



cap



Good Laboratory Practices in Gynecologic Cytopathology

Joseph A. Tworek, M.D.

<http://www.cap.org>
v. 1

Background

- **CAP awarded cooperative agreement from CDC to**
 - Develop an inventory of current practices in gynecologic cytology laboratories
 - Attempt to standardize procedures for quality improvement
- **Changes since CLIA '88**
 - Proficiency testing
 - Liquid based methods with HPV testing
 - Computer assisted screening

Multi-Step Process

- **Survey of QA practices sent to every Laboratory enrolled gynecologic PT**
- **Formation of working groups to analyze data collection**
 - 5 pathologists, 1 cytotechnologist and 1 CAP staff
- **Posting additional queries on web site**
 - Open ended questions
 - Open to cytology community
- **Convene a consensus conference**
 - 100 attendees
 - Open to cytology community

Survey

- **Sent to 1,191 labs with 541 useable anonymous responses**
- **Topics**
 - **Diagnostic rates**
 - **Prospective and retrospective rescreening**
 - **Proficiency testing**
 - **Cytologic-histologic correlation**
 - **Concurrence of cytotechnologist and pathologist**
 - **HPV rates**
 - **Turn-around-time**
 - **General quality**

Strengths

- **541 useable responses**
- **Good Laboratory Practice Statements vetted by working groups, web site and consensus conference.**
- **Literature cited when available.**

Weaknesses

- **Sensitivity to regulatory environment**
- **Literature review not graded on strength of evidence**
- **Literature not always available**
- **Not a prospective study but a survey**
- **How does this apply to laboratories with low volumes of Pap tests.**



cap



Gynecologic Consensus Conference

Working Group 1: Monitoring Interpretive Rates,
Concordance of Interpretations, Turnaround Time

June 4, 2011

<http://www.cap.org>

v. 1

Working Group 1

- **Joseph Tworek, MD (Senior Author)**
- **Karen M. Clary, MD (Chair)**
- **R. Marshall Austin, MD**
- **Diane Davis Davey, MD**
- **Sonya Naryshkin, MD**
- **Chiara Sugrue, SCT**
- **Beth Anne Chmara, CT(ASCP), (CAP Staff)**

Statement: It is most useful to monitor interpretive rates for cytotechnologists individually and in comparison for the entire laboratory.

- **Basis: Respondent Data and Personal Observation**
- **From the survey, strong majorities are currently monitoring all TBS categories for cytotechnologists and also for the laboratory as a whole.**

Voting

Statement: It is most useful to monitor interpretive rates for cytotechnologists individually AND in comparison for the entire laboratory.

4. Do you agree with the consensus statement?

A. Yes *100%*

B. No *-*

Statement: It is currently unclear whether or not monitoring interpretive rates for individual pathologists beyond laboratory rates as a whole is useful.

- **Basis: Respondent Data and Personal Observation**
- **Only a third of respondent laboratories monitor TBS categories for pathologists, perhaps reflecting the more varied volume and more varied case mix of cases reviewed by individual pathologists.**

Voting

Statement: It is currently unclear whether or not monitoring interpretive rates for individual pathologists beyond laboratory rates as a whole is useful.

5. Is monitoring interpretive rates of individual pathologists useful to you?

A. Yes 86%

B. No 13%

C. Other 1%

6. Is this an area that should be explored?

C. Yes 90%

D. No 3%

E. Other 6%

Statement: Providing Feedback of interpretive rates is important.

- **To share individual interpretive rate data laboratory-wide, de-identify individuals**
- **To privately provide personal feedback, compare overall laboratory data in comparison to individual statistics**
- **File/retain such data with other QA documents**
- **58% of respondents believe it is helpful in a quality plan that cytotechnologists and pathologists have access to their personal interpretive rates in comparison with others in the laboratory**

Voting

Statement: Providing Feedback of interpretive rates is important.

- 9. Should individual interpretive statistics be provided to cytotechnologists and pathologists as feedback?**
- A. Yes, regularly 88%*
 - B. No, not at all 1%*
 - C. Only as a part of scheduled employee reviews 11%*

Statement: Actively monitoring rates at which a pathologist UPGRADES a cytotechnologist interpretation prior to sign out may be a useful quality metric.

Basis: Most laboratories monitor these metrics

Table 11. Monitoring change in diagnosis

	Frequency	Percent
The rates at which a pathologist upgrades a cytotechnologist's diagnosis at the time of initial sign out are actively monitored per cytotechnologist:		
Yes	376	73.3
No	137	26.7
The rates at which a pathologist downgrades a cytotechnologist's diagnosis at the time of initial sign out are actively monitored per cytotechnologist:		
Yes	312	62.5
No	187	37.5

Which upgrade rates are monitored for Cytotechnologists?

Table 12. Upgrade rates monitored for cytotechnologists (N=381)*

Cytotechnologist Diagnosis	Pathologist Diagnosis							
	ASC-US		ASC-H		LSIL		HSIL or greater	
	Freq.	Percent	Freq.	Percent	Freq.	Percent	Freq.	Percent
NILM	285	74.8	312	81.9	373	97.9	368	96.6
ASC-US	.	.	179	47.0	216	56.7	317	83.2
LSIL	.	.	169	44.4	.	.	251	65.9
ASC-H	217	57.0

* Multiple responses allowed.

- **Changes from NILM to SIL+ are most critical**
- **Upgrades from ASC-US to HSIL also useful**
- **Upgrades from LSIL to HSIL and ASC-H to HSIL not as frequent**

Voting

Statement: Actively monitor rates at which a pathologist UPGRADES cytotechnologist interpretations prior to sign out. Definition of upgrades should be determined by the laboratory.

Do you agree:

- A. Yes 79.66%**
- B. No 15.25%**
- C. Other 3.39%**
- D. Other 1.69%**

Adjudicating discrepancies and Potential limitations

- **Statement:** Show discrepancies of 2 degrees or more to a third person when possible before sign out.
 - Preferably third person should be blinded to initial result
 - NILM to HSIL, ASC-H, or Atypical glandular cells
 - This applies to both upgrades and downgrades
 - Survey: this is done by 61% of laboratories
- **Small laboratories with only one technologist or pathologist may find this difficult**
- **Some LIS systems may not be able to track cytotechnologist interpretations prior to sign-out**

Other suggestions for adjudication (continued)

- **78% of pathologists in online survey show downgrades of HSIL to another person**
- **54% show downgrades of atypical glandular cells and these lesions also problematic**
- **Consider impact on patient care, follow-up**

Voting

Statement: Laboratories should have policies about which categories of discrepancies should be reviewed by a third individual prior to sign out.

Do you agree:

- A. Yes 73.68%**
- B. No 22.81%**
- C. Other 3.51%**

Statement: Some categories of cases benefit from routine review by 2nd person even if CLIA does not require confirmation by a pathologist.

- **Endometrial cells in women \geq 40years, glandular cells post-hysterectomy (36-44% in online survey)**
- **Herpes: 81%**
- **Unsatisfactory cases: 59%**
- **Rationale: Impact on management, promote interobserver reproducibility, diagnostically difficult areas**

Voting

Statement: Laboratories should have policies as to which cases benefit from review by a second person (cytotechnologist or pathologist), even if not required by CLIA.

These may include:

- **Unsatisfactory**
- **Endometrial cells in women >40**
- **Glandular cells in women post hysterectomy**
- **Herpes**

Do you agree:

A. Yes 90.74%

B. No 9.26%



cap



Gynecologic Consensus Conference

Working Group 2: Prospective and Retrospective Review

June 4, 2011

<http://www.cap.org>

v. 1

Work Group 2

- **Jennifer Brainard MD, FCAP, Chair**
- **Michael Henry MD, FCAP, Senior Author**
- **George Birdsong MD, FCAP**
- **Tarik Elsheikh MD, FCAP**
- **Kalyani Naik MS, SCT(ASCP)**
- **Margaret Neal MD, FCAP**
- **David Andrew Hartley CT(ASCP)^{CM}, CAP
Cytotechnologist Specialist**

Working Group 2: Definitions

- **Prospective rescreen**
 - Review, prior to sign-out, by a second cytotechnologist of a subset of Pap tests interpreted as NILM in the first cytotechnologist review
 - ≠ Prescreen
- **Retrospective rescreen**
 - Review of NILM+ Pap tests that have been signed out – an example is NILM slides from the preceding 5 years in patients with current HSIL+

Statement: Prospective Rescreen

Justification: Survey, Literature and Expert Consensus

- 1. Maximizing the number of high risk cases increases the power of this QA measure.**
- 2. Labs should include all readily identifiable HR cases *in addition* to randomly selected cases.**
- 3. Multiple measures should be used to identify HR cases and to remove patients who no longer meet the criteria.**
- 4. If the information is available prior to sign out, positive hrHPV NILM cases from a HPV DNA Pap test should be prospectively rescreened.**

□ Voting: Prospective Rescreening

Statement or Question	Yes (%)	No (%)
Laboratories should make an effort to maximize the number of high risk cases in their prospective rescreens and multiple measures should be used to identify these patients.	98.61	1.39
Should <u>all</u> readily identifiable high risk cases be included in the prospective rescreen?	89.39	10.61
Should NILM Paps from patients with concurrent positive hrHPV results be rescreened prior to sign-out?	84.48	15.52

Statement: Retrospective Rescreen

Justification: Survey, Literature and Expert Consensus

- 1. Review of UNSAT Paps in addition to NILM Paps should be included in retrospective review**
- 2. Retrospective review based on surgical biopsy results when possible is suggested**
- 3. The monitoring of upgrade rates are very low for pathologists (37.3% for NILM to HSIL+)**

□ Voting: Retrospective Rescreening

Statement or Question	Yes (%)	No (%)
To maximize the power of this measure, should retrospective review based on surgical biopsy results, when possible, be performed?	87.14	12.86
Should pathologists be included when monitoring upgrade rates in a retrospective review?	86.15	13.85

Statement: Prospective and Retrospective Rescreen

- 1. Both CTs and pathologists should get feedback on upgrade/review diagnoses**
- 2. It is important to monitor ASCUS/ASCH upgrades from NILM for CTs, pathologists and the laboratory**
- 3. A major barrier to implementation of enhanced/additional quality measures is limited LIS functionality**

Voting

Question	Yes (%)	No (%)
For both prospective and retrospective reviews, should upgraded diagnoses from NILM to ASC-US/ASC-H be tracked?	63.77	36.23

Additional Voting Questions	Agree	Disagree	Uncertain
For RETROSPECTIVE review, upgraded diagnoses from NILM to ASC-US should be monitored.	39.22	50.98	9.80
For PROSPECTIVE review, upgraded diagnoses from NILM to ASC-US should be monitored.	60.78	31.37	5.88



cap



Gynecologic Consensus Conference

Work Group 3: PAP Proficiency Testing

June 4, 2011

<http://www.cap.org>

v. 1

Working Group 3

- **Joseph Tworek, MD (Senior Author)**
- **Lydia P. Howell, MD (Chair)**
- **Ritu Nayar, MD**
- **Sana O. Tabbara, MD**
- **Barbara Winkler, MD**
- **Lynnette Savaloja, SCT**
- **Nicole E. Thomas, MPH, CT(ASCP), (CAP Staff)**

Statement

- **For a first-time PT failure (CT and Path):**
 - **Re-enrollment for re-testing is a CLIA requirement and is sufficient.**
 - **No other remedial actions required, unless supported by other performance indicators.**

Justification

- Survey:
 - Re-enroll: 83% CT, 86% Path
 - Options for remedial action did not exceed 13% for CTs or 9% for Paths
 - Labs may recognize that a single failure is not a significant finding

Justification, con't

- Testing alters performance (anxiety?)
- Almost everyone passes eventually:
 - 99.6% passed after 3 tests (2006 CAP PT)
 - *Hughes J et al. Arch Pathol Lab Med 2009; 133:279-282.*
 - *Moriarty A et al. Arch Pathol Lab Med 2009; 133: 1757-1760.*

Voting

- **For a first-time PT failure (CT and Path):**
 - **Re-enrollment for re-testing is sufficient.**
 - **No other interventions are required, unless supported by other performance indicators.**
- **Do you agree with this statement?**
 - **A – Yes 94.44%**
 - **B – No 5.56%**

Statement

- **Remedial action policy should not be applied for a passed but non-perfect test (ie, score <100%), even for multiple non-perfect test scores.**
- **Justification**
 - **Survey:**
 - *81.4% report no policy to do so*
 - *Even in 2 consecutive non-perfect exams, only 5.2% take remedial action*
 - **No literature to show that a non-perfect test is a significant finding.**

Voting

- Remedial action policy should not be applied for a passed but non-perfect test (ie, score $<100\%$), even for multiple non-perfect test scores.
- Do you agree with this statement?
 - A – Yes 93.75%
 - B – No 6.25%

Statement

- **Monitoring of incorrect slide diagnoses on passed PT tests:**
 - **Discouraged from inclusion in lab PT policy**
 - **No interventions for this test finding are necessary**

- **Justification**
 - **Survey findings**
 - *This type of monitoring only done by 26% of labs.*
 - **Literature**
 - *Recognition that test-taking alters performance (as in statement #3).*
 - *No literature to show that incorrect results are a significant finding.*

Voting

- **Monitoring of incorrect slide diagnoses on passed PT tests:**
 - **Discouraged from inclusion in lab PT policy**
 - **No interventions for this test finding are necessary**
- **Do you agree with this statement?**
 - **A – Yes 76.56%**
 - **B – No 23.44%**



cap



Gynecologic Consensus Conference

Working Group 4: Cytologic-Histologic Correlations-
Summary of Consensus Conference
June 4, 2011

Working Group 4 Members

- **Barbara A Crothers, DO FCAP, Chair**
- **Bruce A Jones, MD FCAP, Senior Author**
- **Leigh Ann Cahill, CT (ASCP)^{CMIA}**
- **Ann T Moriarty, MD FCAP**
- **Dina R Mody, MD FCAP**
- **William D Tench, MD FCAP**
- **Rhona J Souers, MS, CAP Biostatistician**

Statement

Laboratories should define their cytologic-histologic correlation (CHC) process to address quality issues and account for population variables. CHC benefits from a multilayered approach, employing several processes and metric measurements.

Evidence

- **CHC cited as most valuable QA measure on survey**
 - *4.2 out of 5 points*
- **94% actively monitor the correlation between Pap test and biopsy results (Survey Statistics)**

Multilayered Approach to CHC

- **Multilayered, laboratory-directed approach “drills down” in potential problem areas and can be tailored to laboratory size, issues and practice**
- **Additional QA monitors may be continuous or interval efforts**
 - **Interval efforts may target specific pairs for a pre-defined period (i.e., quarterly) to acquire a “snapshot” of laboratory performance for that indicator**
 - **Continuous efforts may be desirable for laboratories with high personnel turn-over, disruptive environments, or mitigating variables outside of the laboratory’s control**
- **Justification: Professional opinion, survey, literature**

Dual Role of CHC

- **“Real Time” correlation- review of slides prior to issue of biopsy report**
 - Provides critical information for patient follow-up
 - Resolves/confirms discrepancies
 - Timely report to healthcare providers
- **“Retrospective” correlation- review of slides after issuance of both reports**
 - Monitor performance and processes of cytology and biopsy for laboratory quality improvement

Voting

- **88% support lab-defined CHC**
- **97% state that a multi-layered approach to CHC, suited to laboratory size and staffing, optimizes opportunities for quality improvement.**

CHC Statement

The correlation interval between the Pap test and the biopsy should preferably be within 3-4 months, but no greater than 6 months.

Voting

The correlation interval between the Pap test and the biopsy should preferably be within 3-4 months, but no greater than 6 months.

Consensus conference vote:

- **89% agree as stated**

CHC Statement

Standardization of metrics and CHC process is desirable. The PPV of a positive Pap test is the preferred standard CHC metric.

- Allows for interlaboratory comparison

Voting Summary

Consensus conference vote:

- **94% in favor of standardized metrics and CHC process**
- **70% agree that PPV is the best metric**
- **66% agree laboratories should use PPV to develop other QA metrics**

Recommendations

Monitor:

- **Total number of CHC pairs**
- **Number of positive correlations (“true positive,” as defined prior to review)**
- **Number of negative correlations (“false positive,” as defined prior to review)**
- **Calculate Positive Predictive Value (PPV) of a positive Pap test**
- **Tabulate statistics at least annually**

Cytologic-Histologic Calculation

Screening result	Diagnosis as determined by biopsy		Metric
	<i>Positive</i>	<i>Negative</i>	
Positive Pap test	True positive (TP)	False positive (FP)	TP / TP+FP = Positive predictive value
Negative Pap test	False negative (FN)	<i>True negative (TN)</i>	TN / TN+FN = Negative predictive value
Metric	TP / TP+FN = Sensitivity	TN / TN+FP = Specificity	



cap



Gynecologic Consensus Conference

Working Group 5: Monitoring of HPV Rates

June 4, 2011

<http://www.cap.org>

v. 1

Work Group 5

- **Christine Booth MD, FCAP, Chair**
- **Michael Henry MD, FCAP, Senior Author**
- **Carol Filomena MD, FCAP**
- **Marilee Means PhD, SCT(ASCP)**
- **Patricia Wasserman MD, FCAP**
- **Christine Bashleben, MT(ASCP) CAP staff**

Statements: HPV Testing

Justification: Survey, Literature and Expert Consensus

- 1. Laboratories should only offer high-risk HPV testing for GYN specimens.**
- 2. Laboratories should encourage clinicians to consider the latest consensus guidelines in ordering high-risk HPV tests on GYN specimens.**
- 3. Laboratories should be cautious in using HPV test results to change or influence cytologic interpretations.**

Questions: HPV Testing

Statement	Yes (%)	No (%)	Unsure(%)
HR-HPV tests should be ordered by the laboratory to be used as a diagnostic test to aid in morphologic dilemmas and resolve diagnostic discrepancies.	15.2	75.8	9.1
HR-HPV results should aid in down- or upgrading of Pap test interpretations when available prior to sign-out.	8.1	83.9	8.1
It is not appropriate to offer LR-HPV testing on Pap tests.	80.6	17.1	N/A

Statements: Monitoring HPV Rates

Justification: Survey, Literature and Expert Consensus

- 1. While there is significant variability in interinstitutional HPV-positive rates in ASC-US Pap tests, monitoring the HPV-positive rate in ASC-US Pap tests is a valuable broad measure of quality.**
- 2. Performance beyond 2 SD's of the mean should prompt reassessment of diagnostic criteria used in the evaluation of Pap tests and/or investigation of the prevalence of HPV positivity in the population from which the Pap tests are obtained.**

(Tworek et al, *Arch Pathol Lab Med.* 2007;131:1525–1531)

Statements: Monitoring HPV Rates

Justification: Survey, Literature and Expert Consensus

- 3. Monitoring the HPV-positive rate in other diagnostic categories such as LSIL and the comparison of these HR-HPV rates to published benchmarks is also a valuable broad measure of quality for a laboratory and possibly for individuals.**
- 4. When possible, individual pathologist ASC-US/HR-HPV results should be compared to ASC-US/SIL ratios to determine potential trends in over- and under-diagnosis.**

Questions: Monitoring HPV Rates

Statement	Yes (%)	No (%)	Unsure(%)
If possible, ASC-US/HR-HPV results should be compared to ASC-US/SIL ratios per pathologist as a general quality monitor.	58.5	18.9	22.6
ASC-US reflex HR-HPV results should be monitored to determine potential trends in accuracy of diagnoses.	71.9	18.8	9.4
HR-HPV DNA results for other diagnostic categories should be monitored to determine potential trends in accuracy of diagnoses.	32.3	50.8	15.4

Questions

Additional Voting Question	Yes (%)	No (%)	Unsure (%)
Is it appropriate for a lab to order a HR-HPV test as a diagnostic test independent of the clinician?	6	84	10



cap



Gynecologic Consensus Conference

Working Group 3, Topic 6: General Quality

June 4, 2011

<http://www.cap.org>

v. 1

Available methods for monitoring quality data

- Historical data and national benchmarks
 - Useful for smaller labs
 - Historical data will identify trends within lab
 - Published benchmarks may identify lab drift
 - *National benchmarks not always available*

Statement: Selected metrics should be monitored individually, as well as globally for the laboratory.

- Justification: Survey and website
- Monitoring laboratory-wide data against national benchmarks may provide a baseline to identify and stratify lab performance
 - Not valuable for labs with small numbers of primary screeners
 - Taken in context with other factors (eg, high-risk population)
- Comparing individual data to laboratory-wide data may help identify outliers
 - Retain with other QA documents

Voting: Selected metrics should be monitored individually, as well as globally for the laboratory.

- A.** Agree with entire statement 95.92%
- B.** Only individual quality data should be monitored; no global monitoring. 0%
- C.** Only global laboratory monitoring; no individual monitoring. 0%
- D.** Disagree with entire statement (ie, quality data should not be monitored at all). 2.04%

Statement: Results of quality metrics should be shared with individual CTs and pathologists.

- Justification: Survey
- Quality metrics should be shared with each CT and pathologist
 - From survey, 59% and 81% of labs facilitate comparison of CT to other CTs and to laboratory data respectively
 - 48% and 60% of labs facilitate comparison of pathologists to other pathologists and to laboratory data respectively
 - Table 3 page 55
- Lab mean data and/or individual data could be shared openly or privately, identified or de-identified at the discretion of the lab

Voting: Monitoring of selected metrics for individuals should include both CTs and Pathologists

- A. Agree with entire statement. 92.9%
- B. Only cytotechnologist quality data should be monitored. 3.57%
- C. Only pathologist quality data should be monitored. 1.79%
- D. Disagree with entire statement (ie, individual quality data should not be monitored at all). 1.79%

Voting: Results of quality metrics should be shared with individual CTs and pathologists.

- A.** Agree with entire statement. 98.39%
- B.** Quality metrics should only be shared with CTs. 1.61%
- C.** Quality metrics should only be shared with Paths. 0%
- D.** Disagree with the entire statement (ie, quality metrics should not be shared at all). 0%

Statement: Reviewing selected cases for educational purposes is a useful quality tool.

- Justification: Survey
- Multi-head review of difficult cases ranked second most useful quality metric
- 60% of labs conduct in-house review
 - Share interesting cases
 - Review of educational program slides
 - Hone diagnostic criteria
 - Review cases identified from QA
 - Review laboratory generated study material

Voting: Reviewing selected cases for educational purposes is a useful quality tool.

- A. Strongly Agree 86.4%
- B. Agree 13.6%
- C. Disagree 0%
- D. Strong disagree 0%