

# Producing Delay Adjusted Rates and Trends Using SEER\*Stat and Joinpoint

Eric J.(Rocky) Feuer, Ph.D.  
National Cancer Institute

Steve Scoppa  
IMS, Inc.

NAACCR Webinar  
2:00-3:30 pm July 23, 2015

## **Surveillance Research Program, NCI & IMS**

**NCI:** Rocky Feuer, Huann-Sheng Chen, Doug Midthune

**IMS:** Martin Krapcho, Joe Zou, Steve Scoppa, Andy Lake, Danny Miller

## **NPCR, Division of Cancer Prevention and Control, CDC**

**CDC:** Reda Wilson , Trevor Thompson, Jessica King

**CDC Contractors:** Kevin Zhang, Xing Dong

## **Delay Model Sub-Committee, NAACCR**

Kevin Ward, Tom Tucker, Ron Dewar, Brenda Edwards, Frank Boscoe, Betsy Kohler

# What is Reporting Delay?

- Cases are reported approximately 2 years after the end of specific diagnosis year (e.g. NAACCR December 2014 submission reports cases diagnosed through 2012)
  - ◆ Cases are added and deleted in subsequent submissions
  - ◆ Delay modeling predicts case counts after a fixed number of years (e.g. after 12 years of delay) to correct for under-reporting of the most recent years
  
- Even though delay adjustment factors are relatively small, the bias is largest for the most recent data points
  - ◆ Any small change in the recent rates is seen as a potential harbinger of the impact of cancer control activities
  
- In the past delay adjustment factors have been estimated for SEER 9 and SEER 13
  - ◆ In the future we want to facilitate delay adjustment for single registries or any combination of registries

# Reporting Delay Example through 2009 Submission



	Submission Year										
Diagnosis Year	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
1997	2	3	4	5	6	7	8	9	10	11	12
1998		2	3	4	5	6	7	8	9	10	11
1999			2	3	4	5	6	7	8	9	10
2000				2	3	4	5	6	7	8	9
2001					2	3	4	5	6	7	8
2002						2	3	4	5	6	7
2003							2	3	4	5	6
2004								2	3	4	5
2005									2	3	4
2006										2	3
2007											2

# Dark Green – Used to Produce Trends



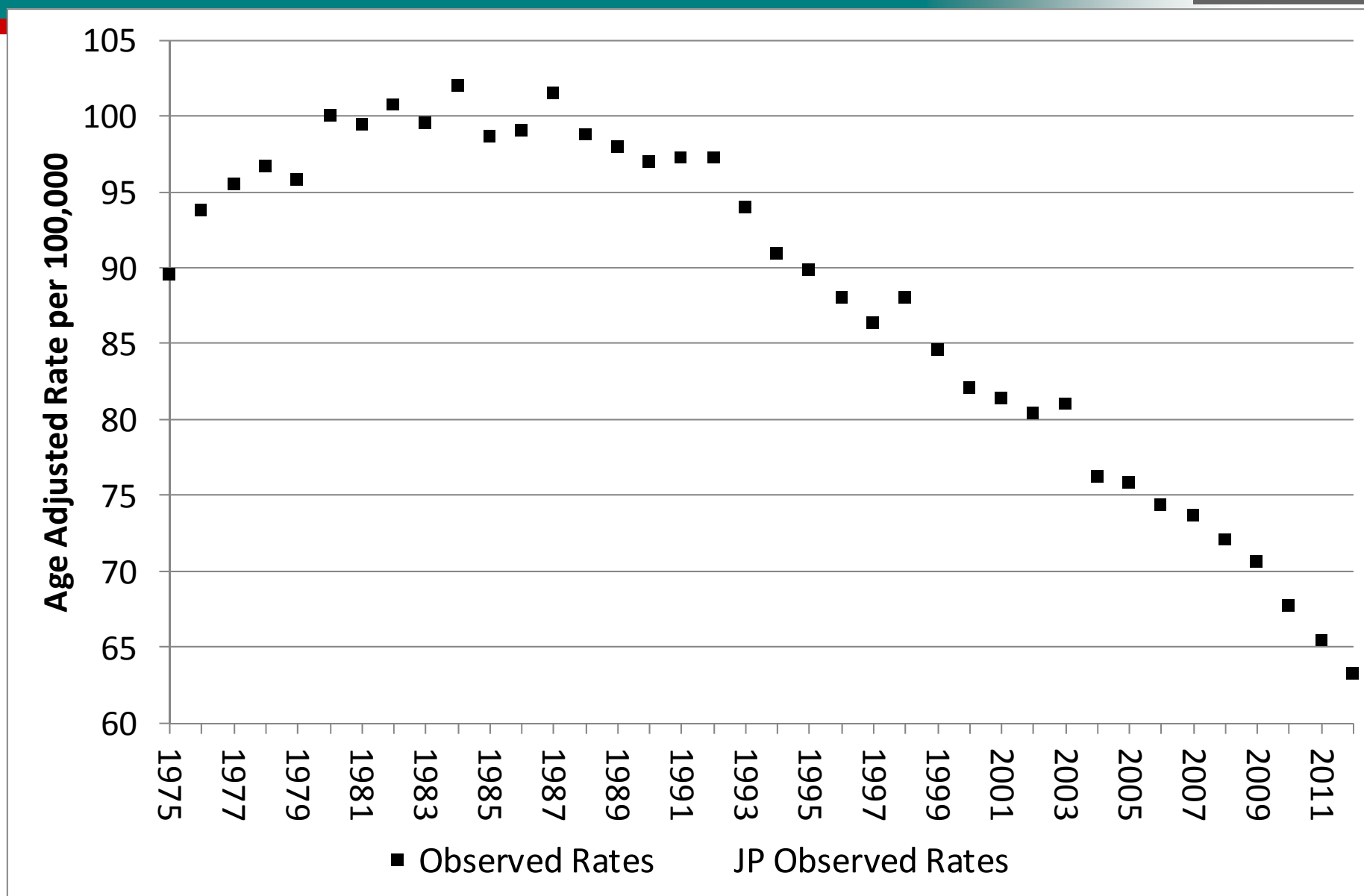
	Submission Year										
Diagnosis Year	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
1997	2	3	4	5	6	7	8	9	10	11	12
1998		2	3	4	5	6	7	8	9	10	11
1999			2	3	4	5	6	7	8	9	10
2000				2	3	4	5	6	7	8	9
2001					2	3	4	5	6	7	8
2002						2	3	4	5	6	7
2003							2	3	4	5	6
2004								2	3	4	5
2005									2	3	4
2006										2	3
2007											2

# Purpose of Delay Modeling: Use the Data in Green to Project to the Yellow

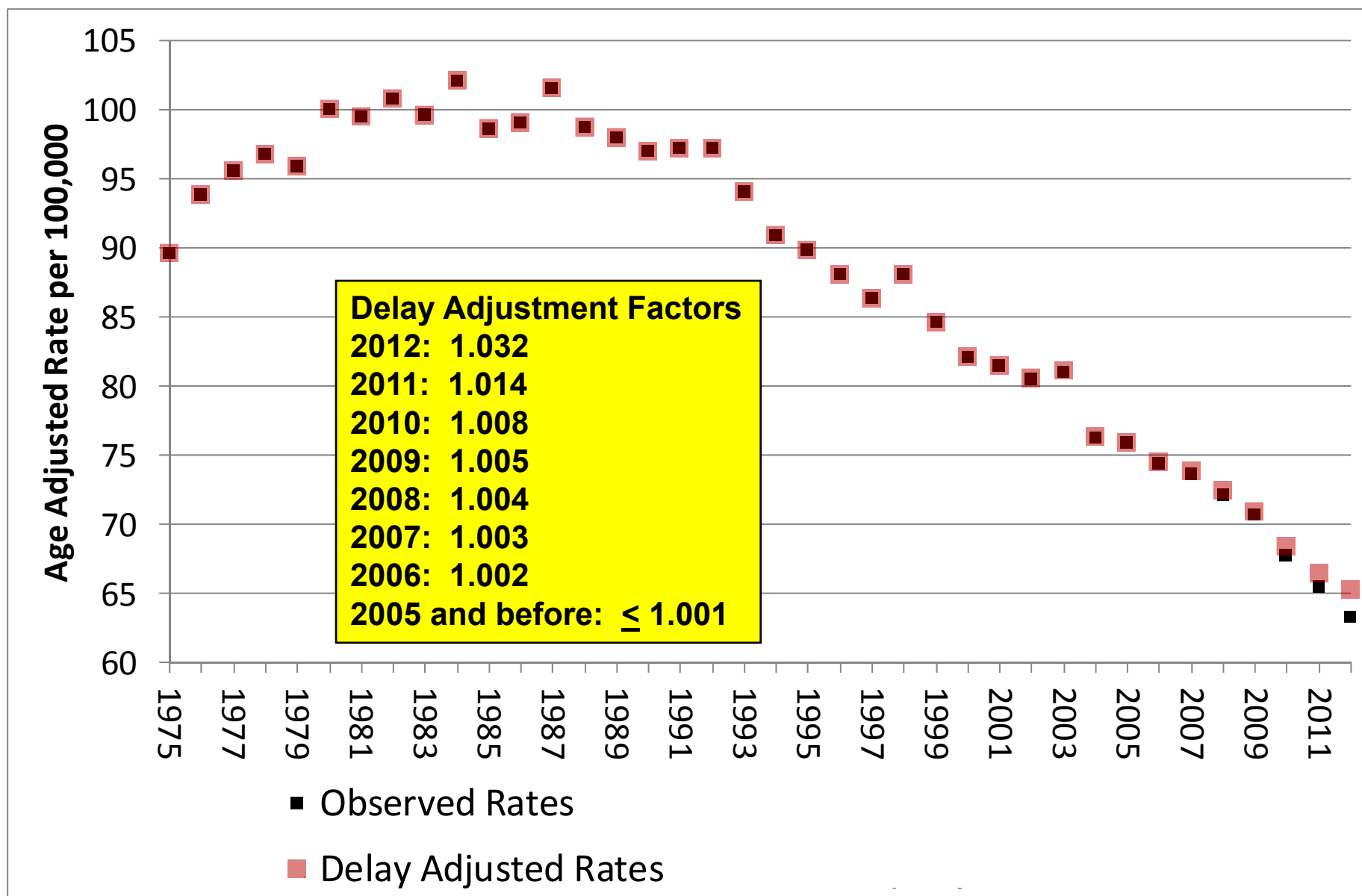


	Submission Year																				
Diagnosis Year	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009										
1997	2	3	4	5	6	7	8	9	10	11	12										
1998		2	3	4	5	6	7	8	9	10	11	12									
1999			2	3	4	5	6	7	8	9	10	11	12								
2000				2	3	4	5	6	7	8	9	10	11	12							
2001					2	3	4	5	6	7	8	9	10	11	12						
2002						2	3	4	5	6	7	8	9	10	11	12					
2003							2	3	4	5	6	7	8	9	10	11	12				
2004								2	3	4	5	6	7	8	9	10	11	12			
2005									2	3	4	5	6	7	8	9	10	11	12		
2006										2	3	4	5	6	7	8	9	10	11	12	
2007											2	3	4	5	6	7	8	9	10	11	12

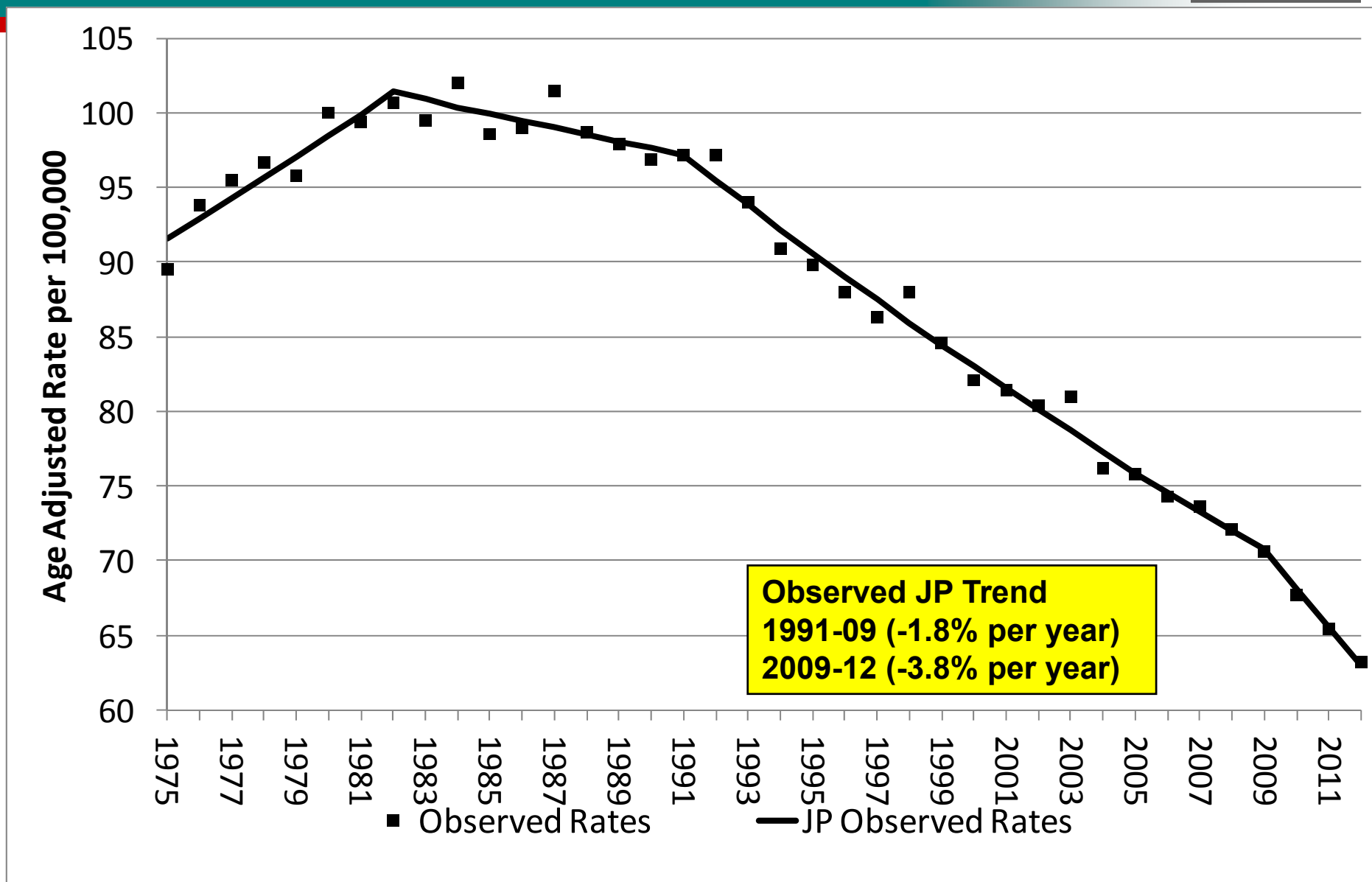
# Lung and Bronchus Cancer SEER 9 Incidence for Males All Races



# Lung and Bronchus Cancer SEER 9 Incidence for Males All Races

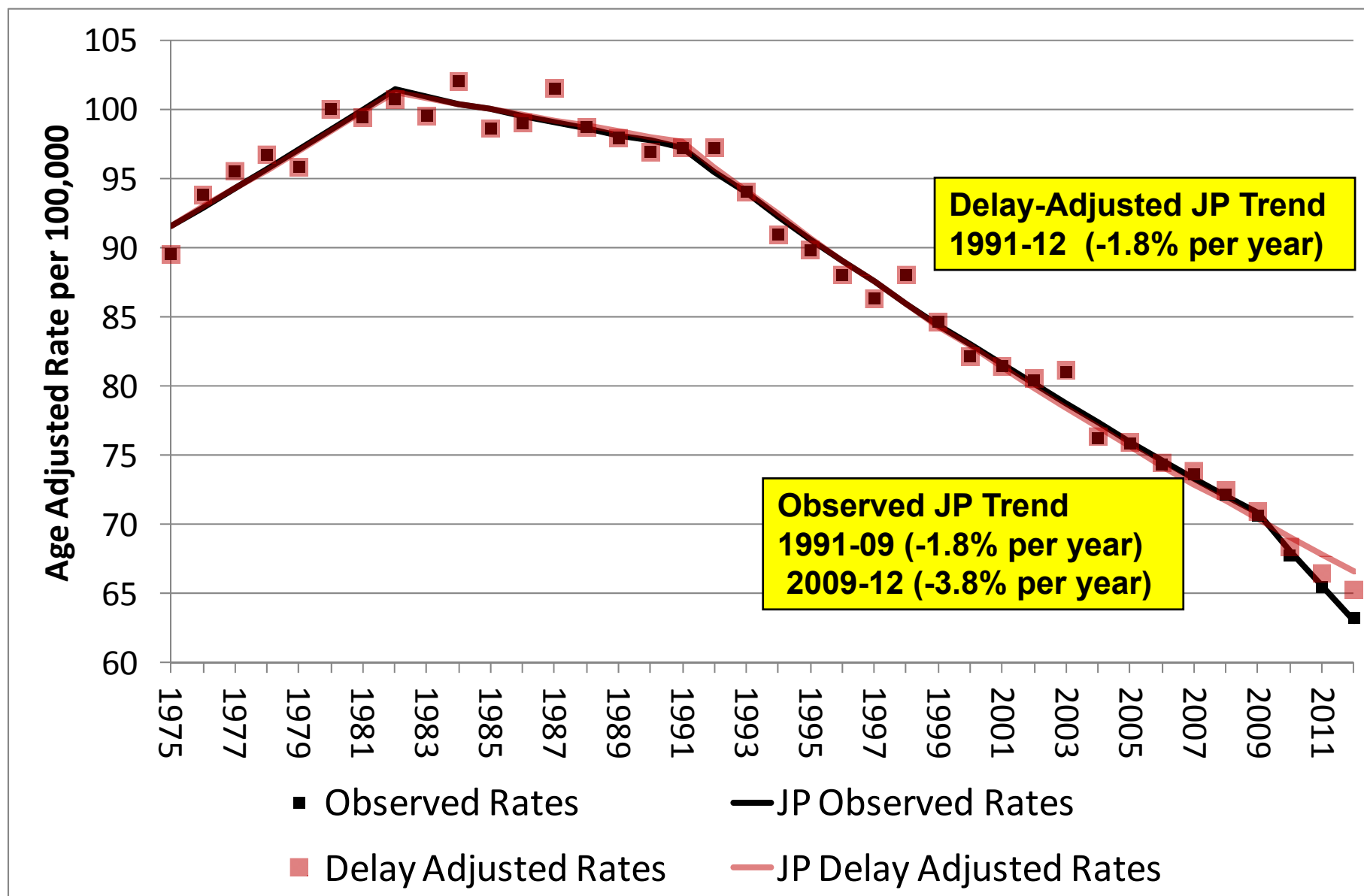


# Lung and Bronchus Cancer SEER 9 Incidence for Males All Races





# Lung and Bronchus Cancer SEER 9 Incidence for Males All Races



# # of Registries Considered for Delay Adjustment

Total:

- 56 US registries
  - ◆ 4 states are divided into 9 sub-state registries
  - ◆ 46 state registries
  - ◆ District of Columbia
- 13 Canadian Registries

# Five Overall Goals to Simplify and Standardize Delay Adjustment for NAACCR Registries



## 1. Delay adjustment factors should be adjusted to the same starting point for all registries

- ◆ SEER 9 is adjusted back to the 1983 submission but it is not possible to adjust NAACCR registries that far back
- ◆ Common starting point for many NAACCR registries is the 1999 submission (1997 diagnosis year)
  - Up to 17 years of reporting delay using the 2014 submission

## Overall Goals (continued)

2. Produce delay factors (and standard errors) for every (or almost every) U.S. and Canadian registry
  - ◆ We would like NAACCR, NPCR, and SEER to use delay factors derived from a common set of models
  - ◆ Estimate delay factors by cancer site and registry as a function of diagnosis year, age group, race, and gender
  - ◆ If the data are too sparse (<50 cases/year per cancer site & registry) model groups of registries with similar delay patterns
  - ◆ Registries that do not have “enough” years of reporting or have reporting patterns that are difficult to model may be excluded
    - From 2014 submission: 12 U.S and 6 Canadian registries excluded

# Overall Goals (continued)

3. These factors should easily be “combinable” across any combination of the factors (cancer site, diagnosis year, registry, age group, race, and gender) so that the analyst can obtain delay adjusted incidence rates for any selected sub-group
  - ◆ Added delay adjustment to SEER\*Stat which takes a weighted average of delay factors (with the weights equal to the # of cases in each subgroup)

Rate Session-1 Matrix-2

Page: All sites

	All races				White				
	Delay Rate	Delay Count	Implicit Delay Factor (Rate)	Implicit Delay Factor (Count)	Rate	Count	Pop	Delay Rate	Delay Count
2000-2012	470.8	5,019,452	1.009333	1.009272	466.4	4,973,341	1,080,680,215	480.4	4,140,526
2000	483.7	353,871	1.000957	1.000951	483.2	353,535	78,996,813	491.8	298,179
2001	487.5	362,628	1.001381	1.001371	486.8	362,131	79,867,817	495.5	304,234
2002	483.7	366,563	1.001871	1.001854	482.8	365,885	80,629,975	491.8	306,440
2003	469.8	362,862	1.002448	1.002422	468.6	361,985	81,347,854	477.3	301,952
2004	471.8	370,662	1.003131	1.003094	470.4	369,519	82,055,585	479.9	307,427
2005	468.1	364,296	1.003712	1.003663	465.4	362,966	80,414,394	478.4	303,150
2006	471.0	382,989	1.005013	1.004945	468.7	381,104	83,099,557	481.9	317,020
2007	478.3	397,359	1.006322	1.006233	475.3	394,897	83,810,676	488.3	326,914
2008	473.3	402,139	1.007973	1.007866	469.6	399,001	84,618,783	484.3	330,116
2009	471.1	409,352	1.010221	1.010085	466.3	405,265	85,402,713	481.0	334,009
2010	461.1	409,844	1.014187	1.013999	454.7	404,186	86,156,841	471.1	333,196
2011	458.3	416,057	1.021518	1.021258	448.7	407,397	86,817,163	469.1	337,445
2012	453.4	420,830	1.038206	1.037883	436.7	405,470	87,462,044	464.6	340,444

Cases were delay-adjusted using Delay factor.  
 Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard.  
 ~ Statistic could not be calculated.  
 @ Statistic could not be calculated due to at least one contributing case with unknown delay factor.

## Overall Goals (continued)

4. At deployment, abandon all previously developed delay adjustment factors, so there is a single agreed upon standard set of factors
  - ◆ Old SEER 9 and SEER 13 were not used this year in the SEER reporting
  
5. Allow (but do not necessarily encourage) analysis of delay adjusted incidence rates by factors beyond what are included in delay models
  - ◆ E.g. delay adjusted rates by stage of disease

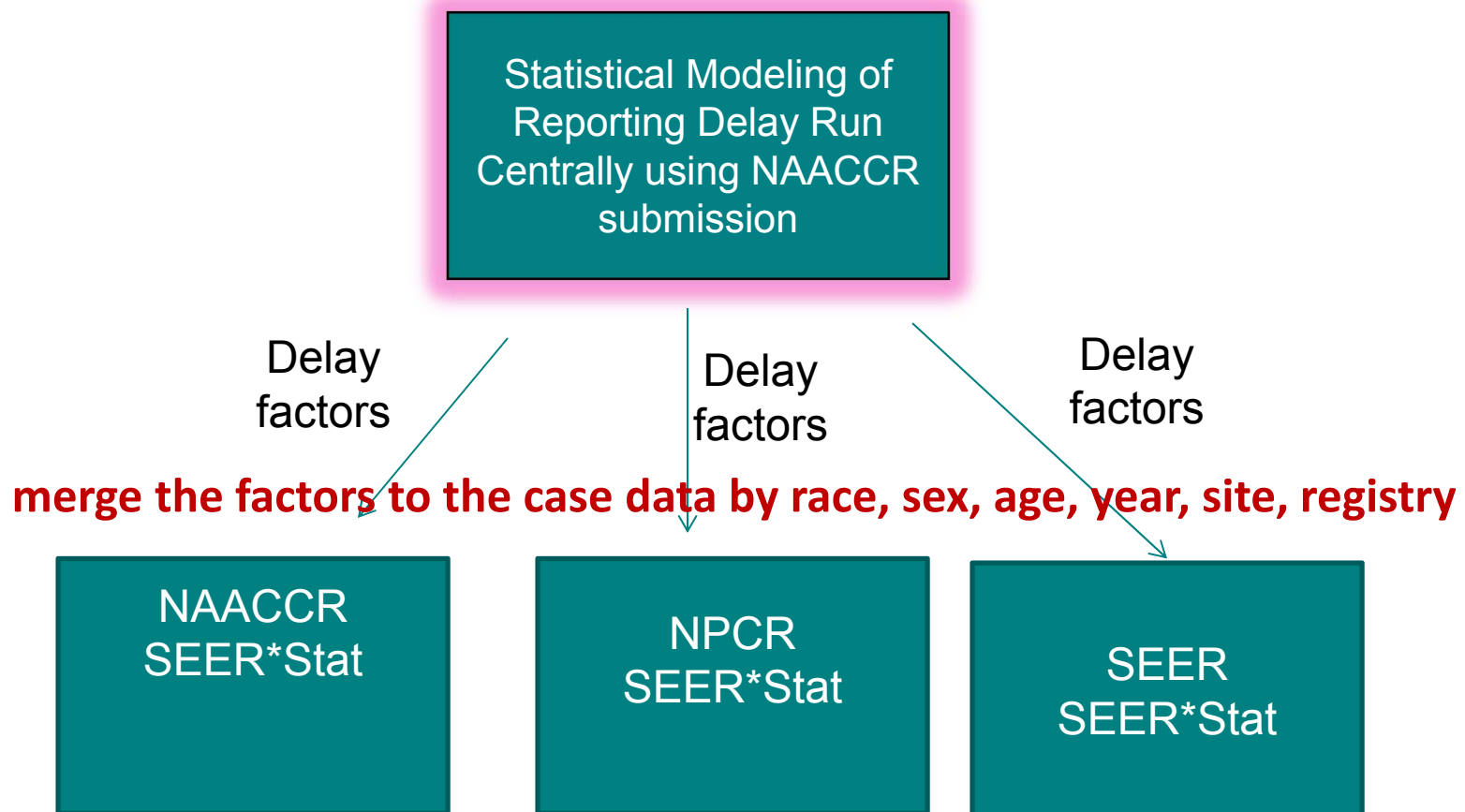
# Four Delay Adjustment Factors per Tumor

- For each tumor, using NAACCR December submission produce delay factors for each combination:

	All races	Race-specific
All sites	X	X
Site-specific	X	X

- Cancer Site
  - ◆ Since not every cancer site is covered, we cannot have SEER\*Stat weight individual sites specific factors to produce delay adjusted rates for All Sites. We need a separate All Sites factor
- Race
  - ◆ Separate all race and race specific factors are needed to account for the fact that some cases first come into the registry without a race designation, and assigned a race in later submissions
  - ◆ Race specific factors for Whites, Blacks, API
    - Not AI/AN (too sparse)

# Implementation: All registries use a single derived set of factors





# Philosophy for First Release of NAACCR-Based Delay Factors

- Factors for aggregations of registries are more stable than individual registry factors
- Release aggregation of registries for general use
  - ◆ SEER 9, SEER 13, SEER 18, US, Canada
  - ◆ SEER\*Stat files with delay adjustment factors will not have individual registry identifiers
- Beta-test year for individual registry use only
  - ◆ Release reports to individual registries with their factors
  - ◆ These should NOT be used for any publication or reporting (including state registry reports) this first year
  - ◆ Encourage feedback during 2015 so we can consider modifications for 2016 release (using December 2015 submission) to allow delay adjustment for individual registries (or any group of registries)
- Reminder – Delay is not a quality measure

# SEER9, SEER 13, SEER18, US, and Canada 2012 Implicit NAACCR-Based Delay Factors



Cancer Site	SEER 9	SEER 13	SEER 18	US	Canada
All Sites	1.032	1.033	1.038	1.054	1.035
Oral Cavity and Pharynx	1.030	1.032	1.038	1.053	1.024
Esophagus	1.019	1.020	1.023	1.030	1.012
Stomach	1.022	1.022	1.025	1.031	1.019
Colon and Rectum	1.020	1.020	1.024	1.034	1.023
Liver and Intrahepatic Bile Duct	1.053	1.064	1.061	1.057	1.091
Pancreas	1.037	1.041	1.041	1.046	1.047
Larynx	1.021	1.022	1.027	1.041	1.016
Lung and Bronchus	1.032	1.033	1.036	1.043	1.038
Melanoma of the Skin	1.030	1.030	1.035	1.058	1.034
Breast (female)	1.017	1.017	1.022	1.042	1.019
Cervix Uteri (female)	1.020	1.020	1.025	1.039	1.030

# SEER9, SEER 13, SEER18, US, and Canada 2012 Implicit NAACCR-Based Delay Factors



	SEER 9	SEER 13	SEER 18	US	Canada
<b>Cancer Site</b>					
Corpus and Uterus, NOS (female)	1.012	1.012	1.016	1.028	1.012
Ovary (female)	1.042	1.040	1.043	1.061	1.039
Prostate (male)	1.038	1.042	1.050	1.082	1.029
Testis (male)	1.018	1.018	1.020	1.035	1.013
Urinary Bladder Kidney and Renal Pelvis	1.032	1.031	1.036	1.052	1.021
	1.037	1.037	1.042	1.054	1.057
Brain and Other Nervous System	1.038	1.040	1.043	1.057	1.019
Thyroid	1.023	1.021	1.027	1.036	1.032
Hodgkin Lymphoma	1.025	1.022	1.027	1.047	1.025
Non-Hodgkin Lymphoma	1.042	1.042	1.049	1.072	1.028
Myeloma	1.108	1.105	1.117	1.152	1.050
Leukemia	1.154	1.146	1.156	1.187	1.133

# Sample Registry Report Dec 2014 Submission



Cancer Site	Sex	Race	Reference Implicit Delay Factors 2012		Registry xxxx Implicit Delay Factor 2012
			United States	SEER 18	
All Sites	Both Sexes	All Races	1.054	1.033	1.088
All Sites	Both Sexes	White	1.065	1.041	1.099
All Sites	Both Sexes	Black	1.066	1.040	1.097
All Sites	Both Sexes	API	1.053	1.040	1.098
All Sites	Male	All Races	1.064	1.040	1.103
All Sites	Male	White	1.077	1.050	1.116
All Sites	Male	Black	1.077	1.049	1.113
All Sites	Male	API	1.066	1.051	1.115
All Sites	Female	All Races	1.046	1.026	1.073
All Sites	Female	White	1.053	1.031	1.082
All Sites	Female	Black	1.055	1.031	1.080
All Sites	Female	API	1.042	1.030	1.082
Oral Cavity and Pharynx	Both Sexes	All Races	1.053	1.032	1.108
Oral Cavity and Pharynx	Both Sexes	White	1.062	1.037	1.097
Oral Cavity and Pharynx	Both Sexes	Black	1.059	1.034	1.018
Oral Cavity and Pharynx	Both Sexes	API	1.045	1.032	1.069

# Goal for the Next Year

Revise the models so they provide the proper balance of two opposing goals

**Capture the unique patterns of reporting delay for each registry**



**Provide stable estimates that are not too noisy**

## Recent Events



- Registry specific reports were sent, or will be sent shortly
- Registry specific SEER\*Stat files with delay adjustment factors will be available upon request
- New delay adjustment factors were included in NCI's Cancer Statistics Review and NAACCR's Cancer Incidence in North America 2008-2012 Volume One