A Study of the Association Between Dietary Iron Intake and Coronary Heart Disease

PH241 Categorical Analysis – Spring 2015

Kurt Anthony, Paul Cho, Daniel Lee, Denise Shieh, Marina Wang

A Study of the Association Between Dietary Iron Intake and Risk of Coronary Heart Disease

Background

Our data derives from a case-control study conducted in Athens, Greece (study period: Jan 1990 to April 1991). The exposure studied is dietary iron intake measured from nutritional survey data and the outcome is coronary heart disease (CHD). There are 329 cases selected on the basis of clinical diagnosis of coronary heart disease and 570 controls selected from patients admitted around the same time period as the case for diseases unrelated to the exposure or outcome. Variables included in the dataset are the following: sex, age, estimated monthly iron intake (mg), and case/control status.

Research Question: What is the association between dietary iron intake and coronary heart disease?

Exposure
Dietary Iron Intake
(mg/month)

Variable Type	Description of **Iron Intake** Categories (mg/month)
Binary	$x^* = 0 \rightarrow \;<= 300$ $x^* = 1 \rightarrow \;> 300$
Categorical	reference $\rightarrow$ <= 250 $x_1 = 1 \rightarrow$ 251-300 $x_2 = 1 \rightarrow$ 301-350 $x_3 = 1 \rightarrow$ > 350
Ordinal	$x = 0 \rightarrow$ <= 250 $x = 1 \rightarrow$ 251-300 $x = 2 \rightarrow$ 301-350 $x = 3 \rightarrow$ > 350

Outcome
Coronary Heart Disease

N = 899
 # Cases: 329 (w/ CHD)
 # Controls: 570 (w/o CHD)

Risk Factors of Interest

Sex

$y = 0 \rightarrow$ male

$y = 1 \rightarrow$ female

Age

$z = 0 \rightarrow$ < 60 years

$z = 1 \rightarrow$ >= 60 years

DAG

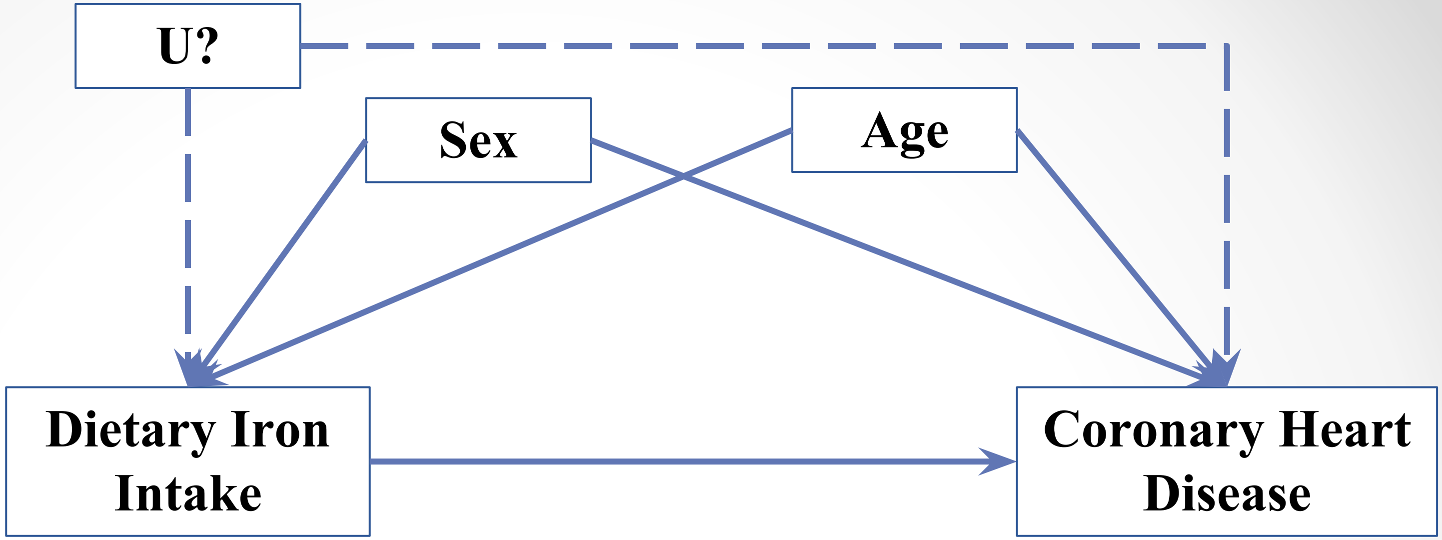
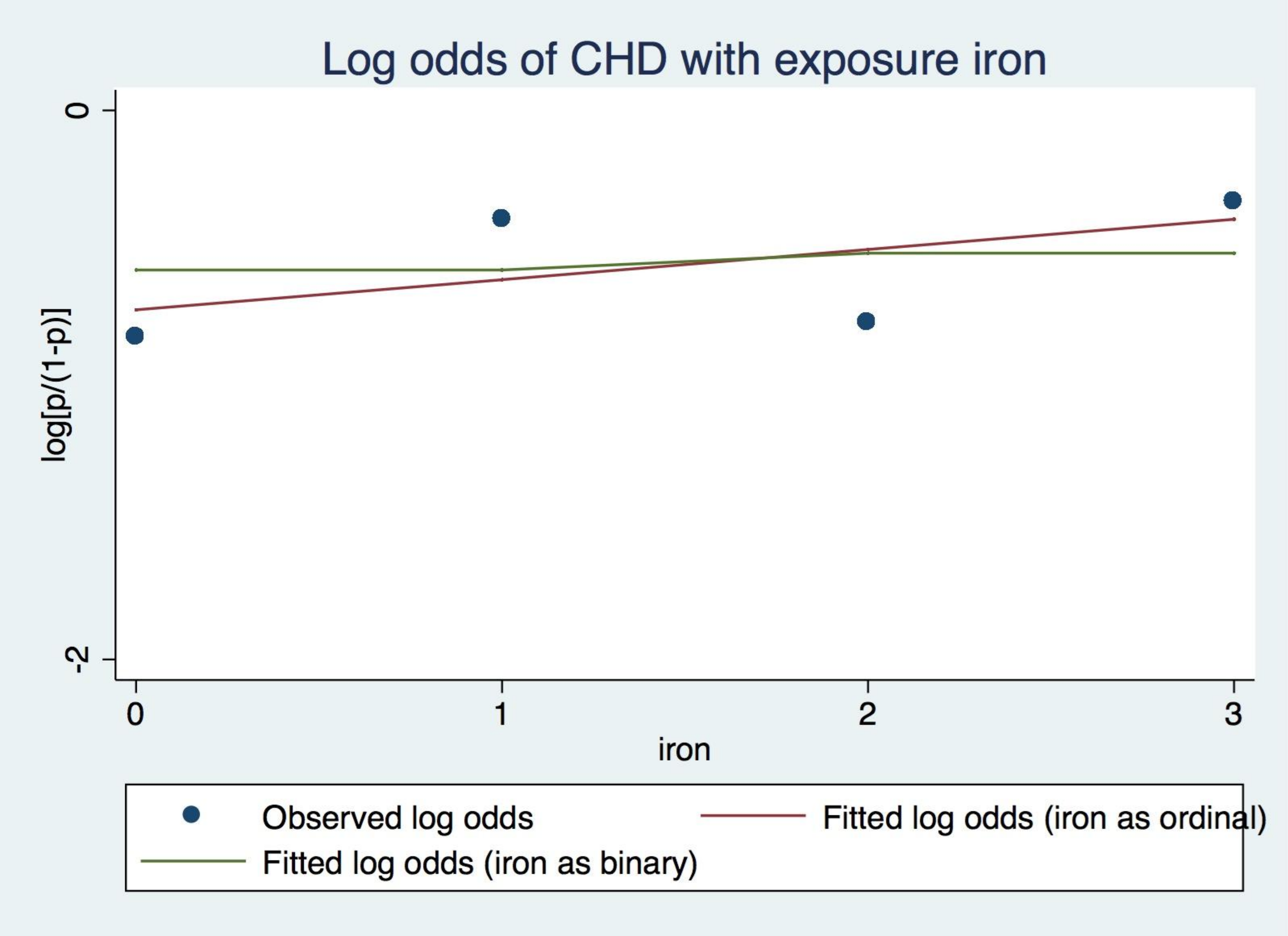


Table 1: Frequency of Cases and Controls by Each Risk Factor Category

		Cases (%)	Controls (%)
Sex	**Female**	45 (15.5%)	246 (84.5%)
	Male	284 (46.7%)	324 (53.3%)
Age	**<60 years**	153 (39.0%)	239 (61.0%)
	>= 60 years	176 (34.7%)	331 (65.3%)

Log Odds Plots of CHD



Model Selection

Table 2: Model selection process and output

#	Model	Para	Estimate	SD	OR	P-value	Max LL
1	$log\left(\frac{p}{1-p}\right) = a$	a	- 0.550	0.069		<0.0001	-590.438
2	$log\left(\frac{p}{1-p}\right) = a + bx^*$ (Binary)	a	- 0.581	0.139		<0.0001	-590.340
		b	0.061	0.100	1.063	0.658	
3	$log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + b_3x_3$	a	-0.828	0.156		<0.0001	-585.487
		b_1	0.433	0.204	1.541	0.034	
		b_2	0.052	0.217	1.054	0.810	
		b_3	0.495	0.200	1.641	0.013	
4	$log\left(\frac{p}{1-p}\right) = a + bx$ (Ordinal)	a	-0.727	0.123		<0.0001	-588.863
		b	0.110	0.062	1.116	0.077	
5	$log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + b_3x_3 + cy$	a	-0.196	0.176		0.266	-543.779
		b_1	0.252	0.216	1.286	0.245	
		b_2	-0.119	0.230	0.888	0.604	
		b_3	0.061	0.214	1.063	0.774	
		c	-1.558	0.186	0.211	<0.0001	
6	$log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + b_3x_3 + cy + d_1(x_1 * y) + d_2(x_2 * y) + d_3(x_3 * y)$	a	2.19×10^{-15}	0.200		1.000	- 540.264
		b_1	0.100	0.255	1.105	0.695	
		b_2	- 0.445	0.270	0.641	0.099	
		b_3	- 0.182	0.242	0.833	0.450	
		c	- 2.245	0.404	0.106	<0.0001	
		d_1	0.491	0.529	1.634	0.353	
		d_2	1.218	0.532	3.381	0.022	
		d_3	1.159	0.569	3.188	0.042	

7	$log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + b_3x_3 + cz$	a	-0.951	0.176		<0.0001	-584.266
		b_1	0.456	0.205	1.577	0.026	
		b_2	0.087	0.219	1.091	0.690	
		b_3	0.534	0.202	1.705	0.008	
		c	0.221	0.141	1.247	0.118	
8	$log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + b_3x_3 + cz + d_1(x_1 * z) + d_2(x_2 * z) + d_3(x_3 * z)$	a	-0.762	0.229		0.001	-581.950
		b_1	0.382	0.289	1.465	0.187	
		b_2	-0.294	0.307	0.745	0.339	
		b_3	0.284	0.280	1.329	0.311	
		c	-0.121	0.313	0.886	0.698	
		d_1	0.089	0.410	1.093	0.829	
		d_2	0.793	0.439	2.210	0.071	
		d_3	0.498	0.406	1.646	0.219	
9	$log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + b_3x_3 + cy + d_2(x_2 * y) + d_3(x_3 * y)$	a	-0.071	0.185		0.700	-540.699
		b_1	0.216	0.223	1.241	0.332	
		b_2	-0.373	0.259	0.688	0.149	
		b_3	-0.111	0.229	0.894	0.628	
		c	-1.972	0.261	0.139	<0.001	
		d_2	0.945	0.434	2.574	0.030	
		d_3	0.886	0.478	2.426	0.064	

Overall Likelihood Ratio Tests:

Model I (restricted)	**Model II**	**LR Test**	**P-value**
$log\left(\frac{p}{1-p}\right) = a$	Model 2b: $log\left(\frac{p}{1-p}\right) = a + bx$	0.20	0.658
$log\left(\frac{p}{1-p}\right) = a$	$log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + b_3x_3$	9.90	0.019
$log\left(\frac{p}{1-p}\right) = a$	$log\left(\frac{p}{1-p}\right) = a + bx^*$	3.15	0.076

Likelihood Ratio Test Assessing Interaction Between Iron and Sex

Model I (restricted)	**Model II**	**LR Test**	**P-value**
$log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + b_3x_3 + cy$	$log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + b_3x_3 + cy + d_1(x_1 * y) + d_2(x_2 * y) + d_3(x_3 * y)$	7.03	0.0709

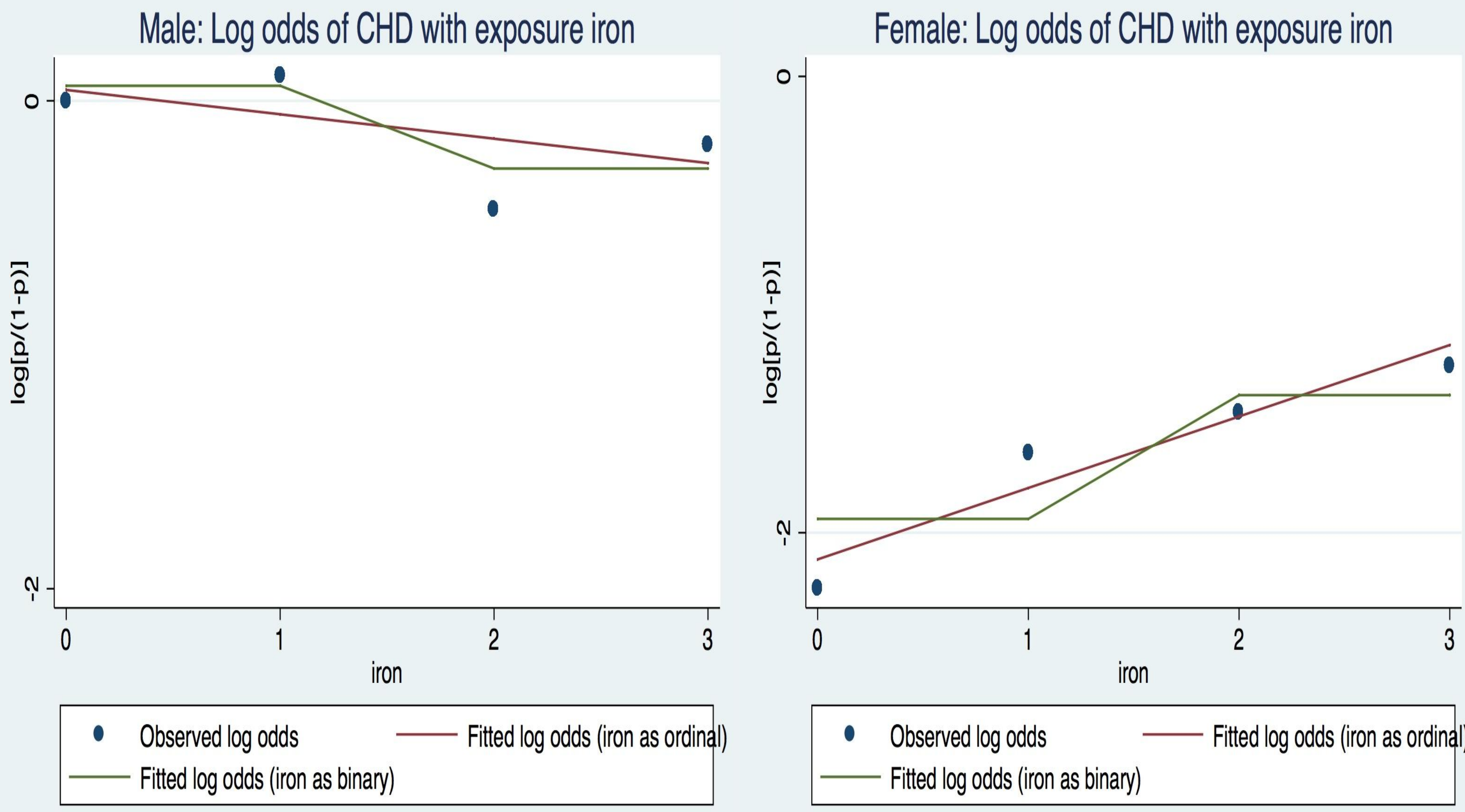
Likelihood Ratio Test Assessing Interaction Between Iron and Age

Model I (restricted)	**Model II**	**LR Test**	**p-value**
$log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + b_3x_3 + cz$	$log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + b_3x_3 + cz + d_1(x_1 * z) + d_2(x_2 * z) + d_3(x_3 * z)$	4.63	0.2008

Likelihood Ratio Test Comparing Model 8 and 9

Model I (restricted)	**Model II**	**LR Test**	**p-value**
$log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + b_3x_3 + cy + d_2(x_2 * y) + d_3(x_3 * y)$	$log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + b_3x_3 + cy + d_1(x_1 * y) + d_2(x_2 * y) + d_3(x_3 * y)$	0.87	0.351

Log Odds Plots of CHD by Sex



Selected Model

$$\log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + b_3x_3 + cy + d_2(x_2 * y) + d_3(x_3 * y)$$

Pearson χ^2 Goodness of Fit

Disease status

		# CHD (# Expected)	# No CHD (# Expected)	Total
Covariate Pattern	1	9 (10.8)	85 (83.2)	94
	2	13 (11.2)	68 (69.8)	81
	3	14 (14)	61 (61)	75
	4	9 (9)	32 (32)	41
	5	50 (50)	78 (78)	128
	6	100 (100)	120 (120)	220
	7	50 (48.2)	50 (51.8)	100
	8	84 (85.8)	76 (74.2)	160
Total		329	570	899

$\chi^2_{\text{Pearson}} = .87$ **p-value** = .3518

Results

In the process of model selection, we fitted the exposure variable, iron, in binary, categorical, and ordinal forms. To assess the association between iron intake and CHD, we calculated the chi-square test statistics for the overall likelihood ratio tests comparing model 1 (the simple model with constant odds of CHD) with models 2, 3, and 4, respectively. The likelihood ratio test comparing models 1 and 3 was the only test that yielded a significant result with p-value = 0.019. Thus, we reject the null hypothesis (nested model) in favor of the full model in which iron is coded as a categorical variable.

We assessed age and sex as possible confounders and/or effect modifiers. A comparison of the odds ratios for each of the indicator variables for iron in models 3 and 5 showed a substantial decrease (>10%) in the model controlling for sex. However, a comparison of the odds ratios for each of the indicator variables for iron in models 3 and 7 did not show a significant difference when controlling for age. Likelihood ratio tests comparing models 5 and 6 assessed sex as an effect modifier of the relationship between iron and CHD. The resulting test statistic was significant (p-value = 0.0709) and suggests that there is a multiplicative interaction between sex and iron. The likelihood ratio tests comparing models 7 and 8 assessed age as an effect modifier in of the relationship between iron and CHD. The resulting test statistic was not significant (p-value = 0.2008), suggesting that there is no interaction between age and iron.

Further, we examined the significance of the coefficients for the interaction between iron and sex for model 6 and removed the interaction term for d_1, where the p-value (0.353) was insignificant. We then then performed the likelihood ratio test comparing model 9 (without d_1) and model 6 (with all interaction terms), which yielded a p-value of 0.351, suggesting that the nested model (model 9) adequately fits our data.

The Pearson chi-square goodness-of-fit test for our selected model resulted in a p-value of 0.3518, which is insignificant. Thus, the frequency distribution of our data does not deviate significantly from that of the expected values from our model. In other words, our selected model is a reasonable overall fit for our data and is represented here:

Final Model

$$log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + b_3x_3 + cy + d_2(x_2 * y) + d_3(x_3 * y)$$

Table 3: Odds ratios comparing each iron intake level to the baseline (level x = 0) amongst men and women.

	Males [95%CI]	**Females [95%CI]**
OR comparing iron level 1 to baseline	1.241 [0.802, 1.921]	1.241 [0.802, 1.921]
OR comparing iron level 2 to baseline	0.688 [0.414, 1.143]	1.772 [0.824, 3.808]
OR comparing iron level 3 to baseline	0.895 [0.571, 1.402]	2.171 [0.890, 5.296]

Conclusion

The model with iron consumption as a categorical variable fits better than those with iron consumption as binary or ordinal variables. There is evidence that sex is both a confounder and effect modifier of the association between iron consumption and CHD; however, the interaction is not significant for the second iron consumption level (251-300 mg/month). There is no evidence for confounding or effect modification by age of this relationship. Iron consumption at any level is not significantly associated with CHD after controlling for sex. We observe in the log odds plots and odds ratio tables that women's CHD odds increases linearly as iron consumption increases, but there is not a linear trend for men.

Reference

Tzonou A, Lagiou P, Trichopoulou A, Tsoutsos V, Trichopoulos D. Dietary iron and coronary heart disease risk: a study from Greece. Am J Epidemiol. 1998;147(2):161-6.