Effects of Changes in U.S. Registry Operations and Standards on the Quality of Cancer Registry Data

Jerri Linn Phillips, MA, CTR
National Cancer Data Base
Commission on Cancer
American College of Surgeons

CoC Registry History

- CoC: registration standards in hospitals since 1955
  - Specified list of registry items to be collected
- Data Acquisition Manual (DAM): 1988
  - Specific definitions of codes for the items
- NCDB: collecting data since 1989
  - Diagnoses back to 1985
- Cancer Program Standards: 1996
  - Required NCDB participation for Approval
- Cancer Program Standards: 2004
  - Required correction and resubmission of cases that failed NCDB submission edits for cases diagnosed in 2002 or subsequently

What has changed since 1988?

- Computerization of registries
  - Standardization of data item codes
  - Standardization of edits
- Medicine has continued to develop
  - What was once important is no longer
  - What was once not known is now important
- Expansion of registration
  - Data exchange
  - Data pooling
  - Broader boundaries for comparability
  - Expanded data uses

Registry Items Already In the Supplement on the Tumor Registry (1981)

- **Facility ID** (5 digits)*
- **Accession Number**
- Sequence Number (1 digit)*
- Medical Record Number
- **Last Name**
- **First Name** & **Initial** (single entry)*
- Number and Street (10 digits)*
- City/town (15 digits)*
- State
- ZIP Code (9 digit)
- **Date of Birth** (6 digit)*
- Age at Diagnosis (2 digit)*
- **Race** (1 digit)*
- **Sex**
- **Accession Year**
- Referring Physician
- Attending Physician (2 of them)
- **Class of Case***
- Referring Hospital(3 digits)*
- **Date of Diagnosis** (4 digit)*
- **Primary Site**
- **Paired Organ Involvement**
- **Histology** (including grade)*
- **Diagnostic Confirmation***
- **Size of Tumor** (whole cm's)*
- **Regional Nodes Positive**
- **Regional Nodes Examined**
- **Site(s) of Distant Mets at DX**
- **General Summary Stage**
- **AJCC Staging Basis** (c vs. p)
- **AJCC-T, AJCC-N, AJCC-N, AJCC Stage**
- **Date of First Course TX** (4 digit)*
- **Surgery** (1 digit)*
- **Radiation***
- **Chemotherapy** (included hormone)*
- **Other Cancer-Directed Treatment** (included BRM)*
- Residual Tumor
- **Date of First Recurrence** (4 digits)*
- **Type of First Recurrence**
- **Distant Sites of First Recurrence**
- [Subsequent Treatments] – **2nd**
- Date of Last Contact/Death (4 digits)*
- Status (patient, cancer)*
- Quality of Survival

* = Changed in 1989 DAM; Boldface = Required in 1994 DAM

Added in 1989 Data Acquisition Manual

- Social Security Number
- Middle Initial (separate)
- Maiden Name and Alias
- County at DX
- Census Tract
- Telephone Number
- Place of Birth
- **Ethnicity**
- Marital Status at DX
- Primary Physician
- Other Physician (4 of them)
- **Grade**
- EOD Tumor Extension
- EOD Lymph Nodes
- Reason for No Cancer-Directed Surgery
- Radiation to Brain, CNS
- **Hormone, endocrine treatment**
- **BRM**
- Follow-Up Method
- Next Follow-Up Method
- Unusual Follow-Up Method
- **Vital Status**
- **Cancer Status**
- Underlying Cause of Death
- ICD (death) Code Version

Boldface = Required in 1994 DAM

ROADS 1996 and 1998

- Little change between item definitions in last DAM and first ROADS
- Additional items became required
- New items generally were optional in ROADS
- Former Surgery item split into dimensions
  - Fully expanded in ROADS 1998
  - Scope and Surgery of Other Reg/Distant sites converted from old Surgery item
- New Radiation items defined (optional)

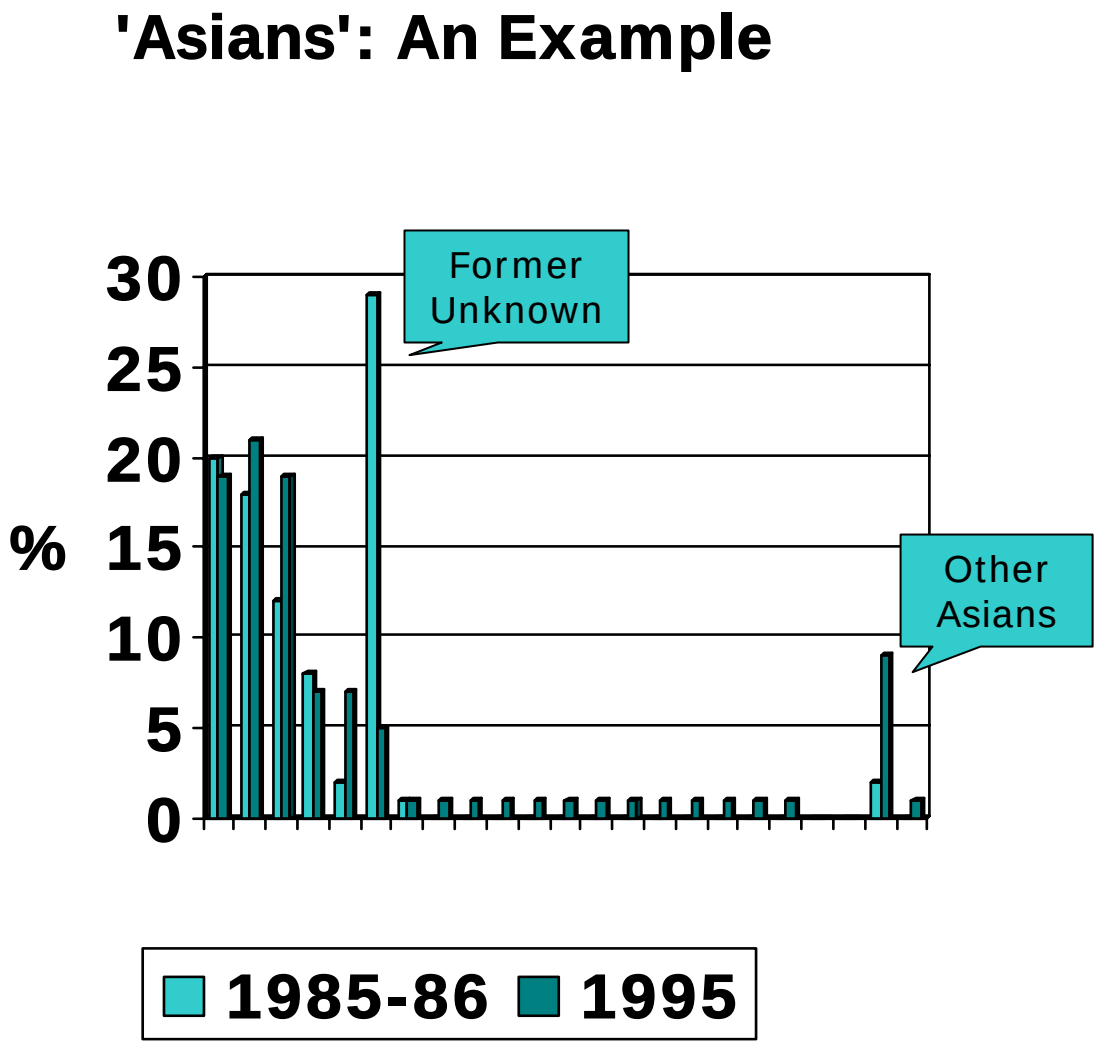
FORDS 2003 and 2004

- Non-required items no longer published; can use NAACCR Volume II for reference or instructions from states or other non-CoC standard setters
- Former “non-cancer-directed surgery” item further split into Palliative and Surgical Diagnostic and Staging Procedure
- Most ROADS radiation items required, boost radiation items added
- Collaborative Staging added in 2004; CoC also requires physician AJCC staging
- Secondary diagnosis items added
- Multiple primary, NPI items will be added in 2007

Effects of Changed Items

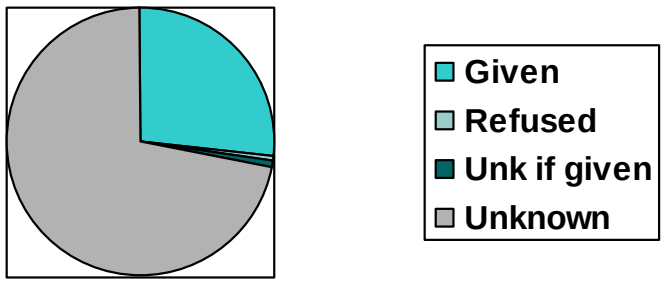
- When new responses are introduced
  - The distribution of all responses may change
  - The meaning of “other” may be altered.
- Examples:
  - Race (Asians)
  - Reason for No … (Radiation)

Effects of Changed Items
Example: Race (Race 1)

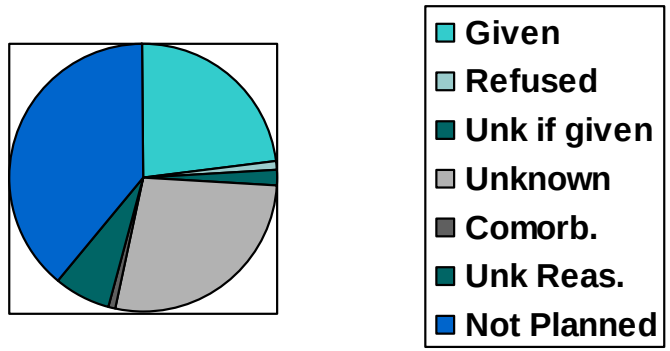


Effects of Changed Items
Example: Reason for No Radiation

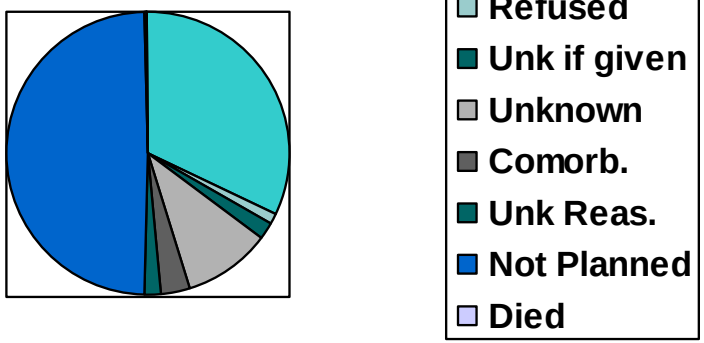
DAM: 1993



ROADS: 1998



FORDS: 2003

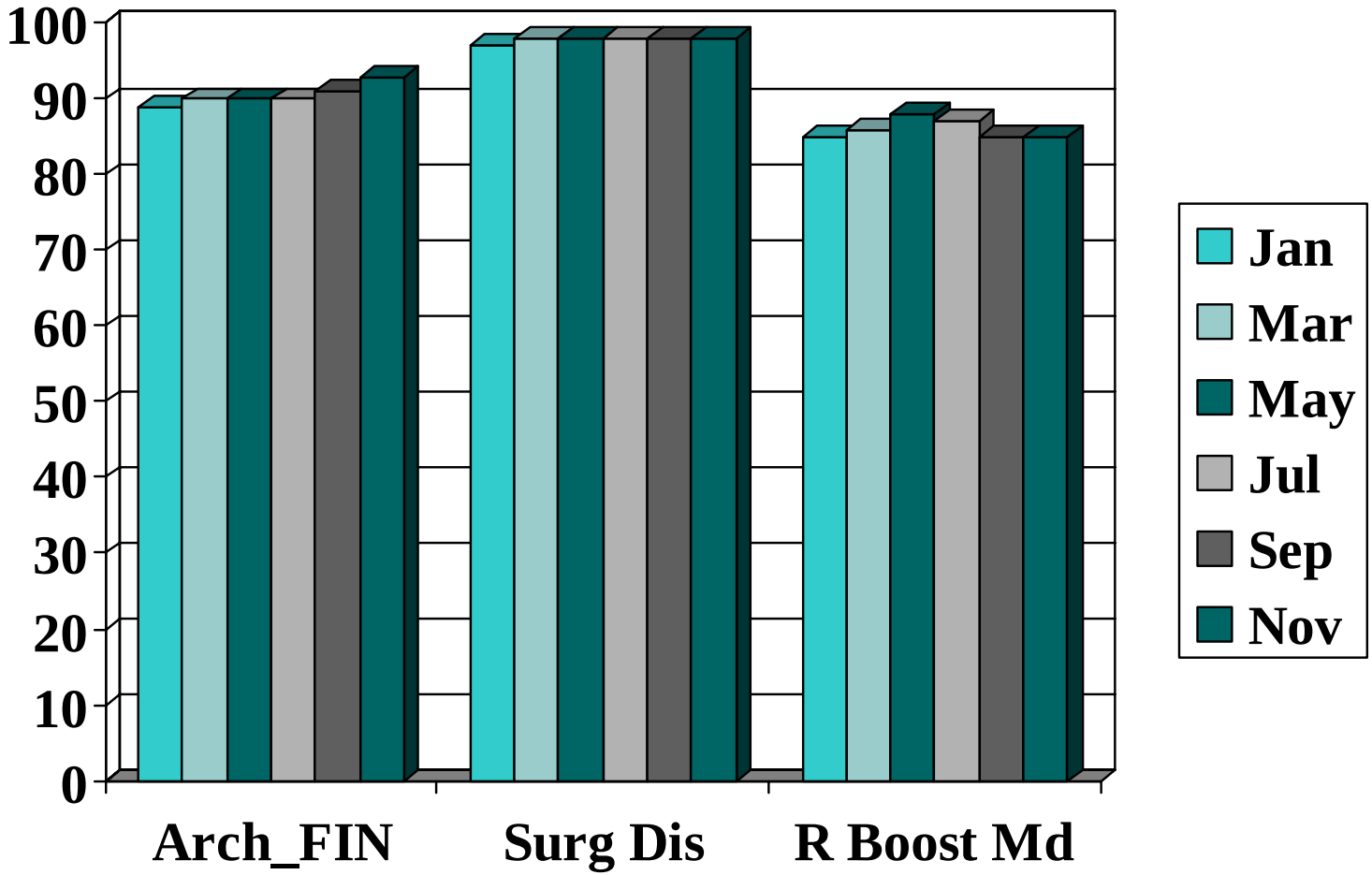


Effects of New Items

- When new items are added that can not be converted from earlier items
  - The content of that item can not be retrieved for older data except through a reabstracting study
  - There may be a delay as software providers and registrars become accustomed to the new item

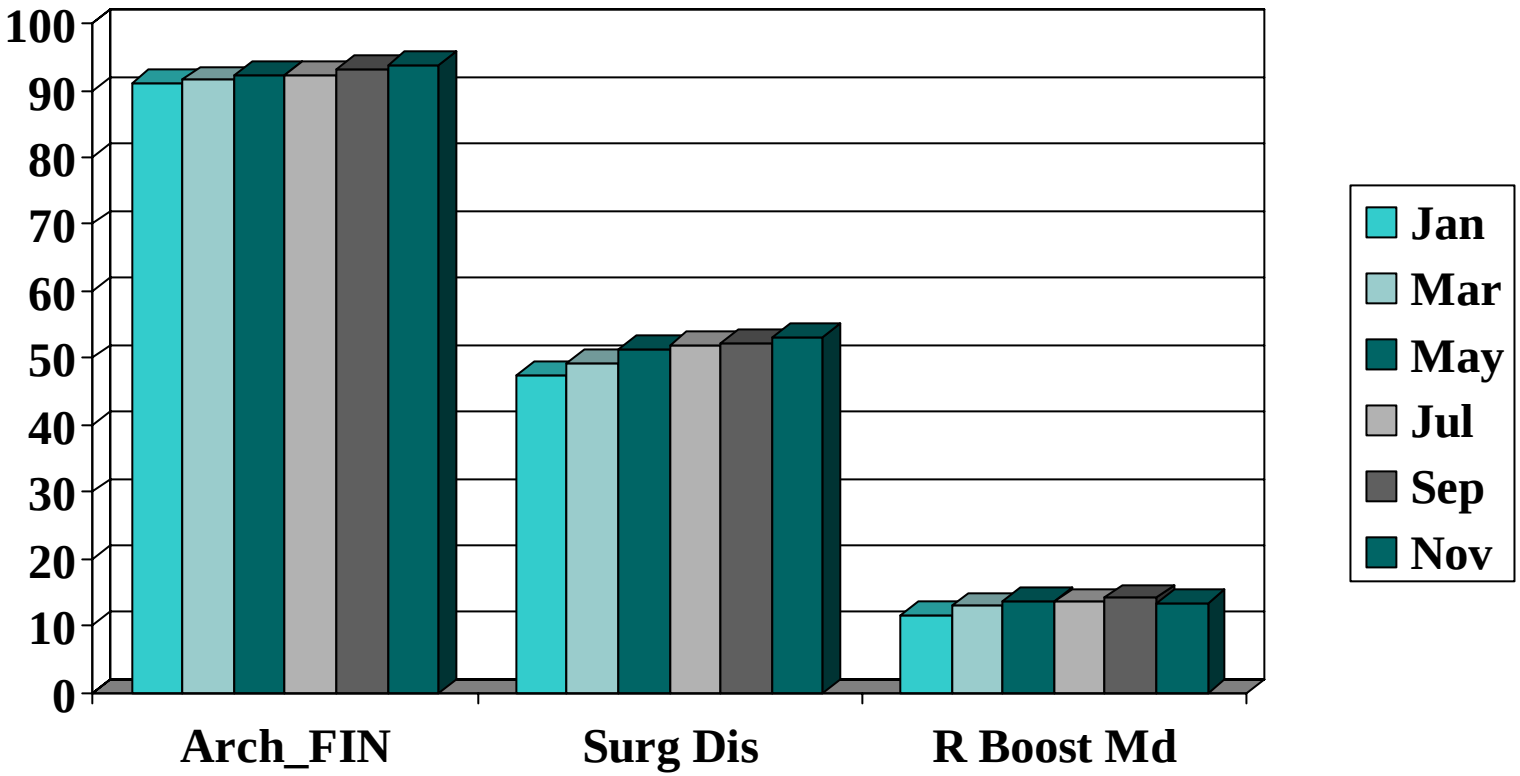
Effects of New Items

Example: % **Facilities** with Known, Valid, Non-0 Content by Month 1st Contact: 2003 Diagnoses



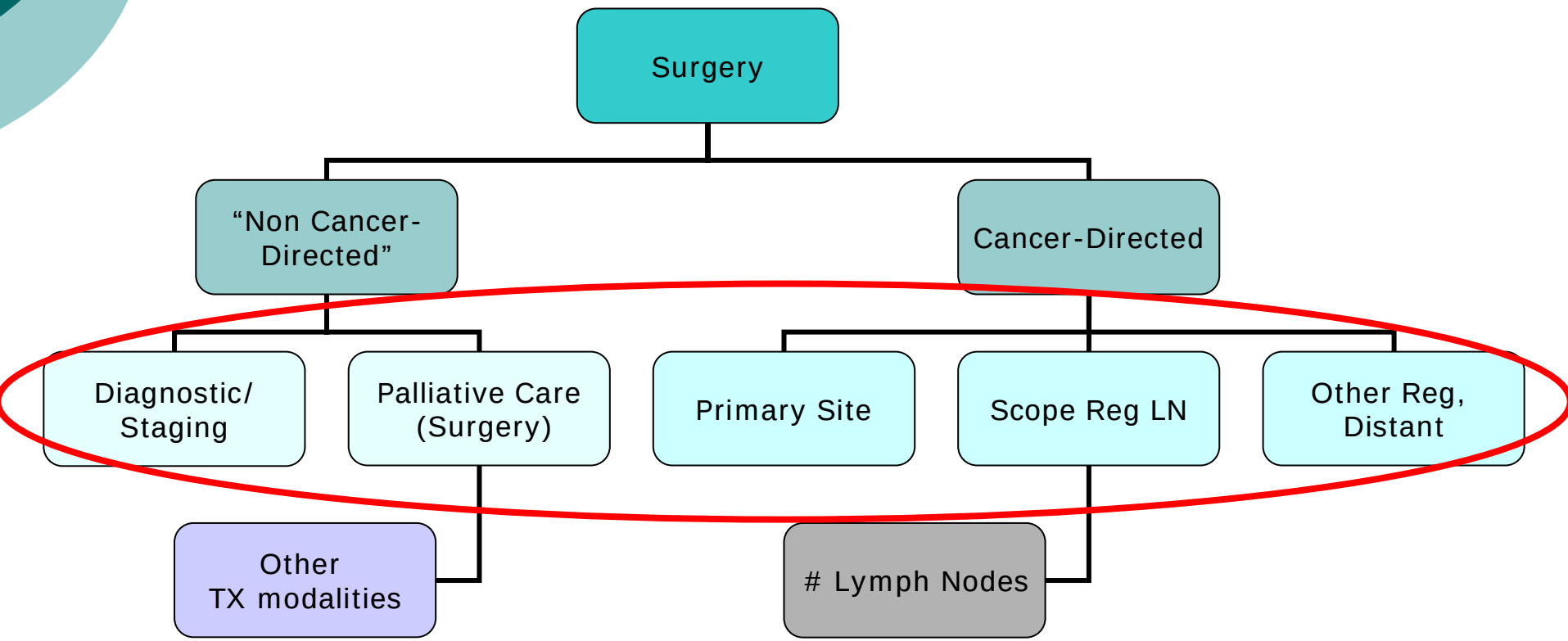
Effects of New Items

Example: % **Reports** with Known, Valid, Non-0 Content by Month 1st Contact



Effects of New Items
Splitting Former Items (Surgery)

- When items are split, lower-priority responses in the earlier item can not be recovered.



Effects of New Items

- When new items are generated by conversion from older items:
  - The amount of information that can be retained by conversion depends on the response sets for the old and new items.
    - It is likely the converted version will not have the complete range of codes represented in the new item
    - Sometimes information is lost in conversion, but not necessarily in comparison to the range collected by the new item

Edit Quality Over Time

Selected 17 Demographic, Diagnostic, Treatment and Outcome Edits That Apply to All Years

- Addr at DX—City (COC)
- Addr at DX—Stage (COC)
- Cancer Status (COC)
- EOD—Tumor Size (COC)
- Grade (COC)
- Laterality (COC)
- Laterality, Primary Site (COC)
- Race 1 (SEER RACE)
- Reason for No Radiation (COC)
- Reason for No Surgery (SEER NCDSURG)
- Regional Nodes Ex, Reg Nodes Pos (COC)
- Regional Nodes Examined (SEER)
- Regional Nodes Positive (SEER)
- Sex, Primary Site (SEER IF17)
- Spanish/Hispanic Origin (SEER SPANORIG)
- Vital Status (COC)

Does not include edits for items that NCDB forces to be valid at input.

Edit Quality Over Time

Based on 17 Selected Edits

Diagnosis Year	Submission Year	# Edit Errors/ 1000 Records
2003	2004-2005	3.4
2002	2003-2004	2.4
1998	2004-2005	4.7
1998	1999-2000	41.2
1993	2004-2005	13.6
1993	1999-2000	63.5

Item Completeness Over Time: Selected Items for Illustration

Item	Percent Unknown by Year of Diagnosis			
	2003	**1998**	**1993**	**1988**
State at Diagnosis	0.2	0.2	0.4	1.4
Birth Place	67.8	66.0	61.3	61.4
Diagnosis **Day**	2.1	2.1	3.6	8.5
Primary Site	2.0	1.9	2.0	2.3
Laterality	1.5	1.3	1.8	2.7
Histology (ICD-O-2, ICD-O-3)	1.1	0.7	0.3	1.3
Grade	34.1	32.2	39.9	39.1
First Contact **Day**	0.1	1.3	1.9	7.2
Tumor Size	47.7	46.5	50.2	52.7
TNM cStage	48.6	50.2	54.7	75.1
TNM pStage	51.3	51.2	52.3	62.8
Reason for No Surgery	3.8	3.9	5.4	7.1
Chemotherapy -- Summary	9.9	5.6	3.4	2.2
Transplant/Endocrine Procedure	8.8	0.4	0.5	0.3

In Summary

- The number of data items collected has increased over the two decades reviewed
  - For core items, the completeness of coverage is fairly consistent
  - Adoption of recently added items was generally timely
- The data reflect CoC Standards
  - AJCC staging coverage improved after its use was required for all cases … but still reflects the old rule that *either* clinical or pathologic stage must be recorded.
  - Edit quality improved for all diagnosis years after correction and resubmission of most recent years' edits became required; but early diagnosis years still have more edit errors within each submission period.
- Item completeness is generally better than in 1988, but more recently has improved little
  - For many items, it has always been good
  - There appears to be a tendency to use "unknown" for some items where the answer may be "none"