Summary Stage 2000:
Directly Coded vs. Derived

NAACCR Data Use and Research Committee –
Summary Stage Work Group

NAACCR Annual Meeting, Quebec, QC, Canada, 2010

Summary Stage Work Group members (alphabetic)

- Umed Ajani, CDC-NPCR
- Patty Andrews, Louisiana
- Baozhen Qiao, New York
- Praveen Ranganath, Louisiana
- Brad Wohler, Florida
- Qingzhao Yu, Louisiana
- Xiao-Cheng Wu, Louisiana (Chair)

Acknowledgement: Zhenzhen Zhang

Background

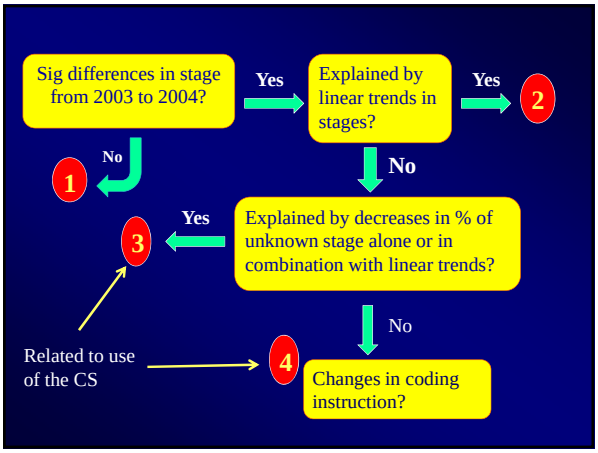
- Historically, there were two major coding systems for staging cancer cases: AJCC TNM and SEER Summary Stage
- To eliminate the duplicate efforts in staging cancer cases, the CS system was developed. With the CS codes, both AJCC TNM and SEER SS can be derived automatically.

Background

- It is unknown whether directly coded summary stage 2000 (SS2000) and CS derived SS2000 (CSDSS2000) produce comparable stage results with "real world" data.
- The NAACCR Data Use and Research Committee formed the Summary Stage Work Group to assess comparability of the two coding systems.

Methods

- CINA Deluxe 1995-2005 dataset (Dec 2007 submission)
- Data from 40 cancer registries that met the standards for high quality incidence data
- Cancer sites defined according to the SEER site recodes. Only invasive cases included
- Likelihood ratio test used



1

4 cancer sites without significant differences in stage distribution between 2003 and 2004

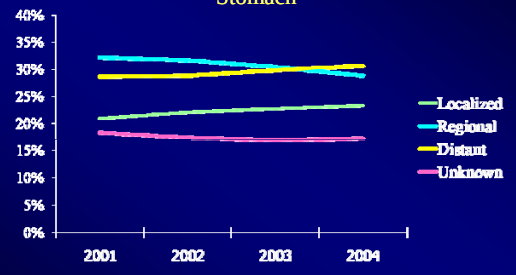
- bone and joints
- intrahepatic bile duct
- NHL
- penis

2

4 cancer sites with significant differences in stage distribution between 2003 and 2004 cases, attributable to 2001-2004 linear trends in stage distribution

- other non-epithelial skin
- stomach
- pancreas
- thyroid

Stage Distributions by Year of Diagnosis
Stomach



Year of Diagnosis

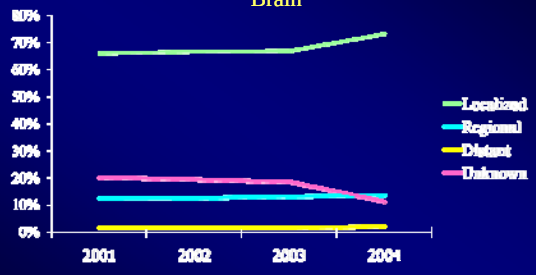
Note: Stages for cases in 2001-2003 were coded according to Seer Summary Staging Manual 2000 (SS2000); stages for cases in 2004 were coded according to collaborative staging manual coding instructions derived SS2000.

3

8 cancer sites with significant differences in stage distribution between 2003 and 2004, attributable to decreases in % of unknown stage from 2003 to 2004 alone or in combination with 2001-2004 linear trends in stage distribution

- breast
- brain
- corpus & uterus, NOS
- esophagus
- liver
- prostate
- soft tissue including heart
- testis

Stage Distributions by Year of Diagnosis
Brain



Year of Diagnosis

Note: Stages for cases in 2001-2003 were coded according to Seer Summary Staging Manual 2000 (SS2000); stages for cases in 2004 were coded according to collaborative staging manual coding instructions derived SS2000.

4

18 cancer sites with significant differences in stage distribution between 2003 and 2004. Decreases in % of unknown stage from 2003 to 2004 alone or in combination with 2001-2004 linear trends do not explain the differences

- bladder
- cervix
- colon
- rectum
- kidney
- larynx
- lung
- melanoma
- oral
- ovary
- vulva
- anus
- HL
- eye
- gallbladder
- small intestine
- vagina
- cranial & other nervous system

4 Comparisons of Coding Instructions

- Differences in coding instructions between SS2000 and CSDSS2000 were identified for all cancer sites except for breast, cervix, and cranial nerves & other nervous sytem
- CS instructions are more detailed; additional anatomic structures used to define direct extension and/or lymph nodes involvement.
- Switched from one stage in SS2000 to another in CSDSS2000 for some cancer sites.

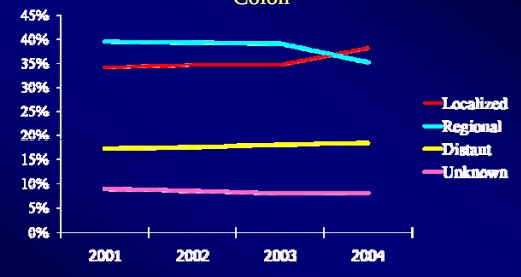
Changes in Extension (SS2000 vs. CS)

- Colon Cancer

Direct extension to	SS2000	CSDSS2000
Non-peritonealized pericolic tissues	-	Localized
Lamina propria, including lamina propira in the stalk of a polyp	-	Localized

Non-peritonealized pericolic tissues are not specified in SS2000 manual. Colon cancers with direct extension to the tissues could be coded as localized, regional, or unknown stage.

Stage Distributions by Year of Diagnosis Colon



Note: Stages for cases in 2001-2003 were coded according to Seer Summary Staging Manual 2000 (SS2000); stages for cases in 2004 were coded according to collaborative staging manual coding instructions derived SS2000.

Changes in Extension (SS2000 vs. CS)

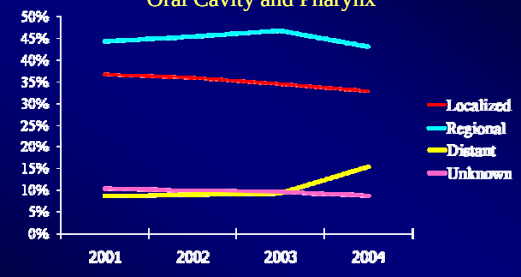
- Tonsil, Oropharynx

Direct extension to	SS2000	CSDSS2000
Retropharyngeal	Regional	Distant

- Nasopharynx

Lymph node extension to	SS2000	CSDSS2000
Posterior cervical (spinal accessory)	Regional	Distant

Stage Distributions by Year of Diagnosis Oral Cavity and Pharynx



Note: Stages for cases in 2001-2003 were coded according to Seer Summary Staging Manual 2000 (SS2000); stages for cases in 2004 were coded according to collaborative staging manual coding instructions derived SS2000.

Discussion

The observed differences in stage distribution between 2003 and 2004 may be attributable to :

- 2001-2004 linear trends in stage distribution
- decreases in unknown stage cases from 2003 to 2004 alone or in combination with 2001-2004 linear trends in stage distribution
- major changes in coding instruction (stage switch)
- additional anatomic structures used to define extensions in the CS manual
- human errors

Discussion

- Similar patterns were observed in a study that examined the stage comparability using double-coded cancer cases (colon & rectum, breast, prostate, and oral cavity only) diagnosed in 2004-2006 (Kahn et al. 2009).
- It supports our speculations that changes in coding instructions may affect stage distributions for some cancer sites.

Limitations

- Kappa statistic method could not be used to assess the agreement of the two coding systems because of the absence of double-coded stage data.
- Impact of changes in coding instructions on stage distributions could not be quantified.
- Not all changes in stage distribution from 2003 to 2004 have explanations. Some could be attributable to human errors.

Limitation

- Linear trends in stage distribution based on short-term data may not reflect true patterns of the trends
- Scales in stage shift are very small for some cancer sites and may not have a significant meaning in practice.

Recommendation

- Because changes in coding instructions may confound the real changes in stage distribution, their impacts should be assessed when analyzing combined pre-CS and post-CS stage data.
- Stage incomparability needs to be included in limitation section.

Recommendation

- Future changes in coding schemes should include provisions to double-code stage data during the transition years to evaluate the impact of the revisions.

Acknowledgement

- Maria J. Schymura
- NAACCR Board
  - Maureen Maclntyne
  - Antoinette Stroup
- Frances E. Rosee