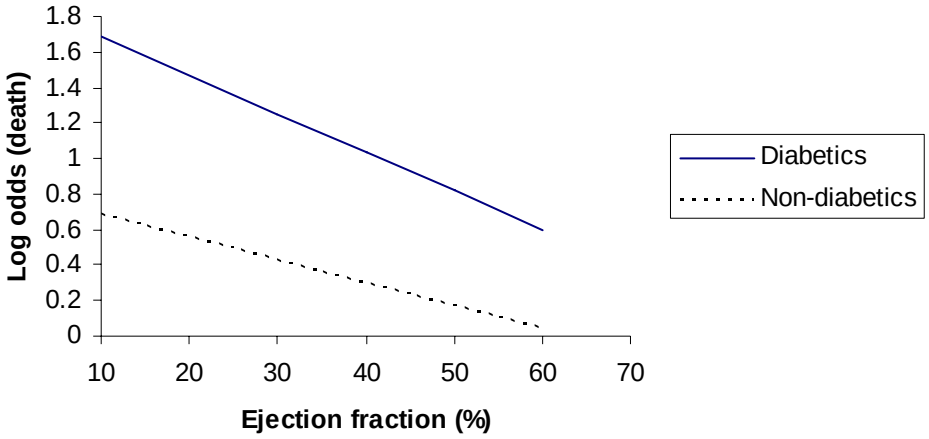

Figure 13–2. Two dose-response functions for the effect of ejection fraction on death: in diabetics (solid line) and in non-diabetics (dashed line)

Complexity need not end here. Suppose we entertain modification of a quadratic dose-response function. In other words, we assume that the relation of ejection fraction to the log-odds of death is curvilinear, but the exact shape differs between diabetics and non-diabetics. The model will contain the variables DIABETES, EF, and EF^2 and two product terms: DIABETES*EF and DIABETES*EF^2.

```
PROC LOGISTIC DESCENDING;
 MODEL death = diabetes ef ef*ef diabetes*ef diabetes*ef*ef;
run;
```

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error
Intercept	1	1.5741	
DIABETES	1	0.2125	0.8930
EF	1	-0.0578	0.0227
EF*EF	1	0.000550	0.000271
DIABETES*EF	1	0.0424	0.0456
DIABETES*EF*EF	1	-0.00062	0.000532

Log odds (DEATH=1) =

= 1.5741 + 0.2125 x DIABETES – 0.0578 x EF + 0.00055 x EF^2 + 0.0424 x DIABETES x EF – 0.00062 x DIABETES x EF^2

The two dose-response functions are reduced to the following:

In non-diabetics (DIABETES=0):

Log odds (DEATH=1) = 1.5741 – 0.0578 x EF + 0.00055 x EF^2

In diabetics (DIABETES=1):

Log odds (DEATH=1) = 1.7866 – 0.0154 x EF – 0.00007 x EF^2

Almost identical results may be obtained by stratified regression (below). Recall, however, that stratified regression and an interaction model may generate very different results when other covariates are included in the model—for example, if we had to remove confounding by age and sex. Omitting confounders, we get the same key coefficients by stratified regression.

```
PROC SORT; BY diabetes;
PROC LOGISTIC DESCENDING;
 MODEL death = ef ef*ef;
 BY diabetes;
run;
```

----------DIABETES=0----------

The LOGISTIC Procedure

Number of Observations Used 667

Response Profile

Ordered Value	DEATH	Total Frequency
1	dead	388
2	alive	279

Probability modeled is DEATH='dead'.

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error
Intercept	1	1.5737	
EF	1	-0.0578	0.0227
EF*EF	1	0.000550	0.000271

----------DIABETES=1----------

The LOGISTIC Procedure

Number of Observations Used 344

Response Profile

Ordered Value	DEATH	Total Frequency
1	dead	257
2	alive	87

Probability modeled is DEATH='dead'.

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error
Intercept	1	1.7866	
EF	1	-0.0154	0.0395
EF*EF	1	-0.00007	0.000458

Figure 13–3 depicts the two dose-response functions. Some departure from linearity is observed in non-diabetic patients who seem to "tolerate" some decline from normal ejection fraction without substantial increase in their log-odds of death. The dose-response function for diabetics seems fairly linear within the observed range of ejection fraction. As you can see, reconciling the inference from different dose-response models may not be easy. Even simple scientific inference may not be simple.

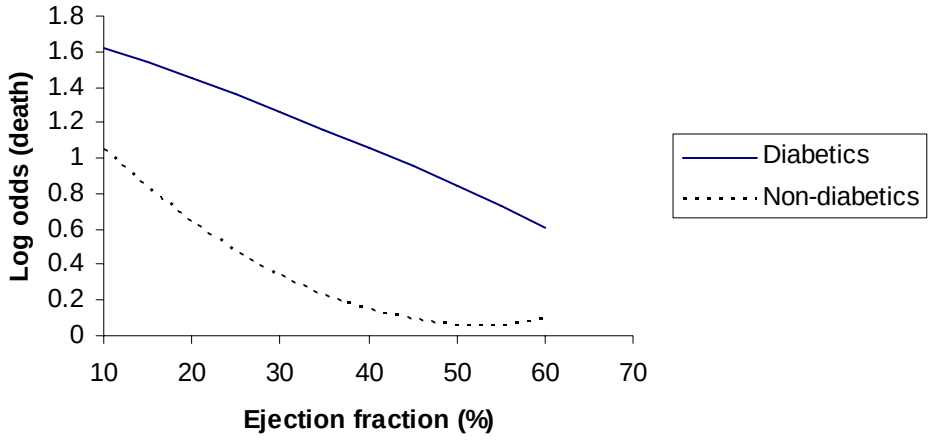

Figure 13–3. Two dose-response functions for the effect of ejection fraction on death: in diabetics (solid line) and in non-diabetics (dashed line)

To sum up, estimating the modified odds ratio is not conceptually different from estimating the modified mean difference (Chapter 10) and is virtually identical to estimating the modified geometric mean ratio (Chapter 11). We just have to remember the crucial, extra step of exponentiation.