

C.F. Turner

September 2, 2011

S:\NIDDK_NoData\CRIC_DSIC_AF\CRIC_AtrialFibrillation_2fin_REV.rtf

Dataset Integrity Check for Chronic Renal Insufficiency Cohort (CRIC): DSIC 2 --- Atrial Fibrillation Dataset (M_37)

As a partial check of the integrity of the CRIC datasets archived in the NIDDK data repository, a set of tabulations was performed to verify that published results from the CRIC study can be reproduced using the archived datasets. Analyses were performed to duplicate published results for two CRIC datasets. The present Dataset Integrity Check (DSIC) reports results for the atrial fibrillation analysis dataset (M_37). This dataset supports findings reported by Soliman et al. in 2010 in the *American Heart Journal* [1]. The results of our dataset integrity check are described below.

Purpose. The intent of this dataset integrity check is to provide confidence that the dataset distributed by the NIDDK repository is a true copy of the study data. Our intent is not to assess the integrity of the statistical analyses reported by study investigators. As with all statistical analyses of complex datasets, complete replication of a set of statistical results should not be expected on a first (or second) exercise in secondary analysis. This occurs for a number of reasons including differences in the handling of missing data, restrictions on cases included in samples for a particular analysis, software coding used to define complex variables, etc. Experience suggests that most discrepancies can ordinarily be resolved by consultation with the study DCC, however this process is labor-intensive for both DCC and Repository staff. It is thus not our policy to resolve every discrepancy that is observed in a dataset integrity check. Thus, we do not attempt to resolve minor or inconsequential discrepancies with published results or discrepancies that involve complex analyses unless staff of the NIDDK Repository suspect that the observed discrepancy suggests that the dataset may have been corrupted in storage, transmission, or processing by repository staff. We do, however, document such discrepancies in footnotes to the dataset integrity check or by highlighting in tables those instances in which

our secondary analyses produced results that were not fully consistent with those reported in the target publication.

Datasets. The dataset used for these analyses was m37_20100104.sas7bdat received from the DCC (with date stamps of 12-23-2010).¹ This SAS dataset was converted to STATA format using Stat/Transfer and output as a STATA data file m37_20100104.dta: (date stamp: 3-16-2011). Although a SAS format file (formats.sas7bcat) was provided by the DCC, incompatible use of upper and lower case letters prevented a simple translation of value labels using Stat/Transfer. Equivalent STATA value labels were constructed, and they were applied to create a new STATA dataset with value labels (m37_w_Labels.dta; time-date stamp: 3-16-2011, 2-42pm).

Comparison of Characteristics of Subjects. Table 1 of the Soliman et al. article reports sample sizes and selected characteristics of CRIC cohort. The archived data were analyzed to produce the same characteristics reported by Soliman et al. Our **Table 1** compares the published results to those obtained from our analysis of the archived data. It will be seen from **Table 1** that our calculations from the archived data are fully equivalent to the published results.

Comparison of Atrial Fibrillation (AF) Prevalence by Selected Characteristics. **Table 2** of the Soliman et al. article reports the prevalence of atrial fibrillation for the total population and subpopulations defined by eGFR status, age, gender, and race/ethnicity. As will be seen from our **Table 2**, calculations from the archived dataset yielded results that were identical to those reported in the published article.

Unadjusted association of AF and selected characteristics. The first panel of Table 3 in Soliman et al. reports unadjusted association between AF and selected sociodemographic and medical characteristics and the prevalence of atrial fibrillation. Our **Table 3a** compares the published results with those obtained by analysis of the archive CRIC data. With one exception, the results of our calculations are identical to the published results. The one exception occurs for gender. The published OR for female gender is the inverse of the result

¹ This same file was re-transmitted in July of 2011 with other data files that included corrections. The AF analysis file in this re-transmission had identical file name, file size, and time-and-date stamp as the original file transmitted in December of 2010.

we obtained. That is to say that rather than females having 0.98 lower odds of reporting atrial fibrillation, we obtain this result for males.

Demographically adjusted association of AF and selected characteristics. Our **Table 3b** repeats the previous analysis (**Table 3a**) but the estimated impact of respondent characteristics are adjusted for: age, gender, race (non-Hispanic Black vs. other), and the study center that enrolled the subject. Here again we fully replicate the published findings except for gender. The published results report that females have lower adjusted odds (0.90) than males of reporting AF. We obtain an identical result but for males not females

Multivariable adjusted association of AF and selected characteristics. Table 4 of Soliman et al. report a final model predicting atrial fibrillation as a function of age, gender, education, total cholesterol level, BMI, physical activity, smoking status, and presence of congestive heart failure and other cardiovascular disease. Our **Table 4** compares the published results to our estimated coefficients for the same multivariable logistic regression model fit to the archived data. As **Table 4** shows, the results are again identical with the exception of those for gender.² The published results for “females” are equivalent to the results we obtained for males.

Conclusion. Our results fully replicate the published findings with the exception of the three instances noted above. The DCC for CRIC has informed us that there is an error in the published tables and that the three calculated results are correct

References.

- [1] Soliman EZ, Prineas RJ, Go AS, Xie D, Lash JP, Rahman M, Ojo A, Teal VL, Jensvold NG, Robinson NL, Dries DL, Bazzano L, Mohler ER, Wright JT, Feldman HI; Chronic Renal Insufficiency Cohort (CRIC) Study Group. Chronic kidney disease and prevalent atrial fibrillation: the Chronic Renal Insufficiency Cohort (CRIC). *Am Heart J.* 2010 Jun;159(6):1102-7. Erratum in: *Am Heart J.* 2010 Dec;160(6):1190.

² There was also a minor deviation in p-values for education (p = 0.06 vs. 0.09).

TABLE 1. Characteristics of the CRIC study population: Published versus tabulated results.

CHARACTERISTIC	Published	Calculated
Age (y)	58.55 (10.81)	58.55 (10.81)
Sex (male)	1775 (54%)	1775 (54%)
Race/ethnicity (non-Hispanic black)	1640 (50%)	1640 (50%)
Education		
< high school grad	523 (16%)	523 (16%)
High school grad	649 (20%)	649 (20%)
Post-high school	1028 (31%)	1028 (31%)
College graduate	606 (19%)	606 (19%)
Postgraduate degree	461 (14%)	461 (14%)
Smoker		
Current	470 (14%)	470 (14%)
Former	1392 (43%)	1392 (43%)
Never	1405 (43%)	1405 (43%)
Alcohol use (drinkers)	1310 (40%)	1310 (40%)
Total physical activity (METh/wk)	204.6 (148.33)	204.6 (148.33)
Hypertension	2807 (86%)	2807 (86%)
Diabetes	1486 (45%)	1486 (45%)
Congestive heart failure	328 (10%)	328 (10%)
Cardiovascular Disease	1125 (34%)	1125 (34%)
eGFR (mL/[min1.73m²])	43.6 (13.4)	43.6 (13.4)
Participants with eGFRb < 45mL/(min1.7	1795 (55%)	1795 (55%)
Body mass index (kg/m²)	32.31 (8.0)	32.31 (8.0)
Total cholesterol (mg/dL)	182.8 (43.8)	182.8 (43.8)
Uric acid (mg/dL)	7.4 (1.9)	7.4 (1.9)
hs-CRP (mg/dL)	5.8 (10.2)	5.8 (10.2)
(N)	(3,267)	(3,267)

TABLE 2. Prevalence of atrial fibrillation by eGFR, age, sex, and race/ethnicity: Comparison of published results to results calculated from archived dataset.

SUBPOPULATIONS	PUBLISHED			CALCULATED		
	N	AF n (%)	P value	N	AF n (%)	P value
All population	3267	602 (18.4%)		3267	602 (18.4%)	
eGFR (mL/[min 1.73 m2])			<.0010			0.001
<45	1795	367 (20.4%)		1795	367 (20.4%)	
45+	1472	235 (16.0%)		1472	235 (16.0%)	
Age (y)			<.0001			<0.001
<40	239	19 (7.9%)		239	19 (7.9%)	
40-49	398	49 (12.3%)		398	49 (12.3%)	
50-59	958	162 (16.9%)		958	162 (16.9%)	
60-69	1217	256 (21.0%)		1217	256 (21.0%)	
70+	455	116 (25.5%)		455	116 (25.5%)	
Sex			0.7807			0.7807
Male	1775	324 (18.3%)		1775	324 (18.3%)	
Female	1492	278 (18.6%)		1492	278 (18.6%)	
Race/ethnicity			0.0156			0.016
Non-Hispanic white	1627	273 (16.8%)		1627	273 (16.8%)	
Non-Hispanic black	1640	329 (20.1%)		1640	329 (20.1%)	

TABLE 3a. Unadjusted association of AF (atrial fibrillation) and selected sociodemographic and biomedical characteristics of patients: Comparison published results (Table 3, columns 1-3) and calculations from dataset archived at NIDDK Central Repository.

Independent Variables	PUBLISHED		CALCULATED	
	OR (95% CI)*	P value	OR (95% CI)*	P value
Age (y)‡	1.03 (1.02-1.04)	<.0001	1.03 (1.02-1.04)	<.001
Male sex**	0.98 (0.82-1.16)**	0.7805	0.975 (0.817-1.164)**	0.781
Race (non-Hispanic black vs. other)	1.24 (1.04-1.49)	0.0157	1.24 (1.04-1.49)	0.0155
Education (reference: post-high school education)				
Less than high school graduate	1.40 (1.09-1.80)	0.0086	1.40 (1.09-1.80)	0.009
High school graduate	0.86 (0.67-1.12)	0.2673	0.86 (0.67-1.12)	0.267
College graduate	0.80 (0.61-1.04)	0.099	0.80 (0.61-1.04)	0.099
Postgraduate degree	0.69 (0.51-0.93)	0.0162	0.69 (0.51-0.93)	0.016
Total cholesterol (mg/dL)‡	1.00 (0.99-1.00)	0.0006	0.996 (0.994 - 0.998)	0.001
Uric acid (mg/dL) ‡	1.04 (1.00-1.09)	0.0716	1.040 (0.996-1.093)	0.072
eGFR (mL/[min 1.73 m ²]) (eGFR<45 vs. 45)	1.35 (1.13-1.62)	0.001	1.35 (1.13-1.62)	0.001
hs-CRP (mg/dL)‡	1.00 (1.00-1.01)	0.2751	1.004 (0.996-1.0125)	0.2751
Body mass index (kg/m ²)‡	1.01 (1.00-1.02)	0.0398	1.01 (1.00-1.02)	0.04
Physical activity (total MET h/wk)‡	1.00 (1.00-1.00)	<.0001	0.998 (0.997-0.999)	<.001
Smoking status (reference: never)				
Current	1.38 (1.04-1.82)	0.0237	1.38 (1.04-1.82)	0.024
Former	1.78 (1.46-2.16)	<.0001	1.78 (1.46-2.16)	<.001
Alcohol use (drinker vs. other)	0.75 (0.62-0.90)	0.0022	0.75 (0.62-0.90)	0.002
Hypertension	1.18 (0.91-1.54)	0.2201	1.18 (0.91-1.54)	0.22
Diabetes	1.25 (1.05-1.49)	0.0139	1.25 (1.05-1.49)	0.014
Congestive heart failure	5.63 (4.43-7.14)	<.0001	5.63 (4.43-7.14)	<.001
Any cardiovascular disease	3.48 (2.90-4.17)	<.0001	3.48 (2.90-4.17)	<.001

* Odds ratios represent unadjusted association of individual variables in the first column with AF.

** As discussed in text, the published article reports results as being for FEMALE subjects. Analysis indicates that published results are for MALE subjects.

‡ The originally published table states that coefficients for continuous variables were odds ratios for a *one standard deviation change* in independent variable. While this is true for Model 2 (see Table 3b) it is not true for the Model 1 (above). Calculations of unadjusted results shown above are for a single unit change in the independent variable (e.g., one year of age) not 1 sd of change in independent variable. The authors issued a correction noting this fact in Am Heart J. 2010 Dec;160(6):1190.

TABLE 3b. Association of AF (atrial fibrillation) and selected sociodemographic and biomedical characteristics of patients controlling for sociodemographic characteristics of sample (age, gender, ethnicity) and clinical site: Comparison of published results (Table 3, columns 1 and 4-5) and calculations from dataset archived at NIDDK Central Repository.

Independent Variables	PUBLISHED		CALCULATED		
	OR (95% CI) [†]	P value	OR [†]	95% CI [†]	P value ^{**}
Age (y) [‡]	1.46 (1.32-1.62)	<.0001	1.464	1.321 - 1.621	<0.001
Male sex*	0.90 (0.75-1.09)	0.2875	0.903	0.748 - 1.090	0.287
Race (non-Hispanic black vs. non-Hispanic white)	1.25 (1.03-1.52)	0.0228	1.252	1.032 - 1.518	0.023
Education (reference: post-high school education)		0.0452			0.045
Less than high school graduate	1.13 (0.86-1.47)		1.126	0.862 - 1.471	
High school graduate	0.78 (0.60-1.01)		0.771	0.598 - 1.010	
College graduate	0.85 (0.65-1.13)		0.854	0.648 - 1.126	
Postgraduate degree	0.73 (0.53-1.00)		0.728	0.529 - 1.001	
Total cholesterol (mg/dL) [‡]	0.85 (0.77-0.94)	0.0014	0.853	0.774 - 0.940	0.001
Uric acid (mg/dL) [‡]	1.01 (0.92-1.11)	0.8366	1.010	0.918 - 1.112	0.837
eGFR (mL/[min 1.73 m ²]) (eGFR<45 vs 45)	1.12 (0.92-1.35)	0.271	1.115	0.920 - 1.353	0.271
hs-CRP (mg/dL) [‡]	1.02 (0.94-1.12)	0.5922	1.024	0.940 - 1.115	0.592
Body mass index (kg/m ²) [‡]	1.09 (1.00-1.20)	0.0599	1.092	0.996 - 1.197	0.060
Physical activity (total MET h/wk) [‡]	0.85 (0.76-0.95)	0.0056	0.853	0.763 - 0.955	0.006
Smoking status (reference: never)		0.0001			
Current	1.30 (0.98-1.73)		1.299	0.976 - 1.727	
Former	1.56 (1.27-1.91)		1.560	1.273 - 1.913	
Alcohol use (drinker vs nondrinkers)	0.87 (0.71-1.06)	0.1678	0.870	0.713 - 1.061	0.168
Hypertension	0.83 (0.63-1.11)	0.2036	0.831	0.625 - 1.105	0.204
Diabetes	1.11 (0.92-1.33)	0.2828	1.106	0.921 - 1.328	0.283
Congestive heart failure	5.20 (4.06-6.67)	<.0001	5.205	4.062 - 6.671	<0.001
Any cardiovascular disease	3.06 (2.53-3.71)	<.0001	3.062	2.529 - 3.708	<0.001

* As discussed in text, the published article reports these results as being for FEMALE subjects. Analysis indicates that published results are for MALE subjects.

** For binary independent variables, p-values are for tests of statistical significance of the estimated logisitic regression coefficient. For categorical independent variables with 3 or more categories (i.e., education and smoking), p-values are for Wald tests of hypotheses that impact of categories of these variables was zero.

† Odds ratios represent age-, sex-, race/ethnicity-, and clinical center-adjusted associations of the individual variables in the first column with AF.

‡ Odds ratios for continuous variables are for 1-SD increase.

TABLE 4. Multivariable adjusted associations of variables with atrial fibrillation: Comparison of published and calculated values.

INDEPENDENT VARIABLES*	PUBLISHED		CALCULATED				
	OR (95% CI)‡	P value	OR‡	95% CI		P values**	
Age (y)†	1.27 (1.13-1.43)	< .0001	1.272	1.131	-	1.430	< 0.001
Male sex***	0.80 (0.65-0.98)	0.0303	0.796	0.648	-	0.979	0.030
Race (non-Hispanic black)	1.07 (0.86-1.34)	0.5283	1.073	0.862	-	1.335	0.528
Education (reference: post-high school education)		0.0662			-		0.093
Less than high school graduate	1.14 (0.86-1.52)		1.141	0.859	-	1.517	
High school graduate	0.76 (0.58-1.01)		0.762	0.578	-	1.009	
College graduate	1.03 (0.77-1.37)		1.025	0.768	-	1.369	
Postgraduate degree	0.84 (0.60-1.17)		0.838	0.600	-	1.170	
Total cholesterol (mg/dL)†	0.93 (0.84-1.02)	0.1243	0.925	0.837	-	1.022	0.124
Body mass index (kg/m ²)†	1.04 (0.94-1.14)	0.4809	1.036	0.939	-	1.142	0.481
Physical activity (total MET h/wk)†	0.93 (0.83-1.04)	0.1957	0.927	0.827	-	1.040	0.195
Smoking status (reference: never)		0.0276					0.029
Current	1.15 (0.84-1.56)		1.145	0.843	-	1.556	
Former	1.34 (1.08-1.66)		1.336	1.078	-	1.656	
Congestive heart failure	3.28 (2.47-4.36)	< .0001	3.280	2.468	-	4.356	<0.001
Any cardiovascular disease	1.94 (1.56-2.43)	< .0001	1.943	1.556	-	2.426	<0.001

* Study geographic center was also included in the multivariable model, but their ORs were omitted from this table.

** For categorical independent variables., p-values for calculated results are Wald tests of hypotheses that impact of categories of individual variables were zero.

*** As discussed in text, the published article reports these results as being for FEMALE subjects. Analysis indicates that published results are for MALE subjects.

† Odds ratios for continuous variables are for 1-SD increase.

‡ Odds ratios represent the multivariable associations of the individual variables with atrial fibrillation.

Appendix A

PUBLISHED ARTICLE

Soliman EZ, Prineas RJ, Go AS, Xie D, Lash JP, Rahman M, Ojo A, Teal VL, Jensvold NG, Robinson NL, Dries DL, Bazzano L, Mohler ER, Wright JT, Feldman HI; Chronic Renal Insufficiency Cohort (CRIC) Study Group. Chronic kidney disease and prevalent atrial fibrillation: the Chronic Renal Insufficiency Cohort (CRIC). *Am Heart J.* 2010 Jun;159(6):1102-7.

Erratum in: *Am Heart J.* 2010 Dec;160(6):1190.

APPENDIX B

STATA output for Analyses

```

-----
name: <unnamed>
log: Z:\CRIC\AnalysisData\MS037\Label_Run_Output.log
log type: text
opened on: 16 Mar 2011, 14:42:44

```

```

.
. clear all

. *pause on
.
. use Z:\CRIC\AnalysisData\MS037\m37_20100104.dta, clear

. *describe
. #delimit ;
delimiter now ;
. label define yesno 0"No" 1"Yes";

. label values mirevasc chf anycvd alcohol_use hypertension hichol diabetes alcohol37 yesno;

. describe mirevasc chf anycvd alcohol_use hypertension hichol diabetes alcohol37;

```

variable name	storage type	display format	value label	variable label
mirevasc	byte	%8.0g	yesno	myocardial infarction/prior revasc (y/n)
chf	byte	%8.0g	yesno	congestive heart failure (y/n)
anycvd	byte	%8.0g	yesno	cardio-vascular disease (y/n)
alcohol_use	byte	%8.0g	yesno	alcohol use (medhx)
hypertension	byte	%8.0g	yesno	hypertension (y/n)
hichol	byte	%8.0g	yesno	high cholesterol
diabetes	byte	%8.0g	yesno	diabetes
alcohol37	byte	%8.0g	yesno	alcohol use

```

. tab1 mirevasc chf anycvd alcohol_use hypertension hichol diabetes alcohol37;

```

```

-> tabulation of mirevasc

```

myocardial infarction/ prior revasc (y/n)	Freq.	Percent	Cum.
No	2,537	77.66	77.66
Yes	730	22.34	100.00
Total	3,267	100.00	

-> tabulation of chf

congestive heart failure (y/n)	Freq.	Percent	Cum.
No	2,939	89.96	89.96
Yes	328	10.04	100.00
Total	3,267	100.00	

-> tabulation of anycvd

cardio-vasc ular disease (y/n)	Freq.	Percent	Cum.
No	2,142	65.56	65.56
Yes	1,125	34.44	100.00
Total	3,267	100.00	

-> tabulation of alcoh_use

alcohol use (medhx)	Freq.	Percent	Cum.
No	1,155	35.35	35.35
Yes	2,112	64.65	100.00
Total	3,267	100.00	

-> tabulation of hypertension

hypertensio n (y/n)	Freq.	Percent	Cum.
No	458	14.03	14.03
Yes	2,807	85.97	100.00
Total	3,265	100.00	

-> tabulation of hichol

high cholesterol	Freq.	Percent	Cum.
No	619	18.95	18.95
Yes	2,648	81.05	100.00
Total	3,267	100.00	

-> tabulation of diabetes

diabetes	Freq.	Percent	Cum.
No	1,781	54.51	54.51
Yes	1,486	45.49	100.00
Total	3,267	100.00	

-> tabulation of alcoh37

alcohol use	Freq.	Percent	Cum.
No	1,957	59.90	59.90
Yes	1,310	40.10	100.00
Total	3,267	100.00	

```
. *pause;  
. label define sex 1"Male" 2"Female" 98"Other";  
. label values sex sex;  
. describe sex;
```

variable name	storage type	display format	value label	variable label
sex	byte	%8.0g	sex	sex

```
. tab sex;
```

sex	Freq.	Percent	Cum.
Male	1,775	54.33	54.33
Female	1,492	45.67	100.00
Total	3,267	100.00	

```
. label define race_ethnicity_cat2a 1"Non-Hispanic White" 2"Non-Hispanic  
> Black" 3"Hispanic" 4"Other";
```

```
. label values race_ethnicity_cat2 race_ethnicity_cat2a;  
. describe race_ethnicity_cat2;
```

variable name	storage type	display format	value label	variable label
race_ethnicit~2	byte	%20.0g	race_ethnicity_cat2a	race ethnicity category 2

```
. tab race_ethnicity_cat2;
```

race ethnicity category 2	Freq.	Percent	Cum.
Non-Hispanic White	1,627	49.80	49.80
Non-Hispanic Black	1,640	50.20	100.00
Total	3,267	100.00	

```
. *pause;
```

```
. label define edu_cat_la
```

```
> 1"0-6th grade"  
> 2"7-12 grade, no diploma"  
> 3"HS grad or equiv"  
> 4"Tech-Voc degree"  
> 5"Some college, no degree"  
> 6"College grad"  
> 7"Prof or grad degree";
```

```
. label values edu_cat_1 edu_cat_la;
```

```
. describe edu_cat_1;
```

variable name	storage type	display format	value label	variable label
edu_cat_1	byte	%23.0g	edu_cat_la	education category 1(categorical)

```
. tab edu_cat_1;
```

education category 1(categorical)	Freq.	Percent	Cum.
0-6th grade	27	0.83	0.83
7-12 grade, no diploma	496	15.18	16.01
HS grad or equiv	649	19.87	35.87
Tech-Voc degree	172	5.26	41.14
Some college, no degree	856	26.20	67.34
College grad	606	18.55	85.89
Prof or grad degree	461	14.11	100.00
Total	3,267	100.00	

```
. *pause;
```

```
. label define age_cat_la
```

```
> 1"< 30"  
> 2"30-39"  
> 3"40-49"  
> 4"50-59"  
> 5"60-69"  
> 6"70+";
```

```
. label values age_cat_1 age_cat_la;
```

```
. describe age_cat_1;
```

variable name	storage type	display format	value label	variable label
age_cat_1	byte	%8.0g	age_cat_1a	age category 1

```
. tab age_cat_1;
```

age category 1	Freq.	Percent	Cum.
< 30	49	1.50	1.50
30-39	190	5.82	7.32
40-49	398	12.18	19.50
50-59	958	29.32	48.82
60-69	1,217	37.25	86.07
70+	455	13.93	100.00
Total	3,267	100.00	

```
. *pause;
```

```
. label define age_cat_2a
```

```
> 1"21-44"
```

```
> 2"45-64"
```

```
> 3"65+"
```

```
> 9"Total ???";
```

```
. label values age_cat_2 age_cat_2a;
```

```
. describe age_cat_2;
```

variable name	storage type	display format	value label	variable label
age_cat_2	byte	%9.0g	age_cat_2a	age category 2

```
. tab age_cat_2;
```

age category 2	Freq.	Percent	Cum.
21-44	410	12.55	12.55
45-64	1,885	57.70	70.25
65+	972	29.75	100.00
Total	3,267	100.00	

```
. *pause;
```

```
. label define smoke100a 0"Non-Smoker" 1"Smoker";
```

```
. label values smoke100 smoke100a;
```

```
. describe smoke100;
```

variable name	storage type	display format	value label	variable label
smoke100	byte	%10.0g	smoke100a	smoked 100 cigarettes

```
. tab smoke100;
```

smoked 100 cigarettes	Freq.	Percent	Cum.
Non-Smoker	1,405	43.01	43.01
Smoker	1,862	56.99	100.00
Total	3,267	100.00	

```
. label define smokenow 0"Not current smoker" 1"Yes current smoker";
. label values smokenow smokenow;
. describe smokenow;
```

variable name	storage type	display format	value label	variable label
smokenow	byte	%18.0g	smokenow	current smoker

```
. tab smokenow;
```

current smoker	Freq.	Percent	Cum.
Not current smoker	2,797	85.61	85.61
Yes current smoker	470	14.39	100.00
Total	3,267	100.00	

```
. label define egfr_roche_cat_baseline
> 1 "<30" 2 "30-39" 3 "40-49" 4 "50-59" 5 "60+" 9 "Total ???";
```

```
. label values egfr_roche_cat_baseline egfr_roche_cat_baseline;
. describe egfr_roche_cat_baseline;
```

variable name	storage type	display format	value label	variable label
egfr_roche_ca~e	byte	%9.0g	egfr_roche_cat_baseline	egfr category baseline

```
. tab egfr_roche_cat_baseline;
```

egfr category baseline	Freq.	Percent	Cum.
<30	579	17.72	17.72
30-39	766	23.45	41.17
40-49	889	27.21	68.38
50-59	669	20.48	88.86
60+	364	11.14	100.00
Total	3,267	100.00	

```
. label define egfr_roche_cat1a 1"<45" 2"45+";
. label values egfr_roche_cat1 egfr_roche_cat1a;
. describe egfr_roche_cat1 ;
```

variable name	storage type	display format	value label	variable label
egfr_roche_cat1	byte	%8.0g	egfr_roche_cat1a	

egfr category 1

```
. tab egfr_roche_cat1 ;
```

egfr category 1	Freq.	Percent	Cum.
<45	1,795	54.94	54.94
45+	1,472	45.06	100.00
Total	3,267	100.00	

```
. label define egfr_roche_cat4a 1"<15" 2"15-29" 3"30-59" 4"60+";
```

```
. label values egfr_roche_cat4 egfr_roche_cat4a;
```

```
. describe egfr_roche_cat4;
```

variable name	storage type	display format	value label	variable label
egfr_roche_cat4	byte	%8.0g	egfr_roche_cat4a	egfr category 4

```
. tab egfr_roche_cat4;
```

egfr category 4	Freq.	Percent	Cum.
<15	4	0.12	0.12
15-29	575	17.60	17.72
30-59	2,324	71.14	88.86
60+	364	11.14	100.00
Total	3,267	100.00	

```
. label define bmi_cat_2a
```

```
> 1"<25 (UnderW/Normal)"  
> 2"25-29 (OverW)"  
> 3"30+ (Obese)";
```

```
. label values bmi_cat_2 bmi_cat_2a;
```

```
. describe bmi_cat_2;
```

variable name	storage type	display format	value label	variable label
bmi_cat_2	byte	%19.0g	bmi_cat_2a	bmi category 2

```
. tab bmi_cat_2;
```

bmi category 2	Freq.	Percent	Cum.
<25 (UnderW/Normal)	452	14.18	14.18
25-29 (OverW)	894	28.04	42.22
30+ (Obese)	1,842	57.78	100.00
Total	3,188	100.00	

```
. #delimit cr  
delimiter now cr
```

```
. save "Z:\CRIC\AnalysisData\MS037\m37_w_Labels.dta", replace
```

file Z:\CRIC\AnalysisData\MS037\m37_w_Labels.dta saved

.
end of do-file

. log close
 name: <unnamed>
 log: Z:\CRIC\AnalysisData\MS037\Label_Run_Output.log
 log type: text
 closed on: 16 Mar 2011, 14:46:41

```

name: <unnamed>
log: Z:\CRIC\AnalysisData\MS037\Analysis_1_Output.log
log type: text
opened on: 16 Mar 2011, 15:37:36

```

```
. clear all
```

```
. pause on
```

```
. set linesize 180
```

```
. use "Z:\CRIC\AnalysisData\MS037\m37_w_Labels.dta", clear
```

```
. tab edu_cat_1 edu
```

education category	edu					Total
1(categorical)	1	2	3	4	5	
0-6th grade	27	0	0	0	0	27
7-12 grade, no diplom	496	0	0	0	0	496
HS grad or equiv	0	649	0	0	0	649
Tech-Voc degree	0	0	172	0	0	172
Some college, no degr	0	0	856	0	0	856
College grad	0	0	0	606	0	606
Prof or grad degree	0	0	0	0	461	461
Total	523	649	1,028	606	461	3,267

```
. label define edu 1"<HS" 2"HSG" 3"Post HS" 4"Col Grad" 5"Postgrad"
```

```
. label values edu edu
```

```
. tab edu_cat_1 edu
```

education category	edu					Total
1(categorical)	<HS	HSG	Post HS	Col Grad	Postgrad	
0-6th grade	27	0	0	0	0	27
7-12 grade, no diplom	496	0	0	0	0	496
HS grad or equiv	0	649	0	0	0	649
Tech-Voc degree	0	0	172	0	0	172
Some college, no degr	0	0	856	0	0	856
College grad	0	0	0	606	0	606
Prof or grad degree	0	0	0	0	461	461
Total	523	649	1,028	606	461	3,267

```
. tab smoke100 smokenow
```

smoked 100	current smoker		Total
cigarettes	Not curre	Yes curre	
Non-Smoker	1,405	0	1,405
Smoker	1,392	470	1,862
Total	2,797	470	3,267

```
. tab smoker
```

smoker	Freq.	Percent	Cum.
1	470	14.39	14.39
2	1,392	42.61	56.99
3	1,405	43.01	100.00
Total	3,267	100.00	

```
. label define smoker 1"Current" 2"Former" 3"Never"
```

```
. label values smoker smoker
```

```
. label var totalmetsum "Total Physical Activity (MET h/wk)"
```

```
. generate egfr_lt_45 = .  
(3267 missing values generated)
```

```
. replace egfr_lt_45 = 1 if egfr_roche < 45  
(1795 real changes made)
```

```
. replace egfr_lt_45 = 0 if egfr_roche >= 45  
(1472 real changes made)
```

```
. * tab egfr_lt_45  
. label var egfr_lt_45 "eGFR < 45"
```

```
. tab1 sex race_ethnicity_cat2 edu smoker alcoh37 hypertension diabetes chf anycvd  
egfr_lt_45
```

```
-> tabulation of sex
```

sex	Freq.	Percent	Cum.
Male	1,775	54.33	54.33
Female	1,492	45.67	100.00
Total	3,267	100.00	

```
-> tabulation of race_ethnicity_cat2
```

race ethnicity category 2	Freq.	Percent	Cum.
Non-Hispanic White	1,627	49.80	49.80
Non-Hispanic Black	1,640	50.20	100.00
Total	3,267	100.00	

```
-> tabulation of edu
```

edu	Freq.	Percent	Cum.
<HS	523	16.01	16.01
HSG	649	19.87	35.87
Post HS	1,028	31.47	67.34
Col Grad	606	18.55	85.89
Postgrad	461	14.11	100.00
Total	3,267	100.00	

-> tabulation of smoker

smoker	Freq.	Percent	Cum.
Current	470	14.39	14.39
Former	1,392	42.61	56.99
Never	1,405	43.01	100.00
Total	3,267	100.00	

-> tabulation of alcoh37

alcohol use	Freq.	Percent	Cum.
No	1,957	59.90	59.90
Yes	1,310	40.10	100.00
Total	3,267	100.00	

-> tabulation of hypertension

hypertensio n (y/n)	Freq.	Percent	Cum.
No	458	14.03	14.03
Yes	2,807	85.97	100.00
Total	3,265	100.00	

-> tabulation of diabetes

diabetes	Freq.	Percent	Cum.
No	1,781	54.51	54.51
Yes	1,486	45.49	100.00
Total	3,267	100.00	

-> tabulation of chf

congestive heart failure (y/n)	Freq.	Percent	Cum.
No	2,939	89.96	89.96
Yes	328	10.04	100.00
Total	3,267	100.00	

-> tabulation of anycvd

cardio-vasc ular disease (y/n)	Freq.	Percent	Cum.
No	2,142	65.56	65.56
Yes	1,125	34.44	100.00
Total	3,267	100.00	

-> tabulation of egfr_lt_45

eGFR < 45	Freq.	Percent	Cum.
0	1,472	45.06	45.06
1	1,795	54.94	100.00
Total	3,267	100.00	

. summarize age totalmetsum egfr_roche bmi tc uric_acid hs_crp

Variable	Obs	Mean	Std. Dev.	Min	Max
age	3267	58.55395	10.80797	21.15733	75.15343
totalmetsum	3259	204.6259	148.3339	0	1692.25
egfr_roche	3267	43.62454	13.37062	7.004488	113.9882
bmi	3259	32.31229	8.039422	14.59011	88.01366
tc	3258	182.8432	43.83615	76	571
uric_acid	3220	7.402422	1.906139	1.9	15.2
hs_crp	3257	5.799183	10.15226	.08	187

.
end of do-file

. exit, clear

Z:\CRIC\AnalysisData\MS037\Analysis_2_Output.log

name: <unnamed>
log: Z:\CRIC\AnalysisData\MS037\Analysis_2_Output.log
log type: text
opened on: 6 Apr 2011, 14:22:13

. clear all

. pause on

. set linesize 180

. use "Z:\CRIC\AnalysisData\MS037\m37_w_Labels.dta", clear

. *tab edu_cat_1 edu

. label define edu 1"<HS" 2"HSG" 3"Post HS" 4"Col Grad" 5"Postgrad"

. label values edu edu

. *tab edu_cat_1 edu

. *tab smoke100 smokenow

. *tab smoker

. label define smoker 1"Current" 2"Former" 3"Never"

. label values smoker smoker

. label var totalmetsum "Total Physical Activity (MET h/wk)"

. generate egfr_lt_45 = .
(3267 missing values generated)

. replace egfr_lt_45 = 1 if egfr_roche < 45
(1795 real changes made)

. replace egfr_lt_45 = 0 if egfr_roche >= 45
(1472 real changes made)

. * tab egfr_lt_45

. label var egfr_lt_45 "eGFR < 45"

. *tab1 sex race_ethnicity_cat2 edu smoker alcoh37 hypertension diabetes
chf anycvd egfr_lt_45

. *summarize age totalmetsum egfr_roche bmi tc uric_acid hs_crp

. tab comp_afib

ecg and/or	Freq.	Percent	Cum.
self report			
0	2,665	81.57	81.57
1	602	18.43	100.00
Total	3,267	100.00	

. tab egfr_lt_45 comp_afib, row chi

```

+-----+
| Key |
+-----+
| frequency |
| row percentage |
+-----+

```

eGFR < 45	ecg and/or self report		Total
	0	1	
0	1,237	235	1,472
	84.04	15.96	100.00
1	1,428	367	1,795
	79.55	20.45	100.00
Total	2,665	602	3,267
	81.57	18.43	100.00

Pearson chi2(1) = 10.8039 Pr = 0.001

```

. gen age5=.
(3267 missing values generated)

. replace age5 = 39 if age<40
(239 real changes made)

. replace age5 = 40 if age>= 40 & age < 50
(398 real changes made)

. replace age5 = 50 if age>= 50 & age < 60
(958 real changes made)

. replace age5 = 60 if age>= 60 & age < 70
(1217 real changes made)

. replace age5 = 70 if age>= 70
(455 real changes made)

. tab age5 comp_afib, row chi

```

```

+-----+
| Key |
+-----+
| frequency |
| row percentage |
+-----+

```


age5	ecg and/or self report		Total
	0	1	
39	220	19	239
	92.05	7.95	100.00
40	349	49	398
	87.69	12.31	100.00
50	796	162	958
	83.09	16.91	100.00
60	961	256	1,217
	78.96	21.04	100.00
70	339	116	455
	74.51	25.49	100.00
Total	2,665	602	3,267
	81.57	18.43	100.00

Pearson chi2(4) = 49.4509 Pr = 0.000

. tab sex comp_afib, row chi

```

+-----+
| Key      |
+-----+
| frequency |
| row percentage |
+-----+

```

sex	ecg and/or self report		Total
	0	1	
Male	1,451	324	1,775
	81.75	18.25	100.00
Female	1,214	278	1,492
	81.37	18.63	100.00
Total	2,665	602	3,267
	81.57	18.43	100.00

Pearson chi2(1) = 0.0775 Pr = 0.781

. tab race_ethnicity_cat2 comp_afib, row chi

```

+-----+
| Key      |
+-----+
| frequency |
| row percentage |
+-----+

```

race ethnicity category 2	ecg and/or self report		Total
	0	1	
Non-Hispanic White	1,354	273	1,627
	83.22	16.78	100.00
Non-Hispanic Black	1,311	329	1,640
	79.94	20.06	100.00
Total	2,665	602	3,267
	81.57	18.43	100.00

Pearson chi2(1) = 5.8515 Pr = 0.016

.
.

end of do-file

. log close

name: <unnamed>

log: Z:\CRIC\AnalysisData\MS037\Analysis_2_Output.log

log type: text

closed on: 6 Apr 2011, 14:24:41

```

name: <unnamed>
log: Z:\CRIC\AnalysisData\MS037\Table_3a_3b_4_Output_rev1.log
log type: text
opened on: 16 Apr 2011, 21:52:14

```

```
. clear all
```

```
. *pause on
. set linesize 180
```

```
. use "Z:\CRIC\AnalysisData\MS037\m37_w_Labels.dta", clear
```

```
. generate NonHispanic_Black = race_ethnicity_cat2
```

```
. recode NonHispanic_Black (1=0) (2=1)
(NonHispanic_Black: 3267 changes made)
```

```
. tab race_ethnicity_cat2 NonHispanic_Black
```

race ethnicity category 2	NonHispanic_Black		Total
	0	1	
Non-Hispanic White	1,627	0	1,627
Non-Hispanic Black	0	1,640	1,640
Total	1,627	1,640	3,267

```
. generate female = sex
```

```
. recode female (2=1) (1=0)
(female: 3267 changes made)
```

```
. tab female sex
```

female	sex		Total
	Male	Female	
0	1,775	0	1,775
1	0	1,492	1,492
Total	1,775	1,492	3,267

```
. generate male = sex
```

```
. recode male (2=0)
(male: 1492 changes made)
```

```
. tab male sex
```

male	sex		Total
	Male	Female	
0	0	1,492	1,492
1	1,775	0	1,775
Total	1,775	1,492	3,267

```

.
. *tab edu_cat_1 edu
. label define edu 1"<HS" 2"HSG" 3"Post HS" 4"Col Grad" 5"Postgrad"

. label values edu edu

. *tab edu_cat_1 edu
.
. *tab smoke100 smokenow
. *tab smoker
. label define smoker 1"Current" 2"Former" 3"Never"

. label values smoker smoker

.

. label var totalmetsum "Total Physical Activity (MET h/wk)"

.

. generate egfr_lt_45 = .
(3267 missing values generated)

. replace egfr_lt_45 = 1 if egfr_roche < 45
(1795 real changes made)

. replace egfr_lt_45 = 0 if egfr_roche >= 45
(1472 real changes made)

. * tab egfr_lt_45
. label var egfr_lt_45 "eGFR < 45"

.

. *tab1 sex race_ethnicity_cat2 edu smoker alcoh37 hypertension diabetes chf
anycvd egfr_lt_45
. *summarize age totalmetsum egfr_roche bmi tc uric_acid hs_crp
.
. * tab comp_afib
. * tab egfr_lt_45 comp_afib, row chi
. gen age5=.
(3267 missing values generated)

. replace age5 = 39 if age<40
(239 real changes made)

. replace age5 = 40 if age>= 40 & age < 50
(398 real changes made)

. replace age5 = 50 if age>= 50 & age < 60
(958 real changes made)

. replace age5 = 60 if age>= 60 & age < 70
(1217 real changes made)

. replace age5 = 70 if age>= 70
(455 real changes made)

. * tab age5 comp_afib, row chi
. * tab sex comp_afib, row chi
. * tab race_ethnicity_cat2 comp_afib, row chi
.
.
. logistic comp_afib age

```

Logistic regression

Number of obs = 3267

```

Log likelihood = -1533.101
LR chi2(1) = 55.76
Prob > chi2 = 0.0000
Pseudo R2 = 0.0179

```

```

-----+-----
comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
age | 1.034323     .004899       7.12  0.000     1.024765    1.043969
-----+-----

```

```
. logistic comp_afib age_1sd
```

```

Logistic regression
Number of obs = 3267
LR chi2(1) = 55.76
Prob > chi2 = 0.0000
Pseudo R2 = 0.0179
Log likelihood = -1533.101

```

```

-----+-----
comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
age_1sd | 1.440129     .0737225     7.12  0.000     1.302648    1.59212
-----+-----

```

```
. logistic comp_afib female
```

```

Logistic regression
Number of obs = 3267
LR chi2(1) = 0.08
Prob > chi2 = 0.7807
Pseudo R2 = 0.0000
Log likelihood = -1560.9414

```

```

-----+-----
comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
female | 1.02553      .0928456     0.28  0.781     .8587874    1.224648
-----+-----

```

```
. logistic comp_afib male
```

```

Logistic regression
Number of obs = 3267
LR chi2(1) = 0.08
Prob > chi2 = 0.7807
Pseudo R2 = 0.0000
Log likelihood = -1560.9414

```

```

-----+-----
comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
male | .9751052     .0882804    -0.28  0.781     .816561     1.164433
-----+-----

```

```
. xi: logistic comp_afib NonHispanic
```

```

Logistic regression
Number of obs = 3267
LR chi2(1) = 5.86
Prob > chi2 = 0.0155
Pseudo R2 = 0.0019
Log likelihood = -1558.0507

```

```

-----+-----
comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
NonHispanic | 1.244656     .1127348     2.42  0.016     1.042201    1.486438
-----+-----

```

```

. char _dta[omit] "prevalent"

. xi: logistic comp_afib i.edu, or
i.edu          _Iedu_1-5          (naturally coded; _Iedu_3 omitted)

Logistic regression                                Number of obs   =       3267
                                                    LR chi2(4)      =       23.89
                                                    Prob > chi2     =       0.0001
Log likelihood = -1549.0339                        Pseudo R2      =       0.0077

```

```

-----+-----
      comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      _Iedu_1 |   1.400871   .1796173     2.63  0.009     1.089579     1.8011
      _Iedu_2 |   .8648755   .1131734    -1.11  0.267     .6692209     1.117732
      _Iedu_4 |   .798853    .1087416    -1.65  0.099     .6117859     1.04312
      _Iedu_5 |   .6880676   .1069612    -2.41  0.016     .5073533     .9331505
-----+-----

```

```

. test _Iedu_1 _Iedu_2 _Iedu_4 _Iedu_5

```

```

( 1) [comp_afib]_Iedu_1 = 0
( 2) [comp_afib]_Iedu_2 = 0
( 3) [comp_afib]_Iedu_4 = 0
( 4) [comp_afib]_Iedu_5 = 0

      chi2( 4) =    24.24
      Prob > chi2 =    0.0001

```

```

. char _dta[omit]

```

```

. logistic comp_afib tc

Logistic regression                                Number of obs   =       3258
                                                    LR chi2(1)      =       12.33
                                                    Prob > chi2     =       0.0004
Log likelihood = -1550.0064                        Pseudo R2      =       0.0040

```

```

-----+-----
      comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      tc |   .9962555   .001084     -3.45  0.001     .9941331     .9983825
-----+-----

```

```

. logistic comp_afib tc_1sd

Logistic regression                                Number of obs   =       3258
                                                    LR chi2(1)      =       12.33
                                                    Prob > chi2     =       0.0004
Log likelihood = -1550.0064                        Pseudo R2      =       0.0040

```

```

-----+-----
      comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      tc_1sd |   .8483581   .0404657    -3.45  0.001     .7726413     .9314949
-----+-----

```

```

. logistic comp_afib uric_acid

Logistic regression                                Number of obs   =       3220

```

```

Log likelihood = -1533.4183
LR chi2(1) = 3.24
Prob > chi2 = 0.0719
Pseudo R2 = 0.0011

```

```

-----+-----
comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
uric_acid |   1.043772   .0248188     1.80  0.072     .9962442    1.093567
-----+-----

```

```
. logistic comp_afib uric_acid_lsd
```

```

Logistic regression
Number of obs = 3220
LR chi2(1) = 3.24
Prob > chi2 = 0.0719
Pseudo R2 = 0.0011
Log likelihood = -1533.4183

```

```

-----+-----
comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
uric_acid_~d | 1.085088   .0491807     1.80  0.072     .992853    1.185891
-----+-----

```

```
. logistic comp_afib egfr_lt_45
```

```

Logistic regression
Number of obs = 3267
LR chi2(1) = 10.89
Prob > chi2 = 0.0010
Pseudo R2 = 0.0035
Log likelihood = -1555.533

```

```

-----+-----
comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
egfr_lt_45 | 1.352819   .1246416     3.28  0.001     1.129313    1.620559
-----+-----

```

```
. logistic comp_afib hs_crp
```

```

Logistic regression
Number of obs = 3257
LR chi2(1) = 1.13
Prob > chi2 = 0.2883
Pseudo R2 = 0.0004
Log likelihood = -1555.4041

```

```

-----+-----
comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
hs_crp | 1.004461   .0040996     1.09  0.275     .9964576    1.012528
-----+-----

```

```
. logistic comp_afib hs_crp_lsd
```

```

Logistic regression
Number of obs = 3257
LR chi2(1) = 1.13
Prob > chi2 = 0.2883
Pseudo R2 = 0.0004
Log likelihood = -1555.4041

```

```

-----+-----
comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
hs_crp_lsd | 1.046221   .0433506     1.09  0.275     .964614    1.134732
-----+-----

```

```
. logistic comp_afib bmi
```

```
Logistic regression                Number of obs   =       3259
                                   LR chi2(1)         =         4.14
                                   Prob > chi2         =       0.0418
Log likelihood = -1552.8146         Pseudo R2      =       0.0013
```

```
-----+-----
      comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      bmi       |   1.011263   .0055099     2.06   0.040     1.000522     1.02212
-----+-----
```

```
. logistic comp_afib bmi_1sd
```

```
Logistic regression                Number of obs   =       3259
                                   LR chi2(1)         =         4.14
                                   Prob > chi2         =       0.0418
Log likelihood = -1552.8146         Pseudo R2      =       0.0013
```

```
-----+-----
      comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      bmi_1sd   |   1.094223   .0479304     2.06   0.040     1.004201     1.192316
-----+-----
```

```
. logistic comp_afib totalmetsum
```

```
Logistic regression                Number of obs   =       3259
                                   LR chi2(1)         =       25.43
                                   Prob > chi2         =       0.0000
Log likelihood = -1545.1501         Pseudo R2      =       0.0082
```

```
-----+-----
      comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      totalmetsum | .9982108   .0003822    -4.68   0.000     .9974619     .9989603
-----+-----
```

```
. logistic comp_afib totalmetsum_1sd
```

```
Logistic regression                Number of obs   =       3259
                                   LR chi2(1)         =       25.43
                                   Prob > chi2         =       0.0000
Log likelihood = -1545.1501         Pseudo R2      =       0.0082
```

```
-----+-----
      comp_afib | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      totalmetsum_1sd | .7667199   .043549    -4.68   0.000     .685945     .8570067
-----+-----
```

```
. char _dta[omit] "prevalent"
```

```
. xi: logistic comp_afib i.smoker
i.smoker      _ismoker_1-3      (naturally coded; _ismoker_3 omitted)
```

```
Logistic regression                Number of obs   =       3267
                                   LR chi2(2)         =       34.17
                                   Prob > chi2         =       0.0000
```


Log likelihood = -1543.8971 Pseudo R2 = 0.0109

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
_Ismoker_1	1.376622	.1944867	2.26	0.024	1.043659	1.815812
_Ismoker_2	1.779791	.1775179	5.78	0.000	1.463759	2.164057

```
. test _Ismoker_1 _Ismoker_2

( 1) [comp_afib]_Ismoker_1 = 0
( 2) [comp_afib]_Ismoker_2 = 0

       chi2( 2) =    33.43
       Prob > chi2 =    0.0000
```

```
. char _dta[omit]
```

```
. logistic comp_afib alcoh37
```

```
Logistic regression                      Number of obs        =                      3267
LR chi2(1)                                =                      9.58
Prob > chi2                                =                      0.0020
Log likelihood = -1556.1882                Pseudo R2            =                      0.0031
```

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
alcoh37	.7487632	.0706104	-3.07	0.002	.6224062	.9007725

```
. logistic comp_afib hypertension
```

```
Logistic regression                      Number of obs        =                      3265
LR chi2(1)                                =                      1.55
Prob > chi2                                =                      0.2138
Log likelihood = -1559.7999                Pseudo R2            =                      0.0005
```

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
hypertension	1.180359	.1595907	1.23	0.220	.9055807	1.538512

```
. logistic comp_afib diabetes
```

```
Logistic regression                      Number of obs        =                      3267
LR chi2(1)                                =                      6.05
Prob > chi2                                =                      0.0139
Log likelihood = -1557.9562                Pseudo R2            =                      0.0019
```

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
diabetes	1.248945	.1128488	2.46	0.014	1.046244	1.490919

```
. logistic comp_afib chf
```

```
Logistic regression                      Number of obs        =                      3267
LR chi2(1)                                =                      189.01
```

Log likelihood = -1466.4746 Prob > chi2 = 0.0000
Pseudo R2 = 0.0605

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
chf	5.625733	.6860969	14.16	0.000	4.429649 7.144783

. logistic comp_afib anycvd

Logistic regression Number of obs = 3267
LR chi2(1) = 182.68
Log likelihood = -1469.6413 Prob > chi2 = 0.0000
Pseudo R2 = 0.0585

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
anycvd	3.476231	.3237289	13.38	0.000	2.896272 4.172324

```

-----
name: <unnamed>
log: Z:\CRIC\AnalysisData\MS037\Table_3b_4_Output_rev1.log
log type: text
opened on: 21 Apr 2011, 15:26:28

. clear all

. pause on

. set linesize 180

.
. use "Z:\CRIC\AnalysisData\MS037\m37_w_Labels.dta", clear
.
. generate NonHispanic_Black = race_ethnicity_cat2

. recode NonHispanic_Black (1=0) (2=1)
(NonHispanic_Black: 3267 changes made)

. tab race_ethnicity_cat2 NonHispanic_Black

      race ethnicity |      NonHispanic_Black
      category 2 |      0      1 |      Total
-----+-----+-----
Non-Hispanic White |      1,627      0 |      1,627
Non-Hispanic Black |      0      1,640 |      1,640
-----+-----+-----
Total |      1,627      1,640 |      3,267

. generate female = sex

. recode female (2=1) (1=0)
(female: 3267 changes made)

. tab female sex

      female |      sex
      Male   Female |      Total
-----+-----+-----
0 |      1,775      0 |      1,775
1 |      0      1,492 |      1,492
-----+-----+-----
Total |      1,775      1,492 |      3,267

. generate male = sex

. recode male (2=0)
(male: 1492 changes made)

. tab male sex

      male |      sex
      Male   Female |      Total
-----+-----+-----
0 |      0      1,492 |      1,492
1 |      1,775      0 |      1,775
-----+-----+-----
Total |      1,775      1,492 |      3,267

.
. *tab edu_cat_1 edu
. label define edu 1"<HS" 2"HSG" 3"Post HS" 4"Col Grad" 5"Postgrad"

```

```

. label values edu edu

. *tab  edu_cat_1 edu
.
. *tab smoke100 smokenow
. *tab smoker
. label define smoker 1"Current" 2"Former" 3"Never"

. label values smoker smoker

.

. label var totalmetsum "Total Physical Activity (MET h/wk)"

.

. generate egfr_lt_45 = .
(3267 missing values generated)

. replace egfr_lt_45 = 1 if egfr_roche < 45
(1795 real changes made)

. replace egfr_lt_45 = 0 if egfr_roche >= 45
(1472 real changes made)

. * tab  egfr_lt_45
. label var egfr_lt_45 "eGFR < 45"

.

. *tab1 sex  race_ethnicity_cat2  edu smoker alcoh37  hypertension  diabetes chf
anycvd egfr_lt_45
. *summarize age  totalmetsum  egfr_roche  bmi tc  uric_acid  hs_crp

.

. * tab comp_afib
. * tab egfr_lt_45 comp_afib, row chi
. gen age5=.
(3267 missing values generated)

. replace age5 = 39 if age<40
(239 real changes made)

. replace age5 = 40 if age>= 40 & age < 50
(398 real changes made)

. replace age5 = 50 if age>= 50 & age < 60
(958 real changes made)

. replace age5 = 60 if age>= 60 & age < 70
(1217 real changes made)

. replace age5 = 70 if age>= 70
(455 real changes made)

. * tab age5 comp_afib, row chi
. * tab sex comp_afib, row chi
. * tab race_ethnicity_cat2 comp_afib, row chi

.

. encode ccidsite, generate(site)

. tab site

```

center-site	Freq.	Percent	Cum.
0101	470	14.39	14.39
0201	266	8.14	22.53
0202	251	7.68	30.21
0302	184	5.63	35.84
0303	156	4.78	40.62

0304		175	5.36	45.97
0401		310	9.49	55.46
0402		120	3.67	59.14
0403		98	3.00	62.14
0501		388	11.88	74.01
0601		382	11.69	85.71
0701		403	12.34	98.04
0702		64	1.96	100.00

Total		3,267	100.00	

. describe site

variable name	storage type	display format	value label	variable label

site	long	%8.0g	site	cric center-site

. tab sex male

sex	male		Total
	0	1	

Male	0	1,775	1,775
Female	1,492	0	1,492

Total	1,492	1,775	3,267

. tab sex female

sex	female		Total
	0	1	

Male	1,775	0	1,775
Female	0	1,492	1,492

Total	1,775	1,492	3,267

. xi: logistic comp_afib age_1sd female NonHispanic_Black i.site
i.site _Isite_1-13 (naturally coded; _Isite_1 omitted)

Logistic regression	Number of obs	=	3267
	LR chi2(15)	=	100.75
	Prob > chi2	=	0.0000
Log likelihood = -1510.6054	Pseudo R2	=	0.0323

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	

age_1sd	1.463735	.0763613	7.30	0.000	1.321467	1.62132
female	1.107426	.1062327	1.06	0.287	.9176165	1.336499
NonHispanic_Black	1.251504	.1233542	2.28	0.023	1.031653	1.518206
_Isite_2	1.447202	.2853398	1.87	0.061	.9833328	2.129892
_Isite_3	1.835154	.3542495	3.15	0.002	1.257073	2.679072
_Isite_4	1.133709	.266161	0.53	0.593	.7155915	1.796132
_Isite_5	.853752	.2320538	-0.58	0.561	.501154	1.454428
_Isite_6	1.368834	.3239388	1.33	0.185	.8608178	2.176658
_Isite_7	1.297846	.2632778	1.29	0.199	.8720682	1.931504
_Isite_8	1.572288	.4119116	1.73	0.084	.9408759	2.627434
_Isite_9	.5189551	.1948611	-1.75	0.081	.2486076	1.083291
_Isite_10	1.654797	.2934477	2.84	0.005	1.168961	2.342554

```

    _Isite_11 | 1.462185 .2661301 2.09 0.037 1.023473 2.088951
    _Isite_12 | .8360953 .1650402 -0.91 0.364 .5678501 1.231056
    _Isite_13 | 1.029914 .3992631 0.08 0.939 .4817463 2.201827

```

```

. xi: logistic comp_afib age_1sd male NonHispanic_Black i.site
i.site      _Isite_1-13      (naturally coded; _Isite_1 omitted)

```

```

Logistic regression      Number of obs   =      3267
                        LR chi2(15)       =      100.75
                        Prob > chi2       =      0.0000
Log likelihood = -1510.6054      Pseudo R2       =      0.0323

```

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
age_1sd	1.463735	.0763613	7.30	0.000	1.321467 1.62132
male	.9029945	.086622	-1.06	0.287	.7482237 1.08978
NonHispanic_Black	1.251504	.1233542	2.28	0.023	1.031653 1.518206
_Isite_2	1.447202	.2853398	1.87	0.061	.9833328 2.129892
_Isite_3	1.835154	.3542495	3.15	0.002	1.257073 2.679072
_Isite_4	1.133709	.266161	0.53	0.593	.7155915 1.796132
_Isite_5	.853752	.2320538	-0.58	0.561	.501154 1.454428
_Isite_6	1.368834	.3239388	1.33	0.185	.8608178 2.176658
_Isite_7	1.297846	.2632778	1.29	0.199	.8720682 1.931504
_Isite_8	1.572288	.4119116	1.73	0.084	.9408759 2.627434
_Isite_9	.5189551	.1948611	-1.75	0.081	.2486076 1.083291
_Isite_10	1.654797	.2934477	2.84	0.005	1.168961 2.342554
_Isite_11	1.462185	.2661301	2.09	0.037	1.023473 2.088951
_Isite_12	.8360953	.1650402	-0.91	0.364	.5678501 1.231056
_Isite_13	1.029914	.3992631	0.08	0.939	.4817463 2.201827

```

. char _dta[omit] "prevalent"

```

```

. xi: logistic comp_afib i.edu age_1sd female NonHispanic_Black i.site
i.edu      _Iedu_1-5      (naturally coded; _Iedu_3 omitted)
i.site     _Isite_1-13   (naturally coded; _Isite_1 omitted)

```

```

Logistic regression      Number of obs   =      3267
                        LR chi2(19)       =      110.55
                        Prob > chi2       =      0.0000
Log likelihood = -1505.7049      Pseudo R2       =      0.0354

```

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
_Iedu_1	1.125698	.15351	0.87	0.385	.8616773 1.470616
_Iedu_2	.7771186	.104065	-1.88	0.060	.5977254 1.010352
_Iedu_4	.854016	.1203861	-1.12	0.263	.6478533 1.125785
_Iedu_5	.7281844	.1183527	-1.95	0.051	.5295349 1.001355
age_1sd	1.455599	.0771006	7.09	0.000	1.312064 1.614836
female	1.096238	.1055685	0.95	0.340	.9076817 1.323964
NonHispanic_Black	1.154838	.1214032	1.37	0.171	.9398048 1.419071
_Isite_2	1.433085	.2832575	1.82	0.069	.9728082 2.111139
_Isite_3	1.743907	.339494	2.86	0.004	1.190738 2.554057
_Isite_4	1.105427	.2606482	0.43	0.671	.6963471 1.754828
_Isite_5	.8350595	.2285173	-0.66	0.510	.488408 1.427749
_Isite_6	1.342549	.3189767	1.24	0.215	.8427384 2.138786
_Isite_7	1.286443	.2616764	1.24	0.216	.8634697 1.916612
_Isite_8	1.519183	.3996253	1.59	0.112	.9071919 2.544023
_Isite_9	.501449	.1884929	-1.84	0.066	.2400287 1.047588
_Isite_10	1.557105	.279515	2.47	0.014	1.095265 2.213689
_Isite_11	1.396231	.2565551	1.82	0.069	.9739806 2.001539
_Isite_12	.8360909	.1659101	-0.90	0.367	.5666891 1.233565
_Isite_13	1.047401	.4070962	0.12	0.905	.4889612 2.243633

```
-----
. test _Iedu_1 _Iedu_2 _Iedu_4 _Iedu_5
```

- (1) [comp_afib]_Iedu_1 = 0
- (2) [comp_afib]_Iedu_2 = 0
- (3) [comp_afib]_Iedu_4 = 0
- (4) [comp_afib]_Iedu_5 = 0

```
      chi2( 4) =      9.73
      Prob > chi2 =    0.0452
```

```
. char _dta[omit]
```

```
. xi: logistic comp_afib tc_1sd age_1sd female NonHispanic_Black i.site
i.site          _Isite_1-13          (naturally coded; _Isite_1 omitted)
```

```
Logistic regression                                Number of obs   =      3258
                                                    LR chi2(16)    =      111.20
                                                    Prob > chi2    =      0.0000
Log likelihood = -1500.5723                       Pseudo R2      =      0.0357
```

```
-----
```

	comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
	tc_1sd	.8533107	.0423187	-3.20	0.001	.7742712	.9404187
	age_1sd	1.452091	.0761003	7.12	0.000	1.310342	1.609175
	female	1.18494	.1160956	1.73	0.083	.9779105	1.435799
NonHispanic_Black		1.260774	.1246572	2.34	0.019	1.038666	1.530377
	_Isite_2	1.463409	.2890991	1.93	0.054	.9935948	2.155371
	_Isite_3	1.863214	.3605012	3.22	0.001	1.275174	2.722427
	_Isite_4	1.128548	.2655125	0.51	0.607	.7116375	1.789705
	_Isite_5	.8400294	.2287576	-0.64	0.522	.4926001	1.432499
	_Isite_6	1.372144	.3251617	1.34	0.182	.8623576	2.183292
	_Isite_7	1.31729	.2676795	1.36	0.175	.8845315	1.961777
	_Isite_8	1.621655	.4267537	1.84	0.066	.9681819	2.71619
	_Isite_9	.5328145	.2004227	-1.67	0.094	.2549114	1.113686
	_Isite_10	1.638777	.2916187	2.78	0.006	1.156244	2.322685
	_Isite_11	1.494767	.2736156	2.20	0.028	1.044147	2.139858
	_Isite_12	.8540288	.16889	-0.80	0.425	.5796177	1.258356
	_Isite_13	1.019967	.3955312	0.05	0.959	.4769798	2.181082

```
-----
```

```
. xi: logistic comp_afib uric_acid_1sd age female NonHispanic_Black i.site
i.site          _Isite_1-13          (naturally coded; _Isite_1 omitted)
```

```
Logistic regression                                Number of obs   =      3220
                                                    LR chi2(16)    =      96.76
                                                    Prob > chi2    =      0.0000
Log likelihood = -1486.6586                       Pseudo R2      =      0.0315
```

```
-----
```

	comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
	uric_acid_~d	1.01017	.0495568	0.21	0.837	.9175643	1.112123
	age	1.034995	.0050407	7.06	0.000	1.025163	1.044922
	female	1.123278	.1105946	1.18	0.238	.926148	1.362367
NonHispanic_Black		1.22885	.1241083	2.04	0.041	1.008164	1.497843
	_Isite_2	1.430341	.2822738	1.81	0.070	.9715327	2.105823
	_Isite_3	1.804302	.3514381	3.03	0.002	1.231726	2.643044
	_Isite_4	1.096476	.2602399	0.39	0.698	.68861	1.745923
	_Isite_5	.8367105	.2275698	-0.66	0.512	.4909803	1.425891
	_Isite_6	1.352516	.3210955	1.27	0.203	.8493018	2.153887
	_Isite_7	1.288069	.2628859	1.24	0.215	.8634058	1.921602
	_Isite_8	1.510297	.4029845	1.55	0.122	.8952424	2.547909
	_Isite_9	.5315409	.1999138	-1.68	0.093	.2543301	1.110901

```
-----
```

_Isite_10		1.619768	.2908665	2.69	0.007	1.139201	2.303062
_Isite_11		1.40436	.2585297	1.84	0.065	.9789944	2.014544
_Isite_12		.7931898	.1580775	-1.16	0.245	.5367082	1.172239
_Isite_13		1.007596	.390708	0.02	0.984	.4712183	2.15452

```
. xi: logistic comp_afib egfr_lt_45 age_1sd female NonHispanic_Black i.site
i.site      _Isite_1-13      (naturally coded; _Isite_1 omitted)
```

```
Logistic regression      Number of obs    =      3267
                        LR chi2(16)       =      101.96
                        Prob > chi2      =      0.0000
Log likelihood = -1509.9983      Pseudo R2      =      0.0327
```

comp_afib		Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
egfr_lt_45		1.115057	.1103239	1.10	0.271	.9184994 1.353677
age_1sd		1.451301	.0764292	7.07	0.000	1.308974 1.609103
female		1.095009	.1056335	0.94	0.347	.9063667 1.322914
NonHispanic_Black		1.254352	.1236783	2.30	0.022	1.033931 1.521764
_Isite_2		1.425985	.2818584	1.80	0.073	.967983 2.100692
_Isite_3		1.794079	.3482259	3.01	0.003	1.226382 2.624566
_Isite_4		1.116642	.262672	0.47	0.639	.7041784 1.770702
_Isite_5		.845934	.2301422	-0.61	0.539	.4963194 1.441822
_Isite_6		1.317353	.3150713	1.15	0.249	.8243667 2.105155
_Isite_7		1.281398	.260405	1.22	0.222	.8604057 1.90838
_Isite_8		1.512266	.3997227	1.56	0.118	.9008205 2.538739
_Isite_9		.5129198	.1926512	-1.78	0.075	.2456635 1.070923
_Isite_10		1.582878	.2877411	2.53	0.012	1.108442 2.260382
_Isite_11		1.446625	.2637013	2.03	0.043	1.012029 2.067851
_Isite_12		.8415089	.1662042	-0.87	0.382	.5713999 1.239302
_Isite_13		1.014162	.3934343	0.04	0.971	.4741242 2.169316

```
. xi: logistic comp_afib hs_crp_1sd age_1sd female NonHispanic_Black i.site
i.site      _Isite_1-13      (naturally coded; _Isite_1 omitted)
```

```
Logistic regression      Number of obs    =      3257
                        LR chi2(16)       =      101.06
                        Prob > chi2      =      0.0000
Log likelihood = -1505.4387      Pseudo R2      =      0.0325
```

comp_afib		Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
hs_crp_1sd		1.023694	.0447583	0.54	0.592	.9396226 1.115287
age_1sd		1.467248	.0767867	7.33	0.000	1.32421 1.625736
female		1.114767	.1071304	1.13	0.258	.9233849 1.345816
NonHispanic_Black		1.236875	.1225762	2.15	0.032	1.018522 1.502039
_Isite_2		1.44943	.2858741	1.88	0.060	.9847208 2.133445
_Isite_3		1.802981	.3492155	3.04	0.002	1.233457 2.635473
_Isite_4		1.136898	.2670519	0.55	0.585	.7174279 1.801626
_Isite_5		.8502197	.2311113	-0.60	0.551	.4990603 1.448469
_Isite_6		1.366896	.3235724	1.32	0.187	.8594857 2.173864
_Isite_7		1.299846	.2638332	1.29	0.196	.8732153 1.934918
_Isite_8		1.589128	.4167449	1.77	0.077	.9504587 2.656956
_Isite_9		.5197533	.1951633	-1.74	0.081	.2489876 1.084968
_Isite_10		1.660966	.2950903	2.86	0.004	1.172559 2.35281
_Isite_11		1.442002	.263323	2.00	0.045	1.008157 2.062543
_Isite_12		.8331372	.1644992	-0.92	0.355	.5657839 1.226824
_Isite_13		1.028801	.3989126	0.07	0.942	.4811515 2.199787

```
. xi: logistic comp_afib bmi_1sd age_1sd female NonHispanic_Black i.site
i.site      _Isite_1-13      (naturally coded; _Isite_1 omitted)
```



```

Logistic regression
Log likelihood = -1504.0247
Number of obs = 3259
LR chi2(16) = 101.72
Prob > chi2 = 0.0000
Pseudo R2 = 0.0327

```

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
bmi_1sd	1.091891	.0510116	1.88	0.060	.9963507	1.196592
age_1sd	1.469319	.0772732	7.32	0.000	1.325411	1.628853
female	1.078075	.104832	0.77	0.439	.891001	1.304427
NonHispanic	1.223158	.121191	2.03	0.042	1.007268	1.48532
_Isite_2	1.427831	.2818022	1.80	0.071	.9697963	2.102196
_Isite_3	1.767123	.3435517	2.93	0.003	1.207208	2.586732
_Isite_4	1.136778	.2669624	0.55	0.585	.7174281	1.801246
_Isite_5	.8265973	.2252505	-0.70	0.485	.4845502	1.410098
_Isite_6	1.364439	.3233847	1.31	0.190	.8574551	2.171185
_Isite_7	1.292482	.2623595	1.26	0.206	.8682408	1.924018
_Isite_8	1.560534	.4097061	1.70	0.090	.9328175	2.610655
_Isite_9	.5194514	.1951954	-1.74	0.081	.2487065	1.084933
_Isite_10	1.644868	.292087	2.80	0.005	1.161393	2.329608
_Isite_11	1.419767	.2594401	1.92	0.055	.9923694	2.031238
_Isite_12	.8382912	.1656062	-0.89	0.372	.569165	1.234672
_Isite_13	1.03908	.4031949	0.10	0.921	.4856871	2.223009

```

. xi: logistic comp_afib totalmetsum_1sd age female NonHispanic i.site
i.site _Isite_1-13 (naturally coded; _Isite_1 omitted)

```

```

Logistic regression
Log likelihood = -1503.2337
Number of obs = 3259
LR chi2(16) = 109.26
Prob > chi2 = 0.0000
Pseudo R2 = 0.0351

```

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
totalmetsum	.8532635	.0488444	-2.77	0.006	.7627054	.9545738
age	1.031698	.0051467	6.26	0.000	1.02166	1.041835
female	1.099683	.1057653	0.99	0.323	.9107533	1.327805
NonHispanic	1.242839	.1227605	2.20	0.028	1.02409	1.508314
_Isite_2	1.44861	.2858452	1.88	0.060	.9839863	2.132621
_Isite_3	1.806399	.3494768	3.06	0.002	1.236332	2.639321
_Isite_4	1.128468	.2655282	0.51	0.608	.7115436	1.789685
_Isite_5	.8539513	.2323276	-0.58	0.562	.5010183	1.455501
_Isite_6	1.385597	.3284802	1.38	0.169	.8706521	2.205105
_Isite_7	1.297269	.2634582	1.28	0.200	.8712893	1.931514
_Isite_8	1.57173	.4129812	1.72	0.085	.9391172	2.630487
_Isite_9	.4442088	.1751636	-2.06	0.040	.2050857	.9621417
_Isite_10	1.58434	.2822059	2.58	0.010	1.117456	2.246292
_Isite_11	1.424214	.2598076	1.94	0.053	.9960872	2.036352
_Isite_12	.8314955	.1642784	-0.93	0.350	.5645315	1.224706
_Isite_13	1.00407	.3892107	0.01	0.992	.4696886	2.146435

```

. char _dta[omit] "prevalent"
. xi: logistic comp_afib i.smoker age_1sd female NonHispanic i.site
i.smoker _Ismoker_1-3 (naturally coded; _Ismoker_3 omitted)
i.site _Isite_1-13 (naturally coded; _Isite_1 omitted)

```

```

Logistic regression
Log likelihood = -1501.2873
Number of obs = 3267
LR chi2(17) = 119.39
Prob > chi2 = 0.0000
Pseudo R2 = 0.0382

```

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
_Ismoker_1	1.298929	.1889252	1.80	0.072	.9767438 1.727388
_Ismoker_2	1.560346	.1621443	4.28	0.000	1.272823 1.91282
age_1sd	1.402452	.0754163	6.29	0.000	1.262162 1.558336
female	1.163566	.1128112	1.56	0.118	.962198 1.407076
NonHispanic_Black	1.267966	.1262095	2.39	0.017	1.043234 1.541111
_Isite_2	1.471872	.2910264	1.95	0.051	.9990014 2.168574
_Isite_3	1.816536	.3518094	3.08	0.002	1.242772 2.655196
_Isite_4	1.0945	.2578582	0.38	0.702	.6897273 1.736818
_Isite_5	.8404764	.2292979	-0.64	0.524	.4923814 1.434661
_Isite_6	1.372514	.3259011	1.33	0.182	.8617883 2.185915
_Isite_7	1.295846	.2637843	1.27	0.203	.8695239 1.931191
_Isite_8	1.549297	.4071781	1.67	0.096	.9256062 2.593242
_Isite_9	.5164511	.1944434	-1.76	0.079	.2469179 1.080204
_Isite_10	1.60452	.2859869	2.65	0.008	1.131432 2.275422
_Isite_11	1.427292	.260744	1.95	0.051	.9977267 2.041804
_Isite_12	.836803	.1657852	-0.90	0.368	.5675255 1.233846
_Isite_13	1.045827	.406241	0.12	0.908	.488449 2.23924

```

. test _Ismoker_1 _Ismoker_1

( 1) [comp_afib]_Ismoker_1 = 0
( 2) [comp_afib]_Ismoker_1 = 0
     Constraint 2 dropped

           chi2( 1) =      3.23
           Prob > chi2 =    0.0721

```

```
. char _dta[omit]
```

```

. xi: logistic comp_afib alcoh37 age_1sd female NonHispanic_Black i.site
i.site      _Isite_1-13      (naturally coded; _Isite_1 omitted)

```

```

Logistic regression              Number of obs   =      3267
                                LR chi2(16)      =      102.66
                                Prob > chi2      =      0.0000
Log likelihood = -1509.649       Pseudo R2   =      0.0329

```

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
alcoh37	.8695494	.0881231	-1.38	0.168	.7129033 1.060615
age_1sd	1.452456	.0762182	7.11	0.000	1.310497 1.609794
female	1.077618	.1055347	0.76	0.445	.8894143 1.305647
NonHispanic_Black	1.219925	.1223416	1.98	0.047	1.002234 1.484899
_Isite_2	1.4347	.2831024	1.83	0.067	.9745355 2.112149
_Isite_3	1.808837	.3497617	3.07	0.002	1.238251 2.642349
_Isite_4	1.127943	.2648941	0.51	0.608	.711844 1.787266
_Isite_5	.8382536	.2282143	-0.65	0.517	.4916274 1.429272
_Isite_6	1.362696	.3226347	1.31	0.191	.8567746 2.16736
_Isite_7	1.282167	.2603841	1.22	0.221	.8611554 1.909008
_Isite_8	1.536963	.4033814	1.64	0.101	.918888 2.570776
_Isite_9	.5098578	.1915717	-1.79	0.073	.2441307 1.064819
_Isite_10	1.622206	.2886439	2.72	0.007	1.144588 2.299126
_Isite_11	1.440088	.2626167	2.00	0.046	1.007309 2.058807
_Isite_12	.8414823	.1661834	-0.87	0.382	.5714026 1.239218
_Isite_13	1.044525	.4050583	0.11	0.911	.4884605 2.233612

```

. xi: logistic comp_afib hypertension age_1sd female NonHispanic_Black i.site
i.site      _Isite_1-13      (naturally coded; _Isite_1 omitted)

```

```

Logistic regression              Number of obs   =      3265

```

```

Log likelihood = -1509.4798
LR chi2(16) = 102.19
Prob > chi2 = 0.0000
Pseudo R2 = 0.0327

```

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
hypertension	.8314069	.1207443	-1.27	0.204	.6254529 1.105179
age_lsd	1.482213	.0787037	7.41	0.000	1.335712 1.644782
female	1.102191	.1057841	1.01	0.311	.9131916 1.330307
NonHispanic_Black	1.280763	.1287515	2.46	0.014	1.051719 1.559688
_1site_2	1.438297	.2837367	1.84	0.065	.9770793 2.117227
_1site_3	1.853435	.358249	3.19	0.001	1.268964 2.707107
_1site_4	1.143567	.2686357	0.57	0.568	.7216151 1.812248
_1site_5	.8633699	.234773	-0.54	0.589	.5066789 1.471164
_1site_6	1.390177	.3295449	1.39	0.165	.8735559 2.212328
_1site_7	1.298895	.2635945	1.29	0.198	.872636 1.933369
_1site_8	1.602332	.4208025	1.80	0.073	.9576589 2.680983
_1site_9	.5289073	.1987754	-1.69	0.090	.2532087 1.104792
_1site_10	1.679536	.2986411	2.92	0.004	1.185321 2.379814
_1site_11	1.472609	.2682078	2.13	0.034	1.030522 2.104348
_1site_12	.8185045	.1622741	-1.01	0.312	.5549638 1.207195
_1site_13	1.031719	.400106	0.08	0.936	.4824599 2.206286

```

. xi: logistic comp_afib diabetes age_lsd female NonHispanic_Black i.site
i.site      _1site_1-13      (naturally coded; _1site_1 omitted)

```

```

Logistic regression      Number of obs = 3267
LR chi2(16) = 101.90
Prob > chi2 = 0.0000
Pseudo R2 = 0.0326
Log likelihood = -1510.0289

```

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
diabetes	1.105662	.1033906	1.07	0.283	.9205058 1.328062
age_lsd	1.457138	.0763749	7.18	0.000	1.314878 1.614789
female	1.108335	.1063841	1.07	0.284	.9182645 1.337747
NonHispanic_Black	1.236665	.1226713	2.14	0.032	1.018162 1.50206
_1site_2	1.460781	.2883664	1.92	0.055	.9920945 2.150884
_1site_3	1.836045	.3544478	3.15	0.002	1.257649 2.680449
_1site_4	1.137974	.2672944	0.55	0.582	.7181201 1.8033
_1site_5	.8573674	.2331161	-0.57	0.571	.5031846 1.460853
_1site_6	1.370973	.324423	1.33	0.182	.86219 2.179991
_1site_7	1.309962	.265992	1.33	0.184	.8798716 1.950284
_1site_8	1.563765	.4097511	1.71	0.088	.935691 2.613428
_1site_9	.5234498	.1966244	-1.72	0.085	.2506898 1.092983
_1site_10	1.65341	.2933094	2.83	0.005	1.167832 2.340889
_1site_11	1.469295	.2675561	2.11	0.035	1.028268 2.099478
_1site_12	.8530228	.1691606	-0.80	0.423	.5783107 1.25823
_1site_13	1.034243	.4009649	0.09	0.931	.4837497 2.211181

```

. xi: logistic comp_afib chf age_lsd female NonHispanic_Black i.site
i.site      _1site_1-13      (naturally coded; _1site_1 omitted)

```

```

Logistic regression      Number of obs = 3267
LR chi2(16) = 263.97
Prob > chi2 = 0.0000
Pseudo R2 = 0.0846
Log likelihood = -1428.9934

```

comp_afib	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
chf	5.205222	.6587374	13.04	0.000	4.061785 6.670548
age_lsd	1.426201	.0771044	6.57	0.000	1.28281 1.58562

female		1.139206	.113248	1.31	0.190	.9375288	1.384267
NonHispanic_Black		1.141406	.1169872	1.29	0.197	.9336775	1.39535
_Isite_2		1.375094	.2820727	1.55	0.120	.9198683	2.055603
_Isite_3		1.810342	.3629308	2.96	0.003	1.222123	2.681677
_Isite_4		1.118726	.2723643	0.46	0.645	.6942098	1.802838
_Isite_5		.9218902	.2564596	-0.29	0.770	.5344221	1.590282
_Isite_6		1.477828	.3611014	1.60	0.110	.9154532	2.385677
_Isite_7		1.460737	.3045066	1.82	0.069	.9708004	2.197933
_Isite_8		1.630032	.4429685	1.80	0.072	.9569256	2.776604
_Isite_9		.5817208	.222201	-1.42	0.156	.2751563	1.229843
_Isite_10		1.595037	.2934608	2.54	0.011	1.11215	2.287589
_Isite_11		1.238082	.2355659	1.12	0.262	.8526954	1.797648
_Isite_12		.9224304	.1871539	-0.40	0.691	.6197713	1.37289
_Isite_13		1.222727	.4779103	0.51	0.607	.5683712	2.63043

```
. xi: logistic comp_afib anycvd age_lsd female NonHispanic_Black i.site
i.site      _Isite_1-13      (naturally coded; _Isite_1 omitted)
```

```
Logistic regression      Number of obs      =      3267
                        LR chi2(16)      =      234.81
                        Prob > chi2      =      0.0000
Log likelihood = -1443.5731      Pseudo R2      =      0.0752
```

comp_afib		Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
anycvd		3.062218	.2990441	11.46	0.000	2.528781 3.708183
age_lsd		1.292058	.0711131	4.65	0.000	1.159902 1.439271
female		1.212925	.1199747	1.95	0.051	.9991679 1.472412
NonHispanic_Black		1.172726	.1188225	1.57	0.116	.9615051 1.430348
_Isite_2		1.450888	.2929328	1.84	0.065	.9767356 2.155215
_Isite_3		1.705413	.3375537	2.70	0.007	1.157046 2.513671
_Isite_4		1.163742	.2791108	0.63	0.527	.727286 1.862122
_Isite_5		.8543107	.2369682	-0.57	0.570	.4960325 1.471369
_Isite_6		1.271803	.3082381	0.99	0.321	.7908971 2.045124
_Isite_7		1.374535	.2848533	1.54	0.125	.9157056 2.063267
_Isite_8		1.458616	.3916321	1.41	0.160	.8617814 2.468794
_Isite_9		.4970436	.189201	-1.84	0.066	.2357123 1.04811
_Isite_10		1.457687	.2655304	2.07	0.039	1.020025 2.083138
_Isite_11		1.247509	.2326051	1.19	0.236	.8656303 1.797857
_Isite_12		.9305798	.1874698	-0.36	0.721	.6270107 1.381123
_Isite_13		1.115705	.440544	0.28	0.782	.5145724 2.419092

```
. * summarize age_lsd tc uric_acid hs_crp bmi totalmetsum totalmetsum_lsd
. char _dta[omit] "prevalent"
```

```
. xi: logistic comp_afib age_lsd female NonHispanic_Black i.edu tc_lsd bmi_lsd
totalmetsum_lsd i.smoker chf anycvd i.site
i.edu      _Iedu_1-5      (naturally coded; _Iedu_3 omitted)
i.smoker   _Ismoker_1-3  (naturally coded; _Ismoker_3 omitted)
i.site     _Isite_1-13   (naturally coded; _Isite_1 omitted)
```

```
Logistic regression      Number of obs      =      3242
                        LR chi2(26)      =      321.48
                        Prob > chi2      =      0.0000
Log likelihood = -1386.2182      Pseudo R2      =      0.1039
```

comp_afib		Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
age_lsd		1.27172	.076162	4.01	0.000	1.130874 1.430109
female		1.256	.1321626	2.17	0.030	1.021932 1.54368
NonHispanic_Black		1.072961	.119803	0.63	0.528	.8620686 1.335445

_Iedu_1		1.141451	.1655623	0.91	0.362	.8590028	1.516772
_Iedu_2		.7624111	.1079263	-1.92	0.055	.5776887	1.006201
_Iedu_4		1.025451	.1511151	0.17	0.865	.7681541	1.368932
_Iedu_5		.8377777	.1427331	-1.04	0.299	.5999412	1.1699
tc_1sd		.9250014	.0469154	-1.54	0.124	.8374715	1.02168
bmi_1sd		1.035837	.0517346	0.70	0.481	.9392439	1.142364
totalmetsu~d		.9270184	.0542663	-1.29	0.195	.8265331	1.03972
_Ismoker_1		1.145412	.1791194	0.87	0.385	.8430453	1.556225
_Ismoker_2		1.336202	.1461347	2.65	0.008	1.078399	1.655634
chf		3.277978	.4755458	8.18	0.000	2.466719	4.356044
anycvd		1.943172	.2201694	5.86	0.000	1.556204	2.426364
_Isite_2		1.405971	.2922553	1.64	0.101	.9354908	2.113067
_Isite_3		1.65442	.3411864	2.44	0.015	1.104345	2.478487
_Isite_4		1.091799	.2708112	0.35	0.723	.6714433	1.775316
_Isite_5		.8984611	.2551446	-0.38	0.706	.5149622	1.567556
_Isite_6		1.425796	.352948	1.43	0.152	.8777021	2.316155
_Isite_7		1.497454	.3169505	1.91	0.056	.9889824	2.26735
_Isite_8		1.642085	.4529227	1.80	0.072	.9563476	2.819522
_Isite_9		.4903336	.1989481	-1.76	0.079	.2213731	1.086071
_Isite_10		1.391995	.2642536	1.74	0.081	.9595055	2.019427
_Isite_11		1.120274	.2190701	0.58	0.561	.7636098	1.643528
_Isite_12		.9574109	.1975486	-0.21	0.833	.6389467	1.434604
_Isite_13		1.204745	.4761351	0.47	0.637	.5552469	2.61399

```

. xi: logistic comp_afib age_1sd  male NonHispanic_Black i.edu tc_1sd bmi_1sd
totalmetsum_1sd i.smoker chf anycvd i.site
i.edu                _Iedu_1-5          (naturally coded; _Iedu_3 omitted)
i.smoker             _Ismoker_1-3       (naturally coded; _Ismoker_3 omitted)
i.site              _Isite_1-13       (naturally coded; _Isite_1 omitted)

```

```

Logistic regression                               Number of obs   =       3242
                                                  LR chi2(26)    =       321.48
                                                  Prob > chi2    =       0.0000
Log likelihood = -1386.2182                    Pseudo R2      =       0.1039

```

comp_afib		Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
age_1sd		1.27172	.076162	4.01	0.000	1.130874 1.430109
male		.7961784	.0837779	-2.17	0.030	.6478027 .9785388
NonHispanic_Black		1.072961	.119803	0.63	0.528	.8620686 1.335445
_Iedu_1		1.141451	.1655623	0.91	0.362	.8590028 1.516772
_Iedu_2		.7624111	.1079263	-1.92	0.055	.5776887 1.006201
_Iedu_4		1.025451	.1511151	0.17	0.865	.7681541 1.368932
_Iedu_5		.8377777	.1427331	-1.04	0.299	.5999412 1.1699
tc_1sd		.9250014	.0469154	-1.54	0.124	.8374715 1.02168
bmi_1sd		1.035837	.0517346	0.70	0.481	.9392439 1.142364
totalmetsu~d		.9270184	.0542663	-1.29	0.195	.8265331 1.03972
_Ismoker_1		1.145412	.1791194	0.87	0.385	.8430453 1.556225
_Ismoker_2		1.336202	.1461347	2.65	0.008	1.078399 1.655634
chf		3.277978	.4755458	8.18	0.000	2.466719 4.356044
anycvd		1.943172	.2201694	5.86	0.000	1.556204 2.426364
_Isite_2		1.405971	.2922553	1.64	0.101	.9354908 2.113067
_Isite_3		1.65442	.3411864	2.44	0.015	1.104345 2.478487
_Isite_4		1.091799	.2708112	0.35	0.723	.6714433 1.775316
_Isite_5		.8984611	.2551446	-0.38	0.706	.5149622 1.567556
_Isite_6		1.425796	.352948	1.43	0.152	.8777021 2.316155
_Isite_7		1.497454	.3169505	1.91	0.056	.9889824 2.26735
_Isite_8		1.642085	.4529227	1.80	0.072	.9563476 2.819522
_Isite_9		.4903336	.1989481	-1.76	0.079	.2213731 1.086071
_Isite_10		1.391995	.2642536	1.74	0.081	.9595055 2.019427
_Isite_11		1.120274	.2190701	0.58	0.561	.7636098 1.643528
_Isite_12		.9574109	.1975486	-0.21	0.833	.6389467 1.434604
_Isite_13		1.204745	.4761351	0.47	0.637	.5552469 2.61399

```
. char _dta[omit]

. test _Iedu_1 _Iedu_2 _Iedu_4 _Iedu_5

( 1) [comp_afib]_Iedu_1 = 0
( 2) [comp_afib]_Iedu_2 = 0
( 3) [comp_afib]_Iedu_4 = 0
( 4) [comp_afib]_Iedu_5 = 0

      chi2( 4) =      7.97
      Prob > chi2 =    0.0925

. test _Ismoker_1 _Ismoker_2

( 1) [comp_afib]_Ismoker_1 = 0
( 2) [comp_afib]_Ismoker_2 = 0

      chi2( 2) =      7.06
      Prob > chi2 =    0.0293

.
. log close
  name: <unnamed>
  log:  Z:\CRIC\AnalysisData\MS037\Table_3b_4_Output_rev1.log
  log type: text
  closed on: 21 Apr 2011, 15:26:41
```
