

Proficiency Testing in Cytopathology

A Personal Perspective

Timothy J. O'Leary

Department of Cellular Pathology
and Genetics

Armed Forces Institute of
Pathology



Purposes of Cytologic Examination

- Diagnose Disease
 - Fine needle aspiration of breast and thyroid
 - Cytopathologic examination of effusion
- Prevent Disease
 - **Screening for cervical cancer precursors**
 - Screening for other cancers
 - Esophagus
 - Anus

Cervical Cancer Incidence and Mortality

Region	Incidence	Mortality
East Africa	44.32	24.24
Central America	40.28	17.03
South America	30.92	11.97
S. Central Asia	26.47	14.95
North Africa	16.77	9.08
Western Europe	10.43	3.74
North America	7.88	3.23
Western Asia	4.77	2.50

Age-Adjusted SEER Incidence and Mortality

Race/Ethnicity	Incidence	Mortality
All Races	8.7	2.7
White Hispanic	15.4	3.6
White non-Hisp.	6.9	2.3
Black	11.0	5.7
Asian/Pac. Isl.	10.3	2.7
Native Amer.	6.4	2.9
Hispanic	14.4	3.3

Age-Adjusted SEER Incidence and Mortality

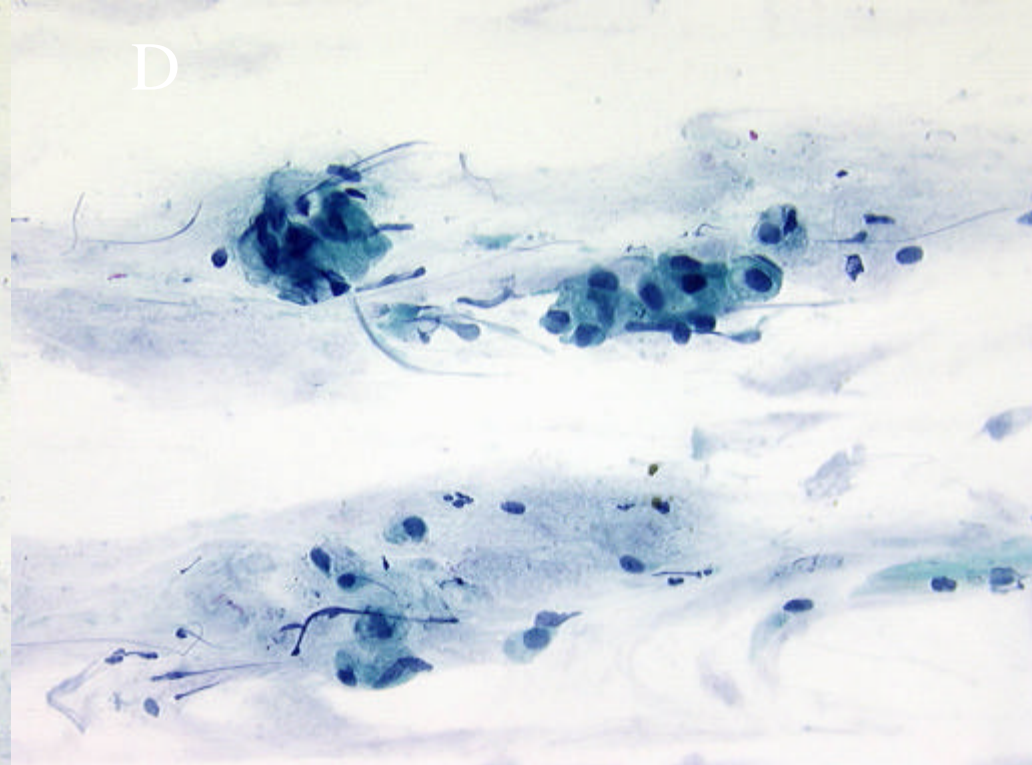
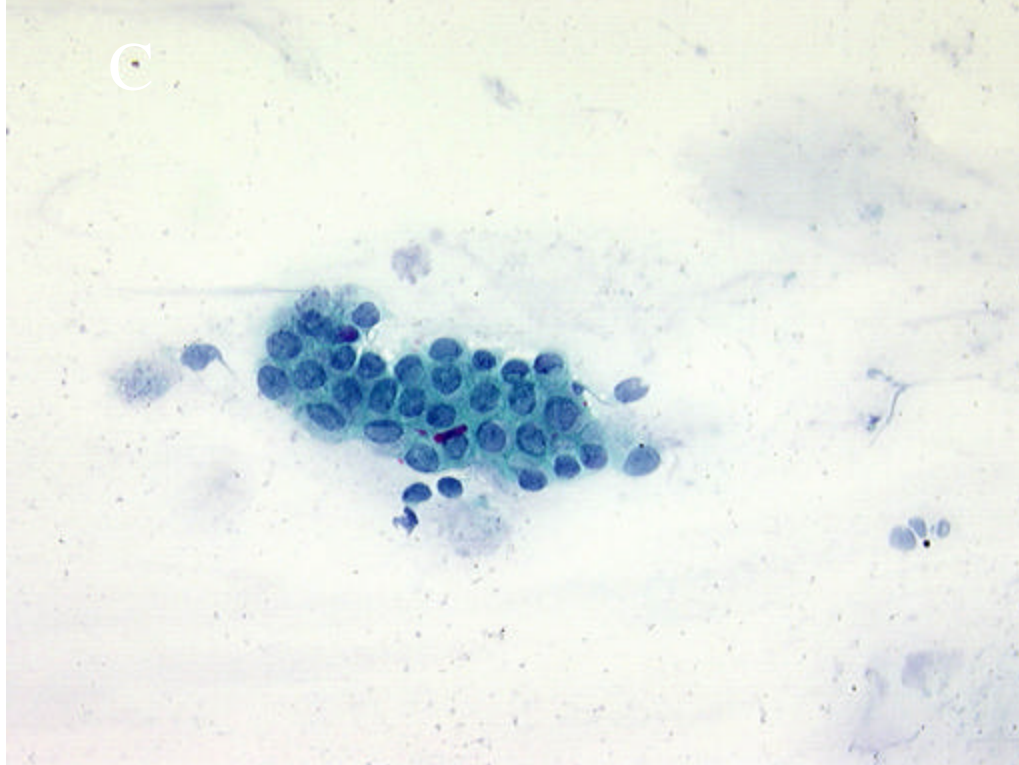
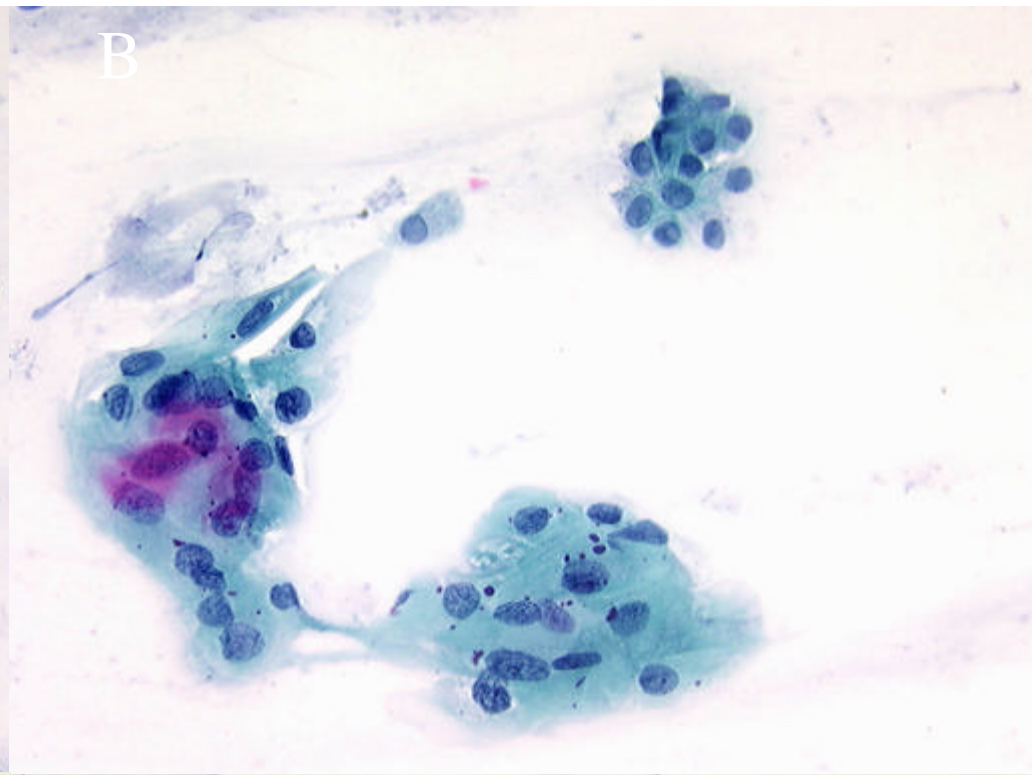
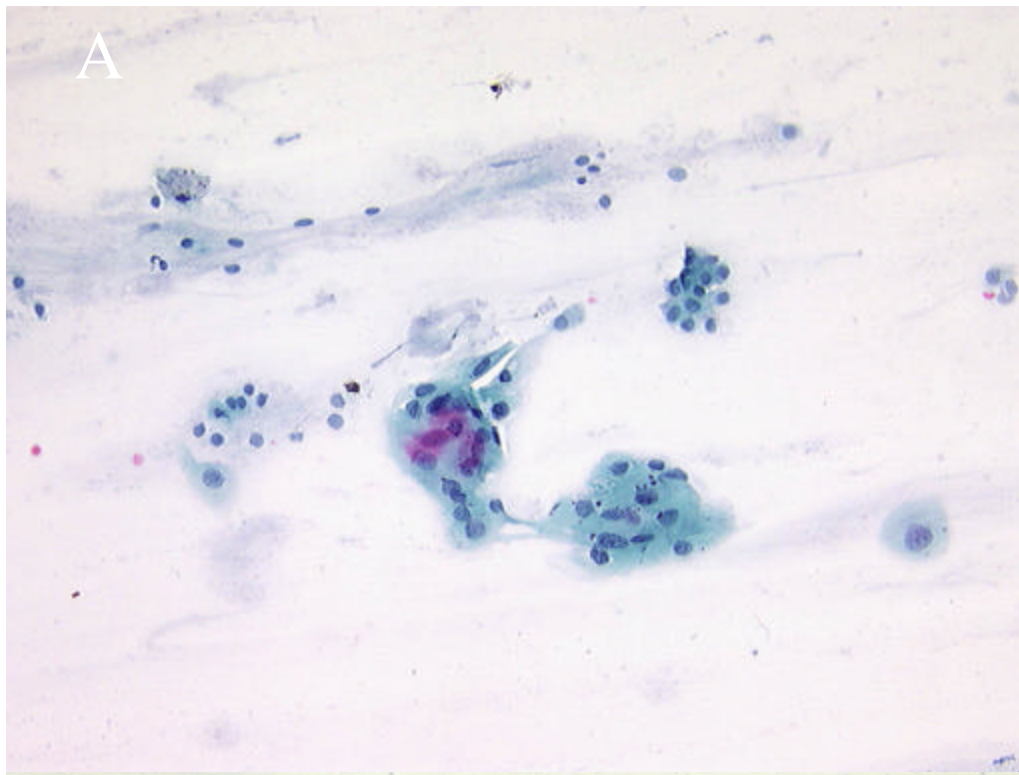
Registry	Incidence	Mortality
Los Angeles	10.8	3.1
Atlanta	9.3	2.3
Hawaii	8.7	1.7
Detroit	8.4	2.5
Utah	6.2	1.9

Average Age-Adjusted Cervical Cancer Mortality Rates, 1994-1998 (per 100,000)

- Highest-rate states
 - DC 4.1
 - Delaware 4.0
 - West Virginia 3.6
 - Louisiana 3.6
 - Kentucky 3.5
- Lowest-rate States
 - Hawaii 1.7
 - Minnesota 1.7
 - South Dakota 1.6
 - Alaska 1.5
 - Wyoming 1.2

Pathogenesis of Cervical Cancer

- Infection with papillomavirus - HPV 16, 18, 31, 33,.....
- Low grade squamous intraepithelial lesion
 - Mild dysplasia
 - Cervical intraepithelial neoplasia (CIN1)
- High grade squamous intraepithelial lesion
 - Moderate to severe dysplasia
 - Cervical intraepithelial neoplasia (CIN2/3)
- Invasive cervical cancer

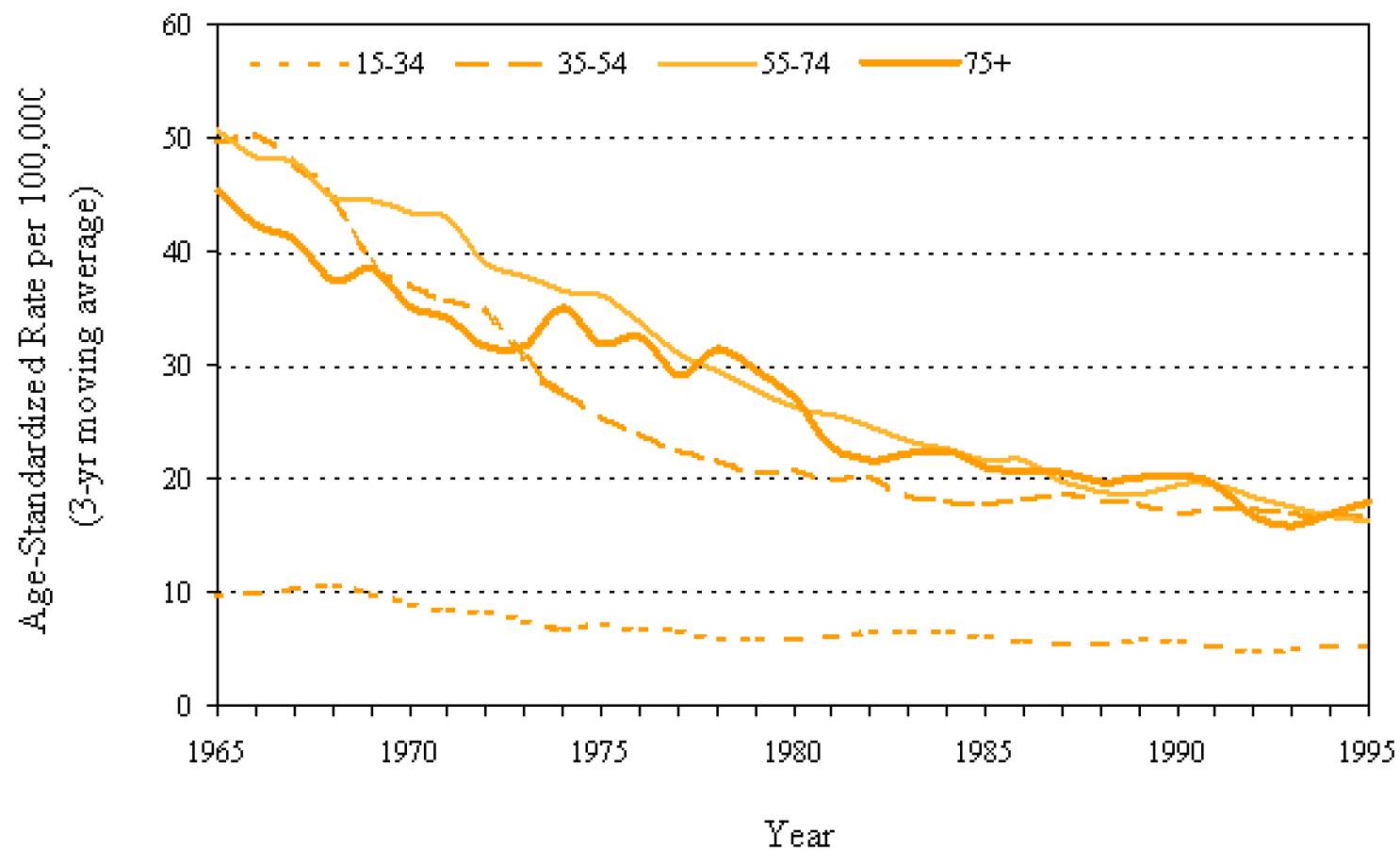


Prevention of Cervical Cancer Mortality

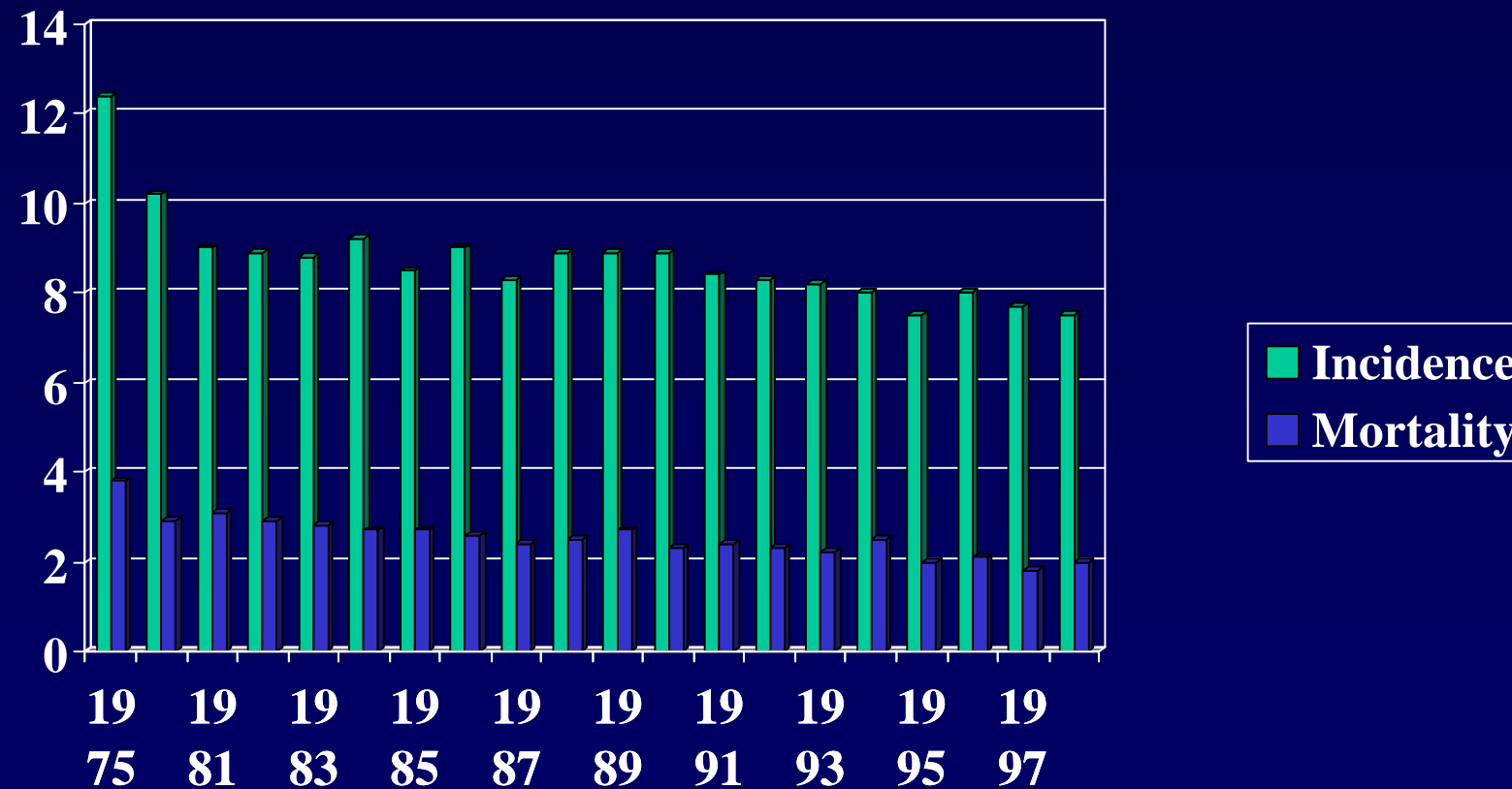
- Prevent transmission of HPV- cheap!
 - Condoms
 - Behavior modification
- Treat HPV infections by removing infected cells - expensive!
 - Identify infected individuals (Pap Smear, HPV testing, colposcopy)
 - Cryosurgery, LEEP, etc.
- Treat invasive cancer - outrageous!

Effects of Screening

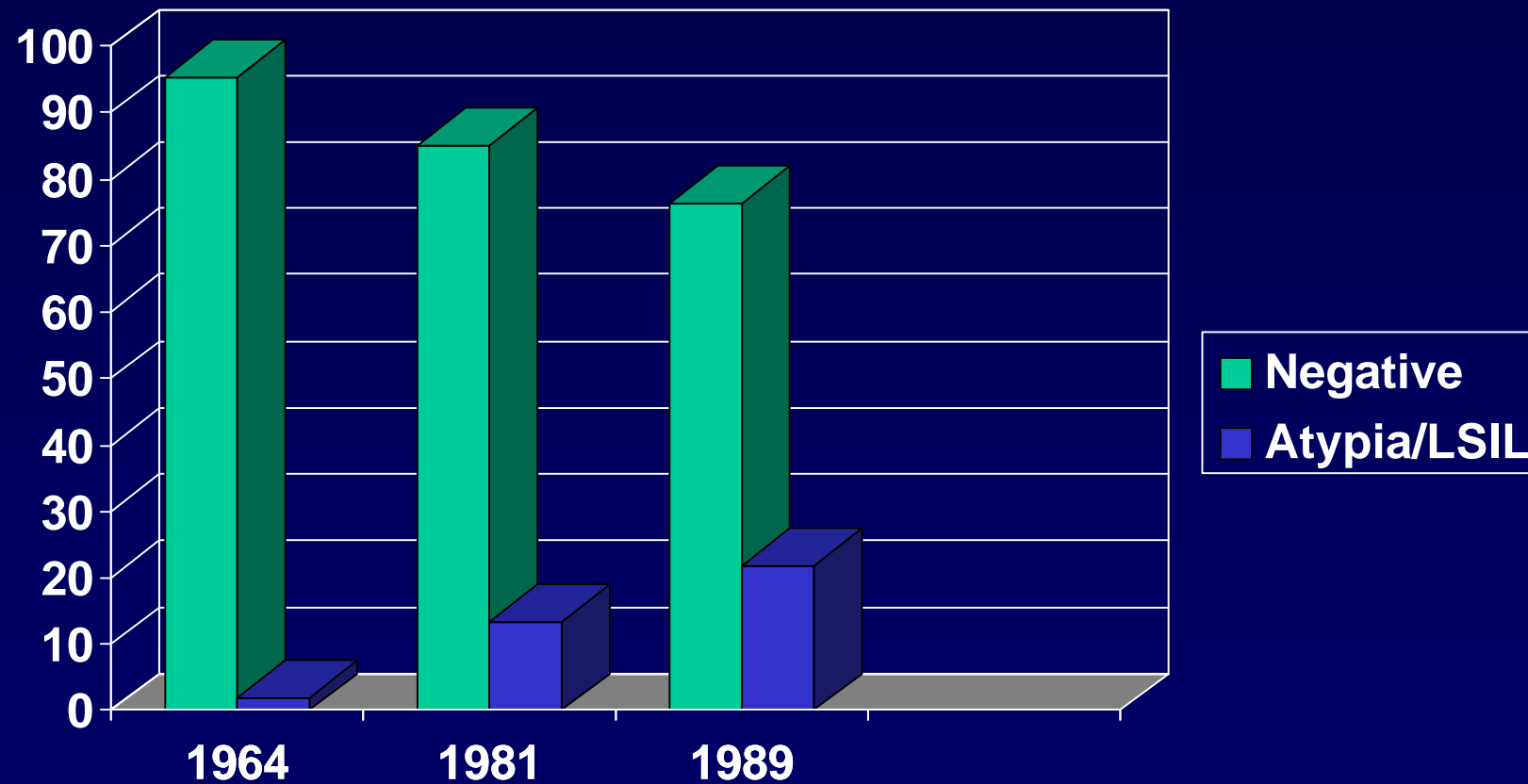
Female Cervical Cancer (ICD9 180) Incidence Rates, Ontario, 1964-1996



Invasive Cervical Cancer (SEER Data)

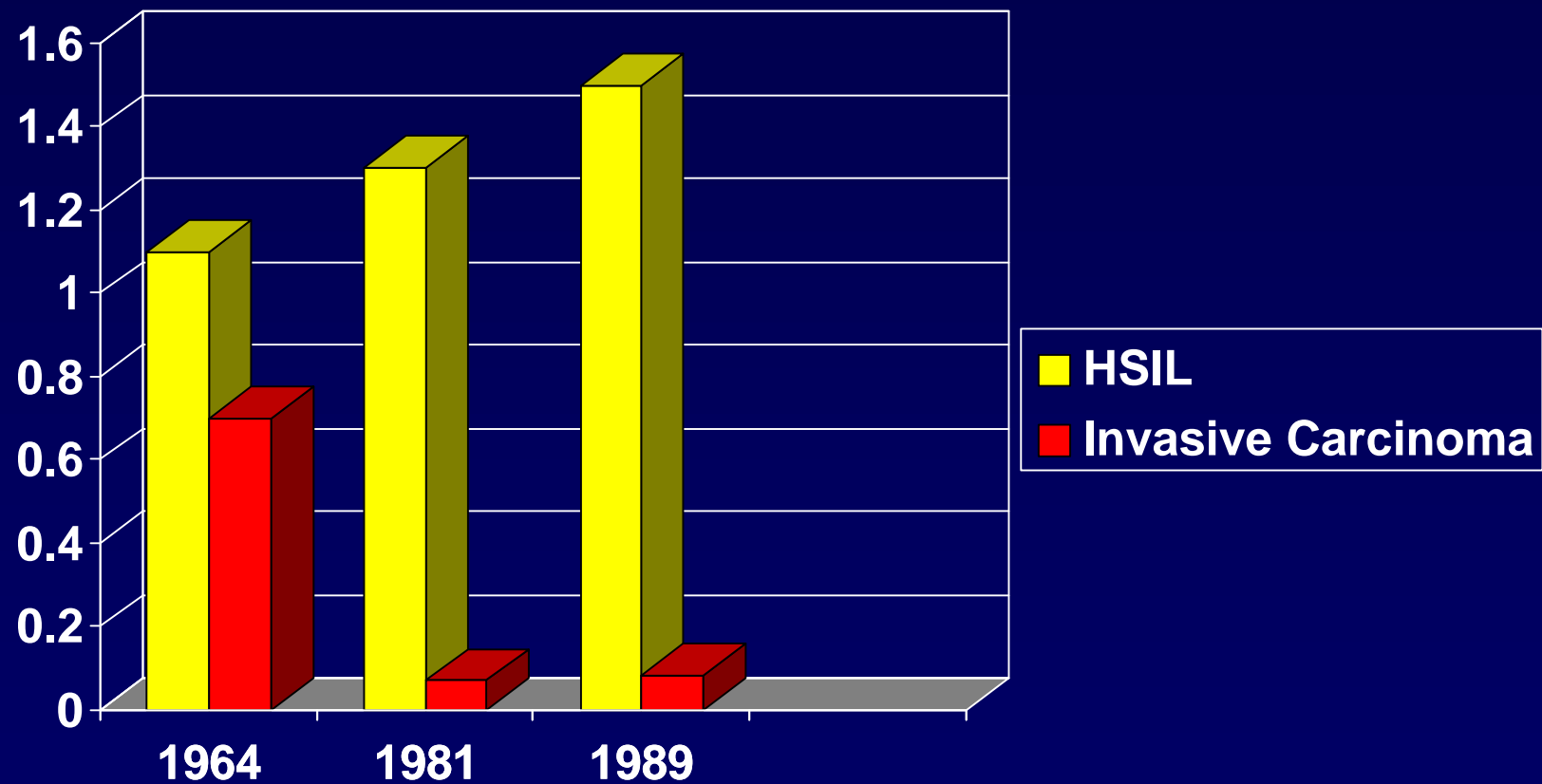


Screening in an Indigent Population



MJ Costa et al, "Cervicovaginal Cytology in an Indigent Population: Comparison of Results for 1964, 1981 and 1989," *Acta Cytol* 35; 1991: 51-56

Screening in an Indigent Population



MJ Costa et al, "Cervicovaginal Cytology in an Indigent Population: Comparison of Results for 1964, 1981 and 1989," *Acta Cytol* 35; 1991: 51-56

Assuring A Successful Cervical Cancer Screening Program

- Women must be screened regularly
 - Most invasive cervical cancer in the United States occurs in women who have not been screened in the last five years - usually older age
- Screening must be competent
- Screening must be followed by appropriate interventions

Approaches to Quality Assurance

- Personnel Standards
- Statistical Quality Control
 - Cytotechnologist/cytopathologist SIL rates
 - Rescreening of “WNL” slides
 - Rescreening of “WNL” slides reported previously on patients with new SIL diagnosis
- Proficiency testing
- Reporting on specimen adequacy

Role of Proficiency Testing

- Assure competent screening
- Should focus on the identification of women harboring HPV disease - most importantly, high-grade lesions
- Should not be overly concerned with other infections, rare diseases (ovarian cancer)

Cytology PT Programs

- Royal College of Pathologists of Australasia
 - Gynecologic and non-Gynecologic
- Cyquest
- Brazilian Society of Cytopathology
 - Gynecologic and non-Gynecologic

Cytology PT Programs

- Labquality (Finland)
 - Gynecologic and non-Gynecologic
 - Digital image based
- Ontario Medical Association
 - Gynecologic and non-Gynecologic
- College of Physicians and Surgeons of Alberta

Cytology PT Programs

- National Working Group for External Quality Control in Cervical Screening - Italy
 - Gynecologic
- UK - Regional Programs
- College of American Pathologists
 - Gynecologic and Non-Gynecologic

Cytology PT Programs

- Department of Veterans Affairs - USA
 - Gynecologic and Non-Gynecologic
- Cytoquest - USA
 - Gynecologic and non-gynecologic
- Maryland and New York- USA
 - Gynecologic

Types of PT Programs

- Validated slides - committee review
- Round-robin slide exchange

False Negative Rates – CAP Surveys- Pap

Lab Specimen Volume

	<5000	5000- 14999	15000- 49999	50000- 99999	>99999
1992	6.2	4.0	2.4	2.8	2.4
1998	5.6	3.6	2.1	2.8	1.9
1999	5.4	3.4	2.1	1.8	2.3
2000*	5.2	3.4	2.6	1.6	2.9

* Includes liquid-based preparations

Wisconsin - 1971

- Program began 1967, reported Pap classes
- Labs compared with 4 reference labs
- Based scoring approach on syphilis serology
- Included cervical, sputum, body fluids, urine

Wisconsin - 1971

- Marked discrepancies reported from reference laboratories in 9 of 53 specimens!

S Inhorn and E Clarke, "A State-wide Proficiency Testing Program in Cytology," *Acta Cytol.* 1971; 15: 351-56.

Wisconsin 1994 - PT Exercise

PT Design

- A Unsatisfactory
- B Normal/BCC
- C LSIL
- D HSIL/Ca

USA Practice

- A Unsatisfactory
- B Normal/BCC
- C ASCUS
- D LSIL/HSIL/Ca

Wisconsin - 1994

- 49 pathologists, Pa 70 cytotechnologists, 10 glass slides
- Tested both locator and interpretive skills
- Pathologist performance significantly worse than that of cytotechnologists
- PT design probably used inappropriate scoring categories

Wisconsin - 1994

Failure	Pathologists	Cytotechnologists
Pathologist grid	11	7
Cytotech Grid	7	1

Maryland Criteria for Government-Mandated PT

- Must evaluate cytology as practiced
- Test material must represent actual patient material.
- Test material must be similar for all examinees.
- Test material evaluators must be experts in cytology
- Must use standard diagnostic terminology
- Must make practical use of limited resources
- Must be enforceable
- Must be legally defensible

JM DeBoy ad BR Jarboe, "Government-Mandated Cytology Proficiency Testing: Practical, Equitable and Defensible Standards," QRB 1991; 17:152-161

Maryland: 1989-present

- All slides evaluated by 5 experts - adopted based upon 4 of 5.
- Diagnostic categories
 - A: Unsatisfactory
 - B: No evidence of cancer
 - C: Significant epithelial atypia
 - D: Cancer present

Maryland: 1989-present

- Must achieve 90% to pass. Uses scoring grid in which missing invasive cancer gives negative score.
- In 1991, cost approximately \$2000 to obtain test set of 10 slides
- In 1991, cost approximately \$68,000 to operate program

Maryland vs. CMS (HCFA)

- Annual
- Announced
- Common terminology
- CIN2 in cat C
- 10 slides/test
- Biannual
- Unannounced
- Bethesda system
- CIN2 in cat D
- 20 slides/test

New York 1968-1982

- Pass rate - single test
 - 1971/1972 - 53%
 - 1982/1984 - 80%
- Failure rate inversely proportional to test volume
- Failing labs (2 events) all “single test labs”, 6-100% of which employed no cytotechnologists.

VA Cytology Program

- All pathologists and cytotechnologists must participate if they are reviewing cytology cases.
- Cases contributed by individuals.
- Slides circulated and evaluate by VA pathologists

VA Cytology Program

- Statistical data accumulated
- Reviewed and validated by committee of cytopathologists.
- “Clinically significant peer outliers” identified. Approximately 5% of diagnoses.

VA Cytology Program

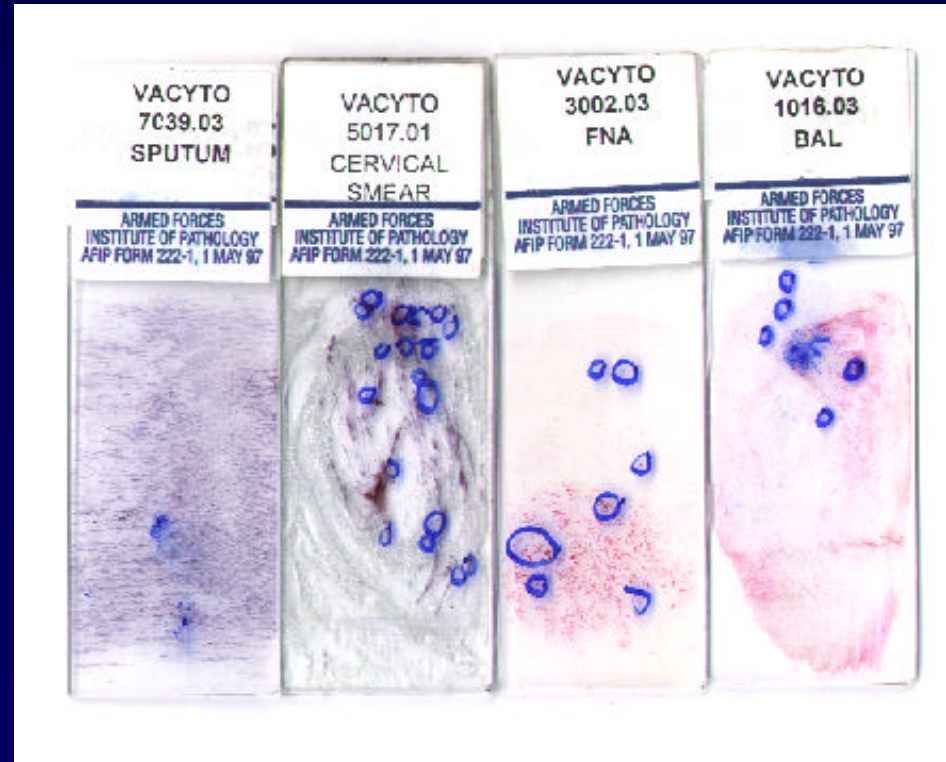
- About 70% of submitted cases judged to have PT value (sufficient concordance among all reviewing pathologists).
- Approximately 5% of diagnoses characterized as “Clinically significant peer outliers”.

Personal communication: Dr. Ted Beals

VA Cytology Program

- Marked as
 - satisfactory
 - unsatisfactory
 - satisfactorybut limited by....
- Categorized as
 - WNL
 - BCC
 - Epithelial cell abnormality
- Given specific diagnosis

VA Cytology Program



West Thames, UK, 1989-94

- Screen 10 slides in 2 hours. A single miss gives a failure.
- 247 cytologists participated in at least 1 round
- 63 Cytologists completed in 7 rounds of PT

West Thames, UK, 1989-94

- Pass rate 96.4%. Maximum score achieved in 65% of tests
- 3 Cytologists had misses in 2 of 3 rounds
- 7 cytologists completed 6 rounds without error, yet missed abnormal on 7th. Each slide missed had been read correctly at least 96% of the time.

West Thames, UK, 1989-94

Conclusions

- Three cytologists with poor performance all chose to retire. Hence, testing capable of eliminating really poor performers.
- Even exceptional screeners will occasionally fail a PT result. This may occur in up to 15% of challenges.

C Green and DV Coleman, "Evaluation of proficiency testing as a method of assessing competence to screen cervical smears,"
Cytopathology 1997; 8: 96-102

Yorkshire, UK

- 15 laboratories
- Each lab contributed 4 slides, for total of 60
- Each lab sends a batch of four slides each week to the next lab, forming a ring, until complete circulation of 60 slides achieved. Slides screened within 40 minutes.
- If 70% (11 of 15) of lab reports agree on an answer, it is considered a reference answer. Slides for which this is not achieved are excluded.

CytoView

- Computerized PT
- Users evaluate 10 cases in 2 hours
- Each case consists of >8,000 image files, representing up to 40x magnification in each of 5 focal planes. Images registered and fused.

RN Taylor et al, "Cytoview: a computer image-based Papanicolaou smear proficiency test," *Acta Cytol* 1999; 43: 1045-1051.

CytoView

<http://www.phppo.cdc.gov/clia/cyto2.asp>

Realty Testing the CMS Scheme

- Workshop PT performance - -three 10-slide PT sets given to 125 pathologists and 270 cytotechnologists.
- Passing grade 90%
 - Pathologists: 46.4% pass rate
 - Cytotechnologists - 69.3% pass rate
- Passing grade 80%
 - Pathologists: 76.4% pass rate
 - Cytotechnologists - 94.6% pass rate

Realty Testing the CMS Scheme

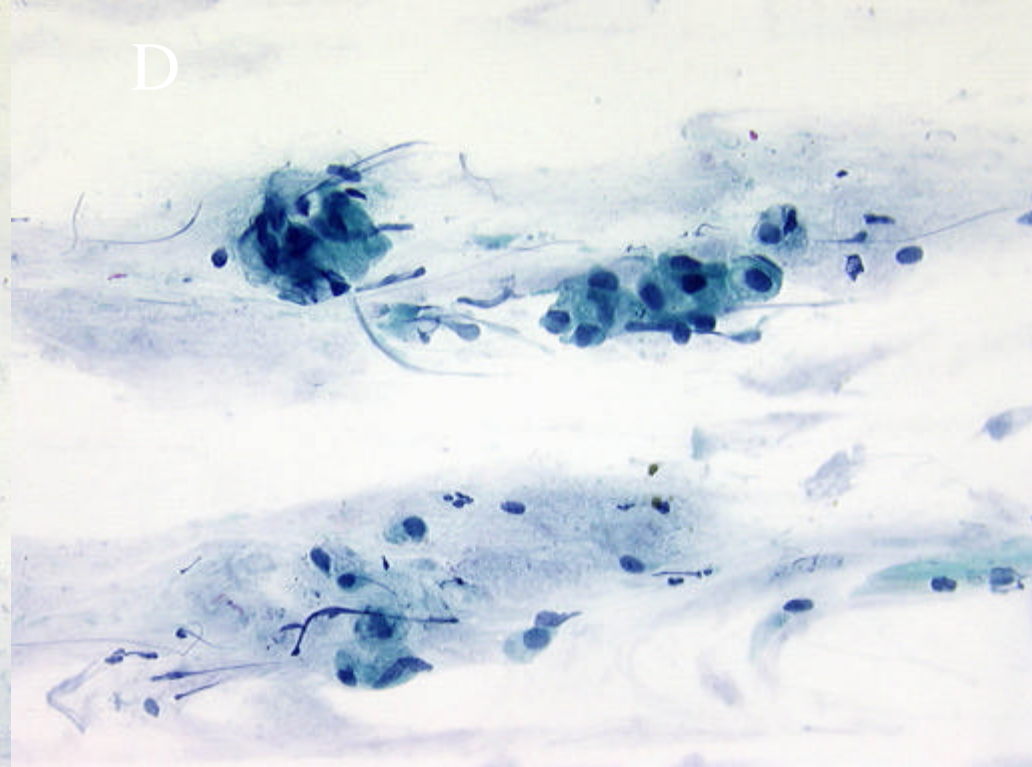
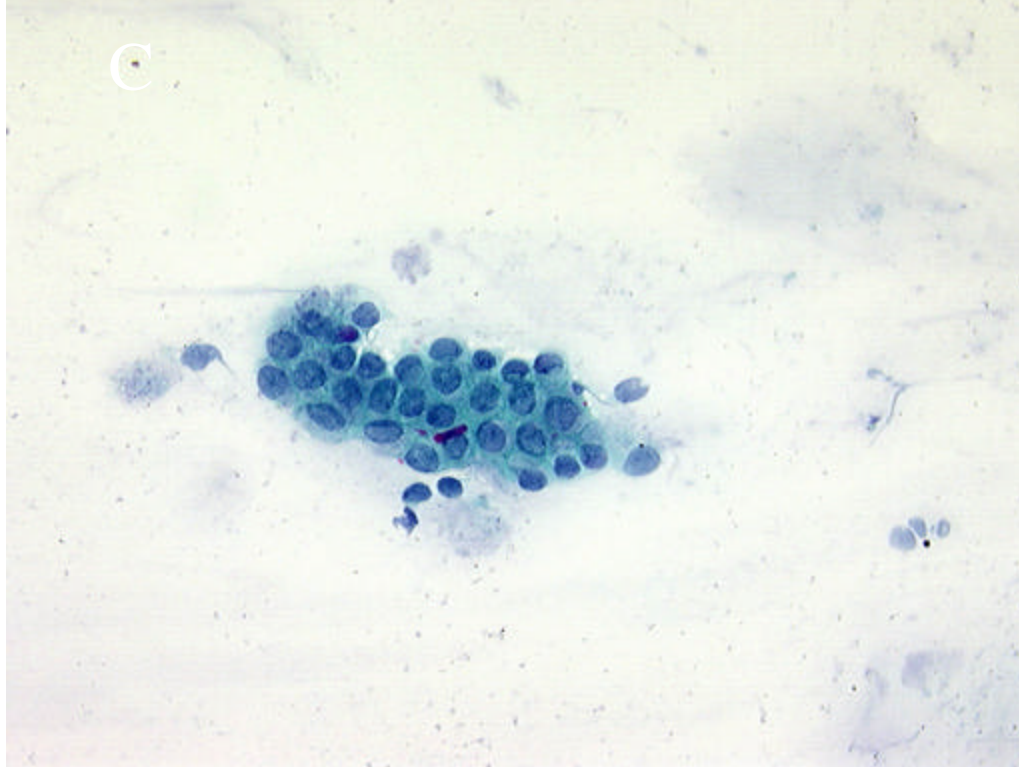
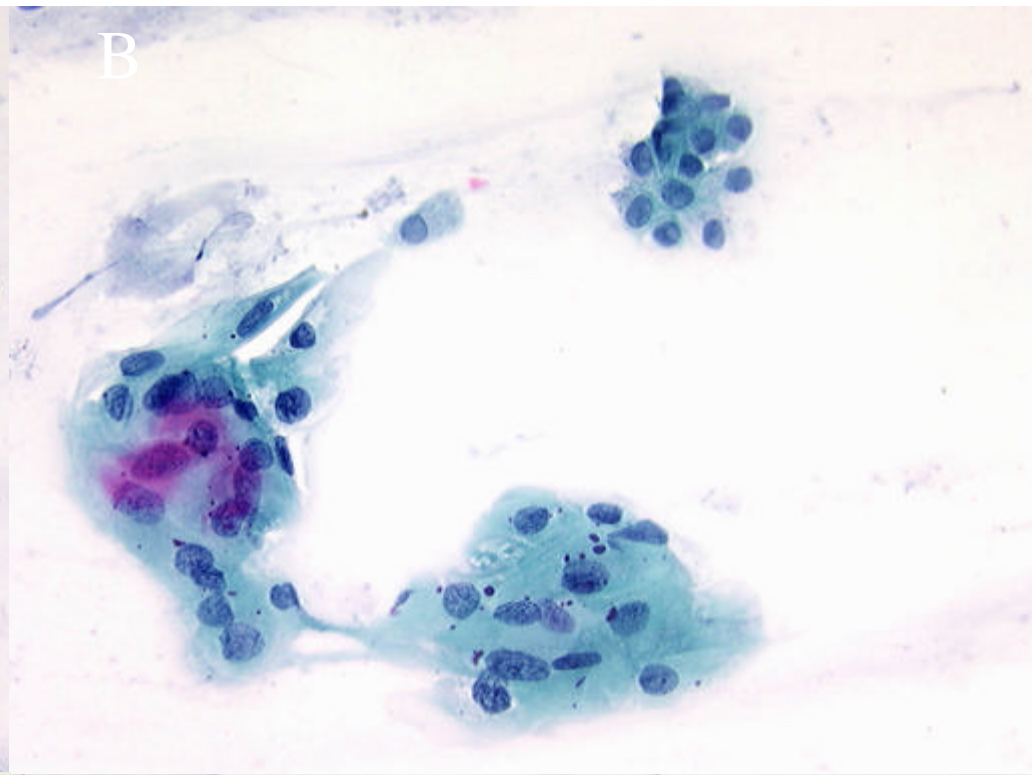
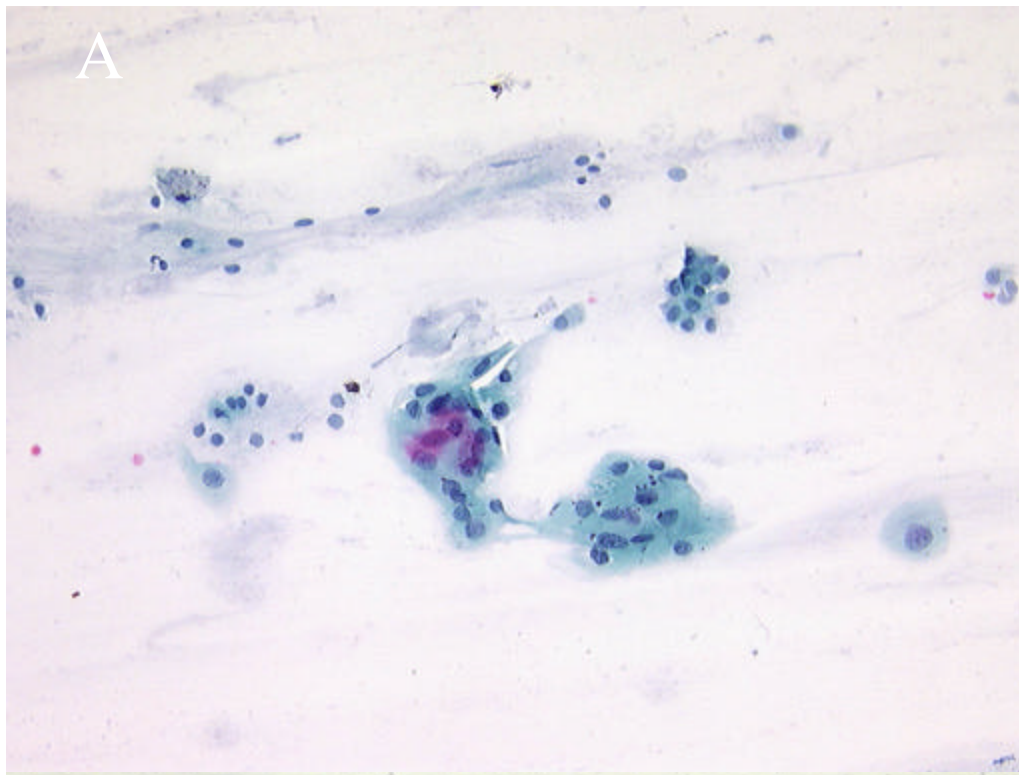
- Biggest issue - the “unsatisfactory” category. 35-40% called “unsatisfactory” slides “negative.”

PT Valente, “Government Mandated Cytology Proficiency Testing: Time for Reality Testing,” *Diag. Cytopathol.* 1994; 10: 105-106

Interobserver Variability

- 20 slides given to 5 experts
- Categories - WNL, RCC, ASCUS, LSIL, HSIL, SCC
- 35% - unanimous agreement
- 35% - 1 category disagreement
- 30% - greater disagreement

Young et al, Diag Cytopathol 11: 352-357, 1994



New York - Effect of Mandatory PT

Years	Number of States with Cervical Cancer Death Rates Lower than New York
1960-69	7
1970-79	16
1980-87	27
1989-81	32

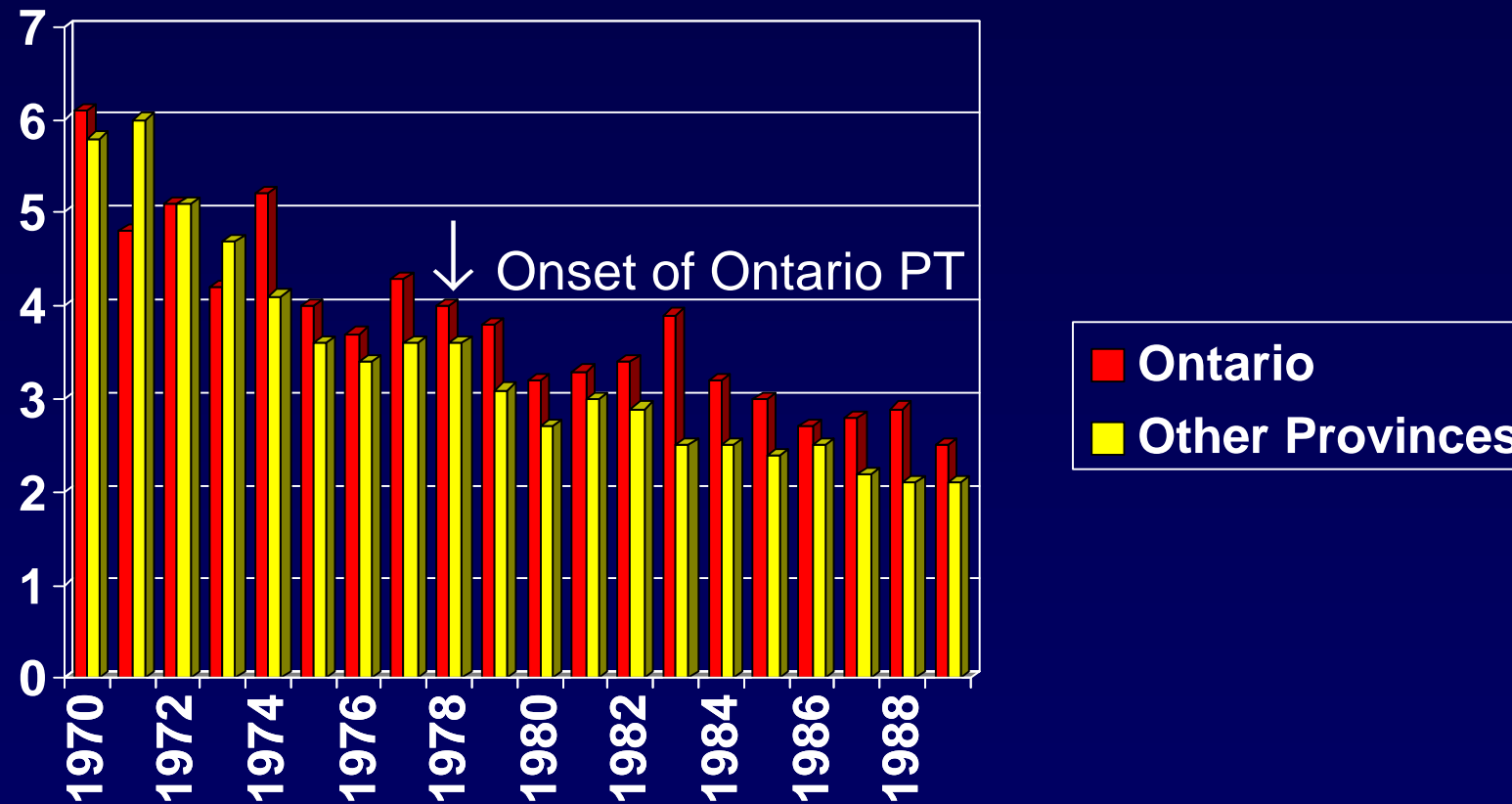
RM Austin, "Can Regulatory Proficiency Testing by the Cytobureaucracy Decrease Both False Negatives and Cervical Cancer Deaths," *Diag Cytopathol* 1994; 11: 109-112

Effect of Mandatory PT

- We note that the data used in this critique of the New York program shows a steadily decline in the cervical cancer death rate in Maryland - a state that also has a PT program.

RM Austin, "Can Regulatory Proficiency Testing by the Cytobureaucracy Decrease Both False Negatives and Cervical Cancer Deaths," *Diag Cytopathol* 1994; 11: 109-112

Cervical Cancer Death Rates



RM Austin, "Can Regulatory Proficiency Testing by the Cytobureaucracy Decrease Both False Negatives and Cervical Cancer Deaths," *Diag Cytopathol* 1994; 11: 109-112

PT vs. Work Performance

“Do proficiency test results correlate with the work performance of screeners who screen Papanicolaou smears,”

RA Keenlyside, *et. al.*, Am J Clin Pathol
1999;112:769-776

Work Performance vs. PT

- Rescreened Pap smears from 40,245 women who had been screened by 81 cytology screeners. All diagnoses placed into CLIA categories.
- Rescreening scores standardize to account for different distributions of abnormalities in PT and in the rescreened slides.
- Computerized and glass slide PT
- Correlation analysis

Categories

- A Unsatisfactory
- B Normal, benign or reactive
- C LSIL
- D HSIL and above
- E ASCUS/AGUS*

* Excluded from correlation analysis

Analysis

- Overall agreement 95.3%
- 98.3% of WNL originally diagnosed correctly
- 52% agreement for unsatisfactory
- 84% agreement on SIL

PT vs Rescreening

- Significantly more screeners scored less than 90 on the coputer PT than on glass slide PT.
- Pathologists scores lower than CT scores on both systems
- Rescreening vs glass slide $r=0.24$ ($p=0.016$)
- Rescreening vs computer $r=0.24$ ($p=0.016$)

PT vs. Work Performance

- The correlation between work correlation and PT results is weak. Correlation coefficient of 0.24 is very, very weak support for the PT concept.
- Study did not directly correlate computer and glass-slide PT results.

False Negative and False-Positive PT Results

- Passing level set between 80 and 90%.
- Only agreement or disagreement allowed.

GK Nagy and DN Collins, "False-Positive and False-Negative Proficiency Test Results in Cytology," *Acta Cytol* 1991; 35: 3-7

False Negative and False-Positive PT Results

Probability of Passing the Test

Competence Levels

Slides in Test Set	<u>Unacceptable</u>			<u>Acceptable</u>		
	75%	80%	85%	90%	95%	97.5 %
10	.24	.38	.54	.74	.91	.98
20	.09	.21	.4	.68	.92	.99
30	.04	.12	.32	.65	.94	.99
40	.02	.08	.26	.63	.95	1.00

False Negative and False-Positive PT Results

Probability of Passing one Test of Two Competence Levels

Slides in Test Set	Unacceptable			Acceptable		
	75%	80%	85%	90%	95%	97.5 %
10	.42	.85	.89	.93	.99	~1
20	.17	.45	.64	.90	.99	~1
30	.08	.23	.54	.88	~1	~1
40	.04	.15	.46	.84	~1	1.00

Challenges in Proficiency Testing

- Base false-negative rate of approximately 5%
- Lack of “gold standard,” leading to lack of interobserver reproducibility
- Failure to replicate “real” screening conditions
- Limited proof that PT is actually useful in reducing cervical cancer mortality

Limitations in Cytology PT

- Suboptimal performance conditions - tests diagnosis more than screening
- Suboptimal slide quality
- **Incomplete clinical information**
- Inability to use normal laboratory terminology/limits on diagnostic categories

Davey et al Acta Cytol 2000; 44:939-943

Limitations in Cytology PT

- PT using only a few slides tests diagnosis more than screening.
- The major source of false-negative screening results is cytotechnologist error - 78%* - not diagnosis

* SE Wang, MJ Ritchie and BF Atkinson, "Cervical Cytologic Smear False Negative Fraction Reduction in a Small Community Hospital," *Acta Cytol.* 1997; 41: 1690-1696

Conclusions - Advances in Cytology PT

1. PT can be carried out on a large scale.
2. PT can meet a political imperative.
3. Introduction of PT program probably drives some people out of the business who should not be there.
3. PT can be probably be carried out using digital image technology.

Conclusions - Advances in Cytology PT

5. Much better understanding of limitations:
 - Little correlation between PT results and performance
 - A single 10-slide PT result can penalize excellent practitioners
 - Stepwise systems not good at identifying poor performers
 - Much work remains to be done.

“We never claimed that Maryland’s, or any other government-mandated, cytology PT program would lead to an improvement in routine clinical practice.”

JM DeBoy, BR Jarboe, “A response to “Can cytology proficiency testing programs discriminate between competent and incompetent practitioners.” QRB 1991; 17 (206)