Interpretive Diagnostic Error Reduction in Surgical Pathology and Cytology

Guideline from the College of American Pathologists (CAP) Pathology and Laboratory Quality Center and the Association of Directors of Anatomic and Surgical Pathology (ADASP)

Raouf E Nakhleh, MD
Mayo Clinic Florida

Disclosure

- None

Agenda

- Factors contributing to accurate diagnosis
- Guideline for Error Reduction
  - 5 guideline statements
  - Limitations
  - Conclusions

Errors in Surgical Pathology

- Pre-analytic
  - Misidentification: 13-38%
  - Defective specimens: 4-14%
- Analytic
  - Diagnostic misinterpretation: 15-29%
- Post-analytic
  - Defective report: 29-58%

Source: Am J Clin Pathol 2008;130:238-246
Arch Pathol Lab Med 2014;138:602-612

The Doctors Company

- 272 surgical pathology claims (1998-2003)
- 166 (61%) false negative
- 73 (27%) false positive
- 10 (4%) frozen section

Analytic

- 22 (8%) operational
  - 13 mix-ups
  - 3 floaters
  - 2 mislabeled biopsy site
  - One transcription error, "no" omitted before malignant cells

Pre-analytic

Analytic

Post-analytic

Source: Am J Surg Pathol 2004;28:1092-1095

Factors Contributing to Interpretive Diagnostic Accuracy

1. The pathologist's knowledge and experience,
2. The use of standardized diagnostic criteria and terminology,
3. Clinical correlation,
4. The use of confirmatory diagnostic ancillary studies,
5. Additional examination of cases in the form of secondary case reviews
   - Many studies demonstrate utility
   - No effort to formalize this strategy to reduce error or discrepancies

Introduction

- The CAP and ADASP convened an expert panel to systematically review published documents and develop an evidence-based guideline to help define the role of case reviews in surgical pathology and cytology.
- The panel focused on the contribution of case reviews to error detection and prevention of interpretive diagnostic errors

7

Introduction

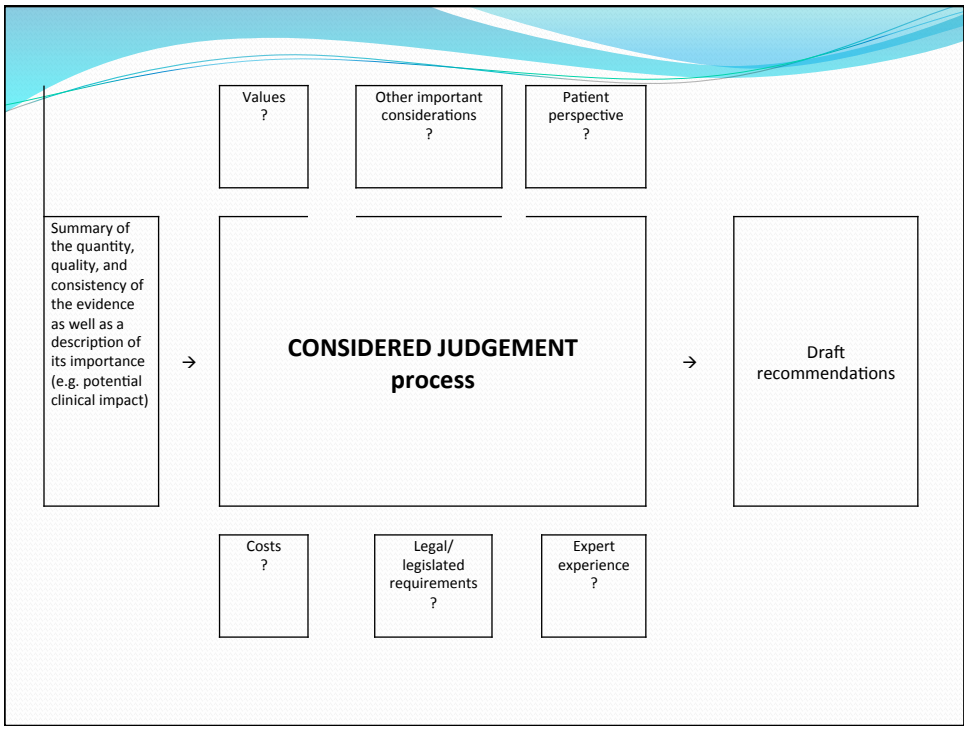
- Closely followed Institute of Medicine *Clinical Practice Guidelines We Can Trust* standards for guideline development

1. **Establish transparency**
2. **Manage conflicts of interest**
3. **Establish a multi-disciplinary panel**
4. **Perform systematic review**
5. **Rate strength of recommendations**
6. **Articulate the recommendations**
7. **Include external review**

Systematic Evidence Review

- Identify Key Questions
- Literature search
- Data extraction
- Develop proposed recommendations
- Open comment period
- Considered judgment process
  - Consider risks and benefits, cost, regulatory requirements, preferences, etc.

9



Defining the types of recommendations and its strength

- **Strong Recommendation**
  - High/Intermediate quality evidence
- **Recommendation**
  - Intermediate/Low quality of evidence
- **Expert Opinion**
  - Low/Insufficient evidence and expert panel uses formal consensus process to reach recommendation
- **No recommendation**
  - Insufficient evidence, confidence, or agreement to provide a recommendation

Adapted from AHRQ Methods Guide for Comparative Effectiveness Reviews 2011

Expert Panel

Co-chairs
Raouf Nakhleh, MD, **CAP**
Vania Nosé, MD, PhD, **ADASP**

Expert Panel
Tamera J. Lillemoe, MD
Douglas C. McCrory, MD, MPH
Frederick A. Meier, MD
Christopher N. Otis, MD
Scott R. Owens, MD
Stephen S. Raab, MD
Andrew A. Renshaw, MD
Roderick R. Turner, MD

Advisory Panel
Timothy C. Allen, MD, JD
Lawrence J. Burgart, MD
Patrick L. Fitzgibbons, MD
Jeffrey L. Myers, MD

Staff (CAP)
Lisa Fatheree, SCT(ASCP), Director
Carol Colasacco, MLIS, SCT(ASCP), Medical Librarian
Christina Ventura, MLS(ASCP), Guideline Development Manager

Purpose

- **The purpose of this guideline is to systematically examine the literature concerning second reviews of cases with the goal of establishing procedures that optimize the use of these additional case reviews in order to reduce interpretive errors or discrepancies.**

Key Questions

1. **Does targeted review (either done at analytic or post-analytic phase) of surgical pathology or cytopathology cases (slides and/or reports) reduce the error rate (often measured as amended reports) or increase the rate of interpretive error detection compared to no review, random review, or usual review procedures?**
2. **What methods of selecting cases for review have been shown to increase/decrease the rate of interpretive error detection compared to no review, random review, or usual review procedures?**

Inclusion/Exclusion Criteria

Inclusion	Exclusion
Surgical pathology or cytopathology studies	Clinical pathology studies
Original research addressing targeted review	Studies focused on pre-analytic specimen processes
English language articles	Non-English studies
All study types were initially included	Animal studies

Literature Review

- 20 year literature search
- Included surgical pathology and cytopathology
- 828 articles included for abstract review
- 299 articles included for full text review
- 148 articles included for data extraction
- Multiple conference calls (2 years)
- Face to face meeting in October 2013
  - Review the literature
  - Formulate recommendations

Literature Review

- 37 Multi-organ studies
  - 30 Surgical pathology
  - 3 Cytology
  - 4 Both
- 111 Single-organ studies
  - Prostate 18
  - Melanoma/skin 11
  - GU 7
  - Liver & GI 7
  - Non-cervical GYN 7
  - Soft tissue and bone 7
  - ENT, lymphoma, neuropath, breast
- Cytology
  - Thyroid8
  - GYN 3
  - Pancreas/biliary 2, lung 1, effusions 1, urine 1
- Both
  - GYN 2, anal 1, Lung 1

Public Comment

- Public comment period December 2, 2013 – January 21, 2014
- 82 respondents
- 295 comments
- Agreement
  - Recommendation 1: 87% agree, 13% disagree
  - Recommendation 2: 92% agree, 8% disagree
  - Recommendation 3: 92% agree, 8% disagree
  - Recommendation 4: 87% agree, 13% disagree
  - Recommendation 5: 90% agree, 10% disagree

Guideline Statement 1

1. Anatomic pathologists should develop procedures for review of pathology cases in order to detect disagreements and potential interpretive errors and to improve patient care.

Guideline Statement 1

- Rationale:
  - All studies show review of cases detect errors
  - Error rates that may affect patient care were variable but significant
  - Should be tailored to the needs of the individual laboratory
  - Ideally case reviews can enhance teamwork and reduce errors

Guideline Statement 1

- Strength of Recommendation: Recommendation
- Quality of Evidence: Low
  - The evidence was inadequate to demonstrate a direct impact on patient safety because few studies reported the clinical impact on patient outcomes that resulted from interpretive errors.
  - The overall quality of evidence was low, but due to consistent findings of a large number of studies of clinically important major discrepancy rates, and the significant impact that a diagnostic error may be expected to have on an affected individual, the panel graded this guideline statement as a "recommendation."

Guideline Statement 1 – Summary of Studies

Study type	Discrepancy rates (%)		Major Discrepancy rates (%)	
	No. of studies	Median (25^{th}-75^{th} percentile)	No. of studies	Median (25^{th} – 75^{th} percentile)
All studies	116	18.3 (7.5-34.5)	78	5.9 (2.1-10.5)
Surgical pathology	84	18.3 (7.5-37.4)	63	6.3 (1.9-10.6)
Cytology	19	24.8 (17.4-38.8)	11	4.3 (2.8 – 7.5)
Both	13	9.1 (6.7 – 15.8)	11	5.9 (3.3 – 8.7)
Multi-organ	43	9.1 (3.8-18.7)	42	3.9 (1.1-7.4)
Single-organ*	73	25.2 (14.0-43.7)	36	8.0 (3.7-15.8)
Internal**	35	10.9 (3.8 – 17.6)	22	1.2 (0.30-3.1)
External	79	23.0 (10.6-40.2)	56	7.4 (4.6-14.7)

*Single-organ refers to studies that focus on one organ or organ system; multi-organ refers to studies that are not limited with regard to organs studied.
**Internal refers to reviews of pathology reports within a single institution; external refers to reviews of cases given a diagnosis at a different institution.o

Guideline Statement 2

2. Anatomic pathologists should perform case reviews in a timely manner to have a positive impact on patient care.

Guideline Statement 2

- Rationale:
  - Reviews should be performed in a timely manner to ensure appropriate treatment decisions and patient care
  - Ideally prospective reviews, before case sign-out reduces rework
  - Retrospective reviews may also be performed, when prospective reviews are not possible due to various lab limitations and constraints, but should occur in a timely manner.
  - Retrospective review examples:
    - clinical correlation conferences
    - correlating cytology/biopsy cases with excision specimens,
    - should not change

Guideline Statement 2

- Strength of Recommendation: Recommendation
- Quality of Evidence: Low

Guideline Statement 2

- The literature review found four moderate-quality comparative studies that show prospective reviews (before sign-out) compared with retrospective review (after sign-out) can reduce disagreement/major disagreement rates and amended report rates
- The evidence was inadequate to demonstrate a direct impact on patient safety because few studies reported patient outcomes that resulted from interpretive errors.
- The quality of evidence is *low* but due to consistent findings in these 4 studies and no contradictory studies, the panel graded this guideline statement as a "recommendation."

Prospective vs. Retrospective Review

Studies	Setting	Comparison	Prospective Rate	Retrospective Rate
Renshaw and Gould, 2006	Single Institution	Subgroup cohort	D 4.8% A 0.0%	7.2% 0.5%
Novis, 2005	Single Institution	Historical cohort	A 0.6%	1.3%
Lind et al, 1995	Single Institution	Historical cohort	D 14.1% SD 1.2%	13.0% 1.7%
Owens et al, 2010	Single Institution	Historical cohort	D 2.3% SD 0.0%	3.4% 0.2%
Nakhleh et al, 1998	Multiple Institutions	Review method	A 0.12%	0.16%

Abbreviations: A, amended reports; D, discordance; SD, significant discordance

Guideline Statement 3

3. Anatomic pathologists should have documented case review procedures that are relevant to their practice setting.

Guideline Statement 3

- Rationale:
  - Many review methods describe with variable results
  - May affect turnaround time, increase workload, and add expense
  - The ideal method may depend on the practice setting
  - Tailor to maximize error detection while minimizing negative impacts
  - Methods to consider include: Targeted review, general review, percentage of cases reviewed, blinded review, review of cases with known high rates of missed lesions and others
  - The laboratory medical director is responsible for determining the policy

Guideline Statement 3

- Strength of Recommendation: Expert Consensus Opinion
- Quality of Evidence: Very Low
  - The quality of evidence was *low* to support using case review procedures compared to no case review procedures and to support targeted reviews versus random case review procedures; however, the evidence was *very low* with regard to distinction between different methods of review.
  - The overall quality of evidence was *very low* leading the panel to rate this guideline statement with the strength of recommendation of “expert consensus opinion.”

Guideline Statement 3

- Review Considerations
  - The reviewing pathologist should independently formulate opinions without influence from others
  - The reviewing pathologist ideally should have sufficient knowledge in the material they are reviewing
  - Case reviews performed prior to sign-out could be used to build collaborative teamwork and are excellent opportunities for pathologists to learn and improve their skills
  - Targeted review of selected organs or diseases leads to detection of more errors compared to review of cases randomly

Random vs. Focused Review

(Raab et al)

- 5% random review vs. focused review
- 5% random review detected 2.6% error (195/7444 cases)
- Focused review detected 13.2% error (50/380 cases)
- *p* value<.001
- Major error rates: Random 27(0.36%) vs. Focused 12 (3.2%)

Source: Am J Clin Pathol 2008;130:905-912

Selection of Material to Review

(Renshaw and Gould)

- In this study different strategies and different combinations were considered
- Data that was considered from the institution:
  - Tissue with highest amended rates: Breast 4.4%, endocrine 4%, GYN 1.8%, cytology 1.3%
  - Specimen types with highest amended rates: Breast core bx 4.0%, Endometrial curettings 2.1%
  - Diagnoses with highest amended rates: non-dx 5%, atypical/ suspicious 2.2%

Source: Am J Clin Pathol 2006;126:736-7.39

Selection of Material to Review

(Renshaw and Gould)

- Different combinations were used to determine types of review
  - Review of nondiagnostic and atypical /suspicious resulted in review of 4% of cases and detect 14% of amended reports
  - Reviewing all breast, GYN, non-GYN cytology and endocrine material resulted in review of 26.9% of cases and detected 88% of amended reports.

Limitations

- Situations where reviews may not be easy or convenient
- Solo Practice and Small group (2-3)
  - Document all outside reviews
  - Document conference cases
- Complete sub-specialization sign-out
  - Document clinico-pathologic conference cases
  - Peer review within the group
  - Share cases across institutions

Guideline Statement 4

4. Anatomic pathologists should continuously monitor and document the results of case review.

Guideline Statement 4

- Rationale:
  - Once established, the process should be monitored, ensuring that the program is functioning as intended and that all anatomic pathologists are compliant.
  - Methods of monitoring include:
    - overall rates of case review before sign-out
    - monitoring amended/revised report rates
    - minor/major discrepancies
    - others
  - Information should be used to assess
    - Local variations
    - Problematic case types with poor agreement

Guideline Statement 4

- Strength of Recommendation: Expert Consensus Opinion
- Quality of Evidence: Very Low

Guideline Statement 4

- The quality of evidence based on agreement studies was *low* for the finding that for several defined diagnoses and/or organ systems interobserver agreement is poor.
- In the panel's literature review there were no studies that directly related continuous monitoring to diagnostic agreement or improvement.
- The quality of evidence was *very low* leading the panel to rate this guideline statement with the strength of recommendation of "expert consensus opinion."

Guideline Statement 4

- Methods of documentation:
  - Documentation of review policy in QA plan
  - Documentation of actual review of cases
    - Body of the report
    - Separate intra-departmental consultation log
    - Consensus conference log
  - Documentation of quality assessment
    - Rate of case reviews
    - Adherence to review policy (diagnosis or organ policy)
    - Amended report rate
    - Periodic assessment of errors or disagreements.

Guideline Statement 5

5. If pathology case reviews show poor agreement within a defined area, anatomic pathologists should take steps to improve agreement.

Guideline Statement 5

- Rationale
  - some diagnoses have inherently higher inter-observer variation, and these differences in achieving diagnostic precision should be acknowledged
  - pathology diagnoses are dynamic and terminology changes, this may lead to the appearance of variation
  - When inter-observer variation is observed:
    - Investigate the cause
    - Identify possible outliers
    - Take steps to improve
      - Consensus conference within department
      - Use calibration slide sets
      - Achieve departmental consensus of the solution

Guideline Statement 5

- Strength of Recommendation: Expert Consensus Opinion
- Quality of Evidence: Not assessed

Guideline Statement 5

- The quality of evidence was *low* regarding the best methods to improve agreement in areas for which agreement is poor. It is likely that best approaches may differ based on features of disease, individual practice patterns and available ancillary diagnostic tests.
- In the panel's literature review there were no studies that directly related continuous monitoring to diagnostic agreement or improvement.
- The quality of evidence was not assessed leading the panel to rate this guideline statement with the strength of recommendation of "expert consensus opinion."

Examples of Studies Addressing Diagnostic Agreement

Author	Organ	Disease	Decision	Kappa
Kerkhof et al., 2007	Esophagus	Barrett's Esophagus	3 cat (ND, IND/LGD HGD/AC)	0.25-0.27
Zaino et al., 2006	Uterus	Atypical endometrial hyperplasia	Atypical hyperplasia vs. others	0.4 (0.34-0.43)
Oyama et al., 2005	Prostate	Adenocarcinoma	Gleason grade	0.49
Davidov et al., 2010	Thyroid	Malignant	Yes/no	0.55
Rakovitch et al., 2004	Breast	DCIS	Nuclear grade Margin status Tumor size	0.7 0.74 0.87

Abbreviations: AC adenocarcnoma; cat, category; DCIS, ductal carcinoma in situ; HGD, high grade displasia;IND, indefinite for displasia; LGD, low grade displasia; ND, no displasia

Limitations of Case Reviews and Rates of Disagreement or Error

- Data should not be used to compare laboratories because:
  - Sources of error may differ
  - Definition of error may differ
  - Clinical significant errors may differ
  - Detection method may differ
  - Review method sensitivity may differ
  - Expected range of performance not well defined

In order to compare quality between groups: we need to:

- Identify and use optimal method of review
- Measure sensitivity of review process
- Standardize criteria for review method,
- Standardize definition of error
- Define expected ranges of discrepancy and error
- Define methods to verify poor performance

Conclusions: Targeted secondary case reviews

- Successfully detect and reduce errors
- Lower error rates vs. no review
- Measure of quality within the group
- Groups that fail to detect discrepancy or error (<1/1000) may not be sensitive enough

Reference to guideline

- *Arch Pathol Lab Med. doi:* 10.5858/arpa.2014-0511-SA

- *Thank you!*
- *Questions?*