Sex Misclassification in Central Cancer Registries:
Misclassification of Sex, the Example of Male Breast Cancer, and the NYSCR Sex Edit

Recinda Sherman, Florida Cancer Registry, NAACCR June 2011

Co-authors:
Jackie Button (Florida Cancer Registry)
Laura Soloway, Frank Boscoe (New York Cancer Registry)

Objectives

- Overview FCDS
- Male Breast Cancer
  - Overview
    - Data Quality; Coding Issues
  - FL "Real Man" Male Breast Cancer Follow-Back Project
- NY Sex Edit
  - Overview
  - Testing FL Data
  - Results
- Conclusion

Overview: FCDS

- Inception year 1981
  - NPCR, Incidence based
  - Contracted to University of Miami (1978)
  - 185,000 cancer reports annually
  - 115,000 incident cancer cases annually
- 2nd largest cancer registry in US
- Limited visual review
- Limited audits
- Extensive automated edits

Male Breast Cancer

- Rare
  - <1% all breast cancers
  - 2010 Estimates: New cases: 1,970 Deaths: 390
- Compared to women
  - Older age, higher stage, lower grade, more ER+/PR+
- Potential risk factors
  - Radiation
  - Genetic Predisposition
  - High estrogen levels
    - Obesity, Cirrhosis, Klinefelter's syndrome
- Incidence rates low but increasing in past 30 years

Data Quality

- Purpose of registry is research to guide public health policy to reduce the burden of cancer
- Research Error
  - Inappropriate public health response
  - Fail to protect the population; waste of public health funds
- Cause
  - Flawed research design, inappropriate assumptions, bias
  - Data quality in registry data
    - Coding changes

Sex Coding

- Male 1; Female 2
  - 3 Other (Hermaphrodite); 4 Transsexual; 9 Not stated
- Site-sex edit
  - 20% of invasive cases
- Visual review & Follow-back
  - Labor intensive
  - Performed only as special project
    - Sub-sets of data

Breast Cancer Coding

- ICD-O codes – cancer surveillance
  - ICD-O
    - 174 series female breast; 175 series male breast—separated code
  - ICD-O-2 & 3
    - 1990+
    - C50 series, combined code
- ICD-9 codes – clinical; reimbursement
  - 174 series female breast; 175 male breast—separated code
- ICD-10 code – mortality 1999+
  - C50 series, combined code

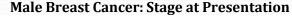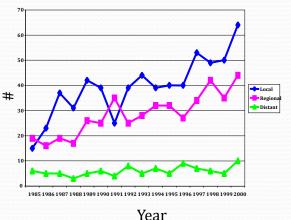# Male Breast Cancer: Why is the Incidence Increasing?

N. C. F. Hodgson, J. Button, J. Powell, F. Moffat, D. Franceschi
Division of Surgical Oncology, Department of Surgery, Department of Epidemiology
University of Miami, Miami, FL

Background: Male breast cancer is rare and little is known about regional population level patterns of incidence. The objective of this study was to determine the incidence of MBC in Florida compared to SEER cancer registry data.
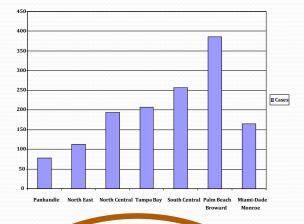
Methods: Study data were obtained from the Florida Cancer Data System (FCDS). All males with pathologically confirmed, invasive breast carcinoma diagnosed from 1985 to 2000 were included. Age-adjusted incidence rates, regional incidence rates and descriptive statistics were calculated. Estimated Annual Percent change (EAPC) using a linear model was calculated for the study period. Results were compared to the SEER Data
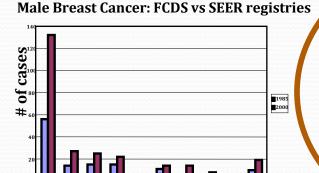
Results: 1386 cases of MBC were identified. Mean age of the study population was 68 ± 13 SD. Age - adjusted incidence rates ranged from 0.9 to 1.5. The highest rates were in the over 85 (11.8), followed by the 80-84 age group (10.6). Infiltrating ductal was the most common histology (960 cases), less common subtypes included mucinous (153 cases), lobular (26 cases), papillary (13 cases) and other (234). Localized disease accounted for 45% of all cases with regional disease in 32%, distant metastases (8%) and unstaged (15%). Most incident cases were diagnosed in the Palm Beach Broward Region (23%) and South Central Florida (19%) The number of cases increased from 56 in 1985 to 132 new cases in 2000. The EAPC for this 16 year period was 2.1 (95% CI: 1.2-4.7, p <0.005). Sensitivity analyses did not change this result. The SEER MBC rates have remained constant (EAPC 0.5, NS).
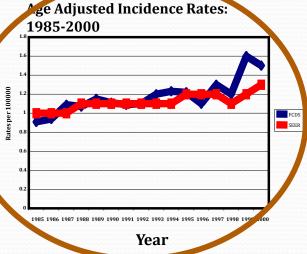

Male Breast Cancer: Stage at Presentation

Male Breast Cancer: Regional Cases

Male Breast Cancer: FCDS vs SEER registries

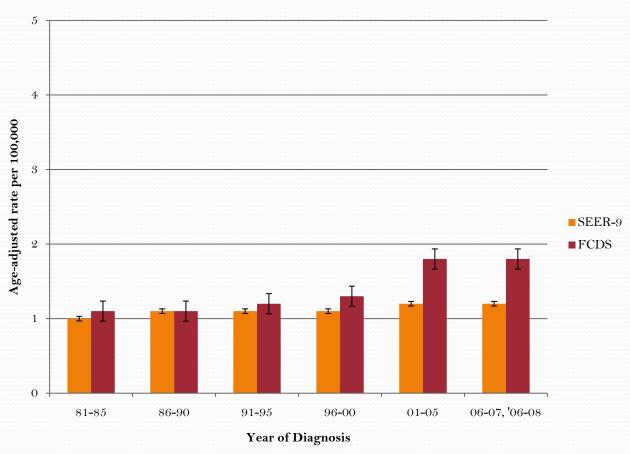Age Adjusted Incidence Rates: 1985-2000

Conclusions: The regional incidence of MBC is discordant from SEER cancer registry data. Population regional variation must be further evaluated to target high risk populations in order to plan future health care directives.

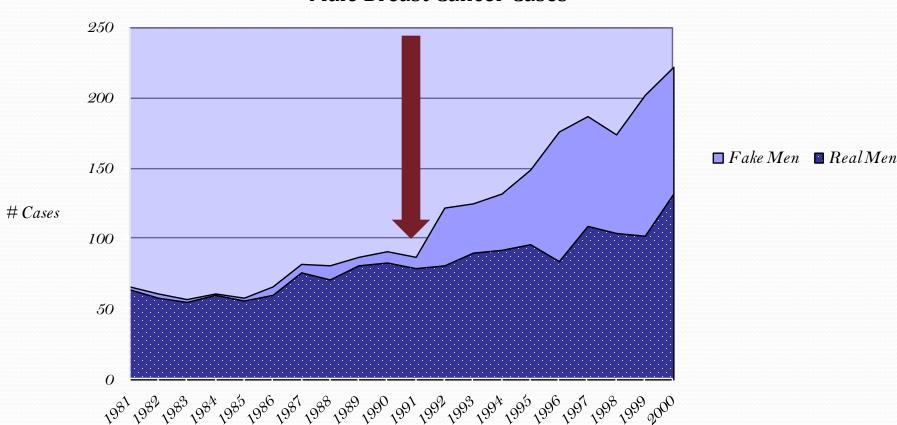# SEER vs Florida

Male Breast Cancer Rates

Florida "Real Man" Project

- Male Breast Cancer
- 2003, dx 1981-2000
- Visual review by first name
  - 3,800 cases
  - 904 manually identified as probably female
- Follow-back with hospital
  - all but 3 were confirmed as female

"Real Man" Follow-Back Results

Male Breast Cancer Cases

NY State Cancer Registry Sex Edit

- Automated
  - Algorithm
- Social Security Administration Data
  - Most popular male & female names
  - Decade of birth
- Flags suspicious name/sex combinations
  - Manual review

Testing NY Edit in Florida

- Sites
  - Breast
    - 100x more common in females
  - Thyroid
    - 3x more common in females
  - Liver
    - More common in minorities
  - Colorectal
- Diagnosed 1981-2008

- "Real" Man Project
  - Male breast cancers
  - Diagnosed 1981 - 2000

NY Edit : FL Site Results

- 68% of the cases agreed with probable sex
- 31% could not be evaluated
  - Name not gender specific or patient not born in decades included in edit
- 0.5% flagged as improbable sex/name combination
  - 0.3% in NY data
- 145 cases unknown sex flagged with probable sex

NY Edit: Site Results

- Overall
  - Varied by year; race/ethnicity
- Breast
  - **21% of males** flagged as improbable (0.2% for females)
- Thyroid
  - **1.3% of males** flagged as improbable (0.4% for female)
- Liver
  - 0.3% of males flagged as improbable (**1.1% for females**)
    - Carmen, Jean, Andrea, Angel
- Colorectal
  - 0.5% of males flagged as improbable (0.6% for female)

NY Edit: Unclassified

By Site

- Breast 31%
- CRC 33%
- Liver 29%
- Thyroid 28%

By Race/Ethnicity

- White 31%
- Black 34%
- AI/AN 37%
- All others 55%

- Non-Hispanic 30%
- Hispanic 46%

NY Edit: "Real"Results

- 2003 "Real Man" Project
  - 904 suspected female
    - 901 confirmed female
    - 3 confirmed as "Real Men"
- NY Edit
  - 81% (729 of 901) correctly identified as female
  - 1 of 3 correctly identified as male
  - Unable to assign probable sex for remaining cases
    - 19%
  - **No cases were incorrectly classified**

Breast Cancer Results

Reported Sex	Probable Sex (Algorithm)		
	Male	Female	Not Determined
Male	2733	1076	1440
Female	984	352216	155032
Unknown	1	77	29
	3718	353369	156501
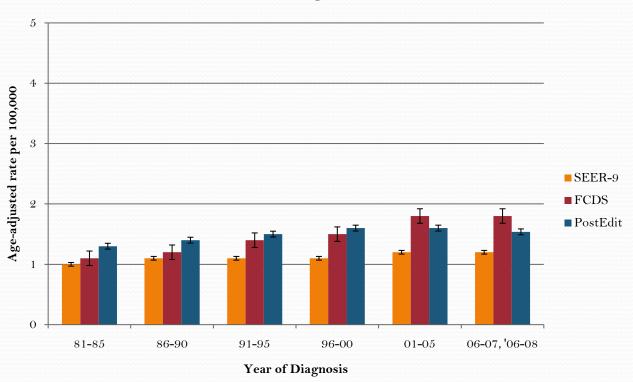
Male						
	1981–1985	1986–1900	1991–1995	1996–2000	2001–2005	2006–2008
PreEdit	1.1	1.2	1.4	1.5	1.8	1.8
PostEdit	1.3	1.4	1.5	1.6	1.6	1.5
% change	14%	11%	9%	6%	-12%	-15%

Female						
	1981–1985	1986–1900	1991–1995	1996–2000	2001–2005	2006–2008
PreEdit	106.9	123.7	125.9	129.4	114.8	109.5
PostEdit	106.8	123.6	125.8	129.4	115.0	109.8
% change	0%	0%	0%	0%	0%	0%

Data Improvement:
78 of 107 previously unknown cases
984 misclassified as Female (0.2%)
1,076 misclassified as Male (21%)

Breast Cancer Results

Male Breast Cancer Rates

Conclusion

- Sex misclassification artificially inflating male breast cancer rates in Florida

- Clinically significant regional differences of Florida male breast cancer rates are unlikely

- Use of NY Edit can improve quality of sex coding in Florida

  - Less sensitive for nonwhites and Hispanics

Recommendations

- For male breast cancer:
  - Change to female those cases NY edit flags as improbable
  - Manually follow-back those cases NY edit flags as un-assessable
- For unknown sex (of any site):
  - Assign probable sex from NY Edit
- Applicability of NY Edit for other sites
  - Needs further investigation
    - Follow-back at a minimum