

### Example of Sampling Strategy for ICD9 480-486

Site 1 has identified 5006 non-matched records for persons aged  $\geq 18$  years, residents of Site 1's catchment area, admitted for hospitalization October 1, 2006 – April 30, 2007, and with at least one discharge ICD9 code between 480-486.

- Stratification of the 5006 persons identifies 7 hospitals.

Hospital 1: 8 persons with ICD9 480-486

Hospital 2: 5 persons “ “

Hospital 3: 3 persons “ “

Hospital 4: 30 persons “ “

Hospital 5: 1000 persons “ “

Hospital 6: 3500 persons “ “

Hospital 7: 460 persons “ “

- The 16 person's medical charts from the first 3 hospitals would all be reviewed. (See first 3 lines of grid)

- Sampling grid

Hospital#	Total nonmatched records	Actual persons 18-49 yrs	Sample used 0.1	Actual persons 50-64 yrs	Sample used 0.08	Actual persons $\geq 65$ yrs	Sample used 0.05	No. charts to review
1	8	2	2*	1	1*	5	5*	8
2	5	0	0	1	1*	4	4*	5
3	3	0	0	0	0	3	3*	3
4	30	3	3	5	5	22	2*	10
5	1000	130	13	200	16	670	33.5	62.5
6	3500	450	45	700	56	2350	117.5	218.5
7	460	45	4.5	80	6.4	236	11.8	22.7

5006

330

\*sampling not used so that total number of charts reviewed at any single hospital should be no less than 10

In this example the total number of charts exceeds 300. Site 1 could reduce the number of charts sampled from person aged  $\geq 65$  years since Hospitals 1-3 contributed a disproportionate number of persons from this age group.

Note: In this example the proportion of persons 18-49 and 50-64 years that were sampled is less than described in the protocol. Sites may have to adjust the proportion by age group sampled, but “oversampling” of the younger age groups is recommended.