



**Data Release Q and A  
2007 Program Directors' Meeting**

Missy Jamison, MPH



**NPCR Data Release Products**

Data Release	Availability
United States Cancer Statistics	2002
County Cancer Incidence Dataset	2002
On-line Dataset in WONDER	2005
Restricted Access Files	In limited release



**NPCR Approach to Data Release**

- The Small Data Release Group
  - Provides guidance to NPCR on data release
- For Restricted Access Files
  - Worked on the application, data release form
  - Will serve as reviewers for the applications
  - Review measures for confidentiality
  - Consider all aspects of the release
- Need to recruit a few new members



**NPCR Restricted Access Files**

- Allows access to individual record level data
  - One record per cancer diagnosis
- Contains only NPCR cancer incidence data
  - Created to be consistent with the SEER data
  - NO treatment data or county level data
  - State is the only geographic identifier
- 1999-2004 data will be available beginning this Fall
  - With an approved proposal
  - With signed data release agreement



**Population Coverage for the Restricted Access Files**

Area	% Coverage	Does not include
National	92%	6 NPCR programs 5 SEER programs
Northeast	90%	2 NPCR programs 1 SEER program
South	94%	2 NPCR programs 2 NPCR programs
Midwest	88%	1 SEER program
West	92%	3 SEER programs



**Changes Since Last Year**

- Addition of a new file
  - A Regional Restricted Access File (RRAF)
  - Only contains the census regions (Northeast, South, Midwest, West)
  - Otherwise identical (almost) to the State Level Restricted Access File
  - Purpose is to create a file with reduced potential for identification of individuals
  - Allows data release with reduced oversight



## Process for Regional Applications

- Regional file application will require
  - Principal Investigator and co-investigator identification and resumes
  - Signed data release agreements
  - A short description of the project and methods to be used
  - Review by CDC staff and member of the Small Data Release Group
  - Only present to whole group if an issue



## Process for State Applications

- State file application will require
  - Principal Investigator and co-investigator identification and resumes
  - Signed data release agreements
  - A more detailed description of the project
    - A description of the problem statement
    - Hypotheses to be tested
    - How the data will be subset



## Process for State Applications

- Project description continues
  - Variables to be used in the analysis
  - Statistical tests to be performed
  - Outcome measurements
  - Report if cell sizes of less than 16 anticipated
- Review by CDC staff and member of the Small Data Release Group
- Present to whole group
  - Discuss and resolve issues
  - A much more thorough review



## Timeline

- The current timeline for the Restricted files is:
  - November, 2007
    - Release of Regional File
  - June, 2008
    - Release State Level File
    - Estimated date to incorporate lessons learned from release of Regional File



## Issues

- Ran NAACCR Record Uniqueness Program on the Regional File
  - Will collapse age in the 85+ group and find a solution for date of birth in the regional file
- Review process has to be finalized for state file
  - How to vote/rank applications
  - Are there any groups/projects we want to exclude from release
    - OR release mostly on merit of application



?? Questions ??

mjamison@cdc.gov