SEER February 2011 Data Submission Requirements and Guidelines
UPDATED: 2/9/2011

Please direct all submission questions to:
SEERSUBM-L@list.nih.gov

1. The data submission is due **Monday, February 28, 2011**. The data submission should include all resident cases through 12/31/2009 diagnoses.

2. **All submission files should be uploaded to the SEER Submission Reports Portal.** A February 2011 folder has been created under each registry's folder. Please add the submission files to this folder.

3. The file should be submitted in NAACCR 12 format and should be named **rr.feb11.txd.gz** where rr represents the two character abbreviation for your registry. A list of the variables to be included in the submission can be downloaded from http://seer.cancer.gov/tools/seer11_Updated 2-3-2011.dataitems.pdf .

   Please Note: All data items included on the list are required for February 2011 submission. Those data items which will be edited are highlighted in yellow and marked with a *.

4. The file should be sorted by the following variables: Registry ID (NAACCR Item #40), Patient ID Number (NAACCR Item #20), Sequence Number Central (NAACCR Item #380). For the Los Angeles, Greater Bay Area and Louisiana registries, the legacy ID should be put in columns 50-57 and the current patient ID number (Eureka ID or SEER*DMS ID) should be put in columns 42-49.

5. The file should be edited prior to submission using the version of the SEER*Edits software designated for this submission. Use the February 2011 SEER Submission edit set. The SEER*Edits software is available on the SEER Submission Reports Portal.

6. **NHAPIIA:** Run NHAPIIA which updates both the API indicator (NAPIIA, NAACCR Item #193) and the Hispanic indicator (NHIA, NAACCR Item #191). Use the NHAPIIA SAS program found on the NAACCR Call for Data website. Every registry must run the algorithm. Please note to set the parameters needed for your registry before running NHAPIIA.

7. **Collaborative Staging:** Please run your 2004-2009 data through the CS algorithm before submission. It needs to be CS version 020200 or higher in CS Version Derived [NAACCR item # 2936].

8. **VA case counts:** To address the backlog of Veterans Affairs Hospital (VA) records, specific data are being requested. The VA tumor case counts are needed for all malignant cancers combined (males only), by race, age, and cancer site. Please restrict case counts to SEER coverage areas. These data are needed for diagnosis years 2000-2009 for the current submission file only.

   For the February submission file, we request the following steps:

   a. Create a temporary NAACCR-format data file containing a VA indicator flag in NAACCR item 2220, Column 2468. The VA indicator may be created using facility codes to distinguish three categories of tumor records: 0) records from non-VA facilities; 1) records from a VA facility only; and 2) records from VA facility and other facilities.

b. Use the temporary NAACCR-format data file as input to the SEER Edits program and generate a VA report. Check the option to create a copy of the file without the VA flag. This copy is the data file that you will submit to SEER. Use the file naming conventions defined in step 3 above.

c. On the SEER*Edits Results tab, export the VA Report to CSV. SEER*Edits will create a gzipped CSV file. Post this file to the SEER Submissions Portal.

9. In-House Record Counts: Use the Word version of Attachment B (http://seer.cancer.gov/tools/Feb2011Submission_AttachmentB.doc) to provide the following information to NCI SEER:

a. Please submit count of 2009 records that are in-house but not yet processed (e.g., abstracts, path reports, death certificates, etc, for reportable cancers). If possible, please provide counts broken down by record types.

b. If possible, provide estimate of likely number of records that will result in 2009 cases. These are records that have not been processed but indicate a new case. The overall count is needed.

Attachment A
SEER Patient Follow-up Calculation
for February 2011 submission only

Note 1: The calendar years of diagnosis used for the follow-up calculation with this submission is as follows.

2009 is the last year of data being submitted. The percent of patients diagnosed during the years prior to 2009 who have current follow-up is defined as

$$P = 100(D + A)/T$$

where D is the number who died prior to January 1, 2010, A is the number with follow-up dates on or after January 1, 2010 (includes both alive and dead patients), and T is equal to A + D + the number of patients who were last known to be alive with follow-up dates prior to January 1, 2010. P can be calculated for individual years of diagnosis up through 2008 and for all years combined prior to 2009

For all invasive cancers and calendar years specified by NCI, P shall be at least 95 percent, but must not be below 90 percent.

For patients ages 20-64, and under age 20, P shall be at least 90 percent in each case, but must not be below 80 percent.

For all in situ cancers (excluding cervix in situ) and calendar years specified by NCI, P shall be at least 90 percent, but must not be below 80 percent.

Note 2: Follow-up is expected on benign and borderline central nervous system tumor cases diagnosed 1/1/2004 and forward.

Attachment B
February 2011 Submission to NCI
Please complete for each submission and e-mail to:
SEERSUBM-L@list.nih.gov

Registry Name:
Registry ID number:
Date of file transfer to SEER:
Data file name:
Years of diagnosis submitted:
Number of cases submitted:
Number of 2009 records in-house but not submitted:
Estimate of likely 2009 cases:

When running NHAPIIA, were any optional exclusions applied? YES or NO

If YES:

1: In counties in which less than 5% of the population is of Hispanic ethnicity, was the surname portion of the algorithm only applied to cases coded as "Spanish surname only" or "Unknown whether Spanish or not" (item 190 – codes 7 or 9)? YES or NO

2: In counties in which less than 5% of the population is of Hispanic ethnicity, was the surname portion of the algorithm only applied to cases coded as "Spanish surname only" (item 190 - code 7)? Were all cases coded to "Unknown whether Spanish or not" (item 190 – code 9) converted to non-Hispanic in NHIA? YES or NO

3. Other; please specify

Attachment B (continued)
February 2011 Submission to NCI
Please complete for each submission and e-mail to:
SEERSUBM-L@list.nih.gov

Technical Contact
Name:
Telephone number:
e-mail address:

Registry Manager Contact
Name:
Telephone number:
e-mail address:

Known data problems including reasons:

Comments: